Author: Aalst, Ted Adrianus Franciscus van der

Title: Conformal invariance and microscopic sensitivity in cosmic inflation

Issue Date: 2012-12-19

2

Cosmic evolution of our universe

The understanding of the evolution of our universe is one of the unabashed successes of modern physics. To a large extent, the evolution of our universe is understood extremely well, from a few moments after its birth to very close to the present day. In this chapter we will review standard big bang cosmology, both theoretically as well as observationally. We will explain how inflation can solve some of the problems associated to the big bang paradigm and, since it is the measurements of the infant stage of the universe that provide the observational backbone for string cosmology, in what way measurements of the infant universe can be related to microscopic models of inflation.

2.1 A short history of big bang cosmology

2.1.1 Theoretical development of a dynamic universe

Modern cosmology is less than 100 years old. Before Einstein had developed his theory of general relativity [3, 4], Newtonian mechanics did not invite to study the universe as a whole. Surely, mankind probably always had an interest for the stars and galaxies that appear on the night sky, but in Newtonian theory this merely results in the study of these objects *within* a fixed arena, not of the black sky itself. The universe itself only features as the stage in which extraterrestrial physical phenomena occur. With the advent of general relativity this all changes, as spacetime itself inevitably becomes dynamic.

To describe the universe as a whole, we rely on the *cosmological principle*, the assumption that no place in the universe is special and that it is the same from any vantage point. This is an extrapolated version of the Copernican principle, that our planet nor our solar system nor our galaxy is the center of the universe. Rather, the laws of physics are the same throughout the universe and no observer can distinguish a preferred location. Consequently, the universe should be homogeneous and isotropic on large scales, a fundamental assumption which enabled Friedmann-Lemaître-Robertson-Walker [5–8] to propose a model for cosmic evolution within general relativity. Homogeneity and isotropy of the cosmological principle translate into the mathematical statement that the metric of spacetime is maximally symmetric in its spatial part,

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2 \left(\frac{\mathrm{d}r^2}{1 - kr^2} + r^2 \mathrm{d}\Omega^2 \right),$$

where the scale factor $a(t)$ describes the overall (spatial) scale of the universe and k corresponds to positively curved, negatively curved or flat spatial slices for $k = 1, -1, 0$ respectively.[1] The matter content must be taken homogeneous and isotropic too, specified by a perfect fluid energy-momentum tensor that only depends on the energy density ρ and pressure p of the fluid,

$$T^{\mu}{}_{\nu} = \begin{pmatrix} -\rho & & & \\ & p & & \\ & & p & \\ & & & p \end{pmatrix}.$$

With these expressions for the metric and energy-momentum tensor, Einstein's equations, including the cosmological constant as a matter contribution, reduce to the Friedmann equations

$$H^2 = \frac{\kappa^2}{3}\rho - \frac{k}{a^2}, \tag{2.1a}$$

$$\frac{\ddot{a}}{a} = -\frac{\kappa^2}{6}(\rho + 3p), \tag{2.1b}$$

where the Hubble parameter, $H = \frac{\dot{a}}{a}$, determines the rate of expansion. The reduced Planck mass $\mathrm{M}_{\mathrm{pl}}^{-2} = \kappa^2 = 8\pi G_N$ acts as the gravitational coupling constant; we will often use natural units, where $\kappa^2 = 1$.

One can combine these equations into the continuity equation

$$\dot{\rho} = -3H(\rho + p),$$

[1] The universe is spatially flat to a very high precision. For this reason, we will be mainly concerned with the metric in case of a flat spatial slicing.

which also follows from energy conservation, $\nabla_\mu T^\mu{}_\nu = 0$. If one specifies the equation of state between energy density and pressure by writing $p = w\rho$, the continuity equation can be integrated to express the evolution of the energy density as a function of the evolution of the scale factor a,

$$\rho \propto a^{-3(1+w)}.$$

In general the evolution will be a mixture of different non-interacting fluids, which are traditionally distinguished as being (pressureless) non-relativistic matter ($w = 0$), relativistic matter and radiation ($w = \frac{1}{3}$) or a contribution from the cosmological constant ($w = -1$). To denote the matter content of the universe, one often uses the dimensionless quantity $\Omega(t) = \frac{1}{3H^2}\rho$. In terms of Ω, the first Friedmann equation, (2.1a), can be written as $\Omega - 1 = k/(aH)^2$.

