Carnegie Mellon University
MELLON COLLEGE OF SCIENCE

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN THE FIELD OF PHYSICS

TITLE: "Formal Topics in Inflation."

PRESENTED BY: Tereza Vardanyan

ACCEPTED BY THE DEPARTMENT OF PHYSICS

Hael Collins	7/25/17
HAEL COLLINS, CHAIR PROFESSOR	DATE
Scott Dodelson	7/27/17
SCOTT DODELSON, DEPT HEAD	DATE

APPROVED BY THE COLLEGE COUNCIL

Rebecca Doerge	8/04/17
REBECCA DOERGE, DEAN	DATE

Three Formal Topics in Inflation

by

Tereza Vardanyan

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Richard Holman

July 26, 2017

Abstract

Inflation can explain why our Universe is so flat, as well as homogeneous and isotropic on large scales, and it can give rise to the inhomogeneities observed in the CMB. For inflation to accomplish all of this, only a limited number of assumptions about the potential of the inflaton field are necessary. In order to have a complete picture about the dynamics and the initial state of the Universe, we need to study correlation functions of the CMB inhomogeneities. The basic inflationary models predict that all of the higher-order correlation functions should be very much suppressed. Because of the limitations of the modern observational tools, only the power-spectrum of the scalar fluctuations has been detected so far.

The first two topics of this thesis (chapter 2 and 3) concentrate on certain properties of two- and three-point correlation functions of the scalar fluctuations that arise in the basic single-field, slow-roll inflationary models. Both of these works are set up to include a general initial state that the system might have had, and not just the traditional Bunch-Davies vacuum. Being able to treat general states is important, because even though current observations are consistent with the fluctuations being initially in the Bunch-Davies state, it is possible that, having more precision, future observations might indicate the presence of small deviations from this state.

The third topic (chapter 4) investigated here is more formal in the sense that we treat not the real curvature perturbations of the inflationary theories, but a toy model of the massless scalar field with quartic interactions in the pure de Sitter background. We construct an effective theory of the long-wavelength part of the field which allows us to study the late-time behavior of this system. We show that at leading order this effective theory matches with the stochastic description. In future work we are going to use this formalism to treat fluctuations in real inflationary models.

Acknowledgments

Financial support was provided by the Department of Energy through the grant DE-FG03-91-ER40682, by the Department of Physics at Carnegie Mellon University, and by a grant from the John Templeton Foundation.

Contents

Chapter 1

A short introduction to inflation

1.1 The expanding Universe

The Universe we inhabit has two important properties that are confirmed by observations: it is homogeneous and isotropic on large scales and it is expanding according to the Hubble law, that is, two distant objects, far enough that their mutual gravitational influence is negligible, run away from each other with a speed proportional to their proper distance. Due to this homogeneity and isotropy, no point in our universe is special, and any observer would see the same overall picture independently of its location.

In general relativity we can choose to work in any coordinate system that we like. There is, however, a coordinatization in which the symmetries of the universe are manifest

$$ds^2 = dt^2 - a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + \sin^2\theta \mathrm{d}\phi^2\right). \tag{1.1}$$

This is called the Friedmann-Robertson-Walker metric, or FRW for short. The positive function $a(t)$ is called the scale factor.

In these coordinates a freely falling particle stays at rest, so these are comoving coordinates. For a comoving observer's coordinates, (r, θ, ϕ) are constant and the coordinate time t is their proper time. The proper distance $d(t)$ between two such observers at a particular moment of time is proportional to the scale factor $a(t)$. Hence, the rate of change of this proper distance is equal to

$$\dot{d} = \frac{\dot{a}}{a} d = Hd,$$

where $H(t) \equiv \dot{a}/a$ is called the Hubble parameter.

In the metric (1.1) k characterizes the spatial geometry of the universe. There are three possibilities:

1) $k = 0$ corresponds to three-dimensional flat space,

2) $k = +1$ is a three-dimensional sphere,

3) $k = -1$ is a three-dimensional pseudo-sphere (a three-dimensional hyperbolic surface embedded in four-dimensional pseudo-Euclidean space).

On large scales, the content of the universe can be approximated by a perfect fluid with proper energy density $\epsilon(t)$ and pressure $p(t)$,

$$T_{\mu\nu} = (p + \epsilon)u_\mu u_\nu - p g_{\mu\nu}. \tag{1.2}$$

Since we are working in comoving coordinates, the 4-velocity u_μ has components (1,0,0,0). General relativistic energy conservation can give us useful information about the evolution of the energy density. It follows from $T^{0\mu}{}_{;\mu} = 0$ that

$$\frac{d\epsilon}{dt} + 3\frac{\dot{a}}{a}(p + \epsilon) = 0. \tag{1.3}$$

We can solve the above equation if we also know the equation of state $p = p(\epsilon)$. There are two cases that are frequently considered:

1) Cold Matter, often just called matter (e.g. dust): $p = 0$, so we have

$$\epsilon \propto a^{-3}. \tag{1.4}$$

2) Hot Matter, or ultra-relativistic matter (e.g. radiation): $p = \epsilon/3$, so we have

$$\epsilon \propto a^{-4}. \tag{1.5}$$

There is another simple case, the vacuum energy, but we will leave it for now and come back to it in Section 1.3.

The dynamics of the expansion of the Universe is governed by two equations, originally derived by Friedmann:

$$\ddot{a} = -\frac{4\pi}{3}G(\epsilon + 3p)a \tag{1.6}$$

and

$$\dot{a}^2 + k = \frac{8\pi G\epsilon a^2}{3}. \tag{1.7}$$

It is convenient to rewrite the last one in terms of the Hubble parameter $H \equiv \dot{a}/a$

$$H^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\epsilon. \tag{1.8}$$

The Friedmann equations are derived from the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = 8\pi G T_{\mu\nu},$$

using the FRW metric (1.1) and the energy-momentum tensor of a perfect fluid (1.2).

Given the equation of state, $p = p(\epsilon)$, we can solve (1.3) to find ϵ as a function of the scale factor a, and then use (1.7) to find a as a function of time t.

The value of the energy density determines the geometry of space. It is useful to define two quantities: the critical density $\epsilon^{cr}(t) \equiv 3H^2/8\pi G$ and the cosmological parameter $\Omega(t) \equiv \epsilon/\epsilon^{cr}$. From (1.8) we can conclude that the value of k depends on the value of the cosmological parameter: for $\Omega(t) > 1$, $k = +1$; for $\Omega(t) < 1$, $k = -1$; and for $\Omega(t) = 1$, $k = 0$, corresponding to flat space. Even though the cosmological parameter is a function of time, $\Omega(t) - 1$ doesn't change its sign; so by measuring the current value, Ω_0, we can determine the spatial geometry of the universe.

For $k = 0$ it is easy to solve (1.7) in the two special cases that we mentioned above:

1) Cold matter: combining (1.4) and (1.7) we can see that

$$a(t) \propto t^{2/3}. \tag{1.9}$$

Taking into account that $p = 0$ it follows from (1.6) that $\epsilon = 1/6\pi G t^2$.

2) Hot matter: combining (1.5) and (1.7) we can see that

$$a(t) \propto t^{1/2}. \tag{1.10}$$

In this case, since $p = \epsilon/3$, it follows from (1.6) that $\epsilon = 3/32\pi G t^2$.

For both of these cases $\epsilon + 3p > 0$ so equation (1.6) tells us that $\ddot{a} < 0$; that is, the expansion is decelerating.

According to the standard Big Bang model, the universe emerged filled with hot matter that was distributed homogeneously and isotropically. This model makes a lot of correct predictions, but also has some problems.

1.2 The initial condition problems of the standard Big Bang

1.2.1 Horizon problem (homogeneity and isotropy problem)

The Universe is homogeneous and isotropic on scales larger than a few hundred megaparsecs. The observations of the Cosmic Microwave background (CMB) show that at the time of recombination it was very homogeneous and isotropic (with an accuracy of 10^{-4}) on all scales up to the present horizon scale. The homogeneity and isotropy scale of the observable Universe is at least as large as the present horizon scale $ct_0 \sim 10^{28}$cm. At some initial time t_i the size of this homogeneous domain was smaller by the factor of a_i/a_0, where a_i and a_0 are the scale factors at some initial time t_i and the present time t_0. If we assume that the expansion can't dissolve the inhomogeneities, then the homogeneous patch from which our observable Universe originated was at least as large as $l_i \sim ct_0(a_i/a_0)$. Let us compare this size to the size of the causal region at the initial time $l_c \sim ct_i$:

$$\frac{l_i}{l_c} \sim \frac{t_0}{t_i}\frac{a_i}{a_0}. \tag{1.11}$$

We can try to obtain a rough estimate of this ratio. If we assume that $t_i \sim t_{Pl}$, where t_{Pl} is the Planck time, $t_{Pl} = (\frac{G\hbar}{c^4})^{1/2} \sim 10^{-43}$s, and if the radiation dominated at that time, then its temperature was of the order $T_{Pl} = \frac{1}{k_B}(\frac{\hbar c^5}{G})^{1/2} \sim 10^{32}K$. As the universe expands, the temperature of the cosmic radiation decreases in the inverse proportion to the scale factor. Hence, taking into account that the present temperature of the CMB is $T_0 = 2.7K$, we estimate that $(a_i/a_0) \sim (T_0/T_{Pl}) \sim 10^{-32}$. From this we can see that

$$\frac{l_i}{l_c} \sim \frac{10^{17}}{10^{-43}}10^{-32} \sim 10^{28}.$$

This would mean that at the Planck time scale, the size of the universe was bigger than the size of the causal patch by 10^{28} orders of magnitude, so this means that there were 10^{84} causally disconnected regions where the energy density was smoothly distributed with a fractional variation $\delta\epsilon/\epsilon \sim 10^{-4}$!

If the scale factor behaves like some power of time then we can use an estimate that $\dot{a} \sim a/t$. Then it follows from (1.11) that

$$\frac{l_i}{l_c} \sim \frac{\dot{a}_i}{\dot{a}_0}.$$

In the standard Big Bang theory with ordinary matter, gravity is always attractive so the expansion of the Universe is always decelerating. Hence, we have $\dot{a}_i/\dot{a}_0 > 1$ and the homogeneity scale always exceeds the causality scale!

1.2.2 Flatness problem

The present value of the cosmological parameter $\Omega(t_0)$ is quite close to unity, corresponding to a flat universe. The current observational bound is $|\Omega(t_0) - 1| < 0.005$. Let us use the definition of the cosmological parameter $\Omega(t)$ and rewrite the second Friedmann equation as follows

$$\Omega(t) - 1 = \frac{k}{(Ha)^2}. \tag{1.12}$$

Taking into account that this equation should hold at some initial time t_i (which we assume to be of the order of the Planck time) as well as at the present time t_0 we conclude that

$$|\Omega(t_i) - 1| = |\Omega(t_0) - 1|\frac{(Ha)_0^2}{(Ha)_i^2} = |\Omega(t_0) - 1|\left(\frac{\dot{a}_0}{\dot{a}_i}\right)^2 \leq 10^{-59}. \tag{1.13}$$

We can see from this that in order for the current value of the cosmological parameter to be as close to 1 as it is, its initial value must have been *extremely* close to unity. This is not really a paradox, there is no reason why $\Omega(t_i)$ couldn't have been this small. This is more of a fine-tuning problem that we would like to avoid if we can.

1.3 Inflation

1.3.1 Solution to the initial condition problems

The ratio of the initial and current rates of expansion, $\dot{a}_i/\dot{a}_0$, enters both initial condition problems. The large number of causally disconnected regions and the extremely high precision with which the cosmological parameter should have been close to unity seem to be the consequence of the fact that in the standard Big Bang theory with ordinary matter and radiation $\dot{a}_i/\dot{a}_0 \gg 1$. The only way we can avoid this condition is to assume that, before the stage when the universe was filled with the ordinary hot matter and was expanding with a deceleration, there was another stage when it was filled with a special type of "matter" that was making it expand with an *acceleration* ($\ddot{a} > 0$), so that we can have $\dot{a}_i/\dot{a}_0 < 1$. This stage of accelerated expansion is referred to as inflation. Depending on the model through which the accelerated expansion is realized, it can be possible to create the entire observable universe from a single causally connected patch.

Let us try to sketch why inflation could solve the horizon problem. One of the important features of an accelerating space-time is the existence of an event horizon

$$r_e(t) = a(t)\int_t^{t_{max}} \frac{dt}{a} = a(t)\int_{a(t)}^{a_{max}} \frac{da}{\dot{a}a}. \tag{1.14}$$

In an accelerating universe $\ddot{a} > 0$, so the above integral is always convergent, even if $a_{max} \to \infty$, since $\dot{a}$ grows with a. Anything at time t that is farther from an observer than $r_e(t)$ can never influence his future. Consider a situation where at some initial time $t = t_i$, the matter was smoothly distributed only inside of a ball of radius $2r_e(t_i)$. Over time the external inhomogeneities can only propagate in the region between spheres that initially had radii $2r_e(t_i)$ and $r_e(t_i)$. The region inside originated from the ball with initial radius $r_e(t_i)$ stays homogeneous. At some later time t_f the size of this smooth region increases and becomes equal to

$$r_h(t_f) = r_e(t_i)\frac{a_f}{a_i}. \tag{1.15}$$

Let us compare this to the size of the particle horizon

$$r_p(t) = a(t)\int_{t_i}^{t} \frac{dt}{a} = a(t)\int_{a_i}^{a} \frac{da}{\dot{a}a}. \tag{1.16}$$

In an accelerating universe the main contribution to (1.16) comes from $a \sim a_i$, so $r_p(t) \sim (a(t)/a_i) r_e(t_i)$. Comparing this with (1.14) we can see that $r_p(t_f) \sim r_h(t_f)$. Thus, inflation can take a small homogeneous causal region and blow it to up to the size of an observable universe.

With inflation it is also possible to resolve the flatness problem. If we look at the Friedmann equation

$$\Omega(t) - 1 = \frac{k}{\dot{a}^2} \ ,$$

we see that in an accelerating universe, where $\dot{a}$ increases, $\Omega(t)$ gets closer to 1 over time. Hence, if the acceleration is big enough $\Omega(t) - 1$ could quickly become negligibly small. Using the equation above we can relate the current value Ω_0 of the cosmological parameter to its value Ω_b at the beginning of inflation

$$|\Omega_0 - 1| = |\Omega_b - 1| \left(\frac{\dot{a}_b}{\dot{a}_0} \right)^2 .$$

Hence, as long as $(\dot{a}_b/\dot{a}_0) < 10^{-2}$, we can have $|\Omega_b - 1|$ be of order one and still get $|\Omega_0 - 1| < 0.005$ for the current value of Ω.

1.3.2 Description of the inflationary background

As we said earlier, in order to solve the horizon and flatness problems we need to have a period during which the expansion of the universe was accelerating. Let us look at one of the Friedmann equations:

$$\ddot{a} = -\frac{4\pi}{3} G(\epsilon + 3p) a. \tag{1.17}$$

We can see from this that to provide an accelerated expansion ($\ddot{a} > 0$) we need to have a type of matter with an energy density and pressure that satisfy the condition $(\epsilon + 3p) < 0$. An example of this type of "matter" is a positive cosmological constant or vacuum energy. In locally inertial coordinate systems the energy-momentum tensor of the vacuum should be proportional to the Minkowski metric $\eta^{\mu\nu}$; so in a general coordinate system it should be proportional to $g^{\mu\nu}$. Comparing this to the expression for the energy-momentum tensor of the perfect fluid, we see that for the vacuum energy the equation of state is $p_V = -\epsilon_V$ and that $T_V^{\mu\nu} = \epsilon g^{\mu\nu}$; so the desired condition is satisfied: $(\epsilon_V + 3p_V) = -2\epsilon_V < 0$. In this case the solution to the Friedmann equations is the de Sitter universe with a scale factor that grows exponentially: $a(t) \propto \exp(Ht)$, where the Hubble parameter $H \equiv \dot{a}/a = \sqrt{8\pi G \epsilon_V/3}$ is now constant, since, as follows from the equation (1.3),

$$\frac{d\epsilon}{dt} + 3\frac{\dot{a}}{a}(p + \epsilon) = 0,$$

ϵ is constant for $p_V = -\epsilon_V$. There is a problem with the cosmological constant being the source of inflation: this type of expansion will always have a positive acceleration. But the standard big bang theory has successful predictions that we would not want to give up, so we would like to have an inflationary era that eventually ends so that thereafter the universe would be decelerating. Hence, the de Sitter solution can only be thought of as a zeroth order approximation to inflation. Using the definition of the Hubble parameter it is easy to show that

$$\frac{\ddot{a}}{a} = H^2 + \dot{H}.$$

It is clear that in order for the $\ddot{a}$ to become negative at the end of inflation we must allow the Hubble parameter to vary in time, and its rate of change, $\dot{H}$, should be negative. The inflation ends when $\ddot{a}$ changes its sign, that is when $H^2 = |\dot{H}|$.

Let us estimate how long inflation should last. If we assume that the hot stage of the universe started immediately after the end of inflation, then, in order to solve the horizon problem we should have that $\dot{a}_e/\dot{a}_0 > 10^{28}$, where the subscript "$e$" denotes the end of inflation, and the subscript "0", the present time. As we saw at the end of the previous subsection, in order to solve the flatness problem we must have $(\dot{a}_b/\dot{a}_0) < 10^{-2}$, with the subscript "$b$" denoting the beginning of inflation. The last condition can be rewritten as

$$\frac{\dot{a}_b}{\dot{a}_e}\frac{\dot{a}_e}{\dot{a}_0} = \frac{a_b}{a_e}\frac{H_b}{H_e}\frac{\dot{a}_e}{\dot{a}_0} < 10^{-2}.$$

Hence,

$$\frac{a_e}{a_b} > 10^{30}\frac{H_b}{H_e}.$$

If we assume that $|\dot{H}_b|/H_b^2 \ll 1$ and neglect the change in the Hubble parameter, so that the expansion is *nearly* exponential, we can make the following estimate,

$$a_e/a_b \sim \exp(H_b\Delta t) > 10^{30}.$$

Hence, in order to solve the initial condition problems, we need to have $\Delta t > 69H_b^{-1}$, that is, inflation should last longer than 69 Hubble times (e-folds). The exact number of e-folds is model-dependent. Using the following estimate: $|\dot{H}_b| \approx (H_b - H_e)/\Delta t < H_b/\Delta t$, we can rewrite the last condition as

$$\frac{|\dot{H}_b|}{H_b^2} < \frac{1}{69}. \tag{1.18}$$

A good candidate to drive the inflation is a scalar field. Let us consider a single-field model with the following action

$$S_\phi = \int d^4x\,\sqrt{-g}\,\Big\{\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\Big\}.$$

The energy-momentum tensor associated with this field equals

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - \tfrac{1}{2}g_{\mu\nu}g^{\lambda\sigma}\partial_\lambda\phi\partial_\sigma\phi + g_{\mu\nu}V.$$

This can be put in a form that resembles the energy-momentum tensor of a perfect fluid with

$$\epsilon \equiv \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + V(\phi), \;\; p \equiv \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + V(\phi),$$

$$u_\alpha \equiv \partial_\alpha\phi/\sqrt{g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi}.$$

In an undisturbed homogeneous, isotropic universe ϵ and p can be functions of time only, so let us consider the case where ϕ doesn't depend on the spatial position, $\phi(t,\vec{x}) = \phi(t)$. Then for its energy density and pressure we have

$$\begin{aligned} \epsilon &= \frac{1}{2}\dot{\phi}^2 + V(\phi) \\ p &= \frac{1}{2}\dot{\phi}^2 - V(\phi). \end{aligned}$$

The metric is the FRW metric (1.1)

$$ds^2 = dt^2 - e^{2\rho(t)}\,\delta_{ij}\,dx^i dx^j, \tag{1.19}$$

where k is set to 0, the spatial part is written in rectangular coordinates, and the scale factor is written in an exponential form $a(t) = e^{\rho(t)}$. In this notation the Hubble parameter is $H(t) = \dot{\rho}$.

From the energy conservation equation (1.3), $\dot{\epsilon} + 3H(p+\epsilon) = 0$, we get

$$\ddot{\phi} + 3\dot{\rho}\dot{\phi} + V'(\phi) = 0. \tag{1.20}$$

The Friedmann equations (1.6) and (1.8) give us the following two equations (we have set $M_{\rm pl} \equiv (8\pi G)^{-1/2} = 1$)

$$3\dot{\rho}^2 = \frac{1}{2}\dot{\phi}^2 + V, \tag{1.21}$$

$$\ddot{\rho} = -\frac{1}{2}\dot{\phi}^2. \tag{1.22}$$

So far we have not said anything about the potential $V(\phi)$. From (1.18) we see that for successful inflation we need to have $|\dot{H}|/H^2 \ll 1$, which, as we can see from (1.21) and (1.22), requires that $\dot{\phi}^2 \ll V(\phi)$. As a consequence the equation of state of the scalar field, $p = -\epsilon + \dot{\phi}^2 \approx -\epsilon$, is close to that of the vacuum energy. To maintain this condition over a sufficiently long time we need to also require that $\ddot{\phi}/(\dot{\rho}\dot{\phi}) \ll 1$.

To summarize, to have enough inflation, the field ϕ should be in a "slow-roll" regime (do not confuse ϵ with energy density), where

$$\epsilon \equiv \frac{d}{dt}\frac{1}{H} = -\frac{\ddot{\rho}}{\dot{\rho}^2} = \frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot{\phi}^2}{\dot{\rho}^2} \ll 1, \qquad\qquad \delta \equiv \frac{1}{H}\frac{\ddot{\phi}}{\dot{\phi}} = \frac{\ddot{\phi}}{\dot{\rho}\dot{\phi}} \ll 1.$$

To get the last equality for ϵ, we used the equation (1.22) and restored $M_{\rm pl}$ for a moment. In the limit when $\epsilon = 0$, we have that $\dot{\phi} = 0$, which means that the energy density and the Hubble parameter are constants, corresponding to de Sitter space. Hence, the parameter ϵ characterizes how much we deviate from the purely exponential expansion.

1.3.3 Fluctuations about the background

Besides solving the horizon and flatness problems, inflation has another very important aspect. It addresses the origin of the primordial inhomogeneities needed to explain the large-scale structure of the universe. Observations of the cosmic microwave background tell us that initially they must have been of order $\delta\epsilon/\epsilon \sim 10^{-5}$. These observations also suggest that the power spectrum of these inhomogeneities is almost scale invariant. The predictions of inflationary theories match with these observations very well.

So far we have been assuming that our space-time is completely homogeneous and isotropic, so that the metric tensor depends only on time and has no spatial dependence. We've been assuming the same thing for the scalar field that drives the inflation. On the quantum mechanical level this can't be true, because there are always quantum fluctuations around, and some amount of spatial dependence in the metric and the scalar field is inevitable. Therefore let us consider adding small perturbations to the background FRW metric (1.19) and to the spatially independent scalar field considered above

$$\begin{gathered} g_{\mu\nu}(t) \to g_{\mu\nu}(t) + \delta g_{\mu\nu}(t,\vec{x}) \\ \phi(t) \to \phi(t) + \delta\phi(t,\vec{x}). \end{gathered}$$

We will be assuming that $|\delta g_{\mu\nu}(t,\vec{x})| \ll |g_{\mu\nu}(t)|$ and $|\delta\phi(t,\vec{x})| \ll |\phi(t)|$.

At any given moment of time, the homogeneous, isotropic background is invariant with respect to spatial rotations and translations. Perturbations in the metric and scalar field can be categorized based on the way they transform under these symmetries. There are three distinct types: scalars, vectors, and tensors. Under 3-rotations and translations $\delta\phi(t,\vec{x})$ and $\delta g_{00}(t,\vec{x})$ transform as scalars. Let us write the latter as $\delta g_{00}(t,\vec{x}) = 2\Phi(t,\vec{x})$.

The components $\delta g_{0i}(t,\vec{x})$ transform as a 3-vector, hence it can be decomposed into the gradient of some scalar $B(t,\vec{x})$ and a divergenceless vector $S_i(t,\vec{x})$:

$$\delta g_{0i}(t,\vec{x}) = -e^{2\rho(t)}\left(\partial_i B + S_i\right).$$

The components δg_{ij} transform as 3-tensors and can be decomposed in the following way

$$\delta g_{ij} = e^{2\rho(t)} \left(-2\zeta\delta_{ij} - \partial_i\partial_j\xi + \partial_j F_i + \partial_i F_j + \gamma_{ij} \right).$$

Here $\zeta(t,\vec{x})$ and $\xi(t,\vec{x})$ are 3-scalars, $F_i(t,\vec{x})$ is a divergenceless 3-vector, and $\gamma_{ij}(t,\vec{x})$ is a traceless and transverse 3-tensor, i.e.

$$\gamma_{ii} = 0, \;\; \partial_i\gamma_{ij} = 0.$$

It can be shown that scalar, vector, and tensor perturbations are not coupled to each other by field equations or conservation equations. Hence they can be studied separately. Throughout this thesis we concentrate only on the scalar perturbations.

From the above analysis it follows that in the presence of small fluctuations about a spatially flat FRW background the metric can be put in the following form

$$ds^2 = (1+2\Phi)^2\, dt^2 - 2e^{2\rho(t)}(\partial_i B)\, dt dx^i - e^{2\rho(t)} \left((1+2\zeta)\delta_{ij} + \partial_i\partial_j\xi \right) dx^i dx^j. \quad (1.23)$$

General relativity is reparametrization invariant, that is, the full action of our theory (scalar field + gravity),

$$S = \int d^4x \, \sqrt{-g} \left\{ \frac{1}{2} M_{pl}^2 R + \frac{1}{2} g^{\mu\nu} \partial_\mu\phi\partial_\nu\phi - V(\phi) \right\} , \quad (1.24)$$

is invariant under arbitrary space-time coordinate transformations. This translates into the gauge invariance of the theory described in terms of the perturbations. Because of this, not all the perturbations we have listed have a physical meaning.

Consider a general coordinate transformation $x^\mu \to x^\mu + \delta x^\mu$. This coordinate transformation contains two 3-scalars: one of them is the transformation δx^0 of the temporal coordinate, the other one is coming from the spatial coordinate transformation δx^i, which is a 3-vector and can be written in the form $\delta x^i \sim \delta^{ij}\partial_j f$, where f is a 3-scalar. Using these coordinate transformations we can get rid of two out of the five scalar perturbations that we have. We are going to choose coordinates (a gauge) where the fluctuations of the scalar field vanish, $\delta\phi = 0$ and also $\xi = 0$. With this choice the metric (1.23) becomes

$$ds^2 = [1+2\Phi(t,\vec{x})]^2\, dt^2 - 2e^{2\rho(t)}[\partial_i B(t,\vec{x})]\, dt dx^i - e^{2\rho(t)+2\zeta(t,\vec{x})}\delta_{ij}\, dx^i dx^j \quad (1.25)$$

and the scalar field depends only on time, $\phi = \phi(t)$. Notice that ζ is slightly redefined here, but to first order in the fluctuations it is the same as in (1.23). This choice of coordinates, where all of the fluctuations are in the metric and none in the scalar field, is often referred to as a unitary gauge. This gauge makes the analysis of the dynamics of the fluctuations more transparent. The fluctuation ζ has two important features: it is related to the three-dimensional scalar curvature associated with the spatial part of the metric and it is conserved for physical wavelengths much longer than the Hubble scale, $\lambda_{phys} \gg H^{-1}$.

The action (1.24) rewritten in terms of the metric (1.25) doesn't contain any time-derivatives of the fields $\Phi(t,\vec{x})$ or of $B(t,\vec{x})$. Thus, they are not dynamical but are simply Lagrange multipliers. Varying the action with respect to these fields will give us two constraints. Solving them allows us to express $\Phi(t,\vec{x})$ and $B(t,\vec{x})$ in terms of $\zeta(t,\vec{x})$. To the first order in ζ the solutions to these constrains are

$$\Phi = \frac{1}{2}\frac{\dot{\zeta}}{\dot{\rho}}\,,$$

$$B = -\frac{e^{-2\rho}}{\dot{\rho}}\zeta + \chi \qquad \text{with} \qquad \partial_k\partial^k\chi = \frac{1}{2}\frac{\dot{\phi}^2}{\dot{\rho}^2}\dot{\zeta}.$$

We are left with only one dynamical variable, $\zeta(t,\vec{x})$. To second order, the action expressed in terms of $\zeta(t,\vec{x})$ has the following simple form,

$$S = \int dt\,\frac{\dot{\phi}^2}{\dot{\rho}^2}\int d^3\vec{x}\,e^{3\rho}\Big\{\tfrac{1}{2}\dot{\zeta}^2 - \tfrac{1}{2}e^{-2\rho}\partial_k\zeta\partial^k\zeta + \cdots\Big\}. \tag{1.26}$$

We would like to quantize this theory. The first thing to do is to put (1.26) into the canonical form. Let us define a new field

$$\varphi(t,\vec{x}) \equiv e^{\rho}\frac{\dot{\phi}}{\dot{\rho}}\,\zeta(t,\vec{x}).$$

It is also convenient to put time and space coordinates on the same footing by defining a *conformal time*

$$\eta(t) \equiv \int dt\,e^{-\rho(t)},$$

where $-\infty < \eta < 0$ for t running forward.

In terms of $\varphi(\eta,\vec{x})$ the action (1.26) can be put in the following form

$$S = \int d\eta d^3\vec{x}\,\big\{\tfrac{1}{2}\varphi'^2 - \tfrac{1}{2}\vec{\nabla}\varphi\cdot\vec{\nabla}\varphi - \tfrac{1}{2}m^2\varphi^2\big\}.$$

Here a prime denotes a derivative with respect to η. This looks like the action of the free field in flat space-time, but the mass here is not constant. It depends on the conformal time and has the following form

$$m^2 = -\rho'^2(2 + 2\epsilon + 3\delta),$$

where only terms leading in the slow-roll parameters have been kept. Since the background is invariant under spatial rotations and translations, we can expand $\varphi(\eta,\vec{x})$ in terms of plane waves

$$\varphi(\eta,\vec{x}) = \int\frac{d^3\vec{k}}{(2\pi)^3}\,\big\{\varphi_k(\eta)e^{i\vec{k}\cdot\vec{x}}a_{\vec{k}} + \varphi_k^*(\eta)e^{-i\vec{k}\cdot\vec{x}}a_{\vec{k}}^\dagger\big\}.$$

The mode functions $\varphi_k^*(\eta)$ then satisfy the Klein-Gordon equation with a time-dependent mass

$$\varphi_k'' + (k^2 + m^2)\varphi_k = 0.$$

Using the fact that $d\eta = e^{-\rho(t)}dt$ and the definition of the ϵ,

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{d}{dt}\frac{1}{\dot{\rho}},$$

it is easy to show that $d(1/\rho') = -(1-\epsilon)\, d\eta$. Integrating this equation and dropping the terms that are higher order in the slow-roll parameters, we get

$$\rho' = -\frac{1+\epsilon}{\eta}.$$

The constant of integration has been fixed such that in the limit where $\epsilon \to 0$ and H is constant (de Sitter space), we have $e^{-\rho(t(\eta))} = -H\eta$. Using the above equation to rewrite the mass in terms of the slow-roll parameters and the conformal time η, the Klein-Gordon equation becomes

$$\varphi_k'' + k^2\left[1 - \frac{2 + 3(2\epsilon + \delta)}{k^2\eta^2}\right]\varphi_k = 0. \tag{1.27}$$

The solution to this equation can be expressed in terms of Hankel functions of the first and second kind

$$\varphi_k(\eta) = N_k(-k\eta)^{1/2}\left[H_\nu^{(1)}(-k\eta) + \theta_k\, H_\nu^{(2)}(-k\eta)\right], \tag{1.28}$$

with the index

$$\nu = \tfrac{3}{2}\sqrt{1 + \tfrac{4}{3}(2\epsilon + \delta)} \approx \tfrac{3}{2} + 2\epsilon + \delta.$$

N_k and θ_k are constants of integration. One of them is fixed by imposing the canonical commutation relations on the field φ and its conjugate momentum π

$$\left[\varphi(\eta, \vec{x}), \pi(\eta, \vec{y})\right] = i\delta^3(\vec{x} - \vec{y}) \qquad\qquad \pi = \frac{\delta\mathcal{L}}{\delta\varphi'} = \varphi'.$$

Rewritten in terms of mode functions the above relation implies

$$\varphi_k\varphi_k'^* - \varphi_k^*\varphi_k' = i.$$

This can be used to fix the N_k:

$$N_k = -\frac{\sqrt{\pi}}{2\sqrt{k}}\frac{1}{\sqrt{1 - \theta_k\theta_k^*}}.$$

How do we choose N_k? The comoving wavelength of any mode is constant, $\lambda = 2\pi/k$, but its physical wavelength changes with time $\lambda_{phys} = \lambda a(t) = 2\pi e^{\rho(t)}/k \approx$

$2\pi/(-kH\eta)$. In the last equality we approximated the scale factor by its de Sitter behavior. At sufficiently early times, when the physical wavelength is much smaller than the Hubble scale, $\lambda \ll 1/H$, so that $-k\eta \gg 1$, the equation (1.27) looks like the mode equation for the massless scalar field in a flat space-time

$$\varphi_k'' + k^2\varphi_k = 0.$$

If we choose its solution to correspond to the usual Minkowski vacuum, $e^{-ik\eta}/\sqrt{2k}$, then, in order for the modes (1.28) to match this in the limit $-k\eta \gg 1$, we need to set $\theta_k = 0$. Thus,

$$\varphi_k(\eta) = -\frac{\sqrt{\pi}}{2}\sqrt{-\eta}\, H_\nu^{(1)}(-k\eta).$$

The solution chosen in this way is called the Bunch-Davies vacuum. Going back to the original field ζ, for its modes we have

$$\zeta_k(\eta) = e^{-\rho}\frac{\dot\rho}{\dot\phi}\varphi_k. \tag{1.29}$$

As was mentioned at the beginning of this subsection, the CMB observations suggest an almost scale invariant power-spectrum. So let us compare these observations with the predictions of the inflationary scenario. The power-spectrum P_k is defined as a Fourier transform (up to some conventional factors) of the two-point correlation function of ζ

$$\langle 0|\zeta(\eta,\vec x)\zeta(\eta,\vec y)|0\rangle = \int \frac{d^3\vec k}{(2\pi)^3}\, e^{i\vec k\cdot(\vec x-\vec y)}\frac{2\pi^2}{k^3}P_k(\eta).$$

From (1.29) it follows that

$$P_k(\eta) = e^{-2\rho}\frac{\dot\rho^2}{\dot\phi^2}\frac{k^3}{2\pi^2}\varphi_k\varphi_k^*.$$

Expanding this in powers of the slow-roll parameters and keeping only the leading piece we get

$$P_k(\eta) = \frac{H^2}{M_{\rm pl}^2}\frac{(-k\eta)^3}{16\pi}\frac{H_\nu^{(1)}(-k\eta)\, H_\nu^{(2)}(-k\eta)}{\epsilon(1+\epsilon)^2}. \tag{1.30}$$

The cosmologically interesting inhomogeneities are those with large physical wavelengths, roughly of the order of the present horizon scale. By the end of inflation the wavelengths of these fluctuations were much larger than the Hubble scale, $\lambda_{phys} \gg 1/H$: during inflation the scale factor and hence the physical wavelength grow more or less exponentially while the Hubble parameter stays approximately constant, so all of the modes (except for *very* short ones) that originally had

$\lambda_{phys} \ll 1/H$, by the end of inflation have $\lambda_{phys} \gg 1/H$. Let us evaluate the power-spectrum for these modes. We have the following: $k_{phys} = k/a(t) \approx -k\eta H \ll H$, so that $-k\eta \ll 1$. In this limit the power-spectrum (1.30) becomes

$$P_k(\eta) = \frac{1}{8\pi^2}\frac{1}{\epsilon}\frac{H^2}{M_{\mathrm{pl}}^2}(-k\eta)^{-4\epsilon-2\delta} + \cdots .$$

As we can see, the power-spectrum is almost scale invariant with a slight growth for the larger physical wavelengths, which is called a red tilt. This red tilt also matches with observations.[1]

[1]The material in this chapter typically follows the approach taken in the following sources: Mukhanov's *The Physical Foundations of Cosmology* [1], Weinberg's *Cosmology* [2] and the paper [3].

Chapter 2

A cosmological consistency relation

2.1 Consistency relations in inflation

The theory of inflation provides an elegant causal explanation for the origin of the primordial fluctuations of the universe. So far, the scalar component of these fluctuations has been the most carefully scrutinised, and its observation has richly added to our understanding of the early universe. Planck and future experiments will continue to observe the more detailed structures in these scalar fluctuations, searching for further consistency between inflation and the observed universe. And there are many other experiments underway working to measure the tensor component of the primordial fluctuations as well. Because we are unlikely to be able to test the dynamics of inflation directly, finding such features in the primordial fluctuations and relations amongst them comes as an immense boon.

One important prediction of inflation is that the Gaussian and non-Gaussian features in these primordial fluctuations should be related to each other. In his analysis of the non-Gaussianities produced by inflation [4], Maldacena presented a simple relation between the correlation functions of these fluctuations, $\zeta_{\vec{k}}(t)$. His original relation states that in the limit where the momentum of one of the fields approaches zero, the three-point function of the fluctuations—the simplest measure of non-Gaussianity—is determined by the amplitude and scaling behaviour of the Gaussian power spectrum $P_k(t)$ through the identity

$$\delta^3(\vec{k}_2+\vec{k}_3)\, P^{-1}_{k\to 0}(t_*)\, \langle \zeta_{\vec{k}\to\vec{0}}(t_*)\zeta_{\vec{k}_2}(t_*)\zeta_{\vec{k}_3}(t_*)\rangle = \left[3+\frac{d}{d\ln k_2}\right] P_{k_2}(t_*). \tag{2.1}$$

Both sides of this expression are evaluated at a late time t_* when the wavelengths responsible for the correlations seen in the cosmic microwave background and large-scale structure have been inflated to a size larger than the inflationary horizon.

Within a wide class of inflationary models—those with a single inflaton field whose potential obeys a set of 'slow-roll' conditions, and where the fluctuations are in the Bunch-Davies state—this relation says that the amplitude of the three-point function in this 'soft' limit should be small, since the basic properties of the power spectrum are already known through the observations made by Planck and its predecessors. The smallness of the three-point function then follows naturally from assumptions that have already been made in these models, and the relation explains why the non-Gaussianities so far have been difficult to detect. And once the three-point function has been measured, this relation will provide an important constraint on the consistency of the minimal inflationary picture.

Building on Maldacena's original insights, others have worked steadily during the past decade to generalise this relation, to extend it and to treat it from deeper perspectives [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. It was shown, for example, that this consistency relation applies even when the slow-roll conditions are relaxed [5], and other work [6, 7] demonstrated that the original consistency relation is just the first in a series of relations between the $n+1$-point and n-point correlation functions predicted by inflation.

The fluctuations about the inflating background are quantum fields. When the tensors are neglected, the scalar fluctuations can be cast, through a suitable choice of coordinates, in a form where they appear within a conformal factor that multiplies an otherwise flat spatial part of the metric. This choice does not quite exhaust the diffeomorphism invariance of the metric. Within this class of coordinates, conformal transformations of the spatial metric remain a residual symmetry of the theory. What could be more natural then, than that these consistency conditions should be the Slavnov-Taylor identities associated with this symmetry of the quantum theory?

This realization was first developed by [8] and then refined and extended in later work [9, 10, 11]. Recently Goldberger, Hui, and Nicolis [12] have followed this approach to derive these relations using one-particle irreducible Green's functions and the effective action appropriate to an inflationary setting. Since their approach relies only on very general properties of quantum field theory, it has great generality and the consistency relations follow as very simple and direct consequences of the residual conformal invariance. Their derivation relies on only a minimal set of assumptions about the inflationary theory—that the initial state, its evolution, and the measure of the path integral are invariant under these conformal symmetries—which apply independently of slow-roll assumptions, the behaviour of the field outside the horizon, *etc*. A further elegant feature of their derivation is that the path integral can be treated as that of a *three-dimensional* field theory defined on the late-time hypersurface, $t = t_*$. Information about the initial state and its evolution influences a probability measure within the path integral; but as long as that measure remains invariant under the residual conformal symmetries, its details are not important for the Slavnov-Taylor identity for the fluctuations at the late-time boundary, $\zeta_{\vec{k}}(t_*)$.

An important element that has so far been missing from all previous treatments of the consistency relation is what happens when the scalar—or the tensor—

fluctuations are *not* in a Bunch-Davies state. One reason for considering states other than the Bunch-Davies state is that while inflation as an overall idea has been very successful in predicting many of the properties of the primordial fluctuations, we still know almost nothing about its details. Cosmological observations are fine enough to be able now to exclude specific models, but it would be prudent to assume as little as we can and let observations constrain and tell us what is consistent with them. Inflationary models are most often studied in terms of their potentials, but that is only half of the picture. Allowing a more interesting initial state could help in explaining some of the observed properties of the primordial fluctuations. Since it is straightforward to introduce more general initial states into a quantum field theory [18, 19], it should be possible to parametrise—and constrain by matching with what is seen—by how much the initial state of inflation can depart from the Bunch-Davies state in much the same way as we now study different actions for inflationary models in an effective theory language.

It is also essential to understand what any departures from the consistency relation could mean if they are seen. If the three-point function is found not to be in accord with the naïve form of the consistency relation, it might be tempting to see this as a failure of the simple single-field, slow-roll picture for inflation. But that is not the only possibility. It could also be that the basic dynamical picture is correct, only that the state contains structures beyond the Bunch-Davies state; these too would modify the consistency relation.

What happens when the initial state breaks some of the residual conformal symmetry? In this article we develop a method for calculating the cosmological Slavnov-Taylor identities for such states. We shall still be following the philosophy of [12], except that now we allow the path integral to include the evolution of the state. Our approach thus keeps the generality of the earlier approaches—which only relied on the symmetries of the action and the functional measure of the path integral—but it broadens them yet further by allowing for states that are not invariant under the full conformal symmetry. Of course, the form of the consistency relation changes for a general state. In fact these changes, though complicated in their expression, have a very familiar meaning: they are the cosmological analogues for how explicitly broken symmetries alter the Ward identities in gauge theories.

The expression for the consistency relation that emerges from our approach can be used to treat much more general situations than those that have been considered so far. To make some of the working details of our formalism clearer, we illustrate how the standard consistency relation, in the Bunch-Davies state, emerges before we undertake an analysis of a completely general initial state. The corrections introduced into the Slavnov-Taylor identities by states that explicitly break the residual conformal symmetry are derived here; their phenomenology will be explored more fully in later work [20].

In the next section we introduce the action for an inflationary theory with a single scalar field. What is most essential for the consistency relation are the symmetries that exist after we have chosen a particular set of coordinates or 'gauge'. When the coordinates are chosen so that the spatial part of the metric remains conformally flat, it is still left invariant by any further conformal transformation of the field.

In section 3 we develop the families of connected and one-particle irreducible Green's functions and the effective action that are needed for a fully time-evolving quantum field theory. Once these theoretical tools have been introduced, we derive the Slavnov-Taylor identity and show that for spatial dilations it generates the consistency relation between the two and three-point functions. While different expressions of this relation might be formally equivalent, it is most easily computed when it is written with a two-point 1PI Green's function acting on a three-point connected Green's function. The necessary contortions needed to put it into this form are sketched in section 4 and explained more fully in a supplementary section of this chapter.

Section 5 derives the form of the Slavnov-Taylor identity for a general initial state. These states can be broadly divided according to whether they are invariant under the same conformal symmetries as the inflationary dynamics or whether they break some or all of these symmetries. For the former states, although the Green's functions change, the consistency relations that relate them do not. For the most general initial states, additional terms occur in the Slavnov-Taylor identity due to the non-invariance of the state.

The complete evaluation of the standard consistency relation for the simple class of inflationary theories analysed by Maldacena contains a few subtleties. The singularities and zeros that occur in the soft momentum limit must be treated with care, so a detailed calculation of this relation for this standard—but extremely important—case is presented in the supplementary section 2.7.

In the conclusions of this chapter we outline the next stages of this work.

2.2 A residual diffeomorphism invariance

We begin by briefly reviewing the classical background used in inflation. The simplest models of inflation contain a single scalar field $\phi(t,\vec{x})$ with a potential $V(\phi)$. The combined dynamics for gravity and this field are then determined by the action

$$S = \int d^4\vec{x}\,\sqrt{-g}\Big\{\tfrac{1}{2}M_{\rm pl}^2 R + \tfrac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\Big\}, \tag{2.2}$$

where $M_{\rm pl}$ is the Planck mass. Here the background space-time is assumed to depend only on the time coordinate. In this case the classical value of the field can be written as just $\phi(t)$ and the background metric can be put into the standard form[1]

$$ds^2 = dt^2 - e^{2\rho(t)}\,\delta_{ij}\,dx^i dx^j. \tag{2.3}$$

[1]Here we are using Maldacena's [4] notation for the background metric and the fluctuations about it.

The quantum fluctuations about this background very naturally introduce some spatial dependence into the universe during inflation. As the time component already has a special role in this background, it is convenient to write the metric in the form

$$ds^2 = \left[N^2 - h_{ij}\, N^i N^j\right] dt^2 - 2h_{ij}\, N^i\, dt dx^j - h_{ij}\, dx^i dx^j \tag{2.4}$$

to treat the fluctuations. The background is spatially flat, so the quantum fluctuations are usually parametrised by how they transform under its symmetries. Those that transform as spatial scalars are the most immediately observable since they are needed to explain the primordial inhomogeneities in the early universe. The tensor fluctuations in h_{ij} correspond to primordial gravity waves. Here we focus just on the scalar fluctuations, leaving the tensors for later work.

We use our freedom to choose our coordinates so that they have two useful properties: first, that the inflaton field has no fluctuations at all, and is given entirely by its background value, $\phi(t, \vec{x}) = \phi(t)$, and second, that the spatial part of the metric remains proportional to δ_{ij}. The second of these conditions eliminates one of the scalar fluctuations in h_{ij}; the remaining scalar field $\zeta(t, \vec{x})$ corresponds to quantum fluctuations in the scale factor itself,[2]

$$h_{ij} = e^{2\rho(t)+2\zeta(t,\vec{x})}\, \delta_{ij}. \tag{2.5}$$

This is an especially convenient choice since $\zeta(t, \vec{x})$ is the quantum analogue of the fluctuations in the classical spatial curvature after inflation. In the standard inflationary models it approaches a constant well outside the horizon. The two further scalar fields present in N and N^i are fixed by the equations of motion for these fields, which are nondynamical Lagrange multipliers. Written in terms of the field $\zeta(t, \vec{x})$, the constraint equations found by varying the action with respect to N and N^i require that

$$N = 1 + \frac{\dot{\zeta}}{\dot{\rho}} \qquad \text{and} \qquad N^i = \delta^{ij}\partial_j\left\{ -\frac{e^{-2\rho}}{\dot{\rho}}\zeta + \frac{1}{2}\frac{\dot{\phi}^2}{\dot{\rho}^2}\partial^{-2}\dot{\zeta} \right\}, \tag{2.6}$$

∂^{-2} being the inverse spatial Laplacian operator.

Is there any freedom left for changing the spatial coordinates further while still remaining within the general form that we have chosen? As one instance of this, we observe that—since our background is flat—it should be possible to absorb a general conformal transformation of the spatial coordinates by a suitable change in $\zeta(t, \vec{x})$.

Let us study this idea more precisely by making a transformation of the spatial coordinates

$$x^i \to x^i + \xi^i(t, \vec{x}). \tag{2.7}$$

[2]If we had wished to include tensor fluctuations $\gamma_{ij}(t, \vec{x})$ as well, we would have replaced δ_{ij} with

$$\delta_{ij} \to \exp[\gamma_{ij}] = \delta_{ij} + \gamma_{ij} + \tfrac{1}{2}\gamma_i{}^k\gamma_{kj} + \cdots .$$

We shall consider this to be a small transformation in the sense that it is consistent to work to linear order in the transformed quantities. This transformation causes the spatial part of the metric to change by the amount

$$\delta h_{ij} = e^{2\rho+2\zeta}\Big\{2\delta_{ij}\xi^k\partial_k\zeta + \delta_{kj}\partial_i\xi^k + \delta_{kj}\partial_i\xi^k\Big\}, \tag{2.8}$$

to first order in ξ^i. Once again we neglected the other fluctuations such as the tensor fluctuations. This change can be absorbed through a corresponding change in the field $\zeta(t,\vec{x})$,

$$\zeta(t,\vec{x}) \to \tilde{\zeta}(t,\vec{x}) = \zeta(t,\vec{x}) + \delta\zeta(t,\vec{x}), \tag{2.9}$$

of the form

$$2\delta\zeta\,\delta_{ij} = 2\delta_{ij}\xi^k\partial_k\zeta + \delta_{kj}\partial_i\xi^k + \delta_{kj}\partial_i\xi^k. \tag{2.10}$$

The fact that the change of coordinates can be related to an overall change in the factor multiplying the flat spatial metric means that $\xi^i(t,\vec{x})$ is generating a conformal transformation. We can make this property still more explicit by taking the trace of this equation, which allows us to solve for $\delta\zeta$ directly,

$$\delta\zeta = \xi^k\partial_k\zeta + \tfrac{1}{3}\partial_k\xi^k, \tag{2.11}$$

and which, when substituted into the previous equation, yields the conformal Killing equation for the flat metric,

$$\partial_i\xi_j + \partial_j\xi_i = \tfrac{2}{3}\delta_{ij}\,\partial_k\xi^k. \tag{2.12}$$

What we have found then is that even within the class of coordinates that we have chosen, there is some additional symmetry available. The fluctuations $\zeta(t,\vec{x})$ that we are considering are quantum fields, so these symmetries will generate constraints on the field and relations amongst its correlation functions. Here we have found that our metric is invariant under the set of conformal transformations of three-dimensional flat space. These transformations correspond to spatial translations and rotations, as well as dilations and special conformal transformations.

Both the spatial translations and rotations have already implicitly been included in the structure of the theory, since $\zeta(t,\vec{x})$ is itself a scalar field under these transformations. And while the correlation functions of $\zeta(t,\vec{x})$ must also be invariant under these symmetries, they do not impose any relations between correlation functions of different orders. This follows from the fact that for both translations and rotations, the corresponding transformation of the field is *linear* and *homogeneous* in $\zeta(t,\vec{x})$, $\delta\zeta = \xi^k\partial_k\zeta$, since $\partial_k\xi^k = 0$.

In contrast, the dilation and the special conformal transformations introduce *inhomogeneous* terms into the infinitesimal transformation of the field. For example, under a dilation,

$$x^i \to x^i + \lambda\, x^i \qquad \text{or} \qquad \xi^i = \lambda\, x^i. \tag{2.13}$$

The effect of this transformation on $\zeta(t, \vec{x})$ is then

$$\delta\zeta = \lambda + \lambda\, \vec{x} \cdot \vec{\nabla}\zeta. \tag{2.14}$$

It is the presence of the inhomogeneous term that leads to relations between the n and $n+1$-point correlation functions of the field.

While we shall not be treating the effects of a special conformal transformation further here, it too introduces an inhomogeneous term in the transformation of the scalar fluctuations. A general special conformal transformation of the spatial coordinates is parametrised by a vector $\vec{b}$,

$$x^i \to \frac{x^i - b^i x_k x^k}{1 - 2b_k x^k + (b_j b^j)(x_k x^k)}. \tag{2.15}$$

Under an infinitesimal transformation then, we have

$$\xi^i = 2b^k\, x_k x^i - x^k x_k b^i. \tag{2.16}$$

Putting this into the general expression for the corresponding transformation of $\zeta(t, \vec{x})$, we arrive at a more complicated expression,

$$\delta\zeta = \tfrac{2}{3}\vec{x} \cdot \vec{\nabla}(\vec{x} \cdot \vec{b}) - \tfrac{1}{3}\|\vec{x}\|^2 \vec{\nabla} \cdot \vec{b} + \tfrac{4}{3}\vec{x} \cdot \vec{b} + 2(\vec{x} \cdot \vec{b})\vec{x} \cdot \vec{\nabla}\zeta - \|\vec{x}\|^2 \vec{b} \cdot \vec{\nabla}\zeta. \tag{2.17}$$

This symmetry also generates relations between different orders of correlation functions of $\zeta(t, \vec{x})$, since we can see that the first pair of terms are inhomogeneous and the second pair are linear in $\zeta(t, \vec{x})$.

A fuller analysis of these residual symmetries, including a treatment of what additional time-dependent information can be gleaned by using the adiabatic properties [21] of the fluctuations at larger scales, is found in [8, 9].

2.3 Evolution

The scalar fluctuation $\zeta(t, \vec{x})$ is a quantum field whose dynamics are determined by the action $S[\zeta]$, which is found by expanding our original inflationary action in powers of $\zeta(t, \vec{x})$. In cosmological settings, the quantities that we should like to compute are the expectation values of operators built from this field. In principle such expectation values could be taken in an arbitrary state, but in this article we shall always have a particular state in mind, the Bunch-Davies state, though we shall set up everything in a way that generalises fairly readily to other states. The Bunch-Davies state is the state that matches with the free Minkowski vacuum in an arbitrarily remote past. It is usually taken to be the natural choice for the ground state for inflation. When the field is not in a free theory, its evolution from this pristine initial state set in an infinite past can be quite complicated and almost always needs to be treated perturbatively.

Let us write the expectation value of an operator $\mathcal{O}$ in the Bunch-Davies state, which we denote by $|0(t_*)\rangle$, as

$$\langle 0(t_*)|\mathcal{O}|0(t_*)\rangle. \tag{2.18}$$

The state has been implicitly evolved from its initial free-theory form at $t = -\infty$ to its form at $t = t_*$, but we have been deliberately ambiguous about the time-dependence of the operator $\mathcal{O}$. One of the most important expectation values for inflation is the equal-time correlation function of n fields,

$$\langle 0(t_*)|\zeta(t_*,\vec{x}_1)\cdots\zeta(t_*,\vec{x}_n)|0(t_*)\rangle. \tag{2.19}$$

However, to treat the full evolution of the theory we shall need more general n-point Green's functions too, where each field has its own independent time.

Before we construct these Green's functions, we ought first to explain how to treat the time-evolution of an expectation value in a little more detail. Later, we shall be using the symmetries described in the previous section to derive relations amongst the expectation values of different n-point functions. As we mentioned in the introduction, these relations are the Slavnov-Taylor identities adapted to a cosmological setting. The path integral formalism is especially well suited for this purpose. The generating functional for an evolving expectation value has the form,[3]

$$Z[J^\pm] = \int \mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\,\exp\Big\{iS[\zeta^+]-iS[\zeta^-]+i\int_{-\infty}^{t_*}dt\int d^3\vec{x}\,\big[J^+\zeta^+-J^-\zeta^-\big]\Big\}. \tag{2.20}$$

The overall structure of $Z[J^\pm]$ should be reminiscent of the generating function used in S-matrix calculations, except that here the fluctuation $\zeta(t,\vec{x})$ has been written in terms of two fields, $\zeta^+(t,\vec{x})$ and $\zeta^-(t,\vec{x})$. These two fields are associated with the evolution of the two states that occur within the expectation value. The '+' index has been added to the field to signal that it is the time coordinate that occurs in the evolution of the state $|0(t)\rangle$ from its initial value, usually defined in some very remote past[4], up to the time t_*. Similarly the ζ^- appears due to the evolution of the state $\langle 0(t)| = (|0(t)\rangle)^\dagger$. This Hermitian conjugation introduces a few further oddities. When time-ordering a product of fields, a ζ^- field always occurs *later* than any ζ^+ field, whatever the numerical values of the times at which they occur might be. This convention places time-ordered ζ^-'s to the left, which is where they should be since they are associated with the $\langle 0(t)|$, which appears left-most in the expectation value.

[3]We shall sometimes use these footnotes to explain what changes when we go from the Bunch-Davies state to a more general state—in preparation for the next stages of this work. As a first comment: for a more general state, the action will not necessarily have the diagonal structure, $S[\zeta^+]-S[\zeta^-]$. It can have cross-terms where ζ^+'s and ζ^-'s couple directly to each other at an initial time t_0.

[4]Another—rather obvious—difference: a general state could have been defined at an arbitrary initial time, t_0.

The conjugation also reverses the evolution of time, so that the time-ordering of two ζ^- fields is the opposite of that of a pair of two ζ^+ fields. We shall write the explicit rules for the propagator below, where the time-ordering should become clearer.

As an example, let us show how the generating functional is used to define an equal-time n-point correlation function,

$$\langle 0(t_*)|\zeta(t_*,\vec{x}_1)\cdots\zeta(t_*,\vec{x}_n)|0(t_*)\rangle. \tag{2.21}$$

Ordinarily, we generate such expectation values by applying n functional derivatives with respect to the source to $Z[J^\pm]$. But here we have two choices: which sorts of fields—ζ^+'s or ζ^-'s—should these be? The ζ's that appear here are meant to be 'external fields.' In that case it does not matter which sign we choose as long as they are *all* chosen to have the *same* sign. What is conventionally done is to choose them to be all '+' fields. The expectation value of n fields at equal times is then evaluated—usually perturbatively—by computing the $n^{\rm th}$ functional derivative of $Z[J^\pm]$,

$$\begin{aligned}
&\langle 0(t_*)|\zeta(t_*,\vec{x}_1)\cdots\zeta(t_*,\vec{x}_n)|0(t_*)\rangle \\
&\quad = \int \mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\,\Big(\zeta^+(t_*,\vec{x}_1)\cdots\zeta^+(t_*,\vec{x}_n)\Big)e^{iS[\zeta^+]-iS[\zeta^-]+i\int_{-\infty}^{t_*}dt\int d^3\vec{x}\,[J^+\zeta^+-J^-\zeta^-]}. \\
&\quad = \left(\frac{1}{i}\frac{\delta}{\delta J^+(t_*,\vec{x}_1)}\right)\cdots\left(\frac{1}{i}\frac{\delta}{\delta J^+(t_*,\vec{x}_n)}\right)Z[J^\pm]\Big|_{J^\pm=0}.
\end{aligned} \tag{2.22}$$

These equal-time correlation functions are what inflation is meant to generate. They are the initial conditions that the inflationary era bequeaths to the subsequent eras.

Once we have introduced the generating functional, we can define further Green's functions which are needed in deriving the cosmological Slavnov-Taylor identities. Consider a family of Green's functions evaluated now at arbitrary space-time points and with arbitrary $\pm$ indices too,

$$G^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n) = \langle 0(t_*)|T\big(\zeta^{\pm_1}(x_1)\cdots\zeta^{\pm_n}(x_n)\big)|0(t_*)\rangle. \tag{2.23}$$

They are defined by taking functional derivatives of the path integral with respect to the appropriate sources, $J^\pm(x_n)$,

$$G^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n) = \left(\frac{1}{\pm i}\frac{\delta}{\delta J^{\pm_1}(x_1)}\right)\cdots\left(\frac{1}{\pm i}\frac{\delta}{\delta J^{\pm_n}(x_n)}\right)Z[J^\pm]\Big|_{J^\pm=0}. \tag{2.24}$$

In principle, in addition to the time-dependence of each field, $x_i=(t_i,\vec{x}_i)$, there is a further time-dependence[5] given by the time to which the states are being evolved,

[5] or rather, a dependence on *two* times, which we could call t_* and t'_*,

$$\langle 0(t'_*)|\mathcal{O}|0(t_*)\rangle = \int\mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\,\mathcal{O}\exp\Big\{i\int_{-\infty}^{t_*}dt\int d^3\vec{x}\,[\mathcal{L}[\zeta^+]+J^+\zeta^+]-i\int_{-\infty}^{t'_*}dt\int d^3\vec{x}\,[\mathcal{L}[\zeta^-]+J^-\zeta^-]\Big\},$$

but we shall refrain from treating anything so perverse here.

which we have called t_*,

$$G^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n) = G^{\pm_1\cdots\pm_n}(t_*;x_1,\ldots,x_n). \tag{2.25}$$

Since the Green's functions here will always be evaluated in a state at the $t = t_*$ hypersurface, this dependence will not be written explicitly. We shall always assume that $t_i \leq t_*$.

Green's functions that contain a mixture of $+$ and $-$ indices and that depend on multiple times do not correspond to graphs where the fields are all 'external'—the analogue of 'on-shell' external fields in an S-matrix calculation. However, such Green's functions naturally occur as the internal subgraphs of a more complicated process. When we integrate the position of an internal vertex over all space-time points, this also means that here we are summing over $\pm$ indices.[6]

The simplest examples of such internal Green's functions are the Feynman propagators themselves. They are derived from the quadratic part of the action and they form the basis of a perturbative treatment of more complicated processes. In taking the Wick contractions of pairs of fields, there are four possibilities for the $\pm$ indices of the two fields that are contracted, which in turn means that there are four types of Feynman propagators,

$$G^{\pm\pm}(x,y) = \int \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})}\, G_k^{\pm\pm}(t,t'), \tag{2.26}$$

with $x = (t,\vec{x})$ and $y = (t',\vec{y})$ and where

$$\begin{aligned}
G_k^{++}(t,t') &= \Theta(t-t')\, G_k^{>}(t,t') + \Theta(t'-t)\, G_k^{<}(t,t') \\
G^{+-}(t,t') &= G_k^{<}(t,t') \\
G^{-+}(t,t') &= G_k^{>}(t,t') \\
G^{--}(t,t') &= \Theta(t'-t)\, G_k^{>}(t,t') + \Theta(t-t')\, G_k^{<}(t,t').
\end{aligned} \tag{2.27}$$

Here, $G_k^{>}(t,t')$ and $G_k^{<}(t,t')$ are the free Wightman functions associated with the two-point functions evaluated in the asymptotic vacuum state,[7]

$$\begin{aligned}
\langle 0(-\infty)|\zeta(t,\vec{x})\zeta(t',\vec{y})|0(-\infty)\rangle &= \int \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})}\, G_k^{>}(t,t') \\
\langle 0(-\infty)|\zeta(t',\vec{y})\zeta(t,\vec{x}))|0(-\infty)\rangle &= \int \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})}\, G_k^{<}(t,t').
\end{aligned} \tag{2.28}$$

[6]A note for the experts: if we think instead of the t in the $\zeta^+(t,\vec{x})$ and $\zeta^-(t,\vec{x})$ as being a *single* time coordinate t_c that runs along a contour from the initial to the final time and then back again, the sum over $\pm$ indices naturally appears when we integrate internal vertices of the fields $\zeta(t_c,\vec{x})$ over the *entire* time contour in $dt_c d\vec{x}$.

[7]For an initial state defined at t_0, they would be evaluated in that state, $|0(t_0)\rangle$, instead.

Notice that the time-ordering follows the pattern that we described before. Times associated with a $-$ field always occur after those of a $+$ field, which explains the absence of Θ-functions in $G_k^{+-}(t,t')$ and $G_k^{-+}(t,t')$, and the fact that the time-ordering of the $-$ fields is the opposite of that of the $+$ fields is a relic of the Hermitian conjugation. This also explains the reversed roles of the Θ-functions in $G_k^{--}(t,t')$.

What are these Wightman functions for an inflationary universe? The free, or quadratic, part of the action for a slow-roll model of inflation is

$$S^{(2)}[\zeta] = \frac{1}{2}\int_{-\infty}^{t_*} dt\, e^{3\rho(t)}\frac{\dot\phi^2}{\dot\rho^2}\int d^3\vec{x}\left\{\dot\zeta^2 - e^{-2\rho(t)}\partial_k\zeta\partial^k\zeta\right\}. \tag{2.29}$$

The slow-roll limit of inflation corresponds to the regime where the dimensionless slow-roll parameters

$$\epsilon = \frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2} \qquad \text{and} \qquad \delta = \frac{\ddot\phi}{\dot\rho\dot\phi} \tag{2.30}$$

are small: $\epsilon, \delta \ll 1$. In this limit, the Wightman functions derived from these quadratic terms are

$$G_k^>(t,t') = G_k^<(t',t') = e^{-\rho(t)}e^{-\rho(t')}\frac{\dot\rho(t)}{\dot\phi(t)}\frac{\dot\rho(t')}{\dot\phi(t')}\frac{\pi}{4}\sqrt{\eta\eta'}H_\nu^{(1)}(-k\eta)H_\nu^{(2)}(-k\eta'), \tag{2.31}$$

where the index of the Hankel functions is

$$\nu = \sqrt{\tfrac{9}{4} + 3(2\epsilon + \delta)}. \tag{2.32}$$

Here some of the time-dependence has been written in terms of the conformal time, $\eta = \eta(t)$ and $\eta' = \eta(t')$, where

$$\eta(t) = \int dt\, e^{-\rho(t)}, \tag{2.33}$$

since the expressions then look simpler. In an inflationary universe the conformal time is usually chosen to be negative, $\eta \in (-\infty, 0)$, since for this choice the coordinate runs forward when the time is also running forward. The leading terms of these Wightman functions in the slow-roll limit have a simpler, approximately de Sitter, form

$$\begin{aligned} G_k^>(t,t') &= \frac{1}{4\epsilon}\frac{H^2}{M_{\rm pl}^2}\frac{1}{k^3}(1+ik\eta)(1-ik\eta')e^{-ik(\eta-\eta')} + \cdots \\ G_k^<(t,t') &= \frac{1}{4\epsilon}\frac{H^2}{M_{\rm pl}^2}\frac{1}{k^3}(1-ik\eta)(1+ik\eta')e^{ik(\eta-\eta')} + \cdots . \end{aligned} \tag{2.34}$$

Just as we can perturbatively treat a general diagram as a graph of $G^{\pm\pm}(x,y)$ propagators connecting vertices composed of just $+$ or $-$ fields, we can similarly imagine more complicated diagrams as being divided into various subgraphs. When a line of a subgraph does not end at an external space-time point, it can be at an arbitrary time with an arbitrary $\pm$ index, though when summing up the graphs that contribute to a process we integrate over the space-time location of the internal point *and* sum over all values of $\pm$. We shall see later that the consistency relation has this structure.

The Green's functions that we have introduced so far correspond to the sums of *all* the graphs that contribute to a process, whether they are connected or not.[8] For the two and three-point functions, we do not need to distinguish between the connected and unconnected Green's functions in this theory. However, for higher point functions, the Green's functions can be separated into a sum of products of lower-order n-point functions plus purely connected components. For example, the four-point function can be written as a sum of three pairs of two-point functions plus a connected four-point function, *e.g.*

$$\begin{aligned} G^{++++}(x_1,x_2,x_3,x_4) &= G^{++}(x_1,x_2)G^{++}(x_3,x_4) + G^{++}(x_1,x_3)G^{++}(x_2,x_4) \\ &\quad + G^{++}(x_1,x_4)G^{++}(x_2,x_3) + G_c^{++++}(x_1,x_2,x_3,x_4). \end{aligned} \tag{2.35}$$

We can extract these latter from the rest by defining a generating functional, $W[J^\pm]$, for just the *connected* graphs,

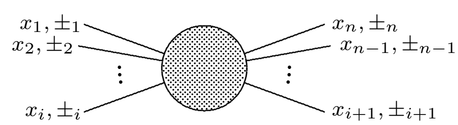
$$Z[J^\pm] = e^{iW[J^\pm]}. \tag{2.36}$$

The connected n-point Green's functions are then defined by taking the appropriate functional derivatives with respect to $W[J^\pm]$,

$$G_c^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n) = \left(\frac{1}{\pm i}\frac{\delta}{\delta J^{\pm_1}(x_1)}\right)\cdots\left(\frac{1}{\pm i}\frac{\delta}{\delta J^{\pm_n}(x_n)}\right) iW[J^\pm]\Big|_{J^\pm=0}. \tag{2.37}$$

Graphically,

$$G_c^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n) = \tag{2.38}$$



where the shaded blob is the sum of all connected diagrams.

Proceeding a step further, we introduce a generating functional for the 'one-particle irreducible' Green's function through a Legendre transform of the $W[J^\pm]$. Let us define the connected expectation value of the field $\zeta(t,\vec{x})$ in the presence of a source to be

$$\bar{\zeta}^\pm(x) \equiv \pm\frac{\delta W}{\delta J^\pm(x)} = \langle 0(t)|\zeta^\pm(x)|0(t)\rangle_{c,J^\pm}. \tag{2.39}$$

[8] We did not need to remove the vacuum-to-vacuum graphs explicitly when we defined these Green's functions. The vacuum-to-vacuum graphs are automatically cancelled due to the $S[\zeta^+] - S[\zeta^-]$ structure that appears in $Z[J^\pm]$.

The functional $\Gamma[\bar{\zeta}^\pm]$ is then defined through the transformation,

$$\Gamma[\bar{\zeta}^\pm] = W[J^\pm] - \int d^4x \left[J^+(x)\bar{\zeta}^+(x) - J^-(x)\bar{\zeta}^-(x) \right]. \tag{2.40}$$

If we take the functional derivative of $\Gamma[\bar{\zeta}^\pm]$ with respect to $\bar{\zeta}^\pm$, we produce the complementary relations

$$J^\pm(x) = \mp \frac{\delta\Gamma}{\delta\bar{\zeta}^\pm(x)}. \tag{2.41}$$

The individual '1PI' n-point functions are defined by differentiating with respect to the fields $\bar{\zeta}^\pm$ an appropriate number of times,

$$\Gamma^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n) = \left. \frac{\delta^n\Gamma[\bar{\zeta}^\pm]}{\delta\bar{\zeta}^{\pm_1}(x_1)\cdots\delta\bar{\zeta}^{\pm_n}(x_n)} \right|_{\bar{\zeta}^\pm=0}. \tag{2.42}$$

We might have defined this equation with some convention for the signs too, but it is simpler in this instance not to do so.

The 1PI functional $\Gamma[\bar{\zeta}^\pm]$ is also called the *effective action*. For a renormalizable theory, it is assumed that the effective vertices, $\Gamma^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n)$, have all been renormalized, when it is necessary to do so. The connected diagrams associated with the Green's functions $G_c^{\pm_1\cdots\pm_n}(x_1,\ldots,x_n)$, and which are calculated from all the tree and loop diagrams generated by the original action $S[\zeta^\pm]$, correspond to the tree-diagrams calculated using the effective action $\Gamma[\bar{\zeta}^\pm]$.

In the Bunch-Davies state, many of the 1PI effective vertices vanish. The structure of the effective action $\Gamma[\bar{\zeta}^\pm]$ mirrors, in part, the structure of the original action

$$S[\zeta^+] - S[\zeta^-], \tag{2.43}$$

which is composed entirely of operators containing just the $\zeta^+(t,\vec{x})$ or just the $\zeta^-(t,\vec{x})$ field, but with no operators coupling the two. Certainly any effective vertex that requires a counterterm of the same order will need also to have this same structure—for otherwise there would not have been the possibility of such a counterterm in the original action. This suggests that the effective vertices should vanish except when all of the indices are $+$ or all are $-$, and that they are related by a single sign,

$$\Gamma^{(n)}(x_1,\ldots,x_n) \equiv \Gamma^{+\cdots+}(x_1,\ldots,x_n) = -\Gamma^{-\cdots-}(x_1,\ldots,x_n). \tag{2.44}$$

Then the effective action becomes

$$\Gamma[\bar{\zeta}^\pm] = \Gamma_{\rm BD}[\bar{\zeta}^+] - \Gamma_{\rm BD}[\bar{\zeta}^-], \tag{2.45}$$

where

$$\Gamma_{\rm BD}[\bar{\zeta}] = \sum_{n=2}^{\infty} \frac{1}{n!} \int d^4x_1\cdots d^4x_n\, \Gamma^{(n)}(x_1,\ldots,x_n)\, \bar{\zeta}(x_1)\cdots\bar{\zeta}(x_n). \tag{2.46}$$

This is *not true* of a more general ground state. An initial state can be defined through an initial action defined on a $t = t_0$ hypersurface [18, 19]. A general initial action *does* contain operators that couple the ζ^+ and ζ^- fields directly, so the action no longer has the simple $S[\zeta^+] - S[\zeta^-]$ form of a Bunch-Davies state. Correspondingly, the effective action will also contain effective vertices with mixed $\pm$ indices.

Apart from the addition of the $\pm$ indices, all of these structures have their familiar analogues in S-matrix calculations. There is, however, a further difference here which makes calculations in this setting slightly more cumbersome. In a typical scattering problem, we are evolving a system from $t = -\infty$ to ∞. When we perform a Fourier transform, it is then in all $3 + 1$ dimensions, and the resulting Green's functions typically have simpler structures when expressed in terms of the four-momenta, rather than in terms of the original space-time coordinates. In a cosmological setting we are evolving only up to a finite time t_*, and we could have started from a finite initial time t_0 too. Therefore we shall only be Fourier transforming in the spatial dimensions, $(t, \vec{x}) \to (t, \vec{k})$. For example, the two-point 1PI Green's function remains a differential operator rather than becoming a purely algebraic function.