2.1.2 Observational confirmation and new challenges

Equation (2.1a) clearly implies that a static universe, $\dot{a} = 0$, only occurs for very specific values of the energy density and spatial curvature. Hence, the cosmological principle in combination with general relativity, seems to tell us that we live in a dynamic universe. Historically, at first a non-static universe was merely a theoretically predicted possibility within general relativity, based on the assumption that the universe is homogeneous and isotropic. By now both the expansion of the universe as well as its homogeneity and isotropy are well established by observations.

Already at the end of the 1920s, very soon after the theory of general relativity and the proposed FLRW solution, first evidence of an expanding universe was obtained by Hubble [9]. He famously discovered that the spectrum of stars is redshifted proportionally to their distance to us. Subsequent experiments have refined his findings to a rate of expansion given by $H_0 = 70.2 \pm 1.4$ $(km/s)/Mpc$ for the present era [10]. A second confirmation of the FLRW model was developed during the following years, when physicists realized that an expanding universe must have had a very hot and very dense early beginning, emerging from an initial singularity called the *big bang*. Such a beginning implies that the universe was so hot that nuclei could not have existed and must have formed as the universe cooled. This epoch is called big bang nucleosynthesis, which ended when the universe cooled down further. The estimated relative production of light elements from protons and neutrons during the epoch of big bang nucleosynthesis accounts for the observed abundances to a very great precision [11] and is therefore also a clear confirmation of the expanding universe model.

The most important confirmation, particularly useful for present-day observations, is the existence of the *cosmic microwave background* (CMB) radiation. In the epoch after big bang nucleosynthesis, electrons were still energetic enough to escape from the pull of ionized nuclei. Only when the universe expanded further and consequently cooled down further, did neutral hydrogen become stable. At this moment of recombination, photons no longer encountered (charged) free electrons and nuclei from which they would scatter. Ever since, they have therefore been traveling (mostly) freely through a neutral universe. As the universe kept expanding, these photons cooled down, redshifting towards microwave radiation, at which frequencies we observe them now. The first observation of the microwave photons was in 1965, when Penzias and Wilson observed an excess microwave background noise in their radio antenna [12], which was quickly realized [13] to be the predicted cosmic microwave background radiation [14, 15].

With following improved observations, the CMB is now the most-precisely measured black body spectrum in nature [16], having a temperature of 2.73 K isotropically across the sky, implying recombination happened approximately 380 000 years after the big bang. The fact that for each local patch across the sky, the variation in the temperature of the CMB is only a remarkable 1 in 10^5 means we have now observationally justified the earlier made assumption of the homogeneity and isotropy of the universe. Each photon on the *surface of last scattering* has the same temperature to an astonishing precision, confirming that recombination and the subsequent expansion happened homogeneously throughout the universe.

Combining different cosmological observations, such as CMB observations [10], the formation of large-scale structures [17, 18], the recessions of type Ia supernovae [19, 20], observed mass distributions through gravitational lensing [21] and the study of peculiar motion of galaxies and clusters [22, 23], we now have an increasingly precise understanding of the content and dynamics of our universe and its evolution after the first fraction of a second. Observationally a huge improvement has been obtained in the last 10–20 years, mainly because of improved measurements of the CMB, which made it possible to estimate the parameters of the FLRW model with ever greater accuracy and firmly established cosmology as a "precision science". We now know we live in a spatially flat universe $\Omega_0 = 1.002 \pm 0.011$, which expands at an accelerated rate [10, 19, 20, 24]. However, this success-story has brought with it a number of new puzzles directly emergent from the data.

One is the discovery of the current accelerated expansion of our universe. It earned its discoverers [19, 20] the 2011 Nobel Prize, which is a recognition of the enormous advances that observational cosmology has seen in the last two decades. However, theoretically the reason for this accelerated expansion is far from clear. At the moment the expansion is accounted for by a dominant energy contribution coming from *dark energy*, such as the cosmological constant Λ or some other energy component which has an equation of state $w < -\frac{1}{3}$. Except for its name, dark energy remains a largely unknown form of “stuff”. We do not know its origin nor its precise characteristics. From the observations it is clear that the universe is currently dominated by dark energy. It constitutes 74% of today’s total energy budget. The remaining 26% of the energy decomposition consists of matter, $w = 0$, although only about 4% (of the total budget) is ordinary visible matter. This means another 22% of the total energy budget is yet unaccounted for. All observations [10, 17, 18, 21–23] indicate the presence of *some* sort of (invisible) matter, dubbed (cold) *dark matter*, a second puzzle.

The ΛCDM-model derives its name from the dominant contributions in our universe, the cosmological constant Λ and cold dark matter. Although we have good indications that these contributions are really there, for the moment their precise nature eludes understanding. For this reason a large branch of present day cosmology focusses on the nature and characteristics of the dominant contributions to the energy decomposition of our universe. However, in this thesis we will focus on yet another mystery of the current cosmological model. This mystery does not focus on the content of our present day universe, but rather on how it has all come to be.

2.2 Cosmic inflation

2.2.1 Initial conditions

Our universe seems to be very special in the way it is very sensitive to its precise initial conditions. In principle the need for such precise initial conditions is not a problem, since cosmology is not claiming to provide a full explanation for the cosmic evolution *including* its starting point. We only need to be able to evolve the universe from a set of given initial conditions to the present day. However, the level of precision for the initial conditions is so high, that one starts wondering why we happen to live in *this* universe. If the initial conditions were only slightly different, standard big bang evolution would lead to a significantly different universe. For this reason it is unsatisfactory to simply take the required initial conditions as a given, without the slightest wondering why it had to be these initial conditions. The strong dependence on the actual initial conditions weakens any claim done by cosmologists, as one can seemingly evolve to *any* universe by simply starting from marginally different initial conditions. For most cosmologists such a sensitive and unstable situation begs for an explanation. Such an explanation exists, it is provided by the inflationary paradigm [25–27].[2]

The problem with the initial conditions consists of two separate problems, called the *horizon problem* and the *flatness problem*. In short, the horizon problem is the observation that the CMB is far more homogeneous than one should naïvely expect. In general, any inhomogeneity will grow bigger and bigger, through gravitational interaction. Indeed, the amount of inhomogeneities today is larger than that in the CMB, but similarly we also expect that the amount of inhomogeneities was even smaller at any time before the CMB. Since the CMB has inhomogeneities of the order of 10^{-5}, one wonders how smooth exactly the initial conditions must have been to provide the smoothness of the CMB. Most importantly, in the standard big bang cosmology, the CMB is homogeneous even across regions which could have never been in causal contact at the time of last scattering. Decoupling occurred 380 000 years after the big bang and so the present-day full-sky observation of the CMB consists of multiple patches, each only 380 000 light years across, in which photons were in causal contact.

Let us explain how this compares with the current causally connected patch. Mathematically we define the *particle horizon* as the (comoving) size of a causally connected region. From (2.1b), it is given by

$$\frac{1}{aH} \propto a^{\frac{1}{2}(1+3w)}. \tag{2.2}$$

The expression takes into account that the universe expands while the light is propagating through space. If the universe would expand too quickly, such that the photons can not “keep up”, the particle horizon decreases. However, for an evolution dominated by ordinary matter, $w \geq 0$, the particle horizon grows with time. This means that, for example those CMB photons that enter the particle horizon now, were not causally connected *before*. Specifically, they were not causally connected at the time the CMB was produced. Yet, the CMB spectrum is consistent over all length scales with a uniform black body spectrum having a homogeneous temperature to 1 part in 10^5. How can the CMB be so homogeneous even across all causally disconnected regions?