Let us illustrate this last point more fully with a particular example, which we shall need later anyway. The 1PI and the connected two-point functions are the functional inverses of each other. This relation follows from

$$\int d^4z \left\{ \frac{\delta J^+(y)}{\delta\bar{\zeta}^+(z)} \frac{\delta\bar{\zeta}^+(z)}{\delta J^+(x)} + \frac{\delta J^+(y)}{\delta\bar{\zeta}^-(z)} \frac{\delta\bar{\zeta}^-(z)}{\delta J^+(x)} \right\} = \sum_{s=\pm} \int d^4z \, \frac{\delta J^+(y)}{\delta\bar{\zeta}^s(z)} \frac{\delta\bar{\zeta}^s(z)}{\delta J^+(x)} = \delta^4(x-y), \tag{2.47}$$

for example. Here we have shown the case when both the sources are J^+'s; but we could have just as well written a similar relation for any choice of the signs,

$$\sum_{s=\pm} \int d^4z \, \frac{\delta J^{r_2}(y)}{\delta\bar{\zeta}^s(z)} \frac{\delta\bar{\zeta}^s(z)}{\delta J^{r_1}(x)} = \delta^{r_2}_{r_1} \, \delta^4(x-y), \tag{2.48}$$

where r_1 and r_2 can each assume either sign. This relation could also have been expressed as the functional derivatives of the appropriate generating functional,

$$-r_2 \sum_{s=\pm} s \int d^4z \, \frac{\delta^2 W}{\delta J^{r_1}(x)\delta J^s(z)} \frac{\delta^2 \Gamma}{\delta\bar{\zeta}^s(z)\delta\bar{\zeta}^{r_2}(y)} = \delta^{r_2}_{r_1} \, \delta^4(x-y), \tag{2.49}$$

which in turn becomes a relation between the two-point functions,

$$\sum_{s=\pm} \int d^4z \, G_c^{r_1 s}(x,z) \Gamma^{s r_2}(z,y) = i\delta^{r_2}_{r_1} \, \delta^4(x-y). \tag{2.50}$$

Fourier transforming in the spatial dimensions,

$$\begin{aligned} G_c^{rs}(t,\vec{x};t',\vec{y}) &= \int \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})} G_k^{rs}(t,t') \\ \Gamma^{rs}(t,\vec{x};t',\vec{y}) &= \int \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \Gamma_k^{rs}(t,t'); \end{aligned} \tag{2.51}$$

the statement that the two types of two-points functions are each other's functional inverse corresponds in this setting to the four equations

$$\begin{aligned} \int_{-\infty}^{t_*} dt'' \left\{ \Gamma_k^{++}(t,t'')\, G_k^{++}(t'',t') + \Gamma_k^{+-}(t,t'')\, G_k^{-+}(t'',t') \right\} &= i\,\delta(t-t') \\ \int_{-\infty}^{t_*} dt'' \left\{ \Gamma_k^{++}(t,t'')\, G_k^{+-}(t'',t') + \Gamma_k^{+-}(t,t'')\, G_k^{--}(t'',t') \right\} &= 0 \\ \int_{-\infty}^{t_*} dt'' \left\{ \Gamma_k^{--}(t,t'')\, G_k^{-+}(t'',t') + \Gamma_k^{-+}(t,t'')\, G_k^{++}(t'',t') \right\} &= 0 \\ \int_{-\infty}^{t_*} dt'' \left\{ \Gamma_k^{--}(t,t'')\, G_k^{--}(t'',t') + \Gamma_k^{-+}(t,t'')\, G_k^{+-}(t'',t') \right\} &= i\,\delta(t-t'). \end{aligned} \tag{2.52}$$

These equations remain true for an arbitrary initial state,[9] with only the tiny modification that if we had fixed the state at t_0, the lower limits would have been replaced by t_0. In the Bunch-Davies state, we can additionally assume that

$$\Gamma_k^{+-}(t,t'') = 0, \qquad \text{and} \qquad \Gamma_k^{-+}(t,t'') = 0. \tag{2.53}$$

These functional relations are between the *full* Green's functions; that is, the sums of *all* connected or 1PI graphs with the appropriate external structures. For the calculation of the consistency relation, we only require their leading behaviour. In that limit, the connected two-point functions are then just the free Feynman propagators and the 1PI two-point functions are the operators derived from the quadratic equation of motion for the field $\zeta_k(t)$,

$$\Gamma_k^{++}(t,t') = -\Gamma_k^{--}(t,t') = -\delta(t-t')\left\{ \frac{\dot\phi^2}{\dot\rho^2} e^{3\rho} \frac{d^2}{dt'^2} + \frac{d}{dt'}\left[\frac{\dot\phi^2}{\dot\rho^2} e^{3\rho} \right] \frac{d}{dt'} + k^2 \frac{\dot\phi^2}{\dot\rho^2} e^{\rho} \right\}. \tag{2.54}$$

2.4 A Slavnov-Taylor identity

Now, we are ready to derive a consistency condition between the two-point and three-point Green's functions by putting these ideas together. As our starting point,

[9]or rather, one that is not quite completely arbitrary—we are still assuming that the spatial translations and rotations are unbroken. Importantly, however, we do not need to assume that the state is invariant under dilations or special conformal transformations.

we use the invariance of the functional measure of the generating functional $Z[J^{\pm}]$ under a transformation of the fields,

$$\zeta(t,\vec{x}) \to \tilde{\zeta}(t,\vec{x}) = \zeta(t,\vec{x}) + \delta\zeta(t,\vec{x}). \tag{2.55}$$

Here, we are assuming that $\delta\zeta(t,\vec{x})$ is linear in the field, with the possibility of an inhomogeneous term too. When $\delta\zeta(t,\vec{x})$ has this form, the measure of the functional integral is invariant $\mathcal{D}\tilde{\zeta} = \mathcal{D}\zeta$. So under this change of the functional integration variable, $Z[J^{\pm}]$ has not changed,

$$\begin{aligned} Z[J^{\pm}] &= \int \mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\, e^{iS[\zeta^+]-iS[\zeta^-]+i\int_{-\infty}^{t_*}dt\int d^3\vec{x}\,[J^+\zeta^+-J^-\zeta^-]} \\ &= \int \mathcal{D}\tilde{\zeta}^+\,\mathcal{D}\tilde{\zeta}^-\, e^{iS[\tilde{\zeta}^+]-iS[\tilde{\zeta}^-]+i\int_{-\infty}^{t_*}dt\int d^3\vec{x}\,[J^+\tilde{\zeta}^+-J^-\tilde{\zeta}^-]}. \end{aligned} \tag{2.56}$$

When the transformation that we have made is additionally a symmetry of the action, as is the case for the conformal transformations mentioned earlier, we have that $S[\tilde{\zeta}] = S[\zeta]$, together with the invariance of the functional measure, yields

$$Z[J^{\pm}] = \int \mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\, e^{iS[\zeta^+]-iS[\zeta^-]+i\int_{-\infty}^{t_*}dt\int d^3\vec{x}\,[J^+(\zeta^++\delta\zeta^+)-J^-(\zeta^-+\delta\zeta^-)]}. \tag{2.57}$$

The only remnant of the transformation appears in its coupling to the sources. Under a small transformation, we can expand the exponential to linear order in $\delta\zeta^{\pm}$,

$$Z[J^{\pm}] = Z[J^{\pm}] + \delta Z[J^{\pm}], \tag{2.58}$$

where

$$\delta Z[J^{\pm}] = i\int_{-\infty}^{t_*} dt \int d^3\vec{x}\,\Big[J^+(t,\vec{x})\,\langle\delta\zeta^+(t,\vec{x})\rangle_{J^{\pm}} - J^-(t,\vec{x})\,\langle\delta\zeta^-(t,\vec{x})\rangle_{J^{\pm}}\Big]. \tag{2.59}$$

The $\langle\delta\zeta^{\pm}(t,\vec{x})\rangle_{J^{\pm}}$ are the expectation values of the infinitesimal change in the field in the presence of the source $J^{\pm}$. Because the generating functional has not changed, we conclude that

$$\delta Z[J^{\pm}] = 0. \tag{2.60}$$

This result is called the *Slavnov-Taylor identity*.[10]

The next step is to use this identity for the residual conformal symmetry to generate relations amongst Green's functions of different orders. The most important of these relations for current observations is the one associated with the spatial dilations. Under a dilation, we found that the field $\zeta(t,\vec{x})$ changes infinitesimally by an amount

$$\delta\zeta = \lambda\big[1 + \vec{x}\cdot\vec{\nabla}\zeta\big]. \tag{2.61}$$

[10]The derivation of the Slavnov-Taylor identity that we have presented here, essentially follows the reasoning of [17].

For a dilation then, the Slavnov-Taylor identity becomes

$$\int_{-\infty}^{t_*} dt \int d^3\vec{x} \left[J^+(t,\vec{x}) - J^-(t,\vec{x}) + J^+(t,\vec{x})\,\vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^\pm(t,\vec{x})} + J^-(t,\vec{x})\,\vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^\pm(t,\vec{x})} \right] = 0. \tag{2.62}$$

Here we have rewritten the one-point function in the presence of a source as

$$\langle \zeta^\pm(t,\vec{x})\rangle_{J^\pm} = \frac{1}{\pm i}\frac{\delta Z[J^\pm]}{\delta J^\pm(t,\vec{x})} = \pm Z[J^\pm]\frac{\delta W[J^\pm]}{\delta J^\pm(t,\vec{x})}, \tag{2.63}$$

and we have removed some non-vanishing prefactors. By differentiating this relation n times with respect to $\bar{\zeta}^\pm$ we generate relations between the $n+1$ and n-point functions.

While the higher-order relations are useful since they place constraints on the (so far) unobserved higher-order correlation functions of the primordial fluctuations, the most immediately important of these relations for observations is that between the three and two-point functions. The two-point function is the only correlation function that has been observed so far, and a decisive measurement of the three-point function is being sought. This relation then places a strong constraint on the amplitude of the three-point function if it is to be consistent with the simplest class of inflationary models with only one inflaton field and where the potential satisfies the slow-roll conditions.

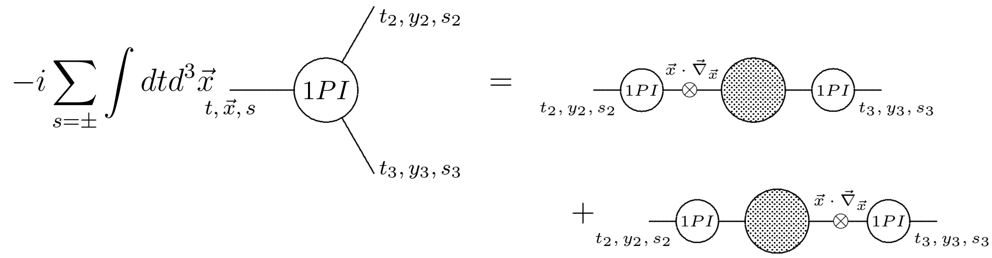
The calculation of this standard consistency relation goes as follows: (1) first differentiate the Slavnov-Taylor dilation identity twice with respect to $\bar{\zeta}^\pm$, setting $\bar{\zeta}^\pm = 0$ after doing so, and then (2) move some of the two-point functions from one set of terms to the other. The second step is more aesthetic—we essentially have the consistency relation after the first step—but it is useful nonetheless since it puts the relation in a form more closely resembling its usual expression elsewhere; and perhaps more importantly, it is only finite and nonzero on both sides of its equation after this second step.

So let us start by differentiating the Slavnov-Taylor identity with respect to the fields $\bar{\zeta}^{s_2}(y_2)$ and $\bar{\zeta}^{s_3}(y_3)$, where we have given them arbitrary $s_2, s_3 = \pm$ indices, to produce

$$\begin{aligned}\int d^4x \Bigg\{ &\frac{\delta^2 J^+(x)}{\delta\bar{\zeta}^{s_2}(y_2)\delta\bar{\zeta}^{s_3}(y_3)} - \frac{\delta^2 J^-(x)}{\delta\bar{\zeta}^{s_2}(y_2)\delta\bar{\zeta}^{s_3}(y_3)} \\ &+ \sum_{r,s=\pm}\int d^4z\, \frac{\delta J^r(x)}{\delta\bar{\zeta}^{s_2}(y_2)}\frac{\delta J^s(z)}{\delta\bar{\zeta}^{s_3}(y_3)}\vec{x}\cdot\vec{\nabla}_{\vec{x}}\frac{\delta W}{\delta J^r(x)\delta J^s(z)} \\ &+ \sum_{r,s=\pm}\int d^4z\, \frac{\delta J^r(x)}{\delta\bar{\zeta}^{s_3}(y_3)}\frac{\delta J^s(z)}{\delta\bar{\zeta}^{s_2}(y_2)}\vec{x}\cdot\vec{\nabla}_{\vec{x}}\frac{\delta W}{\delta J^r(x)\delta J^s(z)} \Bigg\} = 0. \end{aligned} \tag{2.64}$$

To keep the expressions more compact, we have again combined the time and space coordinates and written the arguments as a single four-vector, *e.g.* $x = (t, \vec{x})$.

The first pair of terms is not quite what we want. Cosmological observations are used to constrain the *connected* three-point functions of the primordial fluctuations. What we have here are instead the 1PI three-point functions. If we recall how the functional derivatives of $J^{\pm}$ and W are related to the 1PI and the connected Green's functions, the above equation can be expressed diagrammatically as



(2.65)

On the left side of this equation we are summing over all possible space-time points for the corresponding leg[11]; on the right the $\vec{x}\cdot\vec{\nabla}_{\vec{x}}$ operator is acting on the connected two-point Green's function from both sides.

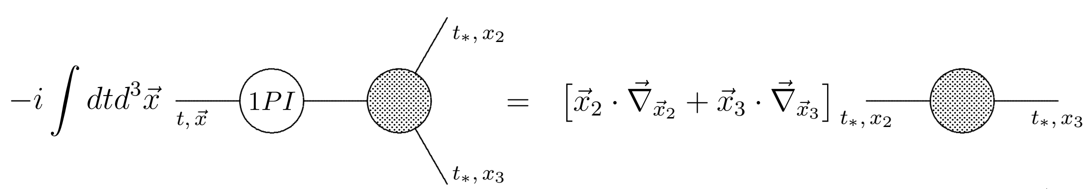
Since the connected and 1PI two-point functions are the functional inverses of each other, if we act on this equation with a pair of propagators, we amputate the 1PI legs on the right side and produce an expression on the left which can be written as a 1PI two-point function acting on the connected three-point function as we wanted,

$$-i\sum_{s=\pm}\int d^4y\int d^4x\left\{\Gamma^{+s}(y,x)\,G_c^{s++}(x,x_2,x_3)+\Gamma^{-s}(y,x)\,G_c^{s++}(x,x_2,x_3)\right\}$$
$$=\left[\vec{x}_2\cdot\vec{\nabla}_{\vec{x}_2}+\vec{x}_3\cdot\vec{\nabla}_{\vec{x}_3}\right]G_c^{++}(x_2,x_3). \quad (2.66)$$

The full details for how to derive this equation appear in the supplementary section 2.8. Since for the Bunch-Davies state $\Gamma^{+-}=\Gamma^{-+}=0$, we have just

$$-i\int d^4y\int d^4x\left\{\Gamma^{++}(y,x)\,G_c^{+++}(x,x_2,x_3)+\Gamma^{--}(y,x)\,G_c^{-++}(x,x_2,x_3)\right\}$$
$$=\left[\vec{x}_2\cdot\vec{\nabla}_{\vec{x}_2}+\vec{x}_3\cdot\vec{\nabla}_{\vec{x}_3}\right]G_c^{++}(x_2,x_3). \quad (2.67)$$

This is the consistency relation between the three and two-point functions expressed in a fully time-evolving formalism. The diagrammatic version of this identity may be written



(2.68)

[11]Another note for the experts: if we think of the times as defined on the full time-contour, the $\pm$ indices (s_2 and s_3) are naturally absorbed into the contour time-coordinate.

This relation holds for the *full* Green's functions evaluated at any points and at any time during the inflationary era. But to understand what it implies for the observations of our universe, we evaluate it at points on a late-time hypersurface at t_*. Only two of the points in the relation are external, so we choose them lie somewhere on the this late-time hypersurface,

$$x_2 = (t_*, \vec{x}_2), \qquad x_3 = (t_*, \vec{x}_3). \tag{2.69}$$

t_* is also meant to be the endpoint of the evolution, so it also appears as the upper limit of the time integrals,

$$\begin{aligned} -i \int_{-\infty}^{t_*} dt' \int d^3\vec{y} \int_{-\infty}^{t_*} dt \int d^3\vec{x} \left\{ \Gamma^{++}(y,x)\, G_c^{+++}(x,x_2,x_3) + \Gamma^{--}(y,x)\, G_c^{-++}(x,x_2,x_3) \right\} \\ = \left[\vec{x}_2 \cdot \vec{\nabla}_{\vec{x}_2} + \vec{x}_3 \cdot \vec{\nabla}_{\vec{x}_3} \right] G_c^{++}(x_2,x_3), \end{aligned}$$

where $y = (t', \vec{y})$ and $x = (t, \vec{x})$. The relation is more conventionally expressed in terms of the spatial momenta,

$$\begin{aligned} -i \int_{-\infty}^{t_*} dt' \int_{-\infty}^{t_*} dt \left\{ \Gamma_{\vec{0}}^{++}(t',t)\, G_c^{+++}(t,\vec{0};t_*,\vec{k}_2;t_*,\vec{k}_3) + \Gamma_{\vec{0}}^{--}(t',t)\, G_c^{-++}(t,\vec{0};t_*,\vec{k}_2;t_*,\vec{k}_3) \right\} \\ = -\left[3 + \vec{k}_2 \cdot \nabla_{\vec{k}_2} \right] G_c^{++}(t_*,\vec{k}_2;t_*,\vec{k}_3). \end{aligned}$$

While we have not written it, both sides are implicitly multiplied by a momentum-conserving $\delta^3(\vec{k}_2+\vec{k}_3)$. The fact that we have integrated out one of the coordinates of the three-point function means that only the constant, $\vec{k} = \vec{0}$, behaviour associated with that coordinate appears in the Fourier-transformed expression.

In this form, both sides of the consistency relation are finite. On its left side, this property is a little more subtle, since the terms are products of something that diverges with something that vanishes in the limit

$$\lim_{\vec{k}\to\vec{0}} \left[\Gamma_{\vec{k}}^{\pm\pm}(t',t)\, G_c^{\pm++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) \right] \to \text{finite and nonzero.} \tag{2.70}$$

What we really mean by the left side of the relation then is this limit.

The consistency relation is a statement about the full, 'all orders in perturbation theory' Green's functions. For the primordial fluctuations there is no need to go beyond the tree-level contributions—very often there is not even a need to go further than the leading expression in the slow-roll limit. Only the amplitude and the leading information about the scale-dependence of the two-point function have been observed. The three-point function has yet to be detected, though observations have constrained it to be quite small. Therefore, using tree-level expressions for the 1PI two-point functions, found at the end of the last section, and integrating over a trivial δ-function, we obtain our final expression for the consistency relation between the connected two and three-point functions,

$$
i \lim_{\vec{k}\to\vec{0}} \int_{-\infty}^{t_*} dt \left\{ \frac{\dot{\phi}^2}{\dot{\rho}^2} e^{3\rho} \frac{d^2}{dt^2} + \frac{d}{dt}\left[\frac{\dot{\phi}^2}{\dot{\rho}^2} e^{3\rho} \right] \frac{d}{dt} + k^2 \frac{\dot{\phi}^2}{\dot{\rho}^2} e^{\rho} \right\}
$$

$$
\left\{ G_c^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_c^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) \right\}
$$

$$
= -\big[3 + \vec{k}_2 \cdot \nabla_{\vec{k}_2}\big] G_c^{++}(t_*,\vec{k}_2;t_*,\vec{k}_3). \quad (2.71)
$$

Once again, there is an implicit factor of $\delta^3(\vec{k}_2 + \vec{k}_3)$ multiplying both sides.

Although it appears a little different from how the relation has often been written, the expression that we have found is in fact *formally* the same. When the correlation functions are derived from a path integral effectively defined only on the late-time hypersurface, as was done in [12], the consistency relation assumes the form

$$
\frac{G_c^{(3)}(\vec{0},\vec{k}_2,-\vec{k}_2)}{P(0)} = -\left[3 + k_2 \frac{\partial}{\partial k_2}\right] P(k_2). \quad (2.72)
$$

Here the Green's functions are the equal-time correlation functions evaluated at t_*;

$$
G_c^{(3)}(\vec{k}_1,\vec{k}_2,\vec{k}_3) = G_c^{+++}(t_*,\vec{k}_1;t_*,\vec{k}_2;t_*,\vec{k}_3) \quad (2.73)
$$

and $P(k) = G_c^{++}(t_*,\vec{k};t_*,-\vec{k})$ is the power spectrum. If we recall that $[P(0)]^{-1}$ in that language is the 1PI two-point function evaluated in the zero-momentum limit, we see that this is structurally the same as the result above which has included the full time evolution.

It is nonetheless an instructive calculation to show how this relation applies to the simple class of slow-roll, single-field, models of inflation that Maldacena originally considered [4]. This calculation is presented in the supplementary section 2.7 of this chapter. It has been provided in full because—as was mentioned—the left side contains a delicate, but finite, balance between diverging and vanishing factors. We have also included it because some of the techniques, though they are becoming more commonly used in inflation, are still perhaps not as familiar as they ought to be.

2.5 The Slavnov-Taylor identity with an initial state

The advantage of following the full time evolution of the Green's functions is that it allows us to treat more general initial states. We are no longer shackled to the Bunch-Davies state, or even to a particular class of symmetric initial states.

Let us consider a universe where the scalar fluctuations $\zeta(t,\vec{x})$ begin in an arbitrary state at $t = t_0$. The information about this initial state can be incorporated into the generating functional by including a density matrix $\rho(t_0)$ for that state [18],

$$Z[J^\pm] = \int \mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\,\rho(t_0)\exp\Big\{iS[\zeta^+] - iS[\zeta^-] + i\int_{t_0}^{t_*} dt\int d^3\vec{x}\,\big[J^+\zeta^+ - J^-\zeta^-\big]\Big\}. \tag{2.74}$$

We write this matrix as an action defined along the initial-time hypersurface at t_0,

$$\rho(t_0) = e^{iS_0[\zeta^+(t_0,\vec{x}),\zeta^-(t_0,\vec{x})]}; \tag{2.75}$$

this trick allows us to put it together with the rest of the action,

$$Z[J^\pm] = \int \mathcal{D}\zeta^+\,\mathcal{D}\zeta^-\,\exp\Big\{iS[\zeta^+]-iS[\zeta^-]+iS_0[\zeta^+,\zeta^-]+i\int_{t_0}^{t_*} dt\int d^3\vec{x}\,\big[J^+\zeta^+-J^-\zeta^-\big]\Big\}. \tag{2.76}$$

Each part of this action has its own meaning in the interaction picture: the quadratic terms in S define the time-dependence of the fields in the free theory, the higher-order interactions in S determine the evolution of the state, and the terms in S_0 determine the initial state. Whether we decide to group the quadratic terms of S_0 with those of S when solving for the time-dependence of the free theory depends on the particular state. When the initial state differs substantially from the Bunch-Davies state, so that a perturbative expansion provides a poor approximation, it is necessary to include all quadratic terms in the free part; but in cosmological examples, where the state does not appear to differ too much in its two-point structure from the Bunch-Davies state, it can sometimes be more convenient to regard all of S_0 as a part of the interactions.

The initial action is arranged as a series of terms according to powers of the fluctuations,[12]

$$S_0[\zeta^\pm] = S_0^{(1)}[\zeta^\pm] + S_0^{(2)}[\zeta^\pm] + S_0^{(3)}[\zeta^\pm] + S_0^{(4)}[\zeta^\pm] + \cdots, \tag{2.77}$$

where $S_0^{(2)}$ is quadratic in $\zeta^\pm$, $S_0^{(3)}$ is cubic, *etc.* As long as the higher-order terms in this series are sufficiently small, their effects can be treated perturbatively. Since the observed universe appears to be consistent with having a rather small non-Gaussian primordial component, we usually assume that we are within this regime. The requirement that the initial density matrix is real, $\rho(t_0) = \rho^\dagger(t_0)$, imposes constraints on the form of S_0. For example, when the state is translationally and rotationally invariant, the quadratic terms without time-derivatives of the field are specified by two functions,

$$\begin{aligned} S_0^{(2)}[\zeta^\pm] \;=\; & -\frac{1}{2}\int d^3\vec{x}\,d^3\vec{y}\,\Big\{\zeta^+(t_0,\vec{x})A(\vec{x}-\vec{y})\zeta^+(t_0,\vec{y}) - \zeta^-(t_0,\vec{x})A^*(\vec{x}-\vec{y})\zeta^-(t_0,\vec{y}) \\ & \qquad + \zeta^+(t_0,\vec{x})iB(\vec{x}-\vec{y})\zeta^-(t_0,\vec{y}) + \zeta^-(t_0,\vec{x})iB(\vec{x}-\vec{y})\zeta^+(t_0,\vec{y})\Big\} \\ & + \cdots, \end{aligned} \tag{2.78}$$

[12] The linear term in $\zeta^\pm$ is included when it is necessary to cancel tadpole graphs involving the initial time. We still are imposing the condition that the expectation value of $\zeta^\pm$ vanishes at all times.

where $A(\vec{x}-\vec{y})$ can be complex but $B(\vec{x}-\vec{y})$ must be real. The '$\cdots$' refer to further terms that could contain time derivatives of ζ. Similarly, the cubic terms, subject to the same assumptions, are fixed by two complex functions,

$$\begin{aligned} S_0^{(3)}[\zeta^\pm] &= -\frac{1}{6}\int d^3\vec{x}\, d^3\vec{y}\, d^3\vec{z} \Big\{ C(\vec{x},\vec{y},\vec{z})\, \zeta^+(t_0,\vec{x})\zeta^+(t_0,\vec{y})\zeta^+(t_0,\vec{z}) \\ &\qquad - C^*(\vec{x},\vec{y},\vec{z})\, \zeta^-(t_0,\vec{x})\zeta^-(t_0,\vec{y})\zeta^-(t_0,\vec{z}) \\ &\qquad + 3D(\vec{x},\vec{y},\vec{z})\, \zeta^+(t_0,\vec{x})\zeta^+(t_0,\vec{y})\zeta^-(t_0,\vec{z}) \\ &\qquad - 3D^*(\vec{x},\vec{y},\vec{z})\, \zeta^+(t_0,\vec{x})\zeta^-(t_0,\vec{y})\zeta^-(t_0,\vec{z}) + \cdots \Big\}, \end{aligned} \tag{2.79}$$

In writing the cubic surface action in this form we have tacitly assumed that the functions $C(\vec{x},\vec{y},\vec{z})$ and $D(\vec{x},\vec{y},\vec{z})$ are invariant under permutations of their coordinates. Again, the cubic action could contain further operators with $\dot{\zeta}^\pm(t_0,\vec{x})$ and higher derivatives. Indeed, in the course of renormalizing the effects of the standard cubic operators that occur in inflation, operators with time derivatives do occur as counterterms in the initial action [22].

Once we have included operators that contain odd numbers of the field, radiative corrections will typically produce tadpole graphs. If they are to be cancelled, there must be linear terms in the action for the initial state,

$$S_0^{(1)}[\zeta^\pm] = -\int d^3\vec{x}\, \Big\{ T(\vec{x})\zeta^+(t_0,\vec{x}) - T^*(\vec{x})\zeta^-(t_0,\vec{x}) \Big\} + \cdots . \tag{2.80}$$

When the initial state is translationally and rotationally invariant, we expect that $T(\vec{x}) = T$ is a constant; but we shall leave it in this more general form for now.

Typically we shall find it more convenient to express the functions describing the state through their Fourier transforms,

$$A(\vec{x}-\vec{y}) = \int \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})} A_k, \tag{2.81}$$

and

$$C(\vec{x},\vec{y},\vec{z}) = \int \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\, e^{i(\vec{k}_1\cdot\vec{x}+\vec{k}_2\cdot\vec{y}+\vec{k}_3\cdot\vec{z})}\, (2\pi)^3\, \delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3)\, C_{\vec{k}_1,\vec{k}_2,\vec{k}_3}, \tag{2.82}$$

for example. Other than requiring that the higher energy modes are not so abundantly populated that they overwhelm the energy density of the inflationary background, these structures can have a more or less arbitrary dependence on the momentum.

What sorts of initial states might we consider? The initial time could be viewed in a variety of ways: it could be simply a theoretical crutch, a cut-off that we impose to avoid considering the very early and correspondingly very short-distance behaviour of the theory. On the other hand, it could genuinely be seen as a time when 'something' happened—the beginning of the inflationary expansion or the end of some other dynamics still within inflation but beyond the minimal inflationary picture. In the latter cases the moments before t_0 might have bequeathed the inflationary era with a state that was different from the standard Bunch-Davies state. If we are taking a completely unprejudiced view of what happened before inflation, it is not *necessary* to assume that the initial state has the same symmetries of the single-field inflationary picture. For computational convenience we shall still restrict to an initial state that is invariant under spatial translations and rotations. But it is possible to allow initial states that break the spatial conformal symmetries. These broken symmetries lead to interesting modifications of the standard consistency relations that we derived in the earlier sections. In some instances too the conformal symmetries might be preserved though the state is not the Bunch-Davies state.

A state that is not invariant under the residual conformal symmetry will introduce additional terms in the primitive Slavnov-Taylor identity. As in the earlier derivation of this identity, suppose that we make an infinitesimal conformal transformation of the field $\zeta(t,\vec{x})$,

$$\zeta(t,\vec{x}) \to \tilde{\zeta}(t,\vec{x}) = \zeta(t,\vec{x}) + \delta\zeta(t,\vec{x}). \tag{2.83}$$

For the purpose of deriving the Slavnov-Taylor identity, it is sufficient if this transformation is a symmetry of the action $S[\tilde{\zeta}] = S[\zeta]$ and if the measure of the functional integral changes at most by a functional constant, which can be absorbed into the normalization of $Z[J^\pm]$. Both of these properties hold for the minimal inflationary picture. When we write the generating functional in terms of $\tilde{\zeta} = \zeta + \delta\zeta$, by expanding the quantities that are not invariant under the symmetry to linear order, we obtain the change in the generating functional

$$\begin{aligned} Z[J^\pm] &= \int \mathcal{D}\tilde{\zeta}^\pm \, e^{iS[\tilde{\zeta}^+]-iS[\tilde{\zeta}^-]+iS_0[\tilde{\zeta}^\pm]+i\int_{t_0}^{t_*}dt\int d^3\vec{x}\,[J^+\tilde{\zeta}^+-J^-\tilde{\zeta}^-]} \\ &= \int \mathcal{D}\zeta^\pm \, e^{iS[\zeta^+]-iS[\zeta^-]+iS_0[\tilde{\zeta}^\pm]+i\int_{t_0}^{t_*}dt\int d^3\vec{x}\,[J^+\tilde{\zeta}^+-J^-\tilde{\zeta}^-]} \\ &= Z[J^\pm] + \delta Z[J^\pm], \end{aligned} \tag{2.84}$$

where

$$\delta Z[J^\pm] = i\int_{t_0}^{t_*} dt \int d^3\vec{x} \left\{ J^+(t,\vec{x})\langle\delta\zeta^+(t,\vec{x})\rangle_{J^\pm} - J^-(t,\vec{x})\langle\delta\zeta^-(t,\vec{x})\rangle_{J^\pm} \right\} + i\langle\delta S_0[\zeta^\pm]\rangle_{J^\pm}. \tag{2.85}$$

δS_0 represents the linear part of $S_0[\tilde{\zeta}^\pm] - S_0[\zeta^\pm]$ when expanded in powers of $\delta\zeta^\pm$. The result,

$$\delta Z[J^\pm] = 0, \tag{2.86}$$

is again the 'ur-statement' of the Slavnov-Taylor identity; however, it now contains additional contributions from the initial state. From this identity we generate relations amongst various Green's functions by differentiating it with respect to $\bar{\zeta}^\pm(t,\vec{x})$ an arbitrary number of times.

The most important class of consistency relations for inflation are those generated by dilations of the spatial part of the metric. In its infinitesimal form, a dilation changes $\zeta(t,\vec{x})$ by

$$\delta\zeta = \lambda\big[1 + \vec{x}\cdot\vec{\nabla}\zeta\big]. \tag{2.87}$$

When this expression is put into δZ, and the one-point functions are written in terms of the generating function for the connected Green's functions, $iW[J^\pm] = \ln Z[J^\pm]$, the Slavnov-Taylor identity becomes

$$\int_{t_0}^{t_*} dt \int d^3\vec{x}\, \biggl\{ J^+(t,\vec{x}) - J^-(t,\vec{x}) + J^+(t,\vec{x})\, \vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^+(t,\vec{x})} + J^-(t,\vec{x})\, \vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^-(t,\vec{x})} \biggr\} \\ = -\frac{1}{\lambda}\frac{1}{Z[J^\pm]}\langle \delta S_0[\zeta^\pm]\rangle_{J^\pm}. \tag{2.88}$$

We can broadly divide the initial states into two classes. We define the first class to include those states that share the same conformal invariance—or here, the invariance under dilations—as as the inflationary metric. The second class, where $\delta S_0 \neq 0$, represents the most general case.

2.5.1 Conformally invariant initial states

So far the form of the initial state $S_0[\zeta^\pm]$ has been left largely arbitrary. One formally simple class of initial states are those that are invariant under the same residual conformal symmetry of the inflationary metric. Requiring that the state be conformally invariant imposes conditions on how the n-point structures of the initial state change under a conformal transformation.

Under a dilation the scalar fluctuation transforms inhomogeneously, so the terms in $\delta S_0^{(n)}$ have either n or $n-1$ factors of the field $\zeta^\pm$. This means that if we wish to have an initial state that is invariant under dilations, the change in the $n^{\rm th}$ order structure function is determined in part by the original function plus the integral of some linear combination of the $(n+1)^{\rm st}$ structure functions. For example, if under a dilation

$$A(\vec{x}-\vec{y}) \to \tilde{A}(\vec{x}-\vec{y}) = A(\vec{x}-\vec{y}) + \lambda\, \delta A(\vec{x}-\vec{y}), \tag{2.89}$$

then $\delta S_0 = 0$ when

$$\delta A(\vec{x}-\vec{y}) = \big[6 + \vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big]\, A(\vec{x}-\vec{y}) - \int d^3\vec{z}\, \big[C(\vec{x},\vec{y},\vec{z}) + D(\vec{x},\vec{y},\vec{z})\big]. \tag{2.90}$$

Similarly, an invariant state should also have

$$\delta B(\vec{x}-\vec{y}) = \big[6 + \vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big]\, B(\vec{x}-\vec{y}) + i\int d^3\vec{z}\, \big[D(\vec{x},\vec{y},\vec{z}) - D^*(\vec{x},\vec{y},\vec{z})\big]. \tag{2.91}$$

The parts of the variation that are linear in the field require

$$\delta T(\vec{x}) = \big[3 + \vec{x}\cdot\vec{\nabla}_{\vec{x}}\big]T(\vec{x}) - \int d^3\vec{y}\, \big[A(\vec{x}-\vec{y}) + iB(\vec{x}-\vec{y})\big]. \tag{2.92}$$

However, if were a imagining that the linear term has only been include to cancel tadpole graphs, so that it vanishes as we turn off the cubic structures and cubic interactions, then we might wish to have the linear and quadratic contributions to this equation vanish separately,

$$\delta T(\vec{x}) = \big[3 + \vec{x}\cdot\vec{\nabla}_{\vec{x}}\big] T(\vec{x}) \quad \text{and} \quad \int d^3\vec{y}\,\big[A(\vec{x}-\vec{y}) + iB(\vec{x}-\vec{y})\big] = 0. \qquad (2.93)$$

Finally, the zeroth order part of the variation of the initial action under a dilation vanishes when

$$\int d^3\vec{x}\,\big[T(\vec{x}) - T^*(\vec{x})\big] = 0, \qquad (2.94)$$

which can be easily satisfied for a real $T(\vec{x})$.

The Fourier-transformed versions of these conditions are

$$\begin{aligned} \delta A_k &= \big[3 - \vec{k}\cdot\vec{\nabla}_{\vec{k}}\big]\, A_k - C_{-\vec{k},\vec{k},\vec{0}} - D_{-\vec{k},\vec{k},\vec{0}} \\ \delta B_k &= \big[3 - \vec{k}\cdot\vec{\nabla}_{\vec{k}}\big]\, B_k + i\big[D_{-\vec{k},\vec{k},\vec{0}} - D^*_{-\vec{k},\vec{k},\vec{0}}\big] \end{aligned} \qquad (2.95)$$

and

$$A_0 + iB_0 = 0, \qquad (2.96)$$

together with

$$\delta T_{\vec{k}} = -\vec{k}\cdot\vec{\nabla}_{\vec{k}} T_{\vec{k}}. \qquad (2.97)$$

The condition $A_0 + iB_0 = 0$ refers to the spatially invariant part of the initial two-point structure; usually we assume that in this long-distance limit the state matches with the state that we have chosen as our reference state, and with respect to which we are defining the excited state. Therefore, it is natural to let $A_0 = 0$ and $B_0 = 0$ and we assume that this is so even when we consider states that break the residual conformal symmetry of the background.

For a conformally invariant initial state, the statement of the Slavnov-Taylor identity *looks* exactly as it did for the Bunch-Davies state,

$$\int_{t_0}^{t_*} dt \int d^3\vec{x}\,\bigg\{ J^+(t,\vec{x}) - J^-(t,\vec{x}) + J^+(t,\vec{x})\,\vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^+(t,\vec{x})} + J^-(t,\vec{x})\,\vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^-(t,\vec{x})}\bigg\} = 0. \qquad (2.98)$$

Taking functional derivatives with respect to $\bar{\zeta}^\pm$ produces exactly the same infinite tower of consistency relations between n and $n+1$-point correlation functions of the fluctuations produced by inflation. However, this appearance is deceptive. The consistency relations *amongst* the Green's functions have not changed, but the Green's functions themselves will be different.[13] They are no longer the Green's functions for the Bunch-Davies state but are rather those for the appropriate initial state. Phenomenologically this can lead to the case where the standard consistency relations

[13] The Green's functions for a general initial state are presented in [18] and [19].

appears to be violated. But this spurious violation has come about not because the wrong consistency relations are being used, but rather because the wrong n-point functions have been assumed. This interesting case is explored in [23].

2.5.2 Broken conformal invariance in the initial state

On the other hand, there is no reason that the initial state must itself be invariant under the residual conformal symmetries of inflation. The dynamics prior to t_0 did not need to have the same symmetries as what happened after t_0. This freedom opens up much richer families of possibilities for the initial state. The degree to which observations are in accord or violate the standard consistency relations can then be used to constrain the possible n-point structures of the initial state and how they transform conformally. This more general case will be treated fully in later work [20], but it is useful to explain the outlines of the calculation here.

An interesting thing happens when we try to consider states that are not invariant under the full conformal symmetry group. When we differentiate the Slavnov-Taylor identity $n > 1$ times with respect to $\bar{\zeta}^{\pm}$ we obtain a non-vanishing result on the left side, which resembles the standard consistency relations except that, as for the conformally invariant initial state, the correlation functions are not those of the Bunch-Davies state. However, unlike the invariant state, the functional derivative of the right side of the generalised Slavnov-Taylor identity does not vanish.

When we differentiate the Slavnov-Taylor identity *exactly* once, we obtain a tadpole condition on the initial state,

$$-\frac{\delta}{\delta\bar{\zeta}^{+}(t,\vec{x})}\frac{\langle\delta S_0[\zeta^{\pm}]\rangle_{J^{\pm}}}{\lambda Z[J^{\pm}]} = 0, \tag{2.99}$$

which can be preserved by choosing the one-point structure, $T(\vec{x})$, appropriately. Once the tadpole has been fixed, it does not appear in any of the higher consistency relations—more than one functional derivative of the tadpole term with respect to $\bar{\zeta}^{\pm}$ annihilates it completely.

Our purpose here has been to develop a formalism that follows the full time-evolution within the Slavnov-Taylor relation, which allows it to be applied to arbitrary initial states. These applications will be treated systematically elsewhere, but the basic recipe is as follows: (1) choose a set of initial state structures and (2) specify how they transform under a conformal transformation. In practice, since the consistency relations will be far more complicated, we might wish to make a few reasonable assumptions about the sizes of the higher order correlation functions relative to the lower order ones to be able to neglect some of the terms in the relation.

The contribution from the variation of the initial state can be rather complicated, even for the simplest structures. For example, the contribution from the quadratic terms of S_0 to the right side of this Slavnov-Taylor identity for a dilation is

$$\begin{aligned}
&\int_{t_0}^{t_*} dt \int d^3\vec{x} \left\{ J^+(t,\vec{x}) - J^-(t,\vec{x}) + J^+(t,\vec{x})\,\vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^+(t,\vec{x})} + J^-(t,\vec{x})\,\vec{x}\cdot\vec{\nabla}\frac{\delta W[J^\pm]}{\delta J^-(t,\vec{x})} \right\} \\
&= \frac{1}{2}\int d^3\vec{x}\, d^3\vec{y} \left\{ A(\vec{x}-\vec{y}) \left[\frac{\delta W[J^\pm]}{\delta J^+(x)} + \frac{\delta W[J^\pm]}{\delta J^+(y)} - i\big[\vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big] \frac{\delta^2 W[J^\pm]}{\delta J^+(x)\delta J^+(y)} \right.\right. \\
&\qquad\qquad \left. + \big[\vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big] \frac{\delta W[J^\pm]}{\delta J^+(x)} \frac{\delta W[J^\pm]}{\delta J^+(y)} \right] \\
&\qquad + \delta A(\vec{x}-\vec{y}) \left[-i\frac{\delta^2 W[J^\pm]}{\delta J^+(x)\delta J^+(y)} + \frac{\delta W[J^\pm]}{\delta J^+(x)} \frac{\delta W[J^\pm]}{\delta J^+(y)} \right] \\
&\qquad + A^*(\vec{x}-\vec{y}) \left[\frac{\delta W[J^\pm]}{\delta J^-(x)} + \frac{\delta W[J^\pm]}{\delta J^-(y)} + i\big[\vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big] \frac{\delta^2 W[J^\pm]}{\delta J^-(x)\delta J^-(y)} \right. \\
&\qquad\qquad \left. - \big[\vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big] \frac{\delta W[J^\pm]}{\delta J^-(x)} \frac{\delta W[J^\pm]}{\delta J^-(y)} \right] \\
&\qquad + \delta A^*(\vec{x}-\vec{y}) \left[i\frac{\delta^2 W[J^\pm]}{\delta J^-(x)\delta J^-(y)} - \frac{\delta W[J^\pm]}{\delta J^-(x)} \frac{\delta W[J^\pm]}{\delta J^-(y)} \right] \\
&\qquad + 2iB(\vec{x}-\vec{y}) \left[\frac{\delta W[J^\pm]}{\delta J^+(x)} - \frac{\delta W[J^\pm]}{\delta J^-(y)} + i\big[\vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big] \frac{\delta^2 W[J^\pm]}{\delta J^+(x)\delta J^-(y)} \right. \\
&\qquad\qquad \left. - \big[\vec{x}\cdot\vec{\nabla}_{\vec{x}} + \vec{y}\cdot\vec{\nabla}_{\vec{y}}\big] \frac{\delta W[J^\pm]}{\delta J^+(x)} \frac{\delta W[J^\pm]}{\delta J^-(y)} \right] \\
&\qquad \left. + 2i\delta B(\vec{x}-\vec{y}) \left[i\frac{\delta^2 W[J^\pm]}{\delta J^+(x)\delta J^-(y)} - \frac{\delta W[J^\pm]}{\delta J^+(x)} \frac{\delta W[J^\pm]}{\delta J^-(y)} \right] + \cdots \right\}, \qquad (2.100)
\end{aligned}$$

and there are, of course, contributions from the cubic operators in $S_0^{(3)}$, the quartic operators, *etc.* Differentiating this identity twice with respect to $\bar{\zeta}^\pm$ yields a relation between two and three-point correlators on the left side and a number of new terms on the right.

2.6 Concluding remarks

The symmetry that remains even after we have made a general choice for the quantum fluctuations about an inflationary background leads to relations amongst their correlation functions. We have seen that these relations are in fact the Slavnov-Taylor identities associated with this residual symmetry. Because the approach that we have presented here only relies on symmetries of the metric and on very general properties of any quantum field theory, our results apply equally generally and are largely independent of the detailed properties of the particular model.

Throughout our derivation we have kept the evolution of the state intact and explicit. This allows us immediately to adapt our approach to inflationary theories that start in more or less arbitrary initial states. Having this freedom allows us to treat more general situations in inflation. It also frees us from the need to make assumptions about the behaviour of the universe at asymptotically distant times and at infinitesimally small scales. Since we do not know what might have occurred before an inflationary expansion, and moreover since there are dangers inherent in extending quantum field theories in inflating backgrounds arbitrarily far back in time—our understanding of nature is inadequate beyond certain scales—having the ability to choose other initial states and having the freedom to start at a finite initial time allows us to explore richer sets of possibilities. Even viewed more conservatively, the formalism we have developed here lets us parametrise by how much the state can differ from a purely Bunch-Davies state—and then constrain these departures experimentally—rather than assume that the universe is in this state *ab initio*. The same analysis equally applies to the tensor fluctuations, about whose state far less is known.

A general initial state modifies a consistency relation in two ways. The Green's functions within the relation change to reflect the influence of the initial density matrix on observables, and the non-invariance of the initial state also alters the basic form of the Slavnov-Taylor identity from which the consistency relations are derived. Because the underlying dynamics are still invariant under the residual conformal symmetries in the metric, it is still possible to derive a Slavnov-Taylor identity as long as we have included the appropriate corrections generated by the change of the initial density matrix under the conformal symmetry transformations. These new terms in the identity are the cosmological analogues of the corrections that appear in the Ward identities of gauge theories in the presence of explicit symmetry breaking.

In this article we have derived the basic effect of an initial state on the Slavnov-Taylor identity. We shall explore the possible observational effects of having non-Bunch-Davies and non-invariant states more fully in [20] using the formalism that we have developed here. Most importantly, it would be interesting to see what the known properties of the power spectrum, and the constraints on the non-Gaussianties, are able to tell us about the state during inflation.

While we have concentrated here on the consistency relation between the two and three-point functions for the scalar field—mainly to illustrate a new method with a familiar example—there are many further quantities to compute. By differentiating the 'ur-form' of the Slavnov-Taylor identity, $\delta Z[J^\pm] = 0$, n times with respect to $\bar\zeta^\pm(t,\vec{x})$, we generate higher-order consistency relations between n and $n+1$-point correlators. Most of these are perhaps of a more formal interest, since thus far it has been difficult even to detect the amplitude of the three-point correlator. Initial states could easily enhance some of these higher-order correlators without disturbing what we already know about the power spectrum. Additionally, it would be interesting to study the consistency relation associated with the special conformal transformations.

In his original derivation of the non-Gaussianities in inflation, Maldacena found further consistency relations satisfied by three-point functions which contained any combination of the tensor and scalar fluctuations. This formalism can also be applied to such correlators.

2.7 Supplement: Evaluating the consistency relation

In this supplementary section we show that the standard slow-roll inflationary models with a single scalar inflaton field satisfy the consistency relation in the form that we have derived it, giving a—by now—standard estimate of the amplitude and scaling of the three-point function in the limit where one of the external momenta is soft. While this conclusion is nothing new in itself, there are many technical points in treating and using the 1PI Green's functions which need to be thoroughly understood.

Before beginning this analysis, we mention a few of these points: (1) Maldacena's trick of shifting the field in order to remove certain terms from the cubic action is not convenient here. It would also entail a shift in the 1PI Green's function which is not so easy to do. Fortunately, the leading parts of the action are simple enough to evaluate without the need to make this shift. (2) In the limit in which the momentum of one of the external legs gets soft, the three-point function diverges. But when we act upon it with a 1PI two-point function, we obtain a finite result. At tree-level, the 1PI two-point function is

$$\Gamma_k^{++}(t',t) = -\Gamma_k^{--}(t',t) = -\delta(t'-t)\left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho}\frac{d^2}{dt^2} + \frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho}\right]\frac{d}{dt} + k^2\frac{\dot\phi^2}{\dot\rho^2}e^{\rho}\right\}. \quad (2.101)$$

In the Bunch-Davies state the parts containing the time-derivatives *vanish* in the late-time limit giving a vanishing result when acting on the three-point function, for reasons that we shall explain. We also should set the k^2 in the spatial derivative term *only after* this operator has acted on the three-point functions.

2.7.1 Determining the leading operators

When the inflationary action is expanded to third order in the fluctuation $\zeta(t,\vec{x})$, the resulting set of operators does not appear to be manifestly second-order in the slow-roll parameters: a fair amount of further effort [3] is required to make this property self-evident. Nearly every operator must be integrated by parts—often multiple times—with respect to its spatial or time derivatives, certain relations of the classical background must be imposed, *etc.* At the end of this lengthy process, the cubic part of the action assumes the following form,

$$
\begin{aligned}
S^{(3)} = \int d^4x \Big\{ & \frac{1}{4}\frac{e^{3\rho}}{M_{\rm pl}^2}\frac{\dot\phi^4}{\dot\rho^4}\dot\zeta^2\zeta + \frac{1}{4}\frac{e^{\rho}}{M_{\rm pl}^2}\frac{\dot\phi^4}{\dot\rho^4}\zeta\partial_k\zeta\partial^k\zeta - \frac{1}{2}\frac{e^{3\rho}}{M_{\rm pl}^2}\frac{\dot\phi^4}{\dot\rho^4}\dot\zeta\partial_k\zeta\partial^k\big(\partial^{-2}\dot\zeta\big) \\
& + \frac{1}{2}e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\dot\zeta\zeta^2\frac{d}{dt}\left[\frac{\ddot\phi}{\dot\phi\dot\rho} + \frac{1}{2M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right] - \frac{1}{16}\frac{e^{3\rho}}{M_{\rm pl}^2}\frac{\dot\phi^6}{\dot\rho^6}\dot\zeta^2\zeta \\
& + \frac{1}{16}\frac{e^{3\rho}}{M_{\rm pl}^4}\frac{\dot\phi^6}{\dot\rho^6}\zeta\partial_k\partial_l\big(\partial^{-2}\dot\zeta\big)\partial^k\partial^l\big(\partial^{-2}\dot\zeta\big) \\
& + \frac{1}{2}\left\{\frac{d}{dt}\left[e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\dot\zeta\right] - e^{\rho}\frac{\dot\phi^2}{\dot\rho^2}\partial_k\partial^k\zeta\right\} \\
& \quad \times\left\{\left[\frac{\ddot\phi}{\dot\phi\dot\rho} + \frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta^2 + 2\frac{1}{\dot\rho}\dot\zeta\zeta - \frac{1}{2}\frac{e^{-2\rho}}{\dot\rho^2}\big[\partial_k\zeta\partial^k\zeta - \partial^{-2}\partial_k\partial_l\big(\partial^k\zeta\partial^l\zeta\big)\big]\right. \\
& \quad\quad \left. + \frac{1}{2}\frac{1}{\dot\rho}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\big[\partial_k\zeta\partial^k\big(\partial^{-2}\dot\zeta\big) - \partial^{-2}\partial_k\partial_l\big(\partial^k\zeta\partial^l\big(\partial^{-2}\dot\zeta\big)\big)\big]\right\}\Big\}. \qquad (2.102)
\end{aligned}
$$

This is the form of the cubic interactions that was derived by Maldacena. Each of its operators either is manifestly second order or higher in the slow-roll parameters,

$$
\epsilon = \frac{1}{2M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2} \qquad \text{and} \qquad \delta = \frac{\ddot\phi}{\dot\phi\dot\rho}, \tag{2.103}
$$

or is proportional to the equation of motion for the quadratic part of the action. A few terms even have both of these properties.

It might seem that we have not quite succeeded in our goal, since not all the terms that are proportional to the equation of motion have enough factors of ϵ or δ. But by performing a nonlinear shift in the field, $\zeta \to \zeta_n + f(\zeta_n)$, where $f(\zeta_n)$ is quadratic in the field and is chosen precisely to remove the terms proportional to the equation of motion, we are left with just the first two lines of $S^{(3)}$. Of course such a shift means that the three-point function of ζ will in turn be replaced by a sum of three and four-point functions of ζ_n, which need to be separately computed. In principle there are higher-point functions of ζ_n that appear too, but they are suppressed in the limits that we are considering. Computing the three-point function in the late-time limit, the only part from these terms that contributes is

$$
\frac{1}{2}\left\{\frac{d}{dt}\left[e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\dot\zeta\right] - e^{\rho}\frac{\dot\phi^2}{\dot\rho^2}\partial_k\partial^k\zeta\right\}\left[\frac{\ddot\phi}{\dot\phi\dot\rho} + \frac{1}{2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta^2. \tag{2.104}
$$

The shift removes it from the three-point function of $\zeta_n(t,\vec x)$, but it reappears through the four-point function of this $\zeta_n(t,\vec x)$ field.

This method for computing the three-point function of the fluctuation $\zeta(t,\vec x)$ is not especially convenient here. The 1PI Green's functions that appear in the consistency relations are those associated with the original field $\zeta(t,\vec{x})$ and not the shifted field $\zeta_n(t,\vec{x})$. We could try to figure out how correspondingly to alter the 1PI Green's functions so that they are compatible with the Green's functions for $\zeta_n(t,\vec{x})$, but it is simpler to avoid the shift altogether. In fact, the piece that contributes in the late-time limit,

$$\frac{1}{2}\left\{\frac{d}{dt}\left[e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\dot\zeta\right]-e^{\rho}\frac{\dot\phi^2}{\dot\rho^2}\partial_k\partial^k\zeta\right\}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta^2, \tag{2.105}$$

is itself a linear combination of operators that are present in the first lines of the action as we wrote it earlier. Integrating a time derivative by parts in the first term, and a spatial derivative in the second, yields

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left[e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\dot\zeta\right]\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta^2-\frac{1}{2}e^{\rho}\frac{\dot\phi^2}{\dot\rho^2}\partial_k\partial^k\zeta\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta^2\\
&\qquad= -e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\dot\zeta^2\zeta+e^{\rho}\frac{\dot\phi^2}{\dot\rho^2}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta\partial_k\zeta\partial^k\zeta\\
&\qquad\quad-\frac{1}{2}e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\dot\zeta\zeta^2\frac{d}{dt}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]+\text{total derivatives.}
\end{aligned}\tag{2.106}$$

Combining their coefficients with those of the first two operators in the cubic action, the set of operators that simultaneously are second-order in ϵ and δ and have contributions that do not vanish in the late-time limit reduces to just three operators,

$$\begin{aligned}
S^{(3)} &= \int d^4x\,\Bigl\{-e^{3\rho}\frac{\dot\phi^2}{\dot\rho^2}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{4}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\dot\zeta^2\zeta+e^{\rho}\frac{\dot\phi^2}{\dot\rho^2}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{3}{4}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]\zeta\partial_k\zeta\partial^k\zeta\\
&\qquad -\frac{1}{2}\frac{e^{3\rho}}{M_{\rm pl}^2}\frac{\dot\phi^4}{\dot\rho^4}\dot\zeta\partial_k\zeta\partial^k\bigl(\partial^{-2}\dot\zeta\bigr)+\cdots\Bigr\}.
\end{aligned}\tag{2.107}$$

Thus, to find the difference of the three-point functions that appear in the consistency relation requires computing the contributions of these three operators. To make some of the intermediate stages of the following calculation a little less cluttered, let us abbreviate these coefficients of the operators by

$$\begin{aligned}
\alpha\,M_{\rm pl}^2 &= -\frac{\dot\phi^2}{\dot\rho^2}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{1}{4}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]=-\epsilon\bigl[\epsilon+2\delta\bigr]\,M_{\rm pl}^2\\
\beta\,M_{\rm pl}^2 &= \frac{\dot\phi^2}{\dot\rho^2}\left[\frac{\ddot\phi}{\dot\phi\dot\rho}+\frac{3}{4}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^2}{\dot\rho^2}\right]=\epsilon\bigl[3\epsilon+2\delta\bigr]\,M_{\rm pl}^2\\
\gamma\,M_{\rm pl}^2 &= -\frac{1}{2}\frac{1}{M_{\rm pl}^2}\frac{\dot\phi^4}{\dot\rho^4}=-2\epsilon^2\,M_{\rm pl}^2,
\end{aligned}\tag{2.108}$$

so that they might be written as

$$\mathcal{O}_1 = \alpha\, M_{\rm pl}^2\, e^{3\rho}\dot{\zeta}^2\zeta, \qquad \mathcal{O}_2 = \beta\, M_{\rm pl}^2\, e^{\rho}\zeta\partial_k\zeta\partial^k\zeta, \qquad \mathcal{O}_3 = \gamma\, M_{\rm pl}^2\, e^{3\rho}\dot{\zeta}\partial_k\zeta\partial^k\big(\partial^{-2}\dot{\zeta}\big). \tag{2.109}$$

Defined in this way, α, β, and γ are dimensionless parameters which are all quadratic in the slow-roll parameters.

2.7.2 A simplification and a subtlety

Having found what we need to compute, let us next determine what each of these operators does in fact contribute to the difference,

$$G_c^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_c^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3), \tag{2.110}$$

which will be evaluated in the limit where $\vec{k}\to\vec{0}$, $\vec{k}_3\to-\vec{k}_2$. The definitions of these Green's functions were explained earlier. Using a perturbative expansion organised in powers of the interaction Hamiltonian, which in this case is

$$H_I(t') = -\int d^3\vec{w}\,\Big\{\mathcal{O}_1(t',\vec{w}) + \mathcal{O}_2(t',\vec{w}) + \mathcal{O}_3(t',\vec{w}) + \cdots\Big\}, \tag{2.111}$$

the leading contribution to the difference comes from the terms with just one power of $H_I(t')$—or equivalently, just one power of each operator. We shall define the contribution from each operator by writing a corresponding subscript on the three-point function,

$$\begin{aligned}
&G_i^{+++}(t,\vec{x};t_*,\vec{x}_2;t_*,\vec{x}_3) - G_i^{-++}(t,\vec{x};t_*,\vec{x}_2;t_*,\vec{x}_3)\\
&= i\int_{-\infty}^{t_*} dt' \int d^3\vec{w}\,\langle 0|T\big(\zeta^+(t,\vec{x})\zeta^+(t_*,\vec{x}_2)\zeta^+(t_*,\vec{x}_3)\big[\mathcal{O}_i^+(t',\vec{w}) - \mathcal{O}_i^-(t',\vec{w})\big]\big)|0\rangle\\
&\quad - i\int_{-\infty}^{t_*} dt' \int d^3\vec{w}\,\langle 0|T\big(\zeta^-(t,\vec{x})\zeta^+(t_*,\vec{x}_2)\zeta^+(t_*,\vec{x}_3)\big[\mathcal{O}_i^+(t',\vec{w}) - \mathcal{O}_i^-(t',\vec{w})\big]\big)|0\rangle + \cdots\\
&= i\int_{-\infty}^{t_*} dt' \int d^3\vec{w}\,\Big\{\langle 0|T\big(\big[\zeta^+(t,\vec{x}) - \zeta^-(t,\vec{x})\big]\zeta^+(t_*,\vec{x}_2)\zeta^+(t_*,\vec{x}_3)\mathcal{O}_i^+(t',\vec{w})\big)|0\rangle\\
&\qquad\qquad - \langle 0|T\big(\big[\zeta^+(t,\vec{x}) - \zeta^-(t,\vec{x})\big]\zeta^+(t_*,\vec{x}_2)\zeta^+(t_*,\vec{x}_3)\mathcal{O}_i^-(t',\vec{w})\big)|0\rangle\Big\} + \cdots .
\end{aligned} \tag{2.112}$$

Once we have performed all of the Wick contractions of the fields, we transform the result to momentum space and proceed to evaluate what it adds to the consistency relation.

The leading contribution of each operator to the left side of the consistency relation is then found by computing

$$
\begin{aligned}
i\int_{-\infty}^{t_*} dt'' \int_{-\infty}^{t_*} dt\,\delta(t''-t) \left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d^2}{dt^2} + \frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\frac{d}{dt} + k^2\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\right\} \\
\left\{G_i^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_i^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3)\right\} \\
= i\int_{-\infty}^{t_*} dt \left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d^2}{dt^2} + \frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\frac{d}{dt} + k^2\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\right\} \\
\left\{G_i^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_i^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3)\right\}
\end{aligned} \tag{2.113}
$$

in the limit where $k \to 0$. Before going further, however, we point out a simplification that occurs when the difference between the three-point functions has the structure,

$$
G_i^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_i^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) = -\int_{-\infty}^{t_*} dt'\,\Theta(t'-t)\,F(t,t'), \tag{2.114}
$$

which is true for each of the three operators, though the function $F(t,t')$ will be different for each. Let us look at the part of the consistency relation that contains the time-derivatives from the 1PI Green's functions acting on something of this form,

$$
\begin{aligned}
-i\int_{-\infty}^{t_*} dt \int_{-\infty}^{t_*} dt' \left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d^2}{dt^2} + \frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\frac{d}{dt}\right\}\left\{\Theta(t'-t)\,F(t,t')\right\} \\
= \quad -i\int_{-\infty}^{t_*} dt \int_{-\infty}^{t_*} dt' \left\{-\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\left[\frac{d}{dt}\delta(t'-t)\right]F(t,t') - 2\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\delta(t'-t)\,\frac{d}{dt}F(t,t')\right. \\
\left. -\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\delta(t'-t)\,F(t,t') + \Theta(t'-t)\,\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d}{dt}F(t,t')\right]\right\}.
\end{aligned} \tag{2.115}
$$

The derivative of the δ-function must be treated carefully. If we *define*

$$
\int_{t_1}^{t_2} dt\,\dot\delta(t-t_0)\,f(t) \equiv -\dot f(t_0), \tag{2.116}
$$

without any boundary terms included, then we obtain the correct result for the consistency relation. Therefore we shall define $\dot\delta(t)$ by this relation, rather than attempt to subtract explicitly the unwanted boundary terms that would otherwise occur.

Applying this definition to the expression above yields,

$$
\begin{aligned}
&-i\int_{-\infty}^{t_*} dt\int_{-\infty}^{t_*} dt'\left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d^2}{dt^2}+\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\frac{d}{dt}\right\}\Big\{\Theta(t'-t)\,F(t,t')\Big\} \\
&= -i\int_{-\infty}^{t_*} dt\int_{-\infty}^{t_*} dt'\left\{-\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\delta(t'-t)\,\frac{d}{dt}F(t,t')+\Theta(t'-t)\,\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d}{dt}F(t,t')\right]\right\} \\
&= i\int_{-\infty}^{t_*} dt'\,\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t')}\left[\frac{d}{dt}F(t,t')\right]_{t=t'}-i\int_{-\infty}^{t_*} dt'\int_{-\infty}^{t'} dt\left\{\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d}{dt}F(t,t')\right]\right\}.
\end{aligned} \tag{2.117}
$$

The second term on the last line is a total derivative, so we can write it as

$$
\int_{-\infty}^{t'} dt\left\{\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d}{dt}F(t,t')\right]\right\}=\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t')}\left[\frac{d}{dt}F(t,t')\right]_{t=t'}, \tag{2.118}
$$

where we have assumed that there is no contribution from the $t\to-\infty$ boundary.[14] Therefore, the two terms exactly cancel, leaving no contribution from the time-derivative parts of the 1PI Green's functions,

$$
\begin{aligned}
&-i\int_{-\infty}^{t_*} dt\int_{-\infty}^{t_*} dt'\left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d^2}{dt^2}+\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\frac{d}{dt}\right\}\Big\{\Theta(t'-t)\,F(t,t')\Big\} \\
&\quad= i\int_{-\infty}^{t_*} dt'\,\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t')}\left[\frac{d}{dt}F(t,t')\right]_{t=t'}-i\int_{-\infty}^{t_*} dt'\,\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t')}\left[\frac{d}{dt}F(t,t')\right]_{t=t'}=0.
\end{aligned} \tag{2.119}
$$

As long as the difference of the three-point functions has the form that we assumed, then to leading order, the only part of the consistency relation that we need to calculate in detail, is the part due to the spatial derivatives,

$$
\begin{aligned}
&i\int_{-\infty}^{t_*} dt''\int_{-\infty}^{t_*} dt\,\delta(t''-t)\left\{\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\frac{d^2}{dt^2}+\frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho(t)}\right]\frac{d}{dt}+k^2\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\right\} \\
&\qquad\qquad\Big\{G_i^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)-G_i^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)\Big\} \\
&=\lim_{k\to0}\left\{ik^2\int_{-\infty}^{t_*} dt\,\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\Big[G_i^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)-G_i^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)\Big]\right\}.
\end{aligned} \tag{2.120}
$$

2.7.3 The second operator

We begin with the second operator since its structure is the simplest, having only spatial derivatives. Substituting this operator into the above expression and taking

[14]This relation will require an initial boundary term, however, when we start the system in a more general initial state at $t=t_0$.

the Wick contractions, what results is

$$\begin{aligned}
&G_2^{+++}(t,\vec{x};t_*,\vec{x}_2;t_*,\vec{x}_3) - G_2^{-++}(t,\vec{x};t_*,\vec{x}_2;t_*,\vec{x}_3) \\
&= \int \frac{d^3\vec{k}}{(2\pi)^3} \int \frac{d^3\vec{k}_2}{(2\pi)^3} \int \frac{d^3\vec{k}_3}{(2\pi)^3}\, e^{i\vec{k}\cdot\vec{x}} e^{i\vec{k}_2\cdot\vec{x}_2} e^{i\vec{k}_3\cdot\vec{x}_3}\, (2\pi)^3\, \delta^3(\vec{k}+\vec{k}_2+\vec{k}_3) \\
&\quad i\beta\, M_{\rm pl}^2 \left[k^2+k_2^2+k_3^2\right] \int_{-\infty}^{t_*} dt'\, e^{\rho(t')} \Big\{ [G_k^{++}(t,t') - G_k^{-+}(t,t')] G_{k_2}^{++}(t_*,t') G_{k_3}^{++}(t_*,t') \\
&\qquad - [G_k^{+-}(t,t') - G_k^{--}(t,t')] G_{k_2}^{+-}(t_*,t') G_{k_3}^{+-}(t_*,t') \Big\}.
\end{aligned} \tag{2.121}$$

To convert the Feynman propagators $G_{k_i}^{\pm\pm}$ into the appropriate time-ordered combinations of Wightman functions, notice that the k_2 and k_3-dependent factors are especially simple since t_* always occurs later than t'. For the k-dependent parts, we find that everything is proportional to a common Θ-function,

$$\begin{aligned}
G_k^{++}(t,t') - G_k^{-+}(t,t') &= \Theta(t-t')\, G_k^{>}(t,t') + \Theta(t'-t)\, G_k^{<}(t,t') - G_k^{>}(t,t') \\
&= -\Theta(t'-t) \left[G_k^{>}(t,t') - G_k^{<}(t,t')\right] \\
G_k^{+-}(t,t') - G_k^{--}(t,t') &= G_k^{<}(t,t') - \Theta(t'-t)\, G_k^{>}(t,t') - \Theta(t-t')\, G_k^{<}(t,t') \\
&= -\Theta(t'-t) \left[G_k^{>}(t,t') - G_k^{<}(t,t')\right].
\end{aligned} \tag{2.122}$$

The Fourier-transform of the contribution from the second operator to the difference in the three-point functions is then

$$\begin{aligned}
&G_2^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_2^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) \\
&\quad = -i\beta\, M_{\rm pl}^2 \left[k^2+k_2^2+k_3^2\right] \int_{-\infty}^{t_*} dt'\, e^{\rho(t')} \Theta(t'-t) \left[G_k^{>}(t,t') - G_k^{<}(t,t')\right] \\
&\qquad\qquad \left[G_{k_2}^{>}(t_*,t') G_{k_3}^{>}(t_*,t') - G_{k_2}^{<}(t_*,t') G_{k_3}^{<}(t_*,t')\right].
\end{aligned} \tag{2.123}$$

We should mention here that this same structure, this combination of Wightman functions,

$$\Theta(t'-t) \left[G_k^{>}(t,t') - G_k^{<}(t,t')\right] \left[G_{k_2}^{>}(t_*,t') G_{k_3}^{>}(t_*,t') - G_{k_2}^{<}(t_*,t') G_{k_3}^{<}(t_*,t')\right], \tag{2.124}$$

occurs in the calculations of the other two operators, though with additional time-derivatives and momentum factors. It has exactly the $\Theta(t'-t)\, F(t,t')$ structure that is needed for the time-derivative parts of the 1PI Green's functions not to contribute.

The only contribution to the consistency relation for the Bunch-Davies state can come from the spatial derivative term,

$$
\begin{aligned}
& ik^2 \int_{-\infty}^{t_*} dt\, \frac{\dot\phi^2}{\dot\rho^2} e^{\rho(t)} \Big\{ G_2^{+++}(t,\vec k; t_*, \vec k_2; t_*, \vec k_3) - G_2^{-++}(t,\vec k; t_*, \vec k_2; t_*, \vec k_3) \Big\} \\
&= ik^2 \int_{-\infty}^{t_*} dt\, \frac{\dot\phi^2}{\dot\rho^2} e^{\rho(t)} \Big\{ -i\beta\, M_{\rm pl}^2 \left[k^2 + k_2^2 + k_3^2 \right] \int_{-\infty}^{t_*} dt'\, e^{\rho(t')} \Theta(t'-t) \\
&\qquad\qquad \left[G_k^>(t,t') - G_k^<(t,t') \right] \left[G_{k_2}^>(t_*,t') G_{k_3}^>(t_*,t') - G_{k_2}^<(t_*,t') G_{k_3}^<(t_*,t') \right] \Big\} \\
&= k^2 \beta\, M_{\rm pl}^2 \frac{\dot\phi^2}{\dot\rho^2} \left[k^2 + k_2^2 + k_3^2 \right] \int_{-\infty}^{t_*} dt'\, e^{\rho(t')} \left[G_{k_2}^>(t_*,t') G_{k_3}^>(t_*,t') - G_{k_2}^<(t_*,t') G_{k_3}^<(t_*,t') \right] \\
&\qquad\qquad \times \int_{-\infty}^{t'} dt\, e^{\rho(t)} \left[G_k^>(t,t') - G_k^<(t,t') \right]. \qquad (2.125)
\end{aligned}
$$

Let us evaluate these two nested time-integrals. As we are only working to leading order in the slow-roll parameters, we can write the scale factor and the Wightman functions in the de Sitter limit. These functions assume a simpler form when expressed in terms of the conformal time (η, η', η_*) rather than the time coordinates that we have been using, (t, t', t_*). For example, the measure and scale factor become

$$
dt\, e^{\rho(t)} = d\eta\, \frac{dt}{d\eta} e^{\rho(t)} = d\eta\, \frac{1}{-H\eta} \frac{1}{-H\eta} = \frac{1}{H^2} \frac{d\eta}{\eta^2}. \qquad (2.126)
$$

The Wightman functions in the de Sitter limit are given by

$$
\begin{aligned}
G_k^>(t,t') &= \frac{1}{4\epsilon} \frac{H^2}{M_{\rm pl}^2} \frac{1}{k^3} (1 + ik\eta)(1 - ik\eta') e^{-ik(\eta-\eta')} + \cdots \\
G_k^<(t,t') &= \frac{1}{4\epsilon} \frac{H^2}{M_{\rm pl}^2} \frac{1}{k^3} (1 - ik\eta)(1 + ik\eta') e^{ik(\eta-\eta')} + \cdots , \qquad (2.127)
\end{aligned}
$$

where we do need to retain the initial $1/\epsilon$ factor since it would otherwise diverge in the strict de Sitter limit. The dt-integral then becomes

$$
\begin{aligned}
& \int_{-\infty}^{t'} dt\, e^{\rho(t)} \left[G_k^>(t,t') - G_k^<(t,t') \right] \\
&\quad = -\frac{i}{2\epsilon} \frac{1}{M_{\rm pl}^2} \frac{1}{k^3} \int_{-\infty}^{\eta'} \frac{d\eta}{\eta^2} \left[(1 + k^2\eta\eta') \sin[k(\eta-\eta')] - k(\eta-\eta') \cos[k(\eta-\eta')] \right] \\
&\quad = -\frac{i}{2\epsilon} \frac{1}{M_{\rm pl}^2} \frac{1}{k^3} \left[-\frac{\sin[k(\eta-\eta')] + k\eta' \cos[k(\eta-\eta')]}{\eta} \right]_{-\infty}^{\eta'} = \frac{i}{2\epsilon} \frac{1}{M_{\rm pl}^2} \frac{1}{k^2}. \qquad (2.128)
\end{aligned}
$$

Though it appeared to depend on t' as well, none of this dependence survives in this leading result. There remains just the dt'-integral,

$$\begin{aligned}
& ik^2\int_{-\infty}^{t_*} dt\,\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\Big\{G_2^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)-G_2^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)\Big\} \\
&= \frac{i\beta}{2\epsilon}\frac{\dot\phi^2}{\dot\rho^2}\left[k^2+k_2^2+k_3^2\right]\int_{-\infty}^{t_*} dt'\,e^{\rho(t')}\left[G^>_{k_2}(t_*,t')G^>_{k_3}(t_*,t')-G^<_{k_2}(t_*,t')G^<_{k_3}(t_*,t')\right] \\
&= i\beta\,M_{\rm pl}^2\left[k^2+k_2^2+k_3^2\right]\int_{-\infty}^{t_*} dt'\,e^{\rho(t')}\left[G^>_{k_2}(t_*,t')G^>_{k_3}(t_*,t')-G^<_{k_2}(t_*,t')G^<_{k_3}(t_*,t')\right].
\end{aligned} \tag{2.129}$$

Here we have used $\dot\phi^2/\dot\rho^2=2\epsilon M_{\rm pl}^2$. In performing this integral, we assume that t_* is being taken in the late-time limit. In terms of the conformal time, this limit corresponds to $\eta_*\to 0$,

$$\begin{aligned}
& \int_{-\infty}^{t_*} dt'\,e^{\rho(t')}\left[G^>_{k_2}(t_*,t')G^>_{k_3}(t_*,t')-G^<_{k_2}(t_*,t')G^<_{k_3}(t_*,t')\right] \\
&= \frac{i}{8\epsilon^2}\frac{H^2}{M_{\rm pl}^4}\frac{1}{k_2^3k_3^3}\int_{-\infty}^{\eta_*}\frac{d\eta'}{\eta'^2}\left[(1-k_2k_3\eta'^2)\sin[(k_2+k_3)\eta']-(k_2+k_3)\eta'\cos[(k_2+k_3)\eta']\right] \\
&= -\frac{i}{8\epsilon^2}\frac{H^2}{M_{\rm pl}^4}\frac{1}{k_2^3k_3^3}\frac{k_2^2+k_3^2+k_2k_3}{k_2+k_3}+\mathcal{O}(\eta_*^2).
\end{aligned} \tag{2.130}$$

Therefore, the leading contribution from the second operator to the consistency relation is

$$\begin{aligned}
& ik^2\int_{-\infty}^{t_*} dt\,\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\Big\{G_2^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)-G_2^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)\Big\} \\
&= \frac{\beta}{8\epsilon^2}\frac{H^2}{M_{\rm pl}^2}\frac{k^2+k_2^2+k_3^2}{k_2^3k_3^3}\frac{k_2^2+k_3^2+k_2k_3}{k_2+k_3},
\end{aligned} \tag{2.131}$$

or rather, since it is meant to be evaluated in the limit where $k\to 0$ and $k_3=k_2$,

$$ik^2\int_{-\infty}^{t_*} dt\,\frac{\dot\phi^2}{\dot\rho^2}e^{\rho(t)}\Big\{G_2^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)-G_2^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3)\Big\}=\frac{3\beta}{8\epsilon^2}\frac{H^2}{M_{\rm pl}^2}\frac{1}{k_2^3}+\cdots. \tag{2.132}$$

Although we have evaluated this result in a slightly different language of Green's functions, the contribution from this particular operator in the cubic action is exactly the same as what appeared in Maldacena's calculation, once we have multiplied his expression in the appropriate 'soft' limit by an inverse factor of the power spectrum.

2.7.4 The first operator

While the combination of Wightman functions that appears in the contribution from the first operator to the difference of the three-point functions is more complicated than what we just encountered, there are many parallels between them. To start, the operator $\mathcal{O}_1$'s contribution is

$$\begin{aligned}
&G_1^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_1^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) \\
&= -2i\alpha\, M_{\rm pl}^2 \int_{-\infty}^{t_*} dt'\, e^{3\rho(t')}\Theta(t'-t) \\
&\quad \Big\{ \big[\dot G_k^>(t,t') - \dot G_k^<(t,t')\big]\big[\dot G_{k_2}^>(t_*,t')G_{k_3}^>(t_*,t') - \dot G_{k_2}^<(t_*,t')G_{k_3}^<(t_*,t')\big] \\
&\quad + \big[\dot G_k^>(t,t') - \dot G_k^<(t,t')\big]\big[G_{k_2}^>(t_*,t')\dot G_{k_3}^>(t_*,t') - G_{k_2}^<(t_*,t')\dot G_{k_3}^<(t_*,t')\big] \\
&\quad + \big[G_k^>(t,t') - G_k^<(t,t')\big]\big[\dot G_{k_2}^>(t_*,t')\dot G_{k_3}^>(t_*,t') - \dot G_{k_2}^<(t_*,t')\dot G_{k_3}^<(t_*,t')\big]\Big\}. \qquad (2.133)
\end{aligned}$$

The appearance of the $\Theta(t'-t)$ and the arrangement of the Wightman functions is exactly as before. All that has changed is the appearance of the time-derivatives, which where inherited from the operator itself, $\mathcal{O}_1 = \alpha\, M_{\rm pl}^2\, e^{3\rho}\dot\zeta^2\zeta$, and which are permuted amongst the various coordinates.