The other problem with initial conditions, the flatness problem, is the observation that the universe is incredibly close to being spatially flat, $\Omega_0 = 1.002 \pm 0.011$ [10, 24]. From $\Omega - 1 = k/(aH)^2$ and (2.2) it follows that for ordinary matter, $w \geq 0$, any deviation away from flatness, $\Omega = 1$, can only be growing. In fact, by taking a derivative of the first Friedmann equation and using (2.1b), we can derive a differential equation

[2] Several excellent books and lecture notes provide a detailed introduction to inflation, cf. [28–31].

for Ω,

$$a\frac{d\Omega}{da} = (1 + 3w)\Omega(\Omega - 1),$$

which tells us that the critical value $\Omega = 1$ is an unstable fixed point for $w > -\frac{1}{3}$. This means that in order to find $\Omega_0 = 1.002$ today, the initial conditions must have been exponentially closer to $\Omega = 1$. How can the universe be so spatially flat if all matter is desperately trying to push it away from flatness?

The horizon problem and the flatness problem have a common origin; in both cases the inconsistency arises because of the growing nature of the particle horizon $(aH)^{-1}$ for ordinary matter or equivalently,

$$\frac{d}{dt}(aH)^{-1} = -\frac{\ddot{a}}{(aH)^2} = \frac{\rho}{6aH^2}(1 + 3w) > 0,$$

because ordinary cosmology is dominated by matter, $w = 0$, or radiation, $w = \frac{1}{3}$. As the intermediate result shows, this is equivalent with an expanding but decelerating universe. Therefore an obvious solution would be to look for a period of *accelerated* expansion, $\ddot{a} > 0$, dominated by some form of matter with $w < -\frac{1}{3}$. Although the current vacuum energy dominated era, with $w = -1$, meets the requirements and seems to make the problem less urgent, the universe has only recently entered the vacuum energy dominated epoch. Radiation and matter dominated for most of its history. To solve the problems with initial conditions, we should consider an accelerating phase *before* the current big bang paradigm, which should at least last for about 60 e-folds to solve the flatness and horizon problems [32, 33]. This phase is called *cosmic inflation*. It is specified by the need to explain the initial conditions, but only in a very coarse manner. Any epoch which is dominated by some matter-component having $w < -\frac{1}{3}$ will be capable of solving the big bang problems. However, the requirement $w < -\frac{1}{3}$ is difficult to meet with ordinary matter and radiation, because it requires a negative pressure. This makes the search for the true microscopic nature of the inflationary epoch a worthwhile and interesting endeavor.

2.2.2 Accelerated expansion

No ordinary matter has negative pressure, but it was the insight of [25] that "order parameter" physics can easily account for this, by considering a (single) scalar field (the order parameter) coupled to gravity,

$$S = \frac{1}{2}\int d^4x\,\sqrt{g}\,\left[R - g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - 2V(\phi)\right]. \tag{2.3}$$

The scalar field is a matter source to the gravitational field, appearing on the right hand side of Einstein's equation, with

$$\rho = \frac{1}{2}\dot{\phi}^2 - V, \qquad p = \frac{1}{2}\dot{\phi}^2 + V,$$

under the assumption that $\phi(t, \boldsymbol{x}) = \phi(t)$ is spatially homogeneous. In a regime in which the potential dominates over the kinetic energy of the field, the equation of state can indeed be negative. The limiting case $w = -1$ is reached by assuming a stationary scalar field. In that situation, the field equations for the field ϕ and for the scale factor a of a (flat) FLRW ansatz for the metric,

$$0 = \ddot{\phi} + 3H\dot{\phi} + V'(\phi), \tag{2.4a}$$

$$H^2 = \frac{1}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right), \tag{2.4b}$$

tell us that $V(\phi)$ should be constant and equal to $3H^2$ and the scale factor is exponentially growing, $a(t) = e^{Ht}$. The resulting accelerated expansion is that of a de Sitter universe,

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + e^{2Ht}\mathrm{d}\boldsymbol{x}^2,$$

corresponding to a maximally symmetric universe with positive cosmological constant $\Lambda > 0$, just as in the current epoch.