The sole contribution to the consistency relation again comes from the spatial derivative, or k^2 part, of the 1PI Green's function acting on the difference of the three-point functions,

$$\begin{aligned}
&ik^2 \int_{-\infty}^{t_*} dt\, \frac{\dot\phi^2}{\dot\rho^2} e^{\rho(t)} \Big\{ G_1^{+++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3) - G_1^{-++}(t,\vec{k};t_*,\vec{k}_2;t_*,\vec{k}_3)\Big\} \\
&= 2\alpha\, M_{\rm pl}^2\, k^2 \frac{\dot\phi^2}{\dot\rho^2} \int_{-\infty}^{t_*} dt'\, e^{3\rho(t')} \Big\{ \int_{-\infty}^{t'} dt\, e^{\rho(t)}\, \big[\dot G_k^>(t,t') - \dot G_k^<(t,t')\big]\Big\} \\
&\qquad \Big\{ \big[\dot G_{k_2}^>(t_*,t')G_{k_3}^>(t_*,t') - G_{k_2}^<(t_*,t')G_{k_3}^<(t_*,t')\big] \\
&\qquad + \big[G_{k_2}^>(t_*,t')\dot G_{k_3}^>(t_*,t') - G_{k_2}^<(t_*,t')\dot G_{k_3}^<(t_*,t')\big]\Big\} \\
&\quad + 2\alpha\, M_{\rm pl}^2\, k^2 \frac{\dot\phi^2}{\dot\rho^2} \int_{-\infty}^{t_*} dt'\, e^{3\rho(t')} \big[\dot G_{k_2}^>(t_*,t')\dot G_{k_3}^>(t_*,t') - \dot G_{k_2}^<(t_*,t')\dot G_{k_3}^<(t_*,t')\big] \\
&\qquad \times \int_{-\infty}^{t'} dt\, e^{\rho(t)}\, \big[G_k^>(t,t') - G_k^<(t,t')\big]. \qquad (2.134)
\end{aligned}$$

One of these integrals—the integral over t—we have already encountered in treating $\mathcal{O}_2$,

$$\int_{-\infty}^{t'} dt\, e^{\rho(t)}\, \big[G_k^>(t,t') - G_k^<(t,t')\big] = \frac{i}{2\epsilon}\frac{1}{M_{\rm pl}^2}\frac{1}{k^2}, \qquad (2.135)$$

while the second of the integrals vanishes altogether,

$$\begin{aligned}
\int_{-\infty}^{t'} dt\, e^{\rho(t)}\big[\dot G_k^>(t,t') - \dot G_k^<(t,t')\big] &= \frac{i}{2\epsilon}\frac{H}{M_{\rm pl}^2}\frac{\eta'^2}{k}\int_{-\infty}^{\eta'}\frac{d\eta}{\eta^2}\Big[\sin[k(\eta-\eta')] - k\eta\cos[k(\eta-\eta')]\Big] \\
&= \frac{i}{2\epsilon}\frac{H}{M_{\rm pl}^2}\frac{\eta'^2}{k}\left[-\frac{\sin[k(\eta-\eta')]}{\eta}\right]_{-\infty}^{\eta'} = 0.
\end{aligned}$$

There remains one integral to perform, which we shall evaluate in the late-time limit ($\eta_* \to 0$) where η_* essentially vanishes,

$$
\begin{aligned}
ik^2 \int_{-\infty}^{t_*} dt\, \frac{\dot\phi^2}{\dot\rho^2} e^{\rho(t)} \Big\{ G_1^{+++}(t,\vec k; t_*, \vec k_2; t_*, \vec k_3) - G_1^{-++}(t,\vec k; t_*, \vec k_2; t_*, \vec k_3) \Big\} & \\
= \; 2i\alpha\, M_{\rm pl}^2 \int_{-\infty}^{t_*} dt'\, e^{3\rho(t')} \big[\dot G_{k_2}^{>}(t_*,t') \dot G_{k_3}^{>}(t_*,t') - \dot G_{k_2}^{<}(t_*,t') \dot G_{k_3}^{<}(t_*,t') \big] & \\
= \; -\frac{\alpha}{4\epsilon^2} \frac{H^2}{M_{\rm pl}^2} \frac{1}{k_2 k_3} \int_{-\infty}^{\eta_*} d\eta'\, \sin[(k_2+k_3)\eta'] & \\
= \; \frac{\alpha}{4\epsilon^2} \frac{H^2}{M_{\rm pl}^2} \frac{1}{k_2 k_3} \frac{1}{k_2+k_3} + \cdots & \qquad (2.136)
\end{aligned}
$$

Going to the limit $k_3 = k_2$ we have a genuine contribution to the left-side of the consistency relation from the first operator,

$$
\langle \mathcal{O}_1 \rangle = \frac{\alpha}{8\epsilon^2} \frac{H^2}{M_{\rm pl}^2} \frac{1}{k_2^3}. \qquad (2.137)
$$

2.7.5 The third operator

The third operator is the most complicated since it has both time and space derivatives,

$$
\mathcal{O}_3 = \gamma\, M_{\rm pl}^2\, e^{3\rho} \dot\zeta \partial_k \zeta \partial^k (\partial^{-2} \dot\zeta). \qquad (2.138)
$$

It contributes the following to the difference of the three-point functions,

$$
\begin{aligned}
& G_3^{+++}(t,\vec k; t_*, \vec k_2; t_*, \vec k_3) - G_3^{-++}(t,\vec k; t_*, \vec k_2; t_*, \vec k_3) \\
&= -i\gamma\, M_{\rm pl}^2 \int_{-\infty}^{t_*} dt'\, e^{3\rho(t')} \Theta(t'-t) \\
&\quad \Bigg\{ \left[\frac{\vec k \cdot \vec k_3}{k^2} + \frac{\vec k_2 \cdot \vec k_3}{k_2^2} \right] \big[\dot G_k^{>}(t,t') - \dot G_k^{<}(t,t') \big] \big[\dot G_{k_2}^{>}(t_*,t') G_{k_3}^{>}(t_*,t') - \dot G_{k_2}^{<}(t_*,t') G_{k_3}^{<}(t_*,t') \big] \\
&\quad + \left[\frac{\vec k \cdot \vec k_2}{k^2} + \frac{\vec k_3 \cdot \vec k_2}{k_3^2} \right] \big[\dot G_k^{>}(t,t') - \dot G_k^{<}(t,t') \big] \big[G_{k_2}^{>}(t_*,t') \dot G_{k_3}^{>}(t_*,t') - G_{k_2}^{<}(t_*,t') \dot G_{k_3}^{<}(t_*,t') \big] \\
&\quad + \left[\frac{\vec k_2 \cdot \vec k}{k_2^2} + \frac{\vec k_3 \cdot \vec k}{k_3^2} \right] \big[G_k^{>}(t,t') - G_k^{<}(t,t') \big] \big[\dot G_{k_2}^{>}(t_*,t') \dot G_{k_3}^{>}(t_*,t') - \dot G_{k_2}^{<}(t_*,t') \dot G_{k_3}^{<}(t_*,t') \big] \Bigg\}.
\end{aligned}
$$

The combinations of Wightman functions, although they are now accompanied by more complicated coefficients depending on the momenta, are nonetheless exactly what we saw while analysing $\mathcal{O}_1$. So to analyse the contribution from $\mathcal{O}_3$ we do not need to reproduce each of the steps of the previous calculation.

The contribution from the part due to the spatial derivatives is quickly evaluated by using results from the $\mathcal{O}_1$ calculation,

$$ik^2 \int_{-\infty}^{t_*} dt\, \frac{\dot\phi^2}{\dot\rho^2} e^{\rho(t)} \Big\{ G_3^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3) - G_3^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3) \Big\}$$
$$= \frac{\gamma}{8\epsilon^2}\frac{H^2}{M_{\rm pl}^2}\left[\frac{\vec k_2\cdot\vec k}{k_2^2} + \frac{\vec k_3\cdot\vec k}{k_3^2}\right]\frac{1}{k_2k_3}\frac{1}{k_2+k_3} + \cdots . \tag{2.139}$$

Notice that in the $\vec k \to \vec 0$ limit, this contribution goes away altogether.

$$ik^2 \int_{-\infty}^{t_*} dt\, \frac{\dot\phi^2}{\dot\rho^2} e^{\rho(t)} \Big\{ G_3^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3) - G_3^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3) \Big\}$$
$$= -\frac{\gamma}{16\epsilon^2}\frac{H^2}{M_{\rm pl}^2}\frac{k^2}{k_2^5} + \cdots = 0. \tag{2.140}$$

Therefore we have no contribution from this operator in the slow-roll, late-time, $k \to 0$ limits,

$$\langle \mathcal{O}_3 \rangle = 0 + \cdots . \tag{2.141}$$

2.7.6 The left side of the consistency relation

If we add up what we have found, we find

$$\langle \mathcal{O}_1 \rangle + \langle \mathcal{O}_2 \rangle + \langle \mathcal{O}_3 \rangle = \frac{\alpha+3\beta}{8\epsilon^2}\frac{H^2}{M_{\rm pl}^2}\frac{1}{k_2^3}. \tag{2.142}$$

Putting in the values of α, β, and γ, we have that

$$\alpha + 3\beta = 4\epsilon[2\epsilon+\delta], \tag{2.143}$$

which yields the following leading result in the slow-roll parameters for the left side of the relation as we have written it,

$$i \lim_{\vec k\to\vec 0} \int_{-\infty}^{t_*} dt \left\{ \frac{\dot\phi^2}{\dot\rho^2}e^{3\rho}\frac{d^2}{dt^2} + \frac{d}{dt}\left[\frac{\dot\phi^2}{\dot\rho^2}e^{3\rho}\right]\frac{d}{dt} + k^2\frac{\dot\phi^2}{\dot\rho^2}e^{\rho} \right\}$$
$$\Big\{ G_c^{+++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3) - G_c^{-++}(t,\vec k;t_*,\vec k_2;t_*,\vec k_3) \Big\} \approx -\frac{2\epsilon+\delta}{2\epsilon}\frac{H^2}{M_{\rm pl}^2}\frac{1}{k_2^3}. \tag{2.144}$$

The two-point function, or power spectrum, in the late time limit is

$$G_c^{++}(t_*,\vec k_2;t_*,-\vec k_2) \equiv P_{k_2}(t_*) = \frac{1}{4\epsilon}\frac{H^2}{M_{\rm pl}^2}\frac{(-k_2\eta)^{-4\epsilon-2\delta}}{k_2^3}, \tag{2.145}$$

to leading order in the slow-roll parameters. Taking the derivative that appears in the consistency relation,

$$-\big[3+\vec{k}_2\cdot\nabla_{\vec{k}_2}\big]G_c^{++}(t_*,\vec{k}_2;t_*,-\vec{k}_2) = -\left[3+k_2\frac{\partial}{\partial k_2}\right]P_{k_2}(t_*) = -\frac{4\epsilon+2\delta}{4\epsilon}\frac{H^2}{M_{\rm pl}^2}\frac{1}{k_2^3}+\cdots\,, \tag{2.146}$$

we do find that the two sides agree.

2.8 Supplement: Transforming the consistency relation into a standard form

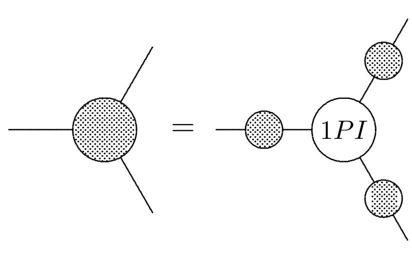
When we differentiated the Slavnov-Taylor identity with respect to the fields $\bar\zeta^{s_2}(y_2)$ and $\bar\zeta^{s_3}(y_3)$, we obtained

$$\begin{aligned}\int d^4x\,\Bigg\{&\frac{\delta^2 J^+(x)}{\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}-\frac{\delta^2 J^-(x)}{\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}\\ &+\sum_{r,s=\pm}\int d^4z\,\frac{\delta J^r(x)}{\delta\bar\zeta^{s_2}(y_2)}\frac{\delta J^s(z)}{\delta\bar\zeta^{s_3}(y_3)}\vec{x}\cdot\vec\nabla_{\vec{x}}\frac{\delta W}{\delta J^r(x)\delta J^s(z)}\\ &+\sum_{r,s=\pm}\int d^4z\,\frac{\delta J^r(x)}{\delta\bar\zeta^{s_3}(y_3)}\frac{\delta J^s(z)}{\delta\bar\zeta^{s_2}(y_2)}\vec{x}\cdot\vec\nabla_{\vec{x}}\frac{\delta W}{\delta J^r(x)\delta J^s(z)}\Bigg\}=0.\end{aligned} \tag{2.147}$$

As a formal expression this is perfectly fine; but it is not yet in a form that is most suited for what is being measured. It is the connected three-point function that is being constrained by observations rather than its one-particle irreducible counterpart,

$$\Gamma^{\pm s_2 s_3}(x,y_2,y_3)=\frac{\delta^3\Gamma}{\delta\bar\zeta^{\pm}(x)\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}=\mp\frac{\delta^2 J^{\pm}(x)}{\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}. \tag{2.148}$$

When we are not evolving from $t=-\infty$ to ∞, the higher-order 1PI Green's functions can be a bit difficult to calculate since we can no longer Fourier transform in all $3+1$ dimensions. But just as in S-matrix calculations, the connected and 1PI three-point functions are related by amputating the propagators associated with the legs of the former to obtain the latter,



(2.149)

The only extra complication is that unlike the analogous S-matrix statement we must sum over the $\pm$ indices in addition to integrating over the intermediate positions where the propagators are attached. By applying the appropriate operator, we can convert the 1PI three-point functions in our expression into connected three-point functions on which the 1PI two-point functions are acting.

Let us apply the following operator to the previous version of the relation,

$$\sum_{s_2,s_3=\pm}\int d^4y_2\int d^4y_3\,\frac{\delta\bar\zeta^{s_2}(y_2)}{\delta J^+(x_2)}\frac{\delta\bar\zeta^{s_3}(y_3)}{\delta J^+(x_3)}. \tag{2.150}$$

If we think of the consistency relation as a relation amongst different graphs, what this operator does is to attach connected two-point functions (propagators) to two of the external legs of each term in the relation. Analytically, the relation becomes

$$\sum_{s_2,s_3=\pm}\int d^4y\int d^4y_2\int d^4y_3\,\frac{\delta\bar\zeta^{s_2}(y_2)}{\delta J^+(x_2)}\frac{\delta\bar\zeta^{s_3}(y_3)}{\delta J^+(x_3)}\left\{\frac{\delta^2J^+(y)}{\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}-\frac{\delta^2J^-(y)}{\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}\right\}$$
$$=-\big[\vec x_2\cdot\vec\nabla_{\vec x_2}+\vec x_3\cdot\vec\nabla_{\vec x_3}\big]\frac{\delta W}{\delta J^+(x_2)\delta J^+(x_3)} \tag{2.151}$$

For later convenience, we have switched one of the integration variables from x to y. Notice that one of the effects of acting with this operator has been to remove the 1PI two-point functions entirely from the terms with the directional derivatives, $\vec x\cdot\vec\nabla_{\vec x}$. This happened because they are inverses of the functional derivatives in the operator that we applied,

$$\sum_{s=\pm}\int d^4z\,\frac{\delta\bar\zeta^s(z)}{\delta J^{r_1}(x)}\frac{\delta J^{r_2}(y)}{\delta\bar\zeta^s(z)}=\delta^{r_2}_{r_1}\,\delta^4(x-y). \tag{2.152}$$

If we differentiate this identity with respect to a source J^+, and set $r=+$ and $s=s_2$, we have a way of rewriting the left side of the consistency relation in terms of connected three-point functions rather than the 1PI ones,

$$\sum_{s_2,s_3}\int d^4y_2\int d^4y_3\left\{\frac{\delta\bar\zeta^{s_2}(y_2)}{\delta J^+(x_2)}\frac{\delta\bar\zeta^{s_3}(y_3)}{\delta J^+(x_3)}\frac{\delta^2J^r(y)}{\delta\bar\zeta^{s_2}(y_2)\delta\bar\zeta^{s_3}(y_3)}\right\}$$
$$=-\sum_{s=\pm}\int d^4x\left\{\frac{\delta J^r(y)}{\delta\bar\zeta^s(x)}\frac{\delta^2\bar\zeta^s(x)}{\delta J^+(x_2)\delta J^+(x_3)}\right\}. \tag{2.153}$$

The consistency relation then assumes the form that we have been seeking,

$$\sum_{s=\pm}\int d^4y\int d^4x\left\{\frac{\delta J^+(y)}{\delta\bar\zeta^s(x)}\frac{\delta^2\bar\zeta^s(x)}{\delta J^+(x_2)\delta J^+(x_3)}-\frac{\delta J^-(y)}{\delta\bar\zeta^s(x)}\frac{\delta^2\bar\zeta^s(x)}{\delta J^+(x_2)\delta J^+(x_3)}\right\}$$
$$=\big[\vec x_2\cdot\vec\nabla_{\vec x_2}+\vec x_3\cdot\vec\nabla_{\vec x_3}\big]\frac{\delta W}{\delta J^+(x_2)\delta J^+(x_3)}. \tag{2.154}$$

All that is left is to convert the functions that appear in this relation into a more conventional form, using the definitions of the connected and 1PI Green's functions in terms of functional derivatives of the appropriate generating functional,

$$-i\sum_{s=\pm}\int d^4y\int d^4x\,\Big\{\Gamma^{+s}(y,x)\,G_c^{s++}(x,x_2,x_3)+\Gamma^{-s}(y,x)\,G_c^{s++}(x,x_2,x_3)\Big\}$$
$$=\big[\vec{x}_2\cdot\vec{\nabla}_{\vec{x}_2}+\vec{x}_3\cdot\vec{\nabla}_{\vec{x}_3}\big]G_c^{++}(x_2,x_3). \qquad (2.155)$$

Chapter 3

Renormalizing an initial state

For calculations in quantum field theory we usually start with the appropriate quadratic action, derive the propagator for this free theory and, based on it, construct Green's functions of the full theory perturbatively. The standard applications assume relatively simple states. In scattering problems, the "in" and "out" states are chosen to be the free theory single particle states in an infinite past and future. In inflationary calculations the "in" state is the free Bunch-Davies state in an infinite past. This is what we do in practice. But in both cases we really mean to be in the eigenstate of the full theory. The reason why using the free eigenstates gives us the correct answer is because the states are being evolved over an infinite time. In this situation we can use mathematical tricks like an $i\epsilon$ prescription or an adiabatic switching on of the interaction to separate the full eigenstate from the free one. For example, the usual logic for calculating cosmological correlation functions in the vacuum state of an interacting theory is to start the evolution from an early enough time $t_0 \to -\infty$. Then it is possible to argue that there are no contributions from the lower end of the time integrals: the fields oscillate rapidly, and after deforming the integration contour ($i\epsilon$ prescription) to project out the full vacuum these terms go to zero.

But let us say that we want to start our evolution from an arbitrary initial time; then we cannot use these procedures to pick out the vacuum state we want. Moreover, if we want to calculate correlation functions not in the full vacuum, but in some arbitrary state of an interacting theory, then even if we started from $-\infty$ we still will not be able to use the $i\epsilon$ prescription since it can only project out a state with the lowest energy, i.e. the vacuum state.

There are several reasons for wanting to start from a finite initial time. First of all, for a lot of states neither an $i\epsilon$ prescription, nor an adiabatic "turning on" of an interaction are useful, so there is no advantage in taking $t_0 \to -\infty$. For instance, the state might not be an equilibrium state of the interacting theory. Starting in the infinite past and "turning on" the interactions, we will not naturally flow into such a state. Another example is a bound state in an interacting theory. This state will not exist in the infinite past once we have "turned off" the interactions. In this case something discrete happens: either particles are bound or they are not; there is no adiabatic transition between these two statements.

Secondly, we will be able to treat interesting excited states that might not necessarily have a reasonable extrapolation all the way back to $t_0 \to -\infty$, but which are sensible enough (non-singular) at a finite time t_0. In this case it is really the state itself that is important, not the particular value of t_0 that we have chosen, as long as it remains finite, since we are not assuming that anything physical is happening at t_0.

Thirdly, there is a danger that by going back to the infinite past we might enter a non-perturbative regime or a regime in which there might be some uncontrolled, poorly understood UV behavior as $t_0 \to -\infty$. The trans-Planckian problem of inflation is an example of this case. Because of the expansion, going to the infinite past is equivalent to going to arbitrarily short distances. But we know that once we reach distances smaller than the Planck scale the contributions from higher order operators will become more and more important and we will end up having an infinite number of unsuppressed nonrenormalizable operators. Thus, we would like to be able to start our evolution from scales far enough from the Planck threshold.

And the last, but most obvious reason is that something is really happening at t_0, so it is a natural choice to use.

In this paper we present a different approach for calculating the expectation values of the products of fields that can be applied in the case of a finite initial time. At this initial time let our fields be in some state, for example, the vacuum state, a thermal state, etc. We can construct such a state through a set of boundary operators on the initial time hypersurface [18]. These operators are implicitly defined with respect to the free theory vacuum. However, what we really want is to calculate correlation functions of an interacting theory in the corresponding interacting theory state, e.g. the interacting vacuum, an interacting thermal state, etc. Therefore we need to renormalize the structures of the initial state perturbatively, order by order in the parameters of the interacting theory, in such a way that this initial state satisfies certain conditions. This is somewhat similar to how operators are renormalized in the dynamical part of a Lagrangian in ordinary quantum field theory. We know how certain n-point functions behave in the free theory case; for example, we know that the one-point function is zero and the pole of the propagator has a residue of 1, and we would like to have the same behavior for these functions in the full theory. As a consequence of imposing this behavior we have to rescale fields and introduce counterterms.

This renormalization is required even in the simplest case—an interacting theory in its vacuum state at a particular time $t_0 \neq -\infty$. We find that the corrections to the n-point functions have an explicit dependence on the initial time. When taking $t_0 \to -\infty$ we see that these functions do not match to the ones that we get when we start evolving from the asymptotic vacuum: they contain additional divergent and oscillatory terms. This means that at t_0 we were in the wrong state, not in the state we intended to be, i.e. not in the interacting vacuum state. To fix this we add operators and structures to the initial state action—these are the "counterterms" of this picture, and they are defined order by order.

In the next section we will show how to specify order by order in perturbation theory the initial state using the eigenstates of the free part of the theory. Section 3 mentions a few details of simple single-field, slow-roll inflationary models that will be used in our calculations. Sections 4 and 5 are the sample calculations of the vacuum state three- and two-point functions of inflation using this method and the fact that we know what we should get for $t_0 \to -\infty$ from the conventional calculations.

3.1 Changing bases

Let the operator $\mathcal{O}$ be a product of fields. In the Schrödinger picture its expectation value at a time t is given by

$$\langle \mathcal{O} \rangle(t) \equiv \langle \Omega(t_0)| U^\dagger(t,t_0) \mathcal{O} U(t,t_0) |\Omega(t_0)\rangle \ , \tag{3.1}$$

where $|\Omega(t_0)\rangle$ is the state of the system at the initial time t_0. The time-evolution operator $U(t,t_0)$ satisfies the Schrödinger equation

$$i\frac{d}{dt} U(t,t_0) = H(t)\, U(t,t_0) \tag{3.2}$$

with $U(t_0,t_0) = \mathbb{I}$ as the initial condition. Here $H(t)$ is the full Hamiltonian of the system.

Suppose that at t_0 the system was in its vacuum state, i.e. $|\Omega(t_0)\rangle$ is such that $E_0 \equiv \langle \Omega(t_0)|H(t_0)|\Omega(t_0)\rangle$ is the lowest energy assumed by any state at t_0. In most cases we are not able to find the explicit form of the full vacuum $|\Omega(t_0)\rangle$, but usually we can solve for the eigenstates of a part of the Hamiltonian, which we call H_0 and which corresponds to the free part of the theory,

$$H(t) = H_0(t) + H'(t). \tag{3.3}$$

Let us suppose that we have solved the eigenvalue problem for $H_0(t_0)$ at the initial time. The set of eigenstates of $H_0(t_0)$ can be used as a basis of our Hilbert space. We label them as

$$\big\{ |0(t_0)\rangle, |n(t_0)\rangle \big\}.$$

The state $|0(t_0)\rangle$ denotes the vacuum state of the free theory at t_0, and $|n(t_0)\rangle$ collectively represents all of the other eigenstates of H_0. We assume that this is a complete set in the sense that we can expand the identity operator in terms of it

$$\mathbb{I} = |0(t_0)\rangle\langle 0(t_0)| + \sum_n |n(t_0)\rangle\langle n(t_0)|.$$

We can use this completeness relation to convert a state in the eigenbasis of the full theory into its expression in the free theory's eigenbasis. The density matrix of the initial state $\rho_0 = \rho(t_0) = |\Omega(t_0)\rangle\langle\Omega(t_0)|$ can be written as

$$
\begin{aligned}
\rho_0 = \mathbb{I}\,|\Omega(t_0)\rangle\langle\Omega(t_0)|\,\mathbb{I} = \; & |0(t_0)\rangle\,\langle 0(t_0)|\Omega(t_0)\rangle\langle\Omega(t_0)|0(t_0)\rangle\,\langle 0(t_0)| \\
& + \sum_n |n(t_0)\rangle\,\langle n(t_0)|\Omega(t_0)\rangle\langle\Omega(t_0)|0(t_0)\rangle\,\langle 0(t_0)| \\
& + \sum_n |0(t_0)\rangle\,\langle 0(t_0)|\Omega(t_0)\rangle\langle\Omega(t_0)|n(t_0)\rangle\,\langle n(t_0)| \\
& + \sum_{n,n'} |n(t_0)\rangle\,\langle n(t_0)|\Omega(t_0)\rangle\langle\Omega(t_0)|n'(t_0)\rangle\,\langle n'(t_0)|.
\end{aligned}
$$

In general, $\langle n(t_0)|\Omega(t_0)\rangle \neq 0$, which means that from the perspective of the free theory, the true vacuum state looks as though it contains multiparticle excitations. But that is only because we are using the "wrong" basis; in the basis of the eigenstates of the full theory, $|\Omega(t_0)\rangle$ does not contain any excitations. It is the lowest energy state.

We have been speaking as though we knew $|\Omega(t_0)\rangle$, $U(t, t_0)$, etc. But if we did, there would be no need ever to resort to the eigenstates of the free theory. So how do we proceed, not knowing ρ_0? Let us make a few observations:

(1) If we really knew ρ_0 in the free eigenbasis, then we could calculate the expectation values of any operator (in principle) in the full vacuum state. Therefore, we should try to determine ρ_0 in this basis somehow.

(2) ρ_0—even though it is a pure state in the full eigenbasis—is a *mixed state* in the free theory's eigenbasis; that is,

$$\rho_{nn'} = \langle n(t_0)|\Omega(t_0)\rangle\langle\Omega(t_0)|n'(t_0)\rangle$$

does not need to be diagonal.

So the problem that we wish to solve is to evaluate an operator in a basis that we *do* understand with an initial state that we *do not* know. When $H'(t)$ is "small" in some sense, we can evaluate the expectation value perturbatively. In fact our approach will be perturbative in a double sense. First, by dividing $H = H_0 + H'$, we can similarly divide the time-evolution operator, $U(t, t_0) = U_0(t, t_0)U_I(t, t_0)$. Thus, we can write the expectation value of $\mathcal{O}$ in the interaction picture as

$$
\begin{aligned}
\langle \mathcal{O}(t)\rangle &= \mathrm{tr}\left[U_I^\dagger(t, t_0)U_0^\dagger(t, t_0)\mathcal{O}U_0(t, t_0)U_I(t, t_0)\rho_0\right] \\
&= \mathrm{tr}\left[U_I^\dagger(t, t_0)\mathcal{O}_I(t)U_I(t, t_0)\rho_0\right]
\end{aligned}
$$

where $\mathcal{O}_I(t) = U_0^\dagger(t, t_0)\mathcal{O}U_0(t, t_0)$ is the operator $\mathcal{O}$ in the interaction picture and $U_0(t, t_0) = Te^{-i\int_{t_0}^{t} dt'\, H_0(t')}$. The idea is that if H'—or the corresponding interaction

Hamiltonian in the interaction picture $H_I = U_0^\dagger(t,t_0)H'U_0(t,t_0)$—is small, we can treat the interactions pertubatively by expanding

$$U_I(t,t_0) = Te^{-i\int_{t_0}^t dt'\, H_I(t')}$$

in powers of H_I.

The second perturbative expansion is based on the idea that if H is close to H_0, $|0(t_0)\rangle$ ought also to be "close to" $|\Omega(t_0)\rangle$ in the sense that the overlap with the multi-particle states is small. If we can establish a few suitable criteria, we can determine ρ_0 in the free theory eigenbasis *perturbatively*. For example,

(1) ρ_0 should have the same symemtries as the full vacuum.

(2) If we believe that the state should match with what we should have obtained by extending back to the $t_0 \to -\infty$, then that requires certain structures in ρ_0.

The only variables around are the fields $\zeta(t,\vec{x})$; therefore, we should have that $\rho_0 = \rho(\zeta(t_0,\vec{x});t_0)$. It is convenient to write the initial density matrix in the following general form

$$\rho_0 = \frac{1}{Z}e^{iS_0} \ , \tag{3.4}$$

where Z is such that $\mathrm{tr}(\rho_0) = 1$. This idea was introduced in [18]. Since a particular configuration of the fields at the initial time t_0 is then weighted by a e^{iS_0} factor, we can think of S_0 as a boundary action on the initial time hypersurface [19]. Hence, the problem of determining the initial density matrix is reduced to the problem of constructing an appropriate initial action.

3.2 Single field inflation

Let us use the method we described in the previous section to calculate several cosmological correlation functions. We will work with a simple single-field, slow-roll inflationary model whose action is given by

$$S = \int d^4x\, \sqrt{-g}\left\{\frac{1}{2}M_{pl}^2 R + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right\} \ . \tag{3.5}$$

The metric for the spatially invariant background can be written as

$$ds^2 = dt^2 - e^{2\rho(t)}\delta_{ij}dx^i dx^j \ . \tag{3.6}$$

To analyze the fluctuations about this background it is convenient to write the metric in the following form

$$ds^2 = \left[N^2 - h_{ij}N^iN^j\right]dt^2 - 2h_{ij}N^i dt dx^j - h_{ij}dx^i dx^j \ . \tag{3.7}$$

Choosing the coordinates in which there are no fluctuations in the inflaton field $\phi(t,\vec{x}) = \phi(t)$ and where the spatial part of the metric is proportional to δ_{ij} and neglecting the tensor fluctuations we can write that

$$h_{ij} = e^{2\rho(t)+2\zeta(t,\vec{x})}\delta_{ij} \ . \tag{3.8}$$

In these coordinates the only scalar fluctuation left is $\zeta(t,\vec{x})$. The quadratic part of its action is

$$S^{(2)} = \frac{1}{2}\int dt\, \frac{\dot{\phi}^2}{\dot{\rho}^2}\int d^3\vec{x}\; e^{3\rho(t)}\left\{\dot{\zeta}^2 - e^{-2\rho(t)}\partial_k\zeta\partial^k\zeta\right\} \ . \tag{3.9}$$

The fields N and N^i are both nondynamical Lagrange multipliers, satisfying constraint equations

$$N = 1 + \frac{\dot{\zeta}}{\dot{\rho}} \ , \tag{3.10}$$

$$N^i = \delta_{ij}\partial^j\left\{-\frac{e^{-2\rho}}{\dot{\rho}}\zeta + \frac{1}{2}\frac{\dot{\phi}^2}{\dot{\rho}^2}\partial^{-2}\dot{\zeta}\right\} \ . \tag{3.11}$$

Expanding the inflationary action to third order in $\zeta(t,\vec{x})$ and going through lots of lengthy manipulations, in particular, doing many integrations by parts, the cubic action can be put into the following form [4, 3]

$$\begin{aligned} S^{(3)} = M_{pl}^2\int d^4x\, \Big\{ & \epsilon(3\epsilon+2\delta)e^{\rho}\zeta\partial_k\zeta\partial^k\zeta - \epsilon(\epsilon+2\delta)e^{3\rho}\dot{\zeta}^2\zeta - 2\epsilon^2 e^{3\rho}\dot{\zeta}\partial_k\zeta\partial^k\zeta\partial^{-2}\dot{\zeta} \\ & -\frac{1}{2}e^{3\rho}\epsilon^3[\dot{\zeta}^2\zeta - \zeta\partial_k\partial_l(\partial^{-2}\dot{\zeta})\partial^k\partial^l(\partial^{-2}\dot{\zeta})] \\ & +\Big\{\frac{d}{dt}[\epsilon e^{3\rho}\dot{\zeta}] - \epsilon e^{\rho}\partial_k\partial^k\zeta\Big\}\Big\{\frac{2}{\dot{\rho}}\dot{\zeta}\zeta - \frac{1}{2}\frac{e^{-2\rho}}{\dot{\rho}^2}[\partial_k\zeta\partial^k\zeta - \partial^{-2}\partial_k\partial_l(\partial^k\zeta\partial^l\zeta)] \\ & \qquad +\frac{1}{\dot{\rho}}\epsilon[\partial_k\partial^k(\partial^{-2}\dot{\zeta}) - \partial^{-2}\partial_k\partial_l(\partial^k\zeta\partial^l(\partial^{-2\dot{\zeta}}))]\Big\}\Big\} \ , \end{aligned} \tag{3.12}$$

where ϵ and δ are small in the slow-roll limit

$$\epsilon = \frac{1}{2}\frac{1}{M_{pl}^2}\frac{\dot{\phi}^2}{\dot{\rho}^2} \ll 1 \ , \tag{3.13}$$

$$\delta = \frac{1}{H}\frac{\ddot{\phi}}{\dot{\phi}} \ll 1 \ . \tag{3.14}$$

Only the first three operators in (3.12) have contributions that don't vanish in the late-time limit.

3.3 The three-point function

For simplicity, we will analyze the correlation functions using an abbreviated version of the standard single-field inflationary theory. We use the quadratic action given in (3.9), but from among the operators in the cubic action we will be only looking at one,

$$S^{(3)} = \int d^4x\; M_{pl}^2 \left\{ \epsilon(3\epsilon+2\delta)e^{\rho(t)}\zeta\partial_k\zeta\partial^k\zeta \right\} \; . \tag{3.15}$$

There are two reasons for doing so. First of all, for what we are trying to illustrate here, adding more cubic terms will not be any more illuminating and will only lengthen and complicate the calculation. Secondly, the standard "late-time", leading slow-roll set of operators is not even sufficient if we really wish to renormalize the single-field inflationary model. The renormalization must be done at an arbitrary time and not just in the late-time limit. All of the operators in (3.12) must be included then.

To calculate the three-point function here, and the two-point function in the next section, we work in the interaction picture and use the "in-in" formalism [24, 25, 26, 27]. In this formalism the three-point function can be written as

$$\begin{aligned}
&\langle\Omega(t)|\zeta(t,\vec{x})\zeta(t,\vec{y})\zeta(t,\vec{z})|\Omega(t)\rangle \qquad (3.16)\\
&= \langle\Omega(t_0)|U_I^\dagger(t,t_0)\zeta(t,\vec{x})\zeta(t,\vec{y})\zeta(t,\vec{z})U_I(t,t_0)|\Omega(t_0)\rangle \\
&= \langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})\zeta^+(t,\vec{z})e^{-i\int_{t_0}^t dt'\,[H_I^+(t')-H_I^-(t')]})|\Omega(t_0)\rangle \\
&= -i\int_{t_0}^t dt'\;\langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})\zeta^+(t,\vec{z})[H_I^+(t')-H_I^-(t')])|\Omega(t_0)\rangle + \cdots \; ,
\end{aligned}$$

where

$$H_I(t) = -M_{pl}^2\epsilon(3\epsilon+2\delta)e^{\rho(t)}\int d^3\vec{x}\;\zeta\partial_k\zeta\partial^k\zeta \tag{3.17}$$

and $H_I^\pm(t) \equiv H_I^+[\zeta^\pm(t,\vec{x})]$. The fields $\zeta^+(t,\vec{x})$ and $\zeta^-(t,\vec{x})$ are associated with $U_I(t,t_0)$ and $U_I^\dagger(t,t_0)$ respectively. The time-ordering operation is extended in the following sense: two "+" fields are ordered in the usual way,

$$T(\zeta^+(t,\vec{x})\zeta^+(t',\vec{y})) = \Theta(t-t')\zeta^+(t,\vec{x})\zeta^+(t',\vec{y}) + \Theta(t'-t)\zeta^+(t',\vec{y})\zeta^+(t,\vec{x}) \; , \tag{3.18}$$

"–" fields always occur after "+" fields,

$$\begin{aligned}
T(\zeta^+(t,\vec{x})\zeta^-(t',\vec{y})) &= \zeta^-(t',\vec{y})\zeta^+(t,\vec{x}) \; , \\
T(\zeta^-(t,\vec{x})\zeta^+(t',\vec{y})) &= \zeta^-(t,\vec{x})\zeta^+(t',\vec{y}) \; ,
\end{aligned}$$

and two "–" fields are ordered in the opposite of the usual sense,

$$T(\zeta^-(t,\vec{x})\zeta^-(t',\vec{y})) = \Theta(t'-t)\zeta^-(t,\vec{x})\zeta^-(t',\vec{y}) + \Theta(t-t')\zeta^-(t',\vec{y})\zeta^-(t,\vec{x}) \; . \tag{3.19}$$

Correspondingly, there are four types of propagators

$$
\begin{aligned}
\langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^+(t',\vec{y}))|\Omega(t_0)\rangle &= G^{++}(t,\vec{x};t',\vec{y}) \\
&= \Theta(t-t')G^{>}(t,\vec{x};t',\vec{y}) + \Theta(t'-t)G^{<}(t,\vec{x};t',\vec{y})\ , \\
\langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^-(t',\vec{y}))|\Omega(t_0)\rangle &= G^{+-}(t,\vec{x};t',\vec{y}) = G^{<}(t,\vec{x};t',\vec{y})\ , \\
\langle\Omega(t_0)|T(\zeta^-(t,\vec{x})\zeta^+(t',\vec{y}))|\Omega(t_0)\rangle &= G^{-+}(t,\vec{x};t',\vec{y}) = G^{>}(t,\vec{x};t',\vec{y})\ , \\
\langle\Omega(t_0)|T(\zeta^-(t,\vec{x})\zeta^-(t',\vec{y}))|\Omega(t_0)\rangle &= G^{--}(t,\vec{x};t',\vec{y}) \\
&= \Theta(t'-t)G^{>}(t,\vec{x};t',\vec{y}) + \Theta(t-t')G^{<}(t,\vec{x};t',\vec{y})\ .
\end{aligned}
$$

Here $G^{>}(t,\vec{x};t',\vec{y})$ and $G^{<}(t,\vec{x};t',\vec{y})$ are Wightman functions

$$
\begin{aligned}
G^{>}(t,\vec{x};t',\vec{y}) &= \langle\Omega(t_0)|\zeta(t,\vec{x})\zeta(t',\vec{y}))|\Omega(t_0)\rangle = \int \frac{d^3\vec{k}}{(2\pi)^3}\ e^{i\vec{k}(\vec{x}-\vec{y})}G^{>}_{k}(t,t')\ , \\
G^{<}(t,\vec{x};t',\vec{y}) &= \langle\Omega(t_0)|\zeta(t',\vec{y})\zeta(t,\vec{x}))|\Omega(t_0)\rangle = \int \frac{d^3\vec{k}}{(2\pi)^3}\ e^{i\vec{k}(\vec{x}-\vec{y})}G^{<}_{k}(t,t')\ .
\end{aligned}
$$

Using these rules to perform the contractions in (3.16), we find that the leading contribution to the three-point function is

$$
\begin{aligned}
&\langle\zeta_{\vec{k}_1}(t)\zeta_{\vec{k}_2}(t)\zeta_{\vec{k}_3}(t)\rangle \qquad (3.20)\\
&\quad = -2iM^2\epsilon(3\epsilon+2\delta)[\vec{k}_1\cdot\vec{k}_2+\vec{k}_1\cdot\vec{k}_3+\vec{k}_2\cdot\vec{k}_3] \\
&\qquad \times \int_{t_0}^{t} dt\ e^{\rho(t')}\ \left\{G^{>}_{k_1}(t,t')G^{>}_{k_2}(t,t')G^{>}_{k_3}(t,t') - G^{<}_{k_1}(t,t')G^{<}_{k_2}(t,t')G^{<}_{k_3}(t,t')\right\}\ .
\end{aligned}
$$

Since $\vec{k}_1+\vec{k}_2+\vec{k}_3=0$, we can rewrite the coefficients in a form that only depends on the magnitudes of the momenta,

$$
\vec{k}_1\cdot\vec{k}_2+\vec{k}_1\cdot\vec{k}_3+\vec{k}_2\cdot\vec{k}_3 = -\frac{1}{2}[k_1^2+k_2^2+k_3^2]\ . \qquad (3.21)
$$

To evaluate the time integral, let us switch from t to the conformal time η. Since we are working at leading order in the slow-roll parameters, we can write the scale factor and the integration measure in the de Sitter limit,

$$
\int_{t_0}^{t} dt'\ e^{\rho(t')}\cdots = \int_{\eta_0}^{\eta} d\eta'\ \frac{dt'}{d\eta'}\ e^{\rho(t')}\cdots = \int_{\eta_0}^{\eta} d\eta'\ e^{2\rho(t')}\cdots = \int_{\eta_0}^{\eta} d\eta'\ \frac{1}{H^2\eta'^2}\cdots\ . \qquad (3.22)
$$

In the standard case, where $t_0\to-\infty$, on the right-hand side of (3.16) one replaces $|\Omega(t_0)\rangle$ with the vacuum state of the free theory $|0\rangle\equiv|0(t_0)\rangle$, which in practice means using the Wightman functions of the free theory to evaluate (3.20). Then t_0 is set to $-\infty(1\pm i\epsilon)$ to project out the vacuum state of the interacting theory $|\Omega(t_0)$

from the vacuum state of the free theory $|0\rangle$. The Wightman functions of the free theory associated with the Bunch-Davies vacuum are

$$
\begin{aligned}
G_k^{>}(t,t') &= \frac{1}{4\epsilon}\frac{H^2}{M_{pl}^2}\frac{1}{k^3}(1+ik\eta)(1-ik\eta')e^{-ik(\eta-\eta')} \\
G_k^{<}(t,t') &= \frac{1}{4\epsilon}\frac{H^2}{M_{pl}^2}\frac{1}{k^3}(1-ik\eta)(1+ik\eta')e^{-ik(\eta-\eta')} \ .
\end{aligned} \tag{3.23}
$$

Substituting (3.23) into (3.20) and using the $i\epsilon$ prescription, which gets rid of the terms coming from the lower limit of the integral, we find that the three-point function is equal to

$$
\begin{aligned}
&\langle \zeta_{\vec{k}_1}(t)\zeta_{\vec{k}_2}(t)\zeta_{\vec{k}_3}(t)\rangle \\
&\quad = \frac{(3\epsilon+2\delta)}{32\epsilon^2}\frac{H^4}{M^4}\frac{(k_1^2+k_2^2+k_3^2)}{k_1^3k_2^3k_3^3} \\
&\qquad \times\left\{K-\frac{k_1k_2+k_1k_3+k_2k_3}{K}-\frac{k_1k_2k_3}{K^2}\right. \\
&\qquad \left. +\left(\frac{(k_1k_2+k_1k_3+k_2k_3)^2}{K}+\frac{(k_1k_2+k_1k_3+k_2k_3)k_1k_2k_3}{K^2}\right)\eta^2+\frac{k_1^2k_2^2k_3^2}{K}\eta^4\right\} ,
\end{aligned} \tag{3.24}
$$

where

$$
K = k_1+k_2+k_3 \ . \tag{3.25}
$$

But what should we do when t_0 is finite? Let us once again try to use (3.23) as our Wightman functions. In this instance, one part of the three-point function is the same as in (3.24), but there is also a piece from the lower limit of the integral in (3.20), which is equal to

$$
\begin{aligned}
\frac{(3\epsilon+2\delta)}{32\epsilon^2}\frac{H^4}{M^4}\frac{(k_1^2+k_2^2+k_3^2)}{k_1^3k_2^3k_3^3}&\left\{\frac{k_1k_2k_3}{K^2}[A\cos K(\eta_0-\eta)+B\sin K(\eta_0-\eta)]\right. \\
&+\frac{k_1k_2+k_1k_3+k_2k_3}{K}[A\cos K(\eta_0-\eta)+B\sin K(\eta_0-\eta)] \\
&+\frac{k_1k_2k_3}{K}\eta_0[B\cos K(\eta_0-\eta)+A\sin K(\eta_0-\eta)] \\
&\left.+\frac{1}{\eta_0}[B\cos K(\eta_0-\eta)+A\sin K(\eta_0-\eta)]\right\} ,
\end{aligned} \tag{3.26}
$$

where

$$
\begin{aligned}
A &= 1-(k_1k_2+k_1k_3+k_2k_3)\eta^2 \\
B &= K\eta-k_1k_2k_3\eta^3 \ .
\end{aligned}
$$

There are terms in (3.26) that either diverge or remain finite as $\eta_0\to-\infty$. The reason for the appearance of these terms is the fact that the free Bunch-Davies state is not the vacuum of the interacting theory. Since we are starting our evolution from a finite t_0, we can't simply use the $i\epsilon$ prescription to project out the vacuum state of the full theory. However, if we want to match smoothly with the interacting vacuum in the $\eta_0 \to -\infty$ limit, another recourse is open to us: to put a cubic term in the initial action. From what we have said earlier, this is equivalent to modifying the initial state, described in terms of the basis of the free theory at t_0, so that it corresponds more closely to the state that we really intended it to be. To do so we use a boundary operator whose structure mirrors the structure of $S^{(3)}$,

$$\begin{aligned} S_0^{(3)} &= M^2\epsilon(3\epsilon+2\delta)e^{2\rho(t_0)}\int d^3\vec{x}\, d^3\vec{y}\, d^3\vec{z}\big\{C(\vec{x},\vec{y},\vec{z})\zeta^+(t_0,\vec{x})\partial_k\zeta^+(t_0,\vec{y})\partial^k\zeta^+(t_0,\vec{z}) \\ &\qquad - C^*(\vec{x},\vec{y},\vec{z})[\zeta^+\to\zeta^-]\big\}\ . \end{aligned} \tag{3.27}$$

For this surface action to cancel the unwanted terms, we need

$$C(\vec{k_1},\vec{k_2},\vec{k_3}) = \frac{1}{K}\Big(\frac{1}{K\eta_0} - i\Big)\ . \tag{3.28}$$

By using $S^{(3)} + S_0^{(3)}$ as our cubic action to calculate the three-point function for a general t_0 we will recover (3.24) when taking $t_0 \to -\infty$. Notice, that for $t_0 \neq -\infty$, the three-point function will not be equal to (3.24). It will have some additional pieces that depend on t_0, but they all vanish when $t_0 \to -\infty$.

3.4 A one-loop correction to the two-point function

If we try to evaluate the two-point function beyond leading order with a finite time, we encounter the same problem as occurred with the three-point function: the lower ends of the integrals associated with the time-evolution of the states will produce pieces that are finite but oscillatory or that are divergent as we take $t_0 \to -\infty$. But here we should be more careful when removing these terms. The reason is that in this case there are other divergences coming from the dynamical part itself: the divergences of the three-momentum integrals in the loop. To take care of them we must supply the usual counterterms in the Lagrangian. These in turn will affect the initial time dependence of the two-point function. Only once we have summed both loop and the counterterm graphs, and isolated the finite oscillatory and divergent parts as $t_0 \to -\infty$ will we be able to determine the appropriate way to modify the state to cancel these effects.

3.4.1 Renormalizing the standard vacuum state

Using the "in-in" formalism we can write the two-point function as

$$\langle\Omega(t)|\zeta(t,\vec{x})\zeta(t,\vec{y})|\Omega(t)\rangle = \langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})e^{-i\int_{t_0}^t dt'\,[H_I^+(t')-H_I^-(t')]})|\Omega(t_0)\rangle\ . \tag{3.29}$$

For the one-loop contribution we have

$$\begin{aligned}
-\frac{1}{2}\int_{t_0}^t dt' \int_{t_0}^t dt''\ &\langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})[H_I^+(t')-H_I^-(t')])[H_I^+(t'')-H_I^-(t'')])|\Omega(t_0)\rangle \\
= -M_{pl}^4\epsilon^2(3\epsilon+2\delta)^2 &\int \frac{d^3\vec{p}}{(2\pi)^3}\ e^{i\vec{p}(\vec{x}-\vec{y})}\int_{t_0}^t dt'\ e^{\rho(t')}\int_{t_0}^{t'} dt''\ e^{\rho(t'')}\{G_p^>(t,t')-G_p^<(t,t')\} \\
&\times\int \frac{d^3\vec{q}}{(2\pi)^3}\ (p^2+q^2+k^2)^2\big\{G_p^>(t,t'')G_q^>(t',t'')G_k^>(t',t'') \\
&\qquad -G_p^<(t,t'')G_q^<(t',t'')G_k^<(t',t'')\big\}\ ,
\end{aligned}$$

where

$$k = |\vec{p}-\vec{q}|\ .$$

Again, for the case where $t_0 = -\infty$ we use the free Bunch-Davies Wightman functions and the $i\epsilon$ prescription for the lower ends of both integrals. Then the zeroth order contribution is just the usual Bunch-Davies propagator and the one-loop contribution is equal to

$$\langle\zeta_{\vec{p}}(t)\zeta_{-\vec{p}}(t)\rangle_{\text{loop}} = \frac{(3\epsilon+2\delta)^2}{256\epsilon^2}\frac{H^4}{M^4}\frac{1}{p^3}\{I_0+p^2\eta^2 I_2+p^4\eta^4 I_4\}\ , \tag{3.30}$$

where

$$\begin{aligned}
I_0 &= \frac{1}{2p^4}\int \frac{d^3\vec{q}}{(2\pi)^3}\ \frac{(p^2+q^2+k^2)^2}{q^3k^3(p+q+k)^2} \\
&\quad\times\{4kp^2(2p^2+2pq+3q^2)+2p^2(p+q)(2p^2+2pq+3q^2) \\
&\qquad +k^3(6p^2+5q^2)+k^2(10p^3+12p^2q+8pq^2+5q^3)\}\ ,
\end{aligned}$$

$$\begin{aligned}
I_2 &= \frac{1}{2p^4}\int \frac{d^3\vec{q}}{(2\pi)^3}\ \frac{(p^2+q^2+k^2)^2}{q^3k^3(p+q+k)^2} \\
&\quad\times\{4kp^2q^2+2p^2q^2(p+q)+k^3(2p^2+5q^2)+k^2(2p^3+4p^2q+8pq^2+5q^3)\}\ ,
\end{aligned}$$

$$I_4 = \frac{1}{p^4}\int \frac{d^3\vec{q}}{(2\pi)^3}\ \frac{(p^2+q^2+k^2)^2}{q^3k^3(p+q+k)^2}\{k^2q^2(k+2p+q)\}\ .$$

By doing power counting we can see that these integrals have divergences. In order to remove them, we introduce the necessary counterterms,

$$\mathcal{L}_{ct} = c_1M_{pl}^2e^{3\rho(t)}\dot{\zeta}^2 - c_2M_{pl}^2e^{\rho(t)}\partial_k\zeta\partial^k\zeta - c_3e^{-\rho(t)}\partial_l\partial_k\zeta\partial^l\partial^k\zeta\ . \tag{3.31}$$

The first two counterterms, which renormalize the operators in the quadratic action (3.9), are not enough to remove all divergences. We need the last four-derivative operator to cancel divergences proportional to $p^4\eta^4$. The $e^{-\rho(t)}$ prefactor is the one appropriate for the geometry: each pair of spatial indices is contracted with an h^{ij}, each of which brings an $e^{-2\rho(t)}$, and there is an overall factor of $\sqrt{-g}$ from the coordinate-invariant measure, which brings $e^{3\rho(t)}$. The corresponding contributions from these counterterms to the two-point function are

$$-\frac{1}{8\epsilon^2}\frac{H^2}{M^2}\frac{c_1}{p^3}(p^2\eta^2-1)\ , \tag{3.32}$$

$$-\frac{1}{8\epsilon^2}\frac{H^2}{M^2}\frac{c_2}{p^3}(p^2\eta^2+3)\ , \tag{3.33}$$

$$-\frac{1}{8\epsilon^2}\frac{H^4}{M^4}\frac{c_3}{2p^3}(2p^4\eta^4+5p^2\eta^2+5)\ . \tag{3.34}$$

To cancel the divergences due to the loop we should choose the coefficients of the counterterms to be

$$\begin{aligned}
c_3 &= \frac{(3\epsilon+2\delta)^2}{32}\Big[\text{infinite part of } I_4\Big]\ ,\\
c_2 &= \frac{H^2}{M^2}\frac{(3\epsilon+2\delta)^2}{128}\Big[\text{infinite part of } (I_0+I_2-5I_4)\Big]\ ,\\
c_3 &= \frac{H^2}{M^2}\frac{(3\epsilon+2\delta)^2}{128}\Big[\text{infinite part of } (3I_2-5I_4-I_0)\Big]\ .
\end{aligned}$$

Hence, for the renormalized loop we have

$$\langle\zeta_{\vec{p}}(t)\zeta_{-\vec{p}}(t)\rangle^{\rm ren}_{\rm loop} = \frac{(3\epsilon+2\delta)^2}{256\epsilon^2}\frac{H^4}{M^4}\frac{1}{p^3}\{I_0^f+p^2\eta^2 I_2^f+p^4\eta^4 I_4^f\}\ , \tag{3.35}$$

where the I^f-s are the finite parts of the corresponding integrals.

3.4.2 Renormalizing the vacuum state with an initial time

To evaluate the correction to the two-point function in the case of a finite t_0 we first replace (3.29) with its renormalized form,

$$\langle\Omega(t)|\zeta(t,\vec{x})\zeta(t,\vec{y})|\Omega(t)\rangle = \langle\Omega(t_0)|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})e^{-i\int_{t_0}^t dt'\,[\bar{H}_I^+(t')-\bar{H}_I^-(t')]})|\Omega(t_0)\rangle\ , \tag{3.36}$$

where

$$\bar{H}_I(t) = H_I(t)+H_{ct}(t) \tag{3.37}$$

and

$$H_{ct} = -\mathcal{L}_{ct}\ . \tag{3.38}$$

To be able to use the free theory Wightman functions we must switch from $|\Omega(t_0)\rangle$ to $|0\rangle$. When making this transition we need to take into account that from the perspective of the free theory the evolution is governed not just by the Hamiltonian $\bar{H}_I$, but also by the initial state cubic action (3.27) that we already included to correct the three-point function. This means that we can replace the right-hand side of (3.36) with

$$\langle 0|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})e^{-i\int_{t_0}^t dt'\,[\bar{H}_I^+(t')-\bar{H}_I^-(t')]+iS_0^{(3)}})|0\rangle \ . \tag{3.39}$$

Since $S_0^{(3)}$ is of the same order in the slow-role parameters as H_I we need to take its contribution into account. Thus, the one-loop correction to the two-point function will be

$$\begin{aligned} -\frac{1}{2}\int_{t_0}^t dt' \int_{t_0}^t dt''\ \langle 0|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})[\bar{H}_I^+(t') - \bar{H}_I^-(t') + H_0^{(3)}(t')]) \\ \times[H_I^+(t'') - H_I^-(t'') + H_0^{(3)}(t'')])|0\rangle \ , \end{aligned} \tag{3.40}$$

where

$$H_0^{(3)}(t) = -\frac{1}{2}\delta(t-t_0)S_0^{(3)} \ . \tag{3.41}$$

The part of (3.40) that is independent of the initial time η_0 will be the same as (3.35). The part that depends on η_0 will have terms that vanish, stay finite (and oscillate) or diverge (linearly and quadratically in η_0) as $\eta_0 \to -\infty$. But when $\eta_0 \to -\infty$ we want (3.40) to match with (3.35); hence, we need to eliminate the last two types of terms. It can be done order by order in η_0. Here we will present the elimination of the quadratically divergent terms. The term from the loop quadratic in η_0 is equal to

$$\begin{aligned} &\frac{(3\epsilon+2\delta)^2}{256\epsilon^2}\frac{H^4}{M^4}\frac{1}{p^3}p^2\eta_0^2 \\ &\times\Bigg\{\Big[(1-p^2\eta^2)\cos 2p(\eta-\eta_0) + 2p\eta\sin 2p(\eta-\eta_0)\Big]\Big[J_1 - 2J_0 - 4J_2 - \frac{32}{(3\epsilon+2\delta)^2}c_3\Big] \\ &\qquad -(1+p^2\eta^2)J_0\Bigg\} \ , \end{aligned} \tag{3.42}$$

where

$$\begin{aligned} J_0 &= \frac{1}{p^3}\int \frac{d^3\vec{q}}{(2\pi)^3}\,\frac{(p^2+q^2+k^2)^2}{qk(p+q+k)^2} \ , \\ J_1 &= \frac{1}{p^4}\int \frac{d^3\vec{q}}{(2\pi)^3}\,\frac{(p^2+q^2+k^2)^2}{qk(q+k-p)} \ , \\ J_2 &= \frac{1}{p^3}\int \frac{d^3\vec{q}}{(2\pi)^3}\,\frac{(p^2+q^2+k^2)^2}{qk(p+q+k)(q+k-p)} \ . \end{aligned}$$

To remove it we add a quadratic term to the initial action

$$\begin{aligned}
S_0^{(2)} &= \frac{1}{2}\int d^3\vec{x}\, d^3\vec{y}\,\Big\{\zeta^+(t_0,\vec{x})A(\vec{x}-\vec{y})\zeta^+(t_0,\vec{y}) - \zeta^-(t_0,\vec{x})A^*(\vec{x}-\vec{y})\zeta^-(t_0,\vec{y}) \\
&\qquad +2i\zeta^+(t_0,\vec{x})B(\vec{x}-\vec{y})\zeta^-(t_0,\vec{y})\Big\} \\
&= \frac{1}{2}\int d^3\vec{x}\, d^3\vec{y}\,\Big\{\mathrm{Re}A(\vec{x}-\vec{y})\big[\zeta^+(t_0,\vec{x})\zeta^+(t_0,\vec{y}) - \zeta^-(t_0,\vec{x})\zeta^-(t_0,\vec{y})\big] \\
&\qquad +\mathrm{Im}A(\vec{x}-\vec{y})\big[\zeta^+(t_0,\vec{x})\zeta^+(t_0,\vec{y}) + \zeta^-(t_0,\vec{x})\zeta^-(t_0,\vec{y})\big] \\
&\qquad +2i\zeta^+(t_0,\vec{x})B(\vec{x}-\vec{y})\zeta^-(t_0,\vec{y})\Big\}\ .
\end{aligned}$$

To first order the contribution to the two-point function coming from this term is

$$i\langle 0|T(\zeta^+(t,\vec{x})\zeta^+(t,\vec{y})S_0^{(2)})|0\rangle\ . \tag{3.43}$$

The part of (3.43) leading in η_0 is equal to

$$\begin{aligned}
\langle S_0^{(2)}\rangle &= -\frac{1}{8\epsilon^2}\frac{H^4}{M^4}\frac{1}{p^6}p^2\eta_0^2\Big\{\big[(1-p^2\eta^2)\sin 2p(\eta-\eta_0) - 2p\eta\cos 2p(\eta-\eta_0)\big]\mathrm{Re}A_p \\
&\qquad -\big[(1-p^2\eta^2)\cos 2p(\eta-\eta_0) + 2p\eta\sin 2p(\eta-\eta_0)\big]\mathrm{Im}A_p \\
&\qquad +(1+p^2\eta^2)B_p\Big\}\ .
\end{aligned} \tag{3.44}$$

Comparing (3.44) to (3.42) we can conclude that for $S_0^{(2)}$ to cancel the quadratically divergent terms we need

$$\begin{aligned}
\mathrm{Re}A_p &= 0\ , \\
\mathrm{Im}A_p &= p^3\Big[c_3 - \frac{(3\epsilon+2\delta)^2}{32}\big[\text{infinite part of } (J_1 - 2J_0 - 4J_2)\big]\Big] \\
&= \frac{(3\epsilon+2\delta)^2}{32}p^3\Big[\text{infinite part of } (I_4 - J_1 + 2J_0 + 4J_2)\Big]\ , \\
B_p &= -\frac{(3\epsilon+2\delta)^2}{16}\Big[\text{infinite part of } J_0\Big]\ .
\end{aligned}$$

To fully renormalize the one-loop correction to the two-point function we also need to extract and eliminate from (3.40) the terms that are zeroth and first order in η_0. Since there is no principal difference between treating these terms and treating the quadratically divergent term, these further calculations are not essential for demonstrating the technique that we are introducing in this paper.

3.5 Conclusions

For the reasons that we talked about in the introduction, it is important to be able to start the evolution of the system from a finite initial time. In this paper we presented a formalism that allows us to calculate correlation functions for states that are defined at some initial time. Using this formalism we can choose a particular state of the interacting theory at an arbitrary time, and not only in the infinite past.

We demonstrated this technique of renormalizing the initial state for the case of the vacuum state of a toy model derived from the standard inflationary theory with a single scalar field. Using the eigenbasis of the free theory and applying matching conditions for the two- and three-point functions we were able to start constructing the initial density matrix order by order in perturbation theory: inclusion of this density matrix eliminated the unwanted finite oscillatory and divergent terms from the two- and three-point functions.

In principle, this method can be used to renormalize other, more complicated, states, although that task might be more challenging. The main difficulty is to determine the conditions that the state should satisfy. We need to be able to translate our ideas about the physical properties of a certain state into conditions on some of its n-point functions. For any non-vacuum state we must start with an initial density matrix that already has some nontrivial structures. If the state we want to consider is such that it has a corresponding state in the free theory, we can start with an initial action that is only quadratic in the fields; otherwise the initial action needs to have structures of higher orders. Since we work in the free theory eigenbasis, the operators in the initial action will be defined with respect to the free theory vacuum. After applying the appropriate conditions these operators will need to be modified.

Chapter 4

The quantum Fokker-Planck equation of stochastic inflation

4.1 Introduction

Certain quantum field theories in an expanding background that involve massless minimally coupled scalar fields, e.g. a simple scalar theory with a quartic interaction ϕ^4 or scalar electrodynamics, have logarithmic infrared divergences in the loop-corrections to their Green's functions [28, 29, 30, 31, 32, 33, 34, 35, 36, 37]. These divergences are associated with the late-time behavior of the theory and come from the modes, whose physical wavelengths have been stretched to scales much larger than the Hubble scale. These logarithmic divergences come from two sources: first, they appear even at the level of the two-point functions of free fields, second, they get enhanced by vertex integration [35].

Infrared divergences that we encounter in theories that are set in the expanding background signal the breakdown of the perturbation theory past a certain point in time. Since the late-time behavior of correlation functions is very important in cosmology, it is crucial to have a non-perturbative technique in order to be able to calculate them. Starobinsky [38, 39] had argued that the dynamics of the long-wavelength modes can be described by the classical stochastic field, whose probability distribution satisfies a Fokker-Planck type equation. He showed that this equation has a late-time static solution, that can be used to calculate correlation functions in the late-time limit. In essence, Starobinsky's Fokker-Planck equations manage to resum the leading log behavior of the perturbative expansion.

Despite the compelling simplicity of the stochastic picture, it would appear to be very difficult to see how it could emerge by following the full quantum evolution of the theory. A recent approach [40, 41] to this problem has been to consider the quantum evolution of the theory from a different perspective by working in the Schrödinger picture.[1] In this picture, the connection between the *classical* probability function of the stochastic description and the *quantum* density matrix of the scalar theory becomes much clearer. The essential element that was missing from these works was the fully quantum treatment of the interactions. In [40, 41], the role of the interactions was only introduced as a background effect on an otherwise quadratic — purely Gaussian — theory.

In this article we add this important missing ingredient to show how to derive the stochastic picture for a genuinely interacting field theory. In particular, we consider here a massless scalar field with a quartic interaction in de Sitter space, and solve for the full time dependence of its density matrix perturbatively in the self-coupling of the field. Once this evolution has been found to a given order, we can then project onto the theory of the long wavelength fluctuations by integrating out the short wavelength parts of the field. The resulting density matrix for this effective theory of the long wavelength fluctuations satisfies a fully quantum version of the Fokker-Planck equation. Essentially, the coarsely grained Liouville equation for the density matrix of the effective theory *is* the quantum version of the Fokker-Planck equation. The parallel between the quantum theory and its stochastic description emerges very naturally in this picture. It becomes a simple matter to read off the stochastic noise and drift from the corresponding quantum version of the equation, as we shall show.

Additionally, by computing the wave-functional for a quartic theory explicitly, we gain a far deeper understanding of the time-evolution of the fluctuations and are able to follow how the structures that depend on the interactions behave both inside and far outside the horizon. Here we construct the wave-functional for the *interacting* Bunch-Davies state perturbatively. When a fluctuation is well inside the horizon, its part of the wave-functional is *close to* Gaussian; but — significantly — it always contains higher-order structures as well. The role of these higher-order parts grows once the momentum associated with a particular fluctuation crosses the horizon. We can then see very clearly how their leading behaviour in the long-wavelength limit leads directly to the drift term of the Fokker-Planck equation.

The purpose of this work is not just to verify the validity of the stochastic picture, but to go further and to lay the groundwork for a more powerful formalism. Having at our disposal a complete derivation that connects the quantum and the classical stochastic theories as we have done here, we can address questions that would be difficult, or otherwise impossible, to approach from the stochastic side. For example, what are the higher order corrections to the standard stochastic picture? How does the stochastic limit arise in the theories of other massless fields and what is the influence of their interactions with other fields? We can even explore, in principle, the degree to which the standard static-limit solution of the stochastic picture is an attractor solution. Such applications of our approach, together with a few others, are mentioned at the end of this article.

[1] A much earlier treatment of a free scalar field in the Schrödinger picture is found in [42].

4.2 The stochastic description of a quantum theory

In a theory of a massless, interacting scalar field, $\Phi(t,\vec{x})$, the simplest quantities that we could calculate are the n-point functions where all the fields are evaluated at exactly the same space-time point and at some suitably late time,

$$\langle \Phi^n(t,\vec{x}) \rangle \equiv \lim_{t\to\infty} \langle \Omega(t) | \Phi^n(t,\vec{x}) | \Omega(t) \rangle. \tag{4.1}$$

$|\Omega(t)\rangle$ denotes the state that we have chosen for our quantum field, which we shall take to be the Bunch-Davies state — the de Sitter invariant state matching the standard Minkowski space vacuum at very short distances. Because we have chosen a de Sitter invariant state and because we are assuming too that we are working in spatial coordinates where the background is invariant under spatial translations, $\langle \Phi^n(t,\vec{x}) \rangle$ cannot depend on the position $\vec{x}$.

Here we are not really interested in the n-point functions of the full theory, which contain information about all scales, but only in the n-point functions of the effective theory of the long wavelength fluctuations of the field, $\langle \Phi_L^n(t,\vec{x}) \rangle$. What we mean by a long or a short wavelength is one whose physical momentum is small or large compared with the Hubble scale, H, associated with the curvature of the de Sitter background. Or, more precisely, we shall use a slightly stricter definition, setting the threshold between 'long' and 'short' to be well outside the horizon, which can be done by introducing a small parameter $\varepsilon \ll 1$.

$$\begin{aligned} \text{long wavelength } (L): &\qquad k < \varepsilon a H \\ \text{short wavelength } (S): &\qquad k > \varepsilon a H, \end{aligned}$$

$a(t)$ is the scale factor associated with the expanding space-time. This definition is more appropriate because, with the extremely rapid expansion during inflation, the physical fluctuations corresponding to the scalar fluctuations of inflation which are needed to explain the primordial fluctuations in the early universe would have been stretched far outside the horizon by the end of the inflationary era. Moreover, if H is meant to be the true cutoff of our effective theory, we should not be including momenta all the way up to this scale.[2] Of course, in an expanding background the threshold for our effective theory also becomes time-dependent. If we divide our scalar field $\Phi(t,\vec{x})$ into two parts,

$$\Phi(t,\vec{x}) = \Phi_L(t,\vec{x}) + \Phi_S(t,\vec{x}) = \int_{k<\varepsilon a H} \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot\vec{x}} \Phi_{\vec{k}}(t) + \int_{k>\varepsilon a H} \frac{d^3\vec{k}}{(2\pi)^3}\, e^{i\vec{k}\cdot\vec{x}} \Phi_{\vec{k}}(t),$$

whether a particular $\Phi_{\vec{k}}(t)$ appears in the first or the second integral depends on the value of $\varepsilon a(t) H$ at that moment. So 'long' and 'short' do not have an absolute

[2] In a similar sense one would not use Fermi's theory of β decay all the way up to the electroweak scale.

physical meaning in de Sitter space, but change over time. A practical consequence — and one that will later prove to be important in our calculation — is that this additional time dependence will mean that derivatives can also act on the limits of integrals once we have restricted to just the long-wavelength momenta.

Now suppose that we have determined all of the values of the n-point functions of this effective theory. We could then introduce a classical variable[3] φ, together with a probability distribution function $p(t, \varphi)$, such that together they reproduce all the information contained in the functions $\langle \Phi_L^n(t, \vec{x}) \rangle$. The weighted average of a power of this variable is defined by the following integral,

$$\langle \varphi^n \rangle(t) \equiv \int_{-\infty}^{\infty} d\varphi\, \varphi^n p(t, \varphi). \tag{4.2}$$

Notice that while φ itself is just a variable without any time dependence, the average $\langle \varphi^n \rangle(t)$ inherits its time dependence from $p(t, \varphi)$. We can then choose the weighting function $p(t, \varphi)$ so that the expectation values of this classical variable φ exactly match with the corresponding n-point functions of our effective theory of long wavelength fluctuations,

$$\langle \varphi^n \rangle(t) = \langle \Phi_L^n(t, \vec{x}) \rangle, \tag{4.3}$$

once, of course, we have formulated a suitable meaning for $\langle \Phi_L^n(t, \vec{x}) \rangle$ derived from our original theory.

The stochastic theory of inflation [38, 39] argues that the probability function for this classical variable should satisfy a Fokker-Planck equation of the form

$$\frac{\partial p}{\partial t} = N \frac{\partial^2 p}{\partial \varphi^2} + D \frac{\partial}{\partial \varphi} \left(\frac{\partial V}{\partial \varphi} p(t, \varphi) \right). \tag{4.4}$$

The coefficients N and D are called the 'noise' and the 'drift' of this stochastic theory. $V(\varphi)$ is a function of the stochastic variable, which is assumed to have the same functional form as the corresponding potential of the quantum theory; that is, one obtains $V(\phi)$ by simply replacing the quantum field $\Phi(t, \vec{x})$ with the stochastic variable φ in the original quantum potential,

$$V(\Phi(t, \vec{x})) \xrightarrow{\Phi(t, \vec{x}) \to \varphi} V(\varphi).$$

The fact that $p(t, \varphi)$ is a solution to the Fokker-Planck equation can then be used to generate a recursion relation[4] amongst the various averages $\langle \varphi^n \rangle$. One starts by taking its time derivative,

$$\frac{\partial}{\partial t} \langle \varphi^n \rangle = \int_{-\infty}^{\infty} d\varphi\, \varphi^n \frac{\partial p}{\partial t} = \int_{-\infty}^{\infty} d\varphi\, \varphi^n \left\{ N \frac{\partial^2 p}{\partial \varphi^2} + D \frac{\partial}{\partial \varphi} \left(\frac{\partial V}{\partial \varphi} p(t, \varphi) \right) \right\}, \tag{4.5}$$

[3]Since we are only analysing n-point functions whose fields are all evaluated at the same space-time point, a *variable* φ suffices. If we had wished to consider the expectation values of fields at different points, we should have needed to generalise to a classical stochastic *field*, $\varphi(\vec{x})$, instead.

[4]This recursion relation was found already in [37].

and integrates by parts as needed — twice for the first term and once for the second term — to produce

$$\frac{\partial}{\partial t}\langle\varphi^n\rangle = n(n-1)N\langle\varphi^{n-2}\rangle - nD\left\langle \varphi^{n-1}\frac{\partial V}{\partial\varphi}\right\rangle. \tag{4.6}$$

For a quartic potential, $V(\varphi) = \frac{1}{4!}\lambda\varphi^4$, this recursion relation has the form

$$\frac{\partial}{\partial t}\langle\varphi^n\rangle = n(n-1)N\langle\varphi^{n-2}\rangle - nD\frac{\lambda}{6}\langle\varphi^{n+2}\rangle. \tag{4.7}$$

So knowing N and D, together with the time-dependence of $\langle\varphi^2\rangle$, is sufficient for calculating all the non-vanishing averages, $\langle\varphi^n\rangle$.

What is the origin of this stochastic description of the theory from the perspective of the quantum theory? The parallels between the stochastic and the quantum versions of the theory emerge most directly when we treat the quantum theory in the Schrödinger picture. In this picture, the evolution of the expectation value of an operator occurs entirely in the state; operators,[5] such as the products of n fields, $\phi(\vec{x})$, have no explicit time dependence. The closest analogue of the probability function in the stochastic description is the density matrix — or rather, its diagonal components — associated with the state that we have chosen. In fact we need to treat two versions of the density matrix: that of the full theory, which we denote by

$$P[\phi] = P[\phi_L, \phi_S] = \Psi[\phi]\Psi^*[\phi], \tag{4.8}$$

as well as the density matrix for the effective theory that just includes the long wavelength fluctuations, which will be denoted by $P_\Omega[\phi_L]$. The evolution of $P[\phi]$ for the full theory is determined entirely by its Liouville equation. The evolution of $P_\Omega[\phi_L]$ is then derived through its relation to $P[\phi]$ together with our knowledge of how $P[\phi]$ itself evolves.

A general, equal-time, expectation value for the product of n fields is given by

$$\begin{aligned} &\langle\Phi_L(t,\vec{x}_1)\cdots\Phi_L(t,\vec{x}_n)\rangle \\ &= \int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\cdots\frac{d^3\vec{k}_n}{(2\pi)^3}\, e^{i\vec{k}_1\cdot\vec{x}_1}\cdots e^{i\vec{k}_n\cdot\vec{x}_n}\,(2\pi)^3\delta^3(\vec{k}_1+\cdots+\vec{k}_n)\,\langle\Phi_{\vec{k}_1}(t)\cdots\Phi_{\vec{k}_n}(t)\rangle. \end{aligned}$$

The notation that we shall adopt here is that an 'L' subscript in an integral indicates that all of the momenta accompanying the integral sign are only those corresponding to physical wavelengths that have been stretched well outside the horizon, $k < \varepsilon aH$. The momentum conserving δ-function follows from the invariance of the background

[5]We shall denote the fields in the Schrödinger picture with a lower case notation, $\phi(\vec{x})$, while the upper case $\Phi(t,\vec{x})$ represents the field more generally, independent of a particular picture.

under spatial translations. When all of the fields are evaluated at the same spatial position, this δ-function causes the exponential factors to vanish,

$$\langle \Phi_L^n(t,\vec{x})\rangle = \int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\cdots\frac{d^3\vec{k}_n}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\cdots+\vec{k}_n)\,\langle \Phi_{\vec{k}_1}(t)\cdots\Phi_{\vec{k}_n}(t)\rangle. \tag{4.9}$$

In the Schrödinger picture, the matrix elements are found by functionally integrating over the relevant degrees of freedom, which in this case are the $\phi_{\vec{k}}$'s whose momentum label $\vec{k}$ corresponds to a long wavelength, weighted by the density matrix $P_\Omega[\phi_L]$ for the effective theory,

$$\langle \Phi_{\vec{k}_1}(t)\cdots\Phi_{\vec{k}_n}(t)\rangle = \int_L \mathcal{D}\phi_{\vec{k}}\;\;\phi_{\vec{k}_1}\cdots\phi_{\vec{k}_n}P_\Omega[\phi_L]. \tag{4.10}$$

The $\phi_{\vec{k}}$'s appearing in this expression are the fields written in the Schrödinger picture.[6] Since the time-dependence is entirely in the density matrix, $\phi_{\vec{k}}$ does not depend on the time.

Now let us imagine for the moment that the density matrix $P_\Omega[\phi_L]$ itself satisfies a functional Fokker-Planck equation of the form

$$\frac{\partial P_\Omega}{\partial t} = \int_L \frac{d^3\vec{k}}{(2\pi)^3}\left\{\mathcal{N}_k\frac{\delta^2 P_\Omega}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} + \mathcal{D}\frac{\delta}{\delta\phi_{\vec{k}}}\left[\frac{\delta\mathcal{V}_\Omega}{\delta\phi_{-\vec{k}}}P_\Omega\right]\right\}. \tag{4.11}$$

The $\mathcal{V}_\Omega$ in this expression is the potential for the long wavelength fluctuations of the fields. For a quartic theory, this potential would be

$$\mathcal{V}_\Omega[\phi_L] = \frac{1}{4!}\lambda\int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\,\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4}. \tag{4.12}$$

We can follow the same procedure that we used in the stochastic description of the theory to generate an analogous recursion relation for the quantum effective theory. When we differentiate $\langle\Phi_L^n(t,\vec{x})\rangle$ with respect to the time and use the appropriate quantum form of the Fokker-Planck equation, we are led to the recursion relation[7]

$$\frac{\partial}{\partial t}\langle\Phi_L^n(t,\vec{x})\rangle = n(n-1)\left(\int_L\frac{d^3\vec{k}}{(2\pi)^3}\,\mathcal{N}_k\right)\langle\Phi_L^{n-2}(t,\vec{x})\rangle - n\mathcal{D}\frac{\lambda}{6}\langle\Phi_L^{n+2}(t,\vec{x})\rangle. \tag{4.13}$$

---

[6] Because of the limits on the integral, the label $\vec{k}$ is always in the region $k < \varepsilon aH$ in this expression. It would be redundant — at least to the order to which we shall be working — and a little cumbersome to write $\phi_{L,\vec{k}}$. Therefore we shall not do so.