An inflationary epoch being driven by a constant (positive) potential $V = \Lambda > 0$ is too simplistic, in that there is no dynamical way for inflation to end. This can be resolved by allowing the field $\phi(t)$ to be dynamical [25–27]. To still maintain a handle on the equations of motion, it is useful to consider a dynamical situation which is still very close to de Sitter, i.e. to study a nearly constant Hubble parameter or equivalently a slowly varying field $\phi(t)$ that is potential energy dominated, $\dot{\phi}^2 \ll V$. In order to quantify the slowness of the variation, we define the slow-roll parameters

$$\epsilon = -\frac{\dot{H}}{H^2} = 2\left(\frac{H'}{H}\right)^2 = \frac{1}{2}\left(\frac{\dot{\phi}}{H}\right)^2, \qquad \eta = 2\frac{H''}{H} = -\frac{\ddot{\phi}}{H\dot{\phi}}, \tag{2.5}$$

where $'$ indicates a derivative of $H(\phi)$ with respect to the field ϕ. The equalities in these expressions are a consequence of (2.4), which imply that $2H' = -\dot{\phi}$. Via $\ddot{a} = aH^2(1 - \epsilon)$, it is clear that the universe undergoes accelerated expansion if and only if ϵ is smaller than unity,

$$\ddot{a} > 0 \Leftrightarrow \epsilon < 1. \tag{2.6}$$

$\epsilon = \eta = 0$ corresponds to the pure de Sitter phase. When both slow-roll parameters are taken to be small but non-vanishing, the field equations (2.4) resemble a quasi-de Sitter phase,

$$H^2(t) \approx \frac{1}{3}V(\phi) \approx \text{constant}, \qquad a \approx e^{H(t)t}, \qquad \dot{\phi} \approx \frac{-V'}{3H(t)} \approx 0.$$

The approximation $\epsilon, |\eta| \ll 1$ is called the *slow-roll approximation*. This name arises in the context of a single scalar field driving the acceleration. The definition of ϵ and η makes clear, however, that the existence of acceleration or quasi-de Sitter evolution is not tied to the existence of a scalar field. Nevertheless, almost all models use a scalar field description and often use a different set of *potential slow-roll parameters* ϵ_V, η_V. These are related to the previous ones *in the slow-roll approximation* by

$$\epsilon_V = \frac{1}{2}\left(\frac{V'}{V}\right)^2 \approx \epsilon, \qquad \eta_V = \frac{V''}{V^2} \approx \epsilon + \eta. \tag{2.7}$$

The potential slow-roll parameters express inflation as a slowly rolling field on a flat potential and have as an advantage over ϵ and η that they provide a direct connection between the potential and the dynamics of the system. However, this connection only holds when the slow-roll approximation is assumed, whereas the field equations can be expressed in terms of ϵ and η exactly. The latter set of parameters is therefore better suited to set up a consistent approximation scheme [34] and are, in this respect, preferred over ϵ_V and η_V. In the slow-roll approximation, $\epsilon, |\eta| \ll 1$, the use of the potential slow-roll parameters may be more convenient. Since there are clear indications that the slow-roll approximation is indeed satisfied during inflation, both sets of slow-roll parameters can be used almost interchangeably.

In summary, inflation is a coarse phenomenon that happens if and only if $\epsilon < 1$. *Realistic* inflation has as additional requirement that $\epsilon, |\eta| \ll 1$ or equivalently $\epsilon_V, |\eta_V| \ll 1$ [10, 35] and describes a quasi-de Sitter evolution.