[7] Had we allowed the drift term in the quantum version of the Fokker-Planck equation to depend on the momentum as well, $\mathcal{D}_k$, we should have arrived at the following quantum recursion relation instead,

$$\frac{\partial}{\partial t}\langle\Phi_L^n(t,\vec{x})\rangle = n(n-1)\left(\int_L\frac{d^3\vec{k}}{(2\pi)^3}\,\mathcal{N}_k\right)\langle\Phi_L^{n-2}(t,\vec{x})\rangle - n\frac{\lambda}{6}\langle\Phi_L^{n+2}(t,\vec{x})\rangle_{\mathcal{D}},$$

where the final $n+2$ point function has been replaced with a 'drift-weighted' version of itself,

$$\langle\Phi_L^{n+2}(t,\vec{x})\rangle_{\mathcal{D}} \equiv \int_L\frac{d^3\vec{k}_1}{(2\pi)^3}\cdots\frac{d^3\vec{k}_{n+2}}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\cdots+\vec{k}_{n+2})\,\mathcal{D}_{\|\vec{k}_n+\vec{k}_{n+1}+\vec{k}_{n+2}\|}\langle\Phi_{\vec{k}_1}(t)\cdots\Phi_{\vec{k}_{n+2}}(t)\rangle.$$

This time we have simply written the result for the particular case of a quartic interaction, rather than for a general polynomial potential. Comparing the two recursion relations, we realise that if the stochastic and the quantum descriptions of the n-point functions are to agree, $\langle\varphi^n\rangle = \langle\Phi_L^n(t,\vec{x})\rangle$, the noise and the drift coefficients of the stochastic Fokker-Planck equation are derived directly from the quantum ones by identifying

$$N = \int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\mathcal{N}_k \qquad \text{and} \qquad D = \mathcal{D}. \tag{4.14}$$

So the path leading from the quantum theory to its stochastic description now becomes clear:

i. We must first solve for the wave-functional and the corresponding density matrix of our full theory, which includes both the long and short wavelength parts of the field. For this purpose, the Schrödinger picture is the best suited, as we shall see.

ii. Once we have determined the density matrix for the state that we have chosen, $P[\phi] = P[\phi_L, \phi_S]$, which here will be the Bunch-Davies state, we project onto the effective theory of the long wavelength part of the field. The most straightforward thing to do is simply to integrate out the short wavelength fluctuations directly and define the density matrix for the effective theory to be

$$P_\Omega[\phi_L] = \int_S \mathcal{D}\phi_{\vec{p}}\,P[\phi_L, \phi_S].$$

iii. The time dependence of $P_\Omega[\phi_L]$ — or rather that of the various functions within it — follows straightforwardly from the time dependence of the functions that appear in the density matrix of the full theory, $P[\phi]$. Their time dependence, in turn, follows from the Schrödinger equation for the wave-functional of the state.

iv. Knowing this time dependence of $P_\Omega[\phi_L]$ then allows us to compute its time derivative explicitly. The resulting equation is a functional Fokker-Planck equation with precisely the form that we claimed that it should have. This functional Fokker-Planck equation for the effective theory could equally be regarded as the coarse-grained version of the Liouville equation derived from the full density matrix $P[\phi]$.

We illustrate these steps by applying them to a familiar example. So we turn next to the case of a scalar field theory with a quartic interaction.

Since the subsequent calculation will show that $\mathcal{D}$ is momentum independent, we shall simply draw upon this foreknowledge here and not consider this more general possibility for the quantum drift for now.

4.3 A quartic interaction

This method for deriving the stochastic description of a quantum theory is best shown through a particular example. For this purpose we choose the theory of a real scalar field in a de Sitter background with a quartic self-interaction, $V(\Phi) = \frac{1}{4!}\lambda\Phi^4$. Provided that the coupling is sufficiently small, this theory can be solved perturbatively in λ. We shall also include a mass for the field for the time being, although we shall ultimately set it to zero. The action for the theory is written as

$$S[\Phi] = \int dt\, L[\Phi] = \int d^4x\, \sqrt{-g}\, \left\{ \tfrac{1}{2} g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi - \tfrac{1}{2}m^2\Phi^2 - \tfrac{1}{24}\lambda\Phi^4 \right\}. \tag{4.15}$$

The metric $g_{\mu\nu}$ for the de Sitter background can be expressed in a spatially flat form, as we had assumed earlier, either in terms of a 'cosmological' time coordinate $t \in (-\infty, \infty)$ or a 'conformal' one $\eta \in (-\infty, 0)$,

$$ds^2 = dt^2 - a^2(t)\,\delta_{ij}\,dx^i dx^j = a^2(\eta)\left[d\eta^2 - \delta_{ij}\,dx^i dx^j\right]. \tag{4.16}$$

We shall later use whichever of these two times best suits our need at the particular moment. These time coordinates are related to each other through the condition $dt = a(\eta)\,d\eta$, and the scale factor a expressed in these two coordinate systems has the form

$$a(t) = e^{Ht} \qquad \text{or} \qquad a(\eta) = -\frac{1}{H\eta}. \tag{4.17}$$

Whereas the final result cannot depend on which picture we have chosen, the interaction picture is not the best suited for drawing the parallels between the stochastic and quantum Fokker-Planck equations. Instead we study the evolution of the theory from a Schrödinger perspective. The time dependence of the state, described in terms of a wave-functional $\Psi[\phi]$, is found by solving the Schrödinger equation,

$$i\frac{\partial\Psi}{\partial t} = H\Psi, \tag{4.18}$$

where $H[\pi(\vec{x}), \phi(\vec{x})]$ is the Hamiltonian written in terms of the time-independent Schrödinger-picture field $\phi(\vec{x})$ and its conjugate momentum $\pi(\vec{x})$. In the space-time coordinates that we have chosen, the Lagrangian for our theory is given by

$$L[\Phi] = \int d^3\vec{x}\, \left\{ \tfrac{1}{2}a^3\dot{\Phi}^2 - \tfrac{1}{2}a\delta^{ij}\partial_i\Phi\partial_j\Phi - \tfrac{1}{2}a^3m^2\Phi^2 - \tfrac{1}{24}a^3\lambda\Phi^4 \right\}, \tag{4.19}$$

and the corresponding canonical momenta are

$$\Pi(t, \vec{x}) = \frac{\delta L}{\delta\dot{\Phi}(t, \vec{x})} = a^3\dot{\Phi}(t, \vec{x}).$$

If we perform the usual Legendre transformation, we are led to the Hamiltonian,

$$H = \int d^3\vec{x}\,\{\Pi\dot{\Phi}\} - L = \int d^3\vec{x}\,\Big\{\tfrac{1}{2}a^{-3}\Pi^2 + \tfrac{1}{2}a\delta^{ij}\partial_i\Phi\partial_j\Phi + \tfrac{1}{2}a^3m^2\Phi^2 + \tfrac{1}{24}a^3\lambda\Phi^4\Big\}. \tag{4.20}$$

It is must easier to describe the truncation of the full theory to its long wavelength parts in terms of the momenta of the fields rather than in terms of their positions. So, after performing the Fourier transformation of the Schrödinger picture fields,

$$\phi_{\vec{k}} = \int d^3\vec{x}\, e^{-i\vec{k}\cdot\vec{x}}\phi(\vec{x}),$$

the Hamiltonian assumes the form

$$\begin{aligned} H \;=\; & \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\vec{k}_2)\Big\{-\frac{1}{2}\frac{1}{a^3}\frac{\delta}{\delta\phi_{\vec{k}_1}}\frac{\delta}{\delta\phi_{\vec{k}_2}} + \frac{1}{2}a^3\Big(m^2+\frac{k_1^2}{a^2}\Big)\phi_{\vec{k}_1}\phi_{\vec{k}_2}\Big\} \\ & + \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\frac{d^3k_3}{(2\pi)^3}\frac{d^3k_4}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\Big\{\frac{1}{24}a^3\lambda\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4}\Big\}. \end{aligned} \tag{4.21}$$

At this stage, it is not possible to find the exact form of the wave-functional in this interacting theory, so we must be content with constructing $\Psi[\phi]$ perturbatively in powers of the coupling λ. One starts by expressing the wave-functional in the form

$$\Psi[\phi] = Ne^{-a^3\,\Gamma[\phi]}, \tag{4.22}$$

where $\Gamma[\phi]$ is a series expanded in powers of the scalar field, $\phi_{\vec{k}}$,

$$\Gamma[\phi] = \sum_{n=2}^{\infty}\frac{1}{n!}\int\frac{d^3k_1}{(2\pi)^3}\cdots\frac{d^3k_n}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\cdots+\vec{k}_n)\,\Gamma_n(t;\vec{k}_1,\ldots,\vec{k}_n)\,\phi_{\vec{k}_1}\cdots\phi_{\vec{k}_n}, \tag{4.23}$$

and N is the normalisation, fixed by the condition,

$$\int \mathcal{D}\phi_{\vec{k}}\,\Psi[\phi]\Psi^*[\phi] = 1. \tag{4.24}$$

The task of solving the Schrödinger equation now becomes the problem of determining the detailed form of the functions $\Gamma_n(t;\vec{k}_1,\ldots,\vec{k}_n)$. The fact that the background is invariant under spatial translations in these coordinates has again allowed us to extract a momentum-conserving δ-function. Furthermore, by a simple relabeling of the momenta over which we are integrating, we can show that these functions are completely symmetric under any permutation of their arguments,

$$\Gamma_n(t;\vec{k}_1,\ldots,\vec{k}_i,\ldots,\vec{k}_j,\ldots,\vec{k}_n) = \Gamma_n(t;\vec{k}_1,\ldots,\vec{k}_j,\ldots,\vec{k}_i,\ldots,\vec{k}_n).$$

When $m^2 \geq 0$, the vacuum state of the theory should have the same $\phi \leftrightarrow -\phi$ symmetry as the potential. This symmetry means that all the odd-order functions vanish,

$$\Gamma_n(t; \vec{k}_1, \ldots, \vec{k}_n) = 0 \qquad \text{for } n \in \text{odd}.$$

To compute the nonvanishing functions, $\Gamma_n(t; \vec{k}_1, \ldots, \vec{k}_n)$ with $n \in$ even, we need to find the appropriate equations of motion. This is done by expanding each side of the Schrödinger equation in powers of $\phi_{\vec{k}}$ and matching the terms that share the same numbers of fields. This process produces a set of coupled differential equations for the functions $\Gamma_n(t; \vec{k}_1, \ldots, \vec{k}_n)$. For example, the left side of the Schrödinger equation is evaluated straightforwardly enough,

$$i\frac{\partial \Psi}{\partial t} = i\left\{\frac{\dot{N}}{N} - a^3 \sum_{n=2}^{\infty} \frac{1}{n!} \int \frac{d^3k_1}{(2\pi)^3} \cdots \frac{d^3k_n}{(2\pi)^3}\, (2\pi)^3 \delta^3(\vec{k}_1 + \cdots + \vec{k}_n) \left[\frac{\partial \Gamma_n}{\partial t} + 3\frac{\dot{a}}{a}\Gamma_n\right] \phi_{\vec{k}_1} \cdots \phi_{\vec{k}_n}\right\}\Psi,$$

but the right side contains a more complicated tower of terms. These are generated when the functional derivatives in the Hamiltonian act on the wave-functional,

$$\begin{aligned}
H\Psi \;=\; & \left\{\int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\, (2\pi)^3 \delta^3(\vec{k}_1 + \vec{k}_2) \left\{\frac{1}{2}a^3\left(m^2 + \frac{k_1^2}{a^2}\right)\phi_{\vec{k}_1}\phi_{\vec{k}_2}\right\}\right. \\
& + \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\frac{d^3k_3}{(2\pi)^3}\frac{d^3k_4}{(2\pi)^3}\, (2\pi)^3 \delta^3(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 + \vec{k}_4)\left\{\frac{1}{24}a^3\lambda\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4}\right\} \\
& + \sum_{n=0}^{\infty}\frac{1}{n!}\int \frac{d^3k_1}{(2\pi)^3} \cdots \frac{d^3k_n}{(2\pi)^3}\, (2\pi)^3\delta^3(\vec{k}_1 + \cdots + \vec{k}_n) \\
& \qquad \times \left[\frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\, \Gamma_{n+2}(t; \vec{k}_1, \ldots, \vec{k}_n, \vec{k}, -\vec{k})\right]\phi_{\vec{k}_1} \cdots \phi_{\vec{k}_n} \\
& + \sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}\frac{1}{n!}\frac{1}{n'!}\int \frac{d^3k_1}{(2\pi)^3} \cdots \frac{d^3k_n}{(2\pi)^3}\frac{d^3k'_1}{(2\pi)^3} \cdots \frac{d^3k'_{n'}}{(2\pi)^3} \\
& \qquad \times \left[-\frac{a^3}{2}\Gamma_{n+1}(t; \vec{k}_1, \ldots, \vec{k}_n, -\sum \vec{k}_i)\Gamma_{n'+1}(t; \vec{k}'_1, \ldots, \vec{k}'_{n'}, -\sum \vec{k}'_i)\right] \\
& \qquad \left. \times (2\pi)^3\delta^3(\vec{k}_1 + \cdots + \vec{k}_n + \vec{k}'_1 + \cdots + \vec{k}'_{n'})\, \phi_{\vec{k}_1} \cdots \phi_{\vec{k}_n}\phi_{\vec{k}'_1} \cdots \phi_{\vec{k}'_{n'}}\right\}\Psi.
\end{aligned}$$

But by collecting and matching the various functions according to the shared factors of $\phi_{\vec{k}_1} \cdots \phi_{\vec{k}_n}$ that accompany them for a given n, we find a differential equation for each of the Γ_n's. The function $\Gamma_2(t; \vec{k}, -\vec{k})$ accompanying the quadratic part of $\Gamma[\phi]$ obviously depends only on a single momentum. Since this function occurs ubiquitously throughout the following calculations, it is advantageous to change our notation slightly and write it a little more succinctly as

$$\alpha_k(t) \equiv \Gamma_2(t; \vec{k}, -\vec{k}). \tag{4.25}$$

The Schrödinger equation then implies the following relations derived from the zeroth, quadratic, and quartic order terms in the fields,

$$\begin{aligned}
\frac{\dot{N}}{N} &= -\frac{i}{2}(2\pi)^3\delta^3(\vec{0})\int\frac{d^3\vec{p}}{(2\pi)^3}\,\alpha_p(t) \\
\frac{\partial\alpha_k}{\partial t}+3\frac{\dot{a}}{a}\alpha_k &= i\Big\{m^2+\frac{k^2}{a^2}-\alpha_k^2+\frac{1}{2}\frac{1}{a^3}\int\frac{d^3\vec{p}}{(2\pi)^3}\,\Gamma_4(t;\vec{k},-\vec{k},\vec{p},-\vec{p})\Big\}, \\
\frac{\partial\Gamma_4}{\partial t}+3\frac{\dot{a}}{a}\Gamma_4(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) &= i\lambda - i\big[\alpha_{k_1}+\alpha_{k_2}+\alpha_{k_3}+\alpha_{k_4}\big]\Gamma_4(t,\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) \\
&\quad+\frac{i}{2}\frac{1}{a^3}\int\frac{d^3\vec{p}}{(2\pi)^3}\,\Gamma_6(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4,\vec{p},-\vec{p}),
\end{aligned}\tag{4.26}$$

and so on for yet higher orders of n. The infinite factor $(2\pi)^3\delta^3(\vec{0})$ that appears in the equation for the time dependence of the normalisation, and in several of the equations that will occur later, is the volume of a spatial hypersurface in de Sitter space. These volume factors always accompany contributions to the normalisation.

It is important to remember that the form of these equations is determined entirely by the dynamical theory that we are considering, that is, by the Hamiltonian of a quartic theory. Because each of these equations is first-order, there is an additional freedom associated with the choice of the constants of integration[8] appearing in the particular solution for the Γ_n's. The collective choice for all of these constants translates into the choice of a particular state $\Psi[\phi]$ in this picture.

4.4 Perturbation theory and the vacuum state

The usual stochastic treatment of inflation always implicitly assumes that the theory is in the Bunch-Davies state. It is therefore important to introduce appropriate conditions on the functions Γ_n at very short wavelengths in order to put the field in the correct state. After we have done so, we can follow the evolution to large wavelengths and see the simplifications that permit a stochastic description of the theory. We construct the Bunch-Davies solution of the Schrödinger equation here by solving the associated functions Γ_n perturbatively to a given order in λ. Fortunately, all that is needed to derive the part of the quantum Fokker-Planck equation that produces the standard stochastic Fokker-Planck equation is to compute these solutions to linear order in λ. In fact, the zeroth order solution — what would exist in the purely quadratic theory — is already enough to find the stochastic noise. By evaluating the order λ parts of the solution as well, we shall obtain the correct drift term. The advantage of this approach is that it is possible to generalise beyond the standard Fokker-Planck equation by simply working to higher orders in λ.

[8]These are constants in *time*. In general they could depend on momenta for particular choices of the state.

Let us begin by expanding each of the functions in the wave-functional as a power series in λ,

$$\begin{aligned}
\alpha_k(t) &= \sum_{n=0}^{\infty} \alpha_k^{(n)}(t) \\
\Gamma_4(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) &= \sum_{n=1}^{\infty} \Gamma_4^{(n)}(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) \\
\Gamma_6(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4,\vec{k}_5,\vec{k}_6) &= \sum_{n=2}^{\infty} \Gamma_6^{(n)}(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4,\vec{k}_5,\vec{k}_6),
\end{aligned} \tag{4.27}$$

and so on. The order in λ is indicated by the corresponding superscript,

$$\alpha_k^{(n)}, \Gamma_4^{(n)}, \Gamma_6^{(n)}, \ldots \propto \lambda^n. \tag{4.28}$$

The higher order functions Γ_n only begin their power series at correspondingly higher order in λ. Because the trivial, Gaussian version of the theory already exists in the absence of any interactions, the leading term in the expansion of $\alpha_k(t)$ starts at zeroth order. Exactly the same reasoning, tells us that Γ_4 and all of the higher order functions must vanish as $\lambda \to 0$. In the quartic theory that we are analysing, Γ_4 itself starts with a linear term in the coupling λ, as is seen directly from its equation of motion. But the equation for Γ_6, which we have not written explicitly here, is quadratic in Γ_4, so the series expansion for Γ_6 only begins with the λ^2 order term. The problem of solving the Schrödinger equation to linear order in λ then reduces to the problem of solving just three functions: the zeroth and first order pieces of $\alpha_k(t)$, which we rename as $\bar{\alpha}_k(t) \equiv \alpha_k^{(0)}(t)$ and $\beta_k(t) \equiv \alpha_k^{(1)}(t)$ to avoid an excessive use of superscripts, and the leading part of Γ_4,

$$\begin{aligned}
\alpha_k(t) &= \bar{\alpha}_k(t) + \beta_k(t) + \mathcal{O}(\lambda^2) \\
\Gamma_4(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) &= \Gamma_4^{(1)}(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) + \mathcal{O}(\lambda^2).
\end{aligned} \tag{4.29}$$

The starting point is the purely Gaussian, noninteracting, theory, which is summarised by the function $\bar{\alpha}_k(t)$. Even when $\alpha_k(t)$ has been shorn of its order λ and higher parts, the differential equation for $\bar{\alpha}_k(t)$ is still nonlinear; so it is convenient to replace it with another function, $u_k(t)$ defined through

$$\bar{\alpha}_k(t) = -i\frac{\dot{u}_k(t)}{u_k(t)}. \tag{4.30}$$

The nonlinear, first-order equation for $\bar{\alpha}_k(t)$ then becomes a linear, second-order equation for $u_k(t)$,

$$\ddot{u}_k + 3\frac{\dot{a}}{a}\dot{u}_k + \left(m^2 + \frac{k^2}{a^2}\right)u_k = 0. \tag{4.31}$$

The vacuum, or Bunch-Davies, solution to this equation has the standard form, which is expressed more simply in terms of the conformal time coordinate η,

$$u_k(\eta) = -\frac{H\sqrt{\pi}}{2}(-\eta)^{3/2}H_\nu^{(2)}(-k\eta) \qquad \text{where} \quad \nu^2 = \frac{9}{4} - \frac{m^2}{H^2}, \tag{4.32}$$

and where $H_\nu^{(2)}(-k\eta)$ is a Hankel function. We have fixed this solution through two conditions. One is the requirement that the solution should reproduce the vacuum solution in Minkowski space at scales $k \gg aH$, or in the limit $k\eta \to -\infty$. This condition fixes the single constant of integration that gives the form of $\bar{\alpha}_k(t)$ that reproduces the familiar positive energy solution far inside the horizon, $\bar{\alpha}_k(t) \to k/a$. The second condition, which was not necessary for $\bar{\alpha}_k(t)$ but which has been used to normalise the function $u_k(t)$, is that we have required it to satisfy the condition,

$$a^3\Big(u_k^*\dot{u}_k - u_k\dot{u}_k^*\Big) = i. \tag{4.33}$$

In the more frequently used Heisenberg picture for a free scalar field theory, this condition naturally emerges as the consequence of the equal-time commutation relation between the field and its conjugate momentum. But in the Schrödinger picture, the overall normalisation always cancels within the ratio $\bar{\alpha}_k(t) = -i\dot{u}_k(t)/u_k(t)$. Nonetheless, since the function $u_k(t)$ assumes a more recognisable form when we do impose this condition, we have chosen to use it here. Taking the massless limit, the function $u_k(t)$ reduces to

$$u_k(\eta) = -\frac{iH}{\sqrt{2}k^{3/2}}(1 - ik\eta)e^{ik\eta}. \tag{4.34}$$

If we then proceed to take the $k\eta \to -\infty$ limit too, we verify that the product $a(\eta)u_k(\eta)$ assumes the form of a Minkowski space vacuum mode for a massless theory in the Schödinger picture,

$$\lim_{k\eta\to-\infty} a(\eta)u_k(\eta) = \frac{e^{ik\eta}}{\sqrt{2k}}.$$

From the perspective of a free, massless theory in Minkowski space, only the positive energies appear in the exponent; the negative energy solutions, $e^{-ik\eta}$, are absent from this limiting form for $a(\eta)u_k(\eta)$.

Once we have found the leading part of the quadratic function $\alpha_k(t)$, we next compute the leading part of the function accompanying the quartic part of $\Gamma[\phi]$. The series expansion for Γ_6 only begins at quadratic order, so we shall not need to include this function when solving for just the leading part of $\Gamma_4^{(1)}$. Without the Γ_6 term, the linear part of the equation for Γ_4 in λ reduces to a first-order inhomogeneous equation,

$$\frac{\partial\Gamma_4^{(1)}}{\partial t} + \frac{\partial}{\partial t}\Big[\ln\Big(a^3u_{k_1}u_{k_2}u_{k_3}u_{k_4}\Big)\Big]\Gamma_4^{(1)} = i\lambda. \tag{4.35}$$

Its general solution is

$$\Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)=\frac{c_4+i\lambda\displaystyle\int_{\eta_0}^{\eta}d\eta' a^4(\eta')u_{k_1}(\eta')u_{k_2}(\eta')u_{k_3}(\eta')u_{k_4}(\eta')}{a^3(\eta)u_{k_1}(\eta)u_{k_2}(\eta)u_{k_3}(\eta)u_{k_4}(\eta)}, \tag{4.36}$$

where the constant of integration, c_4, should be fixed by the requirement that the wave-functional corresponds to the Bunch-Davies state. To find the correct choice for this constant, we again consider the behaviour of Γ_4 when all of the wavelengths are much smaller that the size of the horizon. In this limit, where $k_i\eta\to-\infty$, we can replace

$$a(\eta)u_{k_i}(\eta)\approx\frac{e^{ik_i\eta}}{\sqrt{2k_i}}$$

in the solution for $\Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)$. Being able unambiguously to perform the integral that appears in the solution depends on being able to establish a suitable $i\epsilon$ prescription for the time-integration contour. As one proceeds ever deeper into the horizon by allowing the initial time to reach further back, we should define this $i\epsilon$ prescription so that the initial contribution to the integral in the solution for $\Gamma_4^{(1)}$ vanishes as $\eta_0\to-\infty$. Once this has been done, if we consider times where the momenta are still well within the horizon at η — that is, $-k_i\eta\gg1$ — the leading behaviour that results when performing the integral is

$$\Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)\approx c_4\frac{a(\eta)\sqrt{16k_1k_2k_3k_4}}{e^{i(k_1+k_2+k_3+k_4)\eta}}+\frac{a(\eta)\,\lambda}{k_1+k_2+k_3+k_4}. \tag{4.37}$$

We now realise that the $i\epsilon$ prescription that has successfully suppressed the unwanted contribution from the positive energy fluctuations as $\eta_0\to-\infty$ would correspondingly lead to an exponential growth of the first term as $k_i\eta\to-\infty$. By choosing $c_4=0$, this problem is resolved since the first term has been removed entirely, along with what would appear to be negative energy oscillations from the perspective of an observer only able to measure wavelengths much smaller than the size of the horizon. Thus, the leading part of the quartic function for the Bunch-Davies state is

$$\Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)=\frac{i\lambda\displaystyle\int_{-\infty}^{\eta}d\eta' a^4(\eta')u_{k_1}(\eta')u_{k_2}(\eta')u_{k_3}(\eta')u_{k_4}(\eta')}{a^3(\eta)u_{k_1}(\eta)u_{k_2}(\eta)u_{k_3}(\eta)u_{k_4}(\eta)}. \tag{4.38}$$

Now that we have found the appropriate solution for our state, we can investigate how it behaves in the opposite limit — it is the set of long wavelength fluctuations that are relevant for the stochastic description of the theory. For a massless field, the integral is once again simple enough to evaluate explicitly,

$$\begin{aligned}
&\Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) \\
&\quad = \frac{i\lambda}{3H}\frac{1 - iK\eta - \dfrac{1}{2}K^2\eta^2 + \dfrac{3}{2}\big(k_1^2+k_2^2+k_3^2+k_4^2\big)\eta^2 + \dfrac{3ik_1k_2k_3k_4\eta^3}{K}}{(1-ik_1\eta)(1-ik_2\eta)(1-ik_3\eta)(1-ik_4\eta)} \\
&\qquad - \frac{\lambda}{3H}\frac{\big(k_1^3+k_2^3+k_3^3+k_4^3\big)\eta^3 e^{-iK\eta}\,\mathrm{Ei}(1,-iK\eta)}{(1-ik_1\eta)(1-ik_2\eta)(1-ik_3\eta)(1-ik_4\eta)}.
\end{aligned} \tag{4.39}$$

Here we have abbreviated $K \equiv k_1+k_2+k_3+k_4$ and $\mathrm{Ei}(1,-iK\eta)$ is the standard exponential integral function. The advantage of analysing the theory in the Schrödinger picture is becoming more apparent — this function, which is the one accompanying the quartic term in $\Gamma[\phi]$, is completely free from any divergent behaviour in the long wavelength limit where $k_i\eta \to 0$. The exponential integral diverges logarithmically when its argument approaches zero,

$$\mathrm{Ei}(1,-iK\eta) = -\gamma + \ln(-iK\eta) - iK\eta + \mathcal{O}(K^2\eta^2),$$

but since this only happens when all four of the momenta simultaneously become small, and since the exponential integral is multiplied by $(k_1^3+k_2^3+k_3^3+k_4^3)\eta^3$, there are no long wavelength divergences in $\Gamma_4^{(1)}$.

Later, we shall see that the asymptotic behaviour of this function fixes the drift in the quantum Fokker-Planck equation and — because they are precisely the same — the drift in the stochastic Fokker-Planck equation as well. This function is genuinely produced by the interactions amongst the fields, so it is directly responsible for the existence of the terms in the Fokker-Planck equation that are associated with the potential. For this reason, we need to evaluate $\Gamma_4^{(1)}$ in the limit where all of its momenta have been stretched far outside the horizon, $k_i < \varepsilon aH$. In this case, the function approaches a purely imaginary constant,

$$\lim_{\substack{k_i\eta\to 0\\ m=0}} \Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) = \frac{i}{3}\frac{\lambda}{H} + \frac{i}{2}\frac{\lambda}{H}\big(k_1^2+k_2^2+k_3^2+k_4^2\big)\eta^2 + \cdots. \tag{4.40}$$

In the opposite limit, where the wavelengths of the fluctuations labelled by $\vec{k}_i$ are all well within the horizon, $k_i \gg aH$, we see that this function is more and more suppressed,

$$\lim_{\substack{k_i\eta\to -\infty\\ m=0}} \Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) = -\frac{\lambda}{H}\frac{1}{(k_1+k_2+k_3+k_4)\eta} + \cdots. \tag{4.41}$$

So in a sense, at early times and for short wavelengths the theory assumes a more and more strongly Gaussian character.

Although the quadratic part of the wave-functional $\alpha_k(t)$ also contains parts that scale as λ, which are the same order in the coupling as the leading behaviour of the quartic function $\Gamma_4(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)$, it turns out that for the purpose of treating the static solutions of the Fokker-Planck equation it would not be consistent to include these order λ terms in α_k without having also included the order λ^2 terms of Γ_4 as well. To understand why this is so, let us consider the static limit of the stochastic theory where $\partial p/\partial t = 0$. The Fokker-Planck equation then reduces to the equation

$$N\frac{\partial^2 p}{\partial\varphi^2} + D\frac{\partial}{\partial\varphi}\left(\frac{\partial V}{\partial\varphi}p(\varphi)\right) = 0, \tag{4.42}$$

whose general solution is

$$p(\varphi) = n\left[e^{-\frac{D}{N}V(\varphi)} + ce^{-\frac{D}{N}V(\varphi)}\int^{\varphi} d\varphi'\, e^{\frac{D}{N}V(\varphi')}\right], \tag{4.43}$$

where n is the normalisation of the probability function. Choosing $c=0$, the static solution for a quartic interaction is

$$p(\varphi) = ne^{-\frac{D}{N}V(\varphi)} = \frac{\Gamma\left(\frac{3}{4}\right)}{\pi}\left(\frac{\lambda D}{6N}\right)^{1/4} e^{-\frac{\lambda D}{24N}\varphi^4}. \tag{4.44}$$

The effect of an order λ term in N would only produce an order λ^2 effect in the ratio $\lambda D/N$. This would be exactly the same order as the next term in the series expansion of Γ_4 contributing to the drift.

It is nonetheless useful at this stage to investigate a little of the asymptotic behaviour of $\beta_k(t)$ — the order λ part of the quadratic function — at late times,

$$\alpha_k(t) = \bar{\alpha}_k(t) + \beta_k(t) + \cdots = -i\frac{\dot{u}_k(t)}{u_k(t)} + \beta_k(t) + \mathcal{O}(\lambda^2).$$

By extracting the order λ terms from the differential equation for $\dot{\alpha}_k(t)$, we obtain a differential equation for $\beta_k(t)$,

$$\frac{\partial}{\partial t}\left(u_k^2 a^3\beta_k\right) = iu_k^2\left[a^3\delta m^2 + (Z_1-1)ak^2 + \frac{1}{2}\int\frac{d^3\vec{p}}{(2\pi)^3}\,\Gamma_4^{(1)}(t;\vec{k},-\vec{k},\vec{p},-\vec{p})\right]. \tag{4.45}$$

Since the integral of the quartic function diverges at short wavelengths, we must introduce counterterms, δm^2 and (Z_1-1), which were not necessary for computing the leading behaviour of either $\alpha_k(t)$ or Γ_4. In a massless theory, we already showed that $\Gamma_4^{(1)}$ approaches an imaginary constant for long wavelengths, so there are no divergences in this integral associated with small values of $p=\|\vec{p}\|$. The general solution of this equation is

$$\begin{aligned}\beta_k(t) \;=\; \frac{1}{a^3(t)}\frac{i}{u_k^2(t)}\Bigg\{c_2 + \int^t dt'\, u_k^2(t')\Bigg[&a^3(t')\delta m^2 + (Z_1-1)a(t')k^2 \\ &+\frac{1}{2}\int\frac{d^3\vec{p}}{(2\pi)^3}\,\Gamma_4^{(1)}(t';\vec{k},-\vec{k},\vec{p},-\vec{p})\Bigg]\Bigg\}.\end{aligned}$$

Once again the $i\epsilon$ prescription for the Bunch-Davies state, which suppresses the contribution from the lower end of the time integral associated with fluctuations that are infinitesimally tiny when compared with the size of the horizon, would also cause the term proportional to c_2 to diverge at early times. The appropriate choice for this state is $c_2 = 0$. The order λ part of the quadratic function for the Bunch-Davies state is then

$$\begin{aligned}\beta_k(\eta) \;=\; & \frac{1}{a^3(\eta)}\frac{i}{u_k^2(\eta)}\int_{-\infty}^{\eta} d\eta'\, u_k^2(\eta')\Big[a^4(\eta')\delta m^2 + (Z_1-1)a^2(\eta')k^2 \\ & + \frac{1}{2}a(\eta')\int \frac{d^3\vec{p}}{(2\pi)^3}\,\Gamma_4^{(1)}(\eta';\vec{k},-\vec{k},\vec{p},-\vec{p})\Big], \end{aligned} \tag{4.46}$$

when expressed as a function of the conformal time coordinate.

We now analyse the behaviour of $\beta_k(t)$ at long wavelengths to see that it is well behaved. Looking at the explicit form for $u_k(\eta)$ for a massless theory, it is clear that as long as the result of performing the momentum integral in this solution does not diverge faster than $1/\eta^3$, $\beta_k(\eta)$ will not itself diverge as $k\eta \to 0$. By rescaling the momentum $\vec{p}$ by $\vec{p}\eta'$, we can express the momentum integral in this solution in terms of a dimensionless function,

$$\int \frac{d^3\vec{p}}{(2\pi)^3}\,\Gamma_4^{(1)}(\eta';\vec{k},-\vec{k},\vec{p},-\vec{p}) = \frac{1}{\eta'^3}\frac{\lambda}{3H}\int \frac{d^3(\vec{p}\eta')}{(2\pi)^3}\,\hat{\Gamma}_4^{(1)}(k\eta',p\eta'), \tag{4.47}$$

where

$$\begin{aligned}\hat{\Gamma}_4^{(1)}(k\eta,p\eta) \;=\; & \frac{i + 2(p+k)\eta + i(p^2-4pk+k^2)\eta^2 - \dfrac{3}{2}\dfrac{k^2p^2\eta^3}{(p+k)}}{(1-ik\eta)^2(1-ip\eta)^2} \\ & - \frac{2\big(p^3+k^3\big)\eta^3 e^{-2i(p+k)\eta}\,\mathrm{Ei}(1,-2i(p+k)\eta)}{(1-ik\eta)^2(1-ip\eta)^2}\end{aligned} \tag{4.48}$$

is the dimensionless quantity. Since $\Gamma_4^{(1)}$ behaves well at long wavelengths, it cannot diverge as $k\eta' \to 0$ which in turn means that its integral cannot diverge faster than the $1/\eta'^3$ factor that we have already extracted. In fact, the leading $1/\eta'^3$ scaling has the same power as a mass term. For a massless theory, this term should be absent, as can be arranged — depending on the regularization scheme being used — through a suitable choice for δm^2. Therefore, in a massless theory $\beta_k(\eta)$ should vanish as $k\eta \to 0$.

4.5 The quantum Fokker-Planck equation

We are ready to use what we have learned to derive a quantum version of the Fokker-Planck equation. To do so, we must solve for the evolution of the diagonal part of the density matrix for the Bunch-Davies state, $P[\phi] = \Psi[\phi]\Psi^*[\phi]$, and from it derive the evolution of the density matrix for the coarsely grained version of the theory, $P_\Omega[\phi_L]$. The latter is the density matrix obtained by integrating out the short wavelength fluctuations,

$$P_\Omega[\phi_L] \equiv \int_S \mathcal{D}\phi_{\vec{p}}\, P[\phi] = \int_{p\geq\varepsilon aH} \mathcal{D}\phi_{\vec{p}}\, P[\phi]. \tag{4.49}$$

The time derivative of $P_\Omega[\phi_L]$ will then produce the quantum version of the Fokker-Planck equation that we are seeking. One subtlety that occurs in this effective theory, and which does not usually happen in most standard effective field theories, is that in taking the time derivative of $P_\Omega[\phi_L]$ we must also include the time dependence that occurs in the boundary, εaH, dividing the long wavelength fluctuations that we must keep from the short wavelength ones that we remove.

The important idea here is to match between the two theories. This step allows us to express the functions inside the density matrix of the effective theory in terms of those of the original theory. We can then use the Schrödinger equation of the full theory to compute the time derivative of $P_\Omega[\phi_L]$ directly. In terms of the wave-functional $\Psi[\phi]$ for the Bunch-Davies state, the diagonal part of its density matrix is

$$P[\phi] = \Psi[\phi]\Psi^*[\phi] = |N|^2 e^{-a^3[\Gamma[\phi]+\Gamma^*[\phi]]}. \tag{4.50}$$

The first step in the matching process is to define an analogous expansion for the density matrix of the effective theory,

$$P_\Omega[\phi_L] = |N_\Omega|^2 e^{-a^3[\Gamma_\Omega[\phi_L]+\Gamma^*_\Omega[\phi_L]]}, \tag{4.51}$$

where

$$\Gamma_\Omega[\phi_L] = \sum_{n=2}^{\infty}\frac{1}{n!}\int_L \frac{d^3k_1}{(2\pi)^3}\cdots\frac{d^3k_n}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\cdots+\vec{k}_n)\,\Gamma_{\Omega,n}(t;\vec{k}_1,\ldots,\vec{k}_n)\,\phi_{\vec{k}_1}\cdots\phi_{\vec{k}_n}, \tag{4.52}$$

together with its own normalisation N_Ω which satisfies the condition

$$\int_L \mathcal{D}\phi_{\vec{k}}\, P[\phi_L] = 1. \tag{4.53}$$

In order to calculate the density matrix of the effective theory, it is helpful to separate the full density matrix into three factors,

$$P[\phi] = |N|^2 e^{-a^3[\Gamma+\Gamma^*]} = |N|^2 e^{-a^3[\Gamma_L+\Gamma^*_L]}e^{-a^3[\Gamma_0+\Gamma^*_0]}e^{-a^3[\delta\Gamma_S+\delta\Gamma^*_S]}, \tag{4.54}$$

which we have expressed through an equivalent separation of $\Gamma[\phi_L,\phi_S]$ into three terms: one that only includes the long wavelength modes,

$$\Gamma_L[\phi_L] = \sum_{n=2}^{\infty}\frac{1}{n!}\int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\cdots\frac{d^3\vec{k}_n}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\cdots+\vec{k}_n)\,\phi_{\vec{k}_1}\cdots\phi_{\vec{k}_n}\Gamma_n(\vec{k}_1,\ldots,\vec{k}_n) \tag{4.55}$$

one that is quadratic in the short wavelength modes and zeroth order in λ,

$$\Gamma_0[\phi_S] = \frac{1}{2}\int_S \frac{d^3\vec{p}}{(2\pi)^3}\,\phi_{\vec{p}}\phi_{-\vec{p}}\bar{\alpha}_p, \tag{4.56}$$

and a final term that collects everything else,

$$\begin{aligned}\delta\Gamma_S[\phi_L,\phi_S] \;=\; & \frac{1}{2}\int_S \frac{d^3\vec{p}}{(2\pi)^3}\,\phi_{\vec{p}}\phi_{-\vec{p}}\big(\alpha_p-\bar{\alpha}_p\big) \\ & +\sum_{n=4}^{\infty}\frac{1}{n!}\int_{\not{L}}\frac{d^3\vec{p}_1}{(2\pi)^3}\cdots\frac{d^3\vec{p}_n}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{p}_1+\cdots+\vec{p}_n) \\ & \qquad\qquad\times\phi_{\vec{p}_1}\cdots\phi_{\vec{p}_n}\Gamma_n(\vec{p}_1,\ldots,\vec{p}_n).\end{aligned} \tag{4.57}$$

Here the notation $\not{L}$ means that at least one of the momentum integrals is over just the short wavelength modes; integrals over long wavelengths can appear in this term too.