2.3 Seeds of structure

2.3.1 Primordial perturbations

In the description of inflation above, we have assumed the inflaton field $\phi(t, \boldsymbol{x})$ to be spatially homogeneous $\phi(t, \boldsymbol{x}) = \phi(t)$. This assumption is justified by the observed homogeneity in the universe, but we know it cannot be the end of the story, since we have also observed small anisotropies in the CMB [35, 36]. The inflationary paradigm provides a satisfying explanation for the origin of these anisotropies [37–43]. Like we expect around any classical field, the inflaton is subject to small quantum fluctuations

$$\phi(t, \boldsymbol{x}) = \overline{\phi}(t) + \delta\phi(t, \boldsymbol{x}),$$

which in this case parameterize small deviations from spatial homogeneity in the inflaton field. Qualitatively the consequences are easily understood. Through Einstein's equations, small variations in the inflaton field ϕ generate perturbations in the geometry of spacetime, which lead to gravitional wells and voids in which slight relative overdensities and underdensities of the matter distribution start to form. As a result, CMB photons experience slightly different redshifts and this we observe in our measurements of the CMB.

Quantitatively we also understand the transition from one type of perturbation to the other, providing a powerful bridge between observation and theory. Several excellent books and review papers have been written about this rich topic [28, 29, 44–46]. Here, we only present the very basics in order to provide a flavor of why the theoretical calculations in this thesis are relevant for observations. From observations, we have direct access to the relative temperature anisotropies $\delta T(\hat{\boldsymbol{n}})$ in each direction $\hat{\boldsymbol{n}}$ in the sky. Traditionally the information is encoded in terms of multipole moments a_{lm}, that result from expanding $\delta T(\hat{\boldsymbol{n}})$ on the orthonormal set of spherical harmonic functions $Y_{lm}(\hat{\boldsymbol{n}})$,

$$\delta T(\hat{\boldsymbol{n}}) = \sum_{lm} a_{lm} Y_{lm}(\hat{\boldsymbol{n}}).$$

From these coefficients we can then build an angular n-point function

$$\langle a_{l_1 m_1} \dots a_{l_n m_n} \rangle.$$

In principle, the average is an ensemble average over multiple universes, but since we have only access to one universe, the statistical uncertainty is instead controlled by a (weighted) angular average over the m_j-modes [28, 45]. The multipole modes a_{lm} of the temperature (differences) $\delta T(\hat{\boldsymbol{n}})$ in the CMB are sourced by the primordial scalar curvature perturbations ζ. They are related via a *transfer function* $\Delta_l(k)$,

$$a_{lm} = 4\pi(-i)^l \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \Delta_l(k) \zeta_{\boldsymbol{k}} Y_{lm}(\hat{\boldsymbol{k}}). \tag{2.8}$$

The transfer function is the solution to a set of coupled differential equations, resulting from Einstein's equations and Boltzmann's equations for the interactions among different types of fluids [28]. It can be computed numerically [47], once the background cosmology and the initial spectrum for ζ are specified. Conversely, by scanning over