Integrating out the short wavelength fluctuations of the fields is done the most straightforwardly by further separating the purely noninteracting part of the density matrix for the short-distance degrees of freedom from the rest, defining in the process

$$P_0[\phi_S] \equiv |N_0|\exp\bigg\{-\frac{1}{2}a^3\int_S \frac{d^3\vec{p}}{(2\pi)^3}\,\phi_{\vec{p}}\phi_{-\vec{p}}\big(\bar{\alpha}_p+\bar{\alpha}_p^*\big)\bigg\}, \qquad \int_S \mathcal{D}\phi_{\vec{p}}\,P_0[\phi_S]=1, \tag{4.58}$$

and working perturbatively in the coupling. Since we are only evaluating the coarse-grained density matrix $P_\Omega[\phi_L]$ to linear order in λ, we can afford to be a little sloppy and move the effect of integrating the $\delta\Gamma_S$ term directly into the exponent,

$$\begin{aligned}P_\Omega[\phi_L] &= |N|^2 e^{-a^3[\Gamma_L+\Gamma_L^*]}\int_S \mathcal{D}\phi_{\vec{p}}\,e^{-a^3[\Gamma_0+\Gamma_0^*]}e^{-a^3[\delta\Gamma_S+\delta\Gamma_S^*]} \\ &= \frac{|N|^2}{|N_0|^2}e^{-a^3[\Gamma_L+\Gamma_L^*]}\int_S \mathcal{D}\phi_{\vec{p}}\,P_0[\phi_S]\Big[1-a^3[\delta\Gamma_S+\delta\Gamma_S^*]+\cdots\Big] \\ &= \frac{|N|^2}{|N_0|^2}e^{-a^3[\Gamma_L+\Gamma_L^*]}\bigg[1-a^3\int_S \mathcal{D}\phi_{\vec{p}}\,P_0[\delta\Gamma_S+\delta\Gamma_S^*]+\cdots\bigg] \\ &= \frac{|N|^2}{|N_0|^2}e^{-a^3[\Gamma_L+\Gamma_L^*]-a^3\int_S \mathcal{D}\phi_{\vec{p}}\,P_0[\delta\Gamma_S+\delta\Gamma_S^*]}+\mathcal{O}(\lambda^2).\end{aligned} \tag{4.59}$$

The exponent is not quite yet meant to be identified with $-a^3[\Gamma_\Omega+\Gamma_\Omega^*]$, as it also contains contributions to the normalisation of the coarse-grained density matrix. These contributions are easily recognised since they do not contain any factors of the field, $\phi_{\vec{k}}$, and they are accompanied by the usual infinite factor $(2\pi)^3\,\delta^3(\vec{0})$ associated with the infinite spatial volume. When we perform[9] the functional integrals that occur in the exponent, we obtain

$$\begin{aligned}
&\Gamma_\Omega[\phi_L] - a^{-3}\delta N \\
&= \Gamma_L[\phi_L] + \int_S \mathcal{D}\phi_{\vec{p}}\, P_0[\phi_S]\, \delta\Gamma_S[\phi_L,\phi_S] + \mathcal{O}(\lambda^2) \\
&= \frac{1}{2}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}}\Bigl[\alpha_k + \frac{1}{2}\frac{1}{a^3}\int_S \frac{d^3\vec{p}}{(2\pi)^3}\, \frac{1}{\bar\alpha_p + \bar\alpha_p^*}\Gamma_4^{(1)}(\vec{k},-\vec{k},\vec{p},-\vec{p})\Bigr] \\
&\quad + \frac{1}{4!}\int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\, (2\pi)^3\delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\, \phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4}\Gamma_4^{(1)}(\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) \\
&\quad + \frac{1}{2}\frac{1}{a^3}(2\pi)^3\delta^3(\vec{0})\int_S \frac{d^3\vec{p}}{(2\pi)^3}\, \frac{1}{\bar\alpha_p + \bar\alpha_p^*}\Bigl[\beta_p + \frac{1}{4}\frac{1}{a^3}\int_S \frac{d^3\vec{p}'}{(2\pi)^3}\, \frac{1}{\bar\alpha_{p'} + \bar\alpha_{p'}^*}\Gamma_4^{(1)}(\vec{p},-\vec{p},\vec{p}',-\vec{p}')\Bigr] \\
&\quad + \mathcal{O}(\lambda^2). \qquad (4.60)
\end{aligned}$$

Matching the terms with two, four, or no factors of the field $\phi_{\vec{k}}$ produces the following functions that describe the evolution of the coarse-grained density matrix,

$$\begin{aligned}
\alpha_{\Omega,k} \equiv \Gamma_{\Omega,2}(t;\vec{k},-\vec{k}) &= \bar\alpha_k + \beta_k + \frac{1}{2}\frac{1}{a^3}\int_S \frac{d^3\vec{p}}{(2\pi)^3}\, \frac{1}{\bar\alpha_p + \bar\alpha_p^*}\Gamma_4^{(1)}(t;\vec{k},-\vec{k},\vec{p},-\vec{p}) + \mathcal{O}(\lambda^2) \\
\Gamma_{\Omega,4}(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) &= \Gamma_4^{(1)}(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) + \mathcal{O}(\lambda^2), \qquad (4.61)
\end{aligned}$$

and

$$\begin{aligned}
\delta N &= -\frac{1}{2}(2\pi)^3\delta^3(\vec{0})\int_S \frac{d^3\vec{p}}{(2\pi)^3}\, \frac{1}{\bar\alpha_p + \bar\alpha_p^*}\Bigl[\beta_p + \frac{1}{4}\frac{1}{a^3}\int_S \frac{d^3\vec{p}'}{(2\pi)^3}\, \frac{1}{\bar\alpha_{p'} + \bar\alpha_{p'}^*}\Gamma_4^{(1)}(\vec{p},-\vec{p},\vec{p}',-\vec{p}')\Bigr] \\
&\quad + \mathcal{O}(\lambda^2), \qquad (4.62)
\end{aligned}$$

where the normalisation of the density matrix of the effective theory is

$$|N_\Omega|^2 = \frac{|N|^2}{|N_0|^2}e^{\delta N + \delta N^*}. \qquad (4.63)$$

The density matrix of the coarsely grained theory now inherits its time dependence directly from the original theory. For example, the time derivative of the

[9] We only require terms up to the quadratic order for which the following integrals are sufficient,

$$\begin{aligned}
\int_S \mathcal{D}\phi_{\vec{p}}\, \phi_{\vec{p}_1}\phi_{\vec{p}_2}P_0[\phi_S] &= \frac{1}{a^3}\frac{1}{\bar\alpha_{p_1} + \bar\alpha_{p_1}^*}(2\pi)^3\delta^3(\vec{p}_1+\vec{p}_2) \\
\int_S \mathcal{D}\phi_{\vec{p}}\, \phi_{\vec{p}_1}\phi_{\vec{p}_2}\phi_{\vec{p}_3}\phi_{\vec{p}_4}P_0[\phi_S] &= \frac{1}{a^6}\frac{(2\pi)^3\delta^3(\vec{p}_1+\vec{p}_2)}{\bar\alpha_{p_1} + \bar\alpha_{p_1}^*}\frac{(2\pi)^3\delta^3(\vec{p}_3+\vec{p}_4)}{\bar\alpha_{p_3} + \bar\alpha_{p_3}^*} + \frac{1}{a^6}\frac{(2\pi)^3\delta^3(\vec{p}_1+\vec{p}_3)}{\bar\alpha_{p_1} + \bar\alpha_{p_1}^*}\frac{(2\pi)^3\delta^3(\vec{p}_2+\vec{p}_4)}{\bar\alpha_{p_2} + \bar\alpha_{p_2}^*} \\
&\quad + \frac{1}{a^6}\frac{(2\pi)^3\delta^3(\vec{p}_1+\vec{p}_4)}{\bar\alpha_{p_1} + \bar\alpha_{p_1}^*}\frac{(2\pi)^3\delta^3(\vec{p}_2+\vec{p}_3)}{\bar\alpha_{p_2} + \bar\alpha_{p_2}^*}.
\end{aligned}$$

leading parts of $\alpha_{\Omega,k}$ follows from how $\alpha_k(t)$ and $\Gamma_4^{(1)}$ evolve,

$$\begin{aligned}
\frac{\partial\alpha_{\Omega,k}}{\partial t} + 3\frac{\dot a}{a}\alpha_{\Omega,k} &= -i\alpha_{\Omega,k}^2 + \frac{i}{2}\frac{1}{a^3}\int_L \frac{d^3\vec k'}{(2\pi)^3}\,\Gamma_4^{(1)}(t;\vec k,-\vec k,\vec k',-\vec k') \\
&\quad + \frac{1}{2}\frac{1}{a^3}\int_{\partial S}\frac{d^3\vec p}{(2\pi)^3}\,\frac{1}{\bar\alpha_p+\bar\alpha_p^*}\Gamma_4^{(1)}(t;\vec k,-\vec k,\vec p,-\vec p) \\
&\quad + i\bigg\{m^2 + \frac{k^2}{a^2} + \frac{1}{2}\frac{\lambda}{a^3}\int_S \frac{d^3\vec p}{(2\pi)^3}\,\frac{1}{\bar\alpha_p+\bar\alpha_p^*}\bigg\} + \cdots . \qquad (4.64)
\end{aligned}$$

The terms on the first line are all expressed in terms of the functions of the effective theory. Those appearing on the third line are purely imaginary and cancel within the combinations of $\alpha_{\Omega,k}(t)+\alpha_{\Omega,k}^*(t)$ that occur in the density matrix. The only unfamiliar term is the one on the second line. It arises because when we truncate the momenta, $k \le \varepsilon a(t)H$, the limit of the truncated integral is also time dependent. Introducing this boundary as a step function, its time derivative only contributes at the boundary. This has been denoted about with the following notation,

$$\begin{aligned}
\int_{\partial L}\frac{d^3\vec k}{(2\pi)^3}\,f(\vec k) &\equiv \int \frac{d^3\vec k}{(2\pi)^3}\,f(\vec k)\frac{\partial}{\partial t}\Theta(\varepsilon aH - k) \\
\int_{\partial S}\frac{d^3\vec k}{(2\pi)^3}\,f(\vec k) &\equiv \int \frac{d^3\vec k}{(2\pi)^3}\,f(\vec k)\frac{\partial}{\partial t}\Theta(k - \varepsilon aH),
\end{aligned}$$

where $f(\vec k)$ is a general function of the momentum.

Combining the appropriate derivative and its complex conjugate, the evolution of the quadratic structure in $P_\Omega[\phi_L]$ is summarised by

$$\begin{aligned}
&\frac{\partial}{\partial t}\Big[a^3\left(\alpha_{\Omega,k}+\alpha_{\Omega,k}^*\right)\Big] \\
&\quad = -ia^3\big(\alpha_{\Omega,k}^2 - \alpha_{\Omega,k}^{*2}\big) + \frac{i}{2}\int_L \frac{d^3\vec k'}{(2\pi)^3}\left[\Gamma_4^{(1)}(t;\vec k,-\vec k,\vec k',-\vec k') - \Gamma_4^{(1)*}(t;-\vec k,\vec k,-\vec k',\vec k')\right] \\
&\qquad + \frac{1}{2}\frac{1}{a^3}\int_{\partial S}\frac{d^3\vec p}{(2\pi)^3}\,\frac{1}{\bar\alpha_p+\bar\alpha_p^*}\left[\Gamma_4^{(1)}(t;\vec k,-\vec k,\vec p,-\vec p) + \Gamma_4^{(1)*}(t;\vec k,-\vec k,\vec p,-\vec p)\right] + \cdots \qquad (4.65)
\end{aligned}$$

to linear order in the coupling. At this order the evolution of the quartic term in the fields is precisely the same as in the original theory. Finally, the evolution of the normalisation of the coarsely grained density matrix follows from

$$\frac{\dot N_\Omega}{N_\Omega} + \frac{\dot N_\Omega^*}{N_\Omega^*} = -\frac{i}{2}(2\pi)^3\delta^3(\vec 0)\int_L \frac{d^3\vec k}{(2\pi)^3}\left(\alpha_{\Omega,k}-\alpha_{\Omega,k}^*\right) + \frac{1}{2}(2\pi)^3\delta^3(\vec 0)\int_{\partial L}\frac{d^3\vec k}{(2\pi)^3}. \quad (4.66)$$

We are now able to evaluate the time derivative of $P_\Omega[\phi_L]$ directly,

$$i\frac{\partial P_\Omega}{\partial t} = \bigg\{ i\frac{\dot N_\Omega}{N_\Omega} + i\frac{\dot N_\Omega^*}{N_\Omega^*} - i\frac{\partial}{\partial t}\Big[a^3\big(\Gamma_\Omega+\Gamma_\Omega^*\big)\Big]\bigg\}, \qquad (4.67)$$

through our knowledge of how each of the functions associated with the original theory itself evolves. When we do so, we obtain the following expression,

$$\begin{aligned}
i\frac{\partial P_\Omega}{\partial t} &= \biggl\{ -\frac{1}{2}a^3 \int_L \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}} \bigl(\alpha_{\Omega,k}^2 - \alpha_{\Omega,k}^{*2}\bigr) - \frac{i}{2}a^3 \int_{\partial L} \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}} \bigl(\alpha_{\Omega,k} + \alpha_{\Omega,k}^*\bigr) \\
&\quad + \frac{1}{4}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}} \int_L \frac{d^3\vec{k}'}{(2\pi)^3} \Bigl[\Gamma_4^{(1)}(\vec{k},-\vec{k},\vec{k}',-\vec{k}') - \Gamma_4^{(1)*}(-\vec{k},\vec{k},-\vec{k}',\vec{k}') \Bigr] \\
&\quad - \frac{i}{4}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}} \int_{\partial S} \frac{d^3\vec{p}}{(2\pi)^3}\, \frac{1}{\bar\alpha_p + \bar\alpha_p^*} \Bigl[\Gamma_4^{(1)}(\vec{k},-\vec{k},\vec{p},-\vec{p}) + \Gamma_4^{(1)*}(-\vec{k},\vec{k},-\vec{p},\vec{p}) \Bigr] \\
&\quad - \frac{1}{4!}a^3 \int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\, (2\pi)^3\, \delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\, \phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4} \\
&\qquad\qquad \times \Bigl[\bigl[\bar\alpha_{k_1}+\bar\alpha_{k_2}+\bar\alpha_{k_3}+\bar\alpha_{k_4}\bigr] \Gamma_4^{(1)}(\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) \\
&\qquad\qquad\quad - \bigl[\bar\alpha_{k_1}^*+\bar\alpha_{k_2}^*+\bar\alpha_{k_3}^*+\bar\alpha_{k_4}^*\bigr] \Gamma_4^{(1)*}(-\vec{k}_1,-\vec{k}_2,-\vec{k}_3,-\vec{k}_4) \Bigr] \\
&\quad - \frac{i}{4!}a^3 \int_{\partial L} \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\, (2\pi)^3\, \delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\, \phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4} \\
&\qquad\qquad \times \Bigl[\Gamma_4^{(1)}(\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) + \Gamma_4^{(1)*}(-\vec{k}_1,-\vec{k}_2,-\vec{k}_3,-\vec{k}_4) \Bigr] \\
&\quad + \frac{1}{2}(2\pi)^3\delta^3(\vec{0}) \int_L \frac{d^3\vec{k}}{(2\pi)^3} \bigl(\alpha_{\Omega,k} - \alpha_{\Omega,k}^*\bigr) + \frac{i}{2}(2\pi)^3\delta^3(\vec{0}) \int_{\partial L} \frac{d^3\vec{k}}{(2\pi)^3} + \mathcal{O}(\lambda^2) \biggr\} P_\Omega.
\end{aligned} \tag{4.68}$$

This is essentially the Liouville equation for the effective theory. It is also the quantum Fokker-Planck equation, as we shall now show. The first step is to translate some of these terms into a second functional derivative of $P_\Omega[\phi_L]$ with the appropriate coefficient. For this purpose, the following formula, written for an arbitrary momentum-dependent coefficient, F_k, is very useful,

$$\begin{aligned}
&\frac{1}{a^3}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\, F_k \frac{\delta^2 P_\Omega[\phi_L]}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} \qquad (4.69) \\
&= \biggl\{ -(2\pi)^3\, \delta^3(\vec{0}) \int_L \frac{d^3\vec{k}}{(2\pi)^3}\, F_k \bigl[\alpha_{\Omega,k}(t) + \alpha_{\Omega,k}^*(t)\bigr] + a^3 \int_L \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}}\, F_k \bigl[\alpha_{\Omega,k}(t) + \alpha_{\Omega,k}^*(t)\bigr]^2 \\
&\quad - \frac{1}{2}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\, \phi_{\vec{k}}\phi_{-\vec{k}} \int_L \frac{d^3\vec{k}'}{(2\pi)^3}\, F_{k'} \Bigl[\Gamma_4^{(1)}(t;\vec{k},-\vec{k},\vec{k}',-\vec{k}') + \Gamma_4^{(1)*}(t;-\vec{k},\vec{k},-\vec{k}',\vec{k}') \Bigr] \\
&\quad + \frac{2}{4!}a^3 \int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\, (2\pi)^3\, \delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\, \phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4} \\
&\qquad\qquad \times \Bigl[F_{k_1}\bigl(\bar\alpha_{k_1}+\bar\alpha_{k_1}^*\bigr) + F_{k_2}\bigl(\bar\alpha_{k_2}+\bar\alpha_{k_2}^*\bigr) + F_{k_3}\bigl(\bar\alpha_{k_3}+\bar\alpha_{k_3}^*\bigr) + F_{k_4}\bigl(\bar\alpha_{k_4}+\bar\alpha_{k_4}^*\bigr) \Bigr] \\
&\qquad\qquad \times \Bigl[\Gamma_4^{(1)}(t;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) + \Gamma_4^{(1)*}(t;-\vec{k}_1,-\vec{k}_2,-\vec{k}_3,-\vec{k}_4) \Bigr] + \mathcal{O}(\lambda^2) \biggr\} P_\Omega[\phi_L].
\end{aligned}$$

When we choose the coefficient function F_k to reproduce the zeroth order structures in the coarse-grained Liouville equation, and then gather together what remains, we find that

$$\begin{aligned}
i\frac{\partial P_\Omega}{\partial t} = & -\frac{i}{2}\int_{\partial L}\frac{d^3\vec{k}}{(2\pi)^3}\,\frac{1}{a^3}\frac{1}{\alpha_{\Omega,k}+\alpha^*_{\Omega,k}}\frac{\delta^2 P_\Omega}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} + \frac{i}{2}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\frac{i}{a^3}\frac{\alpha_{\Omega,k}-\alpha^*_{\Omega,k}}{\alpha_{\Omega,k}+\alpha^*_{\Omega,k}}\frac{\delta^2 P_\Omega}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} \\
& -\frac{1}{2}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\phi_{\vec{k}}\phi_{-\vec{k}}\int_L \frac{d^3\vec{k}'}{(2\pi)^3}\,\frac{1}{\bar\alpha_{k'}+\bar\alpha^*_{k'}}\Big[\bar\alpha_{k'}\Gamma_4^{(1)*}(-\vec{k},\vec{k},-\vec{k}',\vec{k}') \\
& \qquad\qquad -\bar\alpha^*_{k'}\Gamma_4^{(1)}(\vec{k},-\vec{k},\vec{k}',-\vec{k}')\Big]P_\Omega[\phi_L] \\
& +\frac{1}{4!}a^3\int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\,(2\pi)^3\,\delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\,\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4} \\
& \qquad\times\Big[\big[\bar\alpha_{k_1}+\bar\alpha_{k_2}+\bar\alpha_{k_3}+\bar\alpha_{k_4}\big]\Gamma_4^{(1)*}(-\vec{k}_1,-\vec{k}_2,-\vec{k}_3,-\vec{k}_4) \\
& \qquad -\big[\bar\alpha^*_{k_1}+\bar\alpha^*_{k_2}+\bar\alpha^*_{k_3}+\bar\alpha^*_{k_4}\big]\Gamma_4^{(1)}(\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)\Big]P_\Omega[\phi_L] \\
& +\mathcal{O}(\lambda^2).
\end{aligned} \tag{4.70}$$

So far we have not used our knowledge of the explicit behaviour of the functions $\alpha_{\Omega,k}$ and $\Gamma_4^{(1)}$. In this coarsely grained version of the Liouville equation, it is only the long wavelength degrees of freedom that appear — all of the momenta are in the region well outside the horizon, $k \le \varepsilon a H$. We are then free to replace $\Gamma_4^{(1)}$ with its asymptotic value, which for a massless field was shown to be

$$\Gamma_4^{(1)}(\eta;\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4) = \frac{i}{3}\frac{\lambda}{H} + \mathcal{O}(\varepsilon^2),$$

up to corrections suppressed by $k_i^2\eta^2 \le \varepsilon^2$, which leaves

$$\begin{aligned}
\frac{\partial P_\Omega}{\partial t} = & -\frac{1}{2}\int_{\partial L}\frac{d^3\vec{k}}{(2\pi)^3}\,\frac{1}{a^3}\frac{1}{\alpha_{\Omega,k}+\alpha^*_{\Omega,k}}\frac{\delta^2 P_\Omega}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} + \frac{1}{2}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\frac{i}{a^3}\frac{\alpha_{\Omega,k}-\alpha^*_{\Omega,k}}{\alpha_{\Omega,k}+\alpha^*_{\Omega,k}}\frac{\delta^2 P_\Omega}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} \\
& +\frac{1}{2}\int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\phi_{\vec{k}}\phi_{-\vec{k}}\int_L \frac{d^3\vec{k}'}{(2\pi)^3}\left[\frac{\lambda}{3H}\right]P_\Omega[\phi_L] \\
& -\frac{1}{4!}a^3\int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\,(2\pi)^3\,\delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\,\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4} \\
& \times\left[\big(\bar\alpha_{k_1}+\bar\alpha^*_{k_1}\big)\frac{\lambda}{3H}+\big(\bar\alpha_{k_2}+\bar\alpha^*_{k_2}\big)\frac{\lambda}{3H}+\big(\bar\alpha_{k_3}+\bar\alpha^*_{k_3}\big)\frac{\lambda}{3H}+\big(\bar\alpha_{k_4}+\bar\alpha^*_{k_4}\big)\frac{\lambda}{3H}\right]P_\Omega[\phi_L] \\
& +\mathcal{O}(\lambda^2).
\end{aligned} \tag{4.71}$$

The reason that we have written the quartic term in this slightly lengthier form becomes clear when we express it in terms of the coarsely grained potential of the effective theory,

$$\mathcal{V}_\Omega[\phi_L] = \frac{1}{4!}\lambda \int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\,(2\pi)^3\delta^3(\vec{k}_1+\vec{k}_3+\vec{k}_3+\vec{k}_4)\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4}. \tag{4.72}$$

A general Fokker-Planck drift term would have the form

$$\begin{aligned}
&\int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\mathcal{D}_k \frac{\delta}{\delta\phi_{\vec{k}}}\left[\frac{\delta\mathcal{V}_\Omega}{\delta\phi_{-\vec{k}}}P_\Omega\right] \\
&\quad = \frac{1}{2}\lambda \int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\phi_{\vec{k}}\phi_{-\vec{k}} \int_L \frac{d^3\vec{k}'}{(2\pi)^3}\,\mathcal{D}_{k'}P_\Omega \\
&\qquad -\frac{1}{4!}a^3\lambda \int_L \frac{d^3\vec{k}_1}{(2\pi)^3}\frac{d^3\vec{k}_2}{(2\pi)^3}\frac{d^3\vec{k}_3}{(2\pi)^3}\frac{d^3\vec{k}_4}{(2\pi)^3}\,(2\pi)^3\,\delta^3(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\,\phi_{\vec{k}_1}\phi_{\vec{k}_2}\phi_{\vec{k}_3}\phi_{\vec{k}_4} \\
&\qquad\qquad \times\Big[(\bar{\alpha}_{k_1}+\bar{\alpha}^*_{k_1})\mathcal{D}_{k_1} + (\bar{\alpha}_{k_2}+\bar{\alpha}^*_{k_2})\mathcal{D}_{k_2} + (\bar{\alpha}_{k_3}+\bar{\alpha}^*_{k_3})\mathcal{D}_{k_3} + (\bar{\alpha}_{k_4}+\bar{\alpha}^*_{k_4})\mathcal{D}_{k_4}\Big]P_\Omega \\
&\qquad +\cdots .
\end{aligned} \tag{4.73}$$

Matching between this general expression and what appears in the Liouville equation of the effective theory, we conclude that $\mathcal{D}_k = 1/3H$ in our theory — the familiar result.

We have now arrived at the quantum version of the Fokker-Planck equation, evaluated to linear order in the coupling,

$$\frac{\partial P_\Omega}{\partial t} = \int_L \frac{d^3\vec{k}}{(2\pi)^3}\left\{\mathcal{N}_k \frac{\delta^2 P_\Omega}{\delta\phi_{\vec{k}}\delta\phi_{-\vec{k}}} + \frac{1}{3H}\frac{\delta}{\delta\phi_{\vec{k}}}\left[\frac{\delta\mathcal{V}_\Omega}{\delta\phi_{-\vec{k}}}P_\Omega\right]\right\} + \cdots , \tag{4.74}$$

where the quantum — momentum dependent — noise term is

$$\mathcal{N}_k = -\frac{1}{2}\frac{1}{a^3}\frac{1}{\alpha_{\Omega,k}+\alpha^*_{\Omega,k}}\left[\frac{\partial}{\partial t}\Theta(\varepsilon aH-k)\right] + \frac{1}{2}\frac{i}{a^3}\frac{\alpha_{\Omega,k}-\alpha^*_{\Omega,k}}{\alpha_{\Omega,k}+\alpha^*_{\Omega,k}}\,\Theta(\varepsilon aH-k)+\cdots . \tag{4.75}$$

In the limit where the wavelengths have all been stretched to be much larger than the horizon, it is actually only the first of these terms that determines the leading form of the noise. Recall that the quantum noise $\mathcal{N}_k$ and the stochastic noise N are related by

$$N = \int_L \frac{d^3\vec{k}}{(2\pi)^3}\,\mathcal{N}_k$$

Let us evaluate the stochastic noise at leading order in the coupling by replacing $\alpha_{\Omega,k} = \bar{\alpha}_k + \cdots$ and using the explicit form for $\bar{\alpha}_k$ in a massless theory,

$$\bar{\alpha}_k(\eta) = iH\frac{k^2\eta^2}{1-ik\eta}. \tag{4.76}$$

The stochastic noise coefficient is then found to be

$$
\begin{aligned}
N &= \int \frac{d^3\vec{k}}{(2\pi)^3} \left\{ -\frac{1}{2}\frac{1}{a^3}\frac{1}{\bar{\alpha}_k+\bar{\alpha}_k^*}\left[\frac{\partial}{\partial t}\Theta(\varepsilon aH-k)\right] + \frac{1}{2}\frac{i}{a^3}\frac{\bar{\alpha}_k-\bar{\alpha}_k^*}{\bar{\alpha}_k+\bar{\alpha}_k^*}\Theta(\varepsilon aH-k)\right\} \\
&= \varepsilon a\frac{H^4}{8\pi^2}\int_0^\infty dk \left\{\frac{1+k^2\eta^2}{k}\delta(k-\varepsilon aH)\right\} - \frac{H^3\eta^2}{4\pi^2}\int_0^{\varepsilon aH} dk\, k \\
&= \frac{H^3}{8\pi^2}\left(1+\varepsilon^2\right) - \frac{H^3}{8\pi^2}\varepsilon^2 \\
&= \frac{H^3}{8\pi^2}.
\end{aligned} \tag{4.77}
$$

In the long wavelength limit, we recover the precisely standard noise term for the stochastic Fokker-Planck equation,

$$
\frac{\partial p}{\partial t} = \frac{H^3}{8\pi^2}\frac{\partial^2 p}{\partial\varphi^2} + \frac{1}{3H}\frac{\partial}{\partial\varphi}\left(\frac{\partial V}{\partial\varphi}p(\varphi)\right), \tag{4.78}
$$

at leading nontrivial order.

4.6 Applications and further refinements of the stochastic picture

We see that the leading form of the quantum version of the Fokker-Planck equation for the effective theory of the long wavelength fluctuations exactly generates the standard Fokker-Planck equation for the stochastic theory. However, now that we can completely follow the derivation between these two pictures, we can — as in any effective theory —refine the basic picture further by evaluating the higher order 'corrections' that should appear on the stochastic side by deriving their analogues directly on the quantum side. For example, we can see that the standard noise and drift,

$$
N(\lambda) = \frac{H^3}{8\pi^2} + \mathcal{O}(\lambda) \qquad \text{and} \qquad D(\lambda) = \frac{1}{3H} + \mathcal{O}(\lambda),
$$

are in fact only the first contributions in a perturbative expansion. What are the forms of the higher order contributions? Are they also free of late-time divergences? Do other terms appear in the Fokker-Planck equation? These last would be the analogues of the higher order operators that appear in the effective Lagrangians in the more familiar applications of effective field theories.

With a means of directly connecting the quantum and stochastic descriptions of the theory, we can — at least in principle — explore the behaviour in the late-time limit more fully. In the static limit of the stochastic theory, the probability function assumes a simple form at leading order in the coupling, e.g. $p(\varphi) \propto e^{-\frac{\lambda D}{24N}\varphi^4}$ for the quartic theory. However, as we mentioned in the introduction, the usual interaction-picture treatments, while consistent with the expectations of the stochastic picture, have late-time divergences that make the approach to this simple, constant, limit difficult to see. In the Schrödinger picture, we have an alternative framework for investigating the behaviour of the quantum theory in this limit. In particular it would be interesting to learn the explicit time-dependence as the probability function approaches its static limit [43].

The technique that we have developed here can also be applied to study the leading behaviour of the stochastic theories associated with other light or massless fields: multiple interacting scalar fields, gauge fields, or the actual scalar and tensor fluctuations of inflationary theories. It should be equally instructive to investigate the probability distribution function, $p(t, \varphi(\vec{x}))$, that is associated with a classical stochastic field. Such fields are used to describe the long wavelength parts of the n-point functions of quantum fields that are evaluated at *different* spatial positions; these are needed to treat the power spectrum and the non-Gaussianities predicted by inflation.

Bibliography

[1] V. Mukhanov, *Physical Foundations of Cosmology*, Cambridge, UK: Cambridge University Press (2005).

[2] S. Weinberg, *Cosmology*, Oxford, UK: Oxford University Press (2008).

[3] H. Collins, "Primordial non-Gaussianities from inflation," arXiv:1101.1308 [astro-ph.CO].

[4] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," JHEP **0305** (2003) 013 [astro-ph/0210603].

[5] P. Creminelli and M. Zaldarriaga, "Single field consistency relation for the 3-point function," JCAP **0410** (2004) 006 [astro-ph/0407059].

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

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

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

[9] 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]].

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

[11] L. Berezhiani, J. Khoury and J. Wang, "Non-Trivial Checks of Novel Consistency Relations," arXiv:1401.7991 [hep-th].

[12] W. D. Goldberger, L. Hui and A. Nicolis, "One-particle-irreducible consistency relations for cosmological perturbations," Phys. Rev. D **87** (2013) 10, 103520 [arXiv:1303.1193 [hep-th]].

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

[14] A. Kehagias and A. Riotto, “Operator Product Expansion of Inflationary Correlators and Conformal Symmetry of de Sitter,” Nucl. Phys. B **864** (2012) 492 [arXiv:1205.1523 [hep-th]].

[15] 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]].

[16] G. L. Pimentel, “Inflationary Consistency Conditions from a Wavefunctional Perspective,” JHEP **1402** (2014) 124 [arXiv:1309.1793 [hep-th]].

[17] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications*, Cambridge, UK: Cambridge University Press (1996)

[18] N. Agarwal, R. Holman, A. J. Tolley and J. Lin, “Effective field theory and non-Gaussianity from general inflationary states,” JHEP **1305** (2013) 085 [arXiv:1212.1172 [hep-th]].

[19] H. Collins, “Initial state propagators,” JHEP **1311** (2013) 077 [arXiv:1309.2656 [hep-th]].

[20] H. Collins, R. Holman and T. Vardanyan, *in progress*.

[21] S. Weinberg, “Adiabatic modes in cosmology,” Phys. Rev. D **67** (2003) 123504 [astro-ph/0302326].

[22] H. Collins, R. Holman and T. Vardanyan, “Renormalizing an initial state,” arXiv:1408.4801 [hep-th].

[23] L. Berezhiani and J. Khoury, “On the Initial State and Consistency Relations,” arXiv:1406.2689 [hep-th].

[24] J. S. Schwinger, “Brownian motion of a quantum oscillator,” J. Math. Phys. **2** (1961) 407.

[25] P. M. Bakshi and K. T. Mahanthappa, “Expectation value formalism in quantum field theory. 1.,” J. Math. Phys. **4** (1963) 1; P. M. Bakshi and K. T. Mahanthappa, “Expectation value formalism in quantum field theory. 2.,” J. Math. Phys. **4** (1963) 12.

[26] L. V. Keldysh, “Diagram technique for nonequilibrium processes,” Zh. Eksp. Teor. Fiz. **47** (1964) 1515 [Sov. Phys. JETP **20** (1965) 1018].

[27] S. Weinberg, “Quantum contributions to cosmological correlations,” Phys. Rev. D **72** (2005) 043514 [hep-th/0506236].

[28] L. H. Ford, “Quantum Instability of de Sitter Space-time,” Phys. Rev. D **31** (1985) 710.

[29] L. H. Ford and A. Vilenkin, “Global Symmetry Breaking in Two-dimensional Flat Space-time and in De Sitter Space-time,” Phys. Rev. D **33** (1986) 2833.

[30] I. Antoniadis, J. Iliopoulos, and T. N. Tomaras, “Quantum Instability of De Sitter Space,” Phys. Rev. Lett. **56** (1986) 1319.

[31] C. P. Burgess, L. Leblond, R. Holman and S. Shandera, “Super-Hubble de Sitter Fluctuations and the Dynamical RG,” JCAP **1003** (2010) 033 [arXiv:0912.1608 [hep-th]].

[32] D. Seery, “Infrared effects in inflationary correlation functions,” Class. Quant. Grav. **27** (2010) 124005 [arXiv:1005.1649 [astro-ph.CO]].

[33] T. Prokopec, N. C. Tsamis, and R. P. Woodard, “Stochastic Inflationary Scalar Electrodynamics,” Annals Phys. **323** (2008) 1324 [arXiv:0707.0847 [gr-qc]].

[34] T. Prokopec, N. C. Tsamis, and R. P. Woodard, “Two Loop Scalar Bilinears for Inflationary SQED,” Class. Quant. Grav. **24** (2007) 201 [gr-qc/0607094].

[35] T. Prokopec, N. C. Tsamis, and R. P. Woodard, “Two loop stress-energy tensor for inflationary scalar electrodynamics,” Phys. Rev. D **78** (2008) 043523 [arXiv:0802.3673 [gr-qc]].

[36] E. O. Kahya, V. K. Onemli, and R. P. Woodard, “The ζ-ζ Correlator is Time Dependent,” Phys. Lett. B **694** (2011) 101 [arXiv:1006.3999 [astro-ph.CO]].

[37] N. C. Tsamis and R. P. Woodard, “Stochastic quantum gravitational inflation,” Nucl. Phys. B **724** (2005) 295 [gr-qc/0505115].

[38] A. A. Starobinsky, “Stochastic de Sitter (inflationary) Stage In The Early Universe,” Lect. Notes Phys. **246** (1986) 107.

[39] A. A. Starobinsky and J. Yokoyama, “Equilibrium state of a selfinteracting scalar field in the De Sitter background,” Phys. Rev. D **50** (1994) 6357 [astro-ph/9407016].

[40] C. P. Burgess, R. Holman, G. Tasinato, and M. Williams, “EFT Beyond the Horizon: Stochastic Inflation and How Primordial Quantum Fluctuations Go Classical,” JHEP **1503** (2015) 090 [arXiv:1408.5002 [hep-th]].

[41] C. P. Burgess, R. Holman, and G. Tasinato, "Open EFTs, IR effects & late-time resummations: systematic corrections in stochastic inflation," JHEP **1601** (2016) 153 [arXiv:1512.00169 [gr-qc]].

[42] M. Mijic, "Stochastic dynamics of coarse grained quantum fields in the inflationary universe," Phys. Rev. D **49** (1994) 6434 [gr-qc/9401030].

[43] H. Collins, R. Holman, and T. Vardanyan, *in progress*.