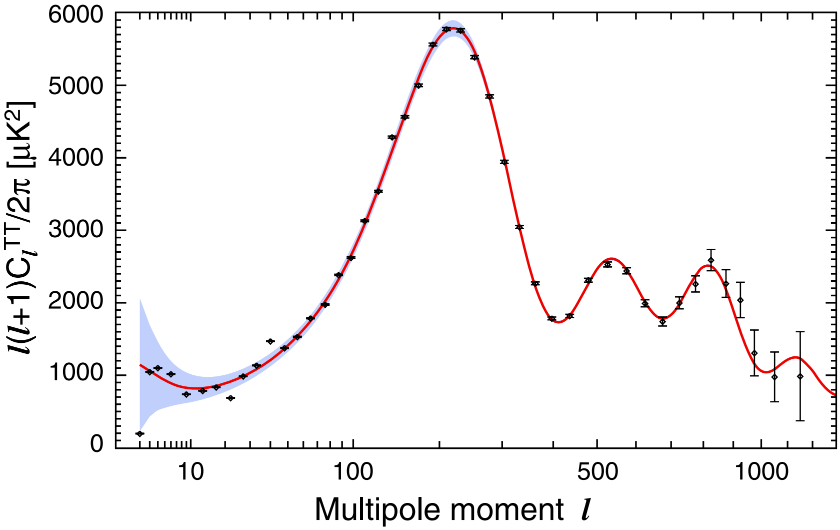

Figure 2.1: *The power spectrum of the temperature anisotropies, expressed in terms of the multipole coefficients* $C_l = \frac{1}{2l+1}\sum_m \langle a^*_{lm} a_{lm}\rangle$. *The curve represents a* Λ*CDM best fit to the 7-year WMAP data with a nearly scale invariant power spectrum* $P_\zeta \sim k^{n_s-1}$ *with* $n_s \approx 0.96$ *[35].*

many results, we can fit the parameters of the background cosmology as well as the primordial spectrum of perturbations to the CMB data.

The upshot of the preceding paragraph is clear: the n-point function of primordial perturbations $\langle \zeta_{\boldsymbol{k}_1} \ldots \zeta_{\boldsymbol{k}_n}\rangle$ is directly related to the correlations of observationally accessible temperature fluctuations, $\langle \delta T(\hat{\boldsymbol{n}}_1) \ldots \delta T(\hat{\boldsymbol{n}}_n)\rangle$. Via (2.8), the temperature two-point function, given in terms of the multipole coefficients $C_l = \frac{1}{2l+1}\sum_m \langle a^*_{lm} a_{lm}\rangle$, is related to the primordial two-point function

$$C_l = 4\pi \int \frac{\mathrm{d}k}{k} P_\zeta(k) \Delta_l(k)^2. \tag{2.9}$$

Here, the relation is given in terms of the power spectrum $P_\zeta(k)$ of the primordial curvature perturbations. For any quantum operator $\hat{f}$, the power spectrum $P_f(k)$ is defined via

$$\langle \hat{f}_{\boldsymbol{k}} \hat{f}_{\boldsymbol{k}'}\rangle = \delta(\boldsymbol{k} + \boldsymbol{k}') \frac{16\pi^5}{k^3} P_f(k).$$

In recent years, the two-point function of temperature anisotropies C_l has been observed to very high precision by the WMAP collaboration [35], cf. figure 2.1. In order to produce the temperature spectrum as shown in figure 2.1, the primordial power spectrum $P_\zeta(k)$ should be almost constant over all scales. On phenomenological grounds, such a scale invariant primordial power spectrum was already proposed by Harrison and Zel'dovich [48, 49]. As we will calculate shortly, one of the great successes of inflation, in addition to solving the flatness and horizon problem, is that it provides a very natural explanation for such a nearly scale invariant spectrum. The precision with which theory matches observation in the CMB temperature two-point function and the way inflation provides us with an explanation for its peaks and valleys [50] and for the underlying scale invariance, lends incredible credence to the existence of a primordial inflationary epoch. With new investigations that focus on subleading effects in the power spectrum, such as small oscillations on top of the near scale invariance [51, 52], more and more details about the inflationary epoch will hopefully soon be revealed.

Another way to probe deeper into the nature of inflation is by studying higher order n-point functions. Similar to the two-point function, the three-point function of temperature anisotropies can be expressed in terms of the three-point functions of the primordial curvature perturbations [44], called the primordial *bispectrum*,

$$\langle \zeta_{\boldsymbol{k}_1} \zeta_{\boldsymbol{k}_2} \zeta_{\boldsymbol{k}_3} \rangle = (2\pi)^3 \delta(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3).$$

The bispectrum is the leading contribution to *non-Gaussian* effects in the spectrum of the CMB, which would be a pure Gaussian distribution if only the two-point functions were non-vanishing [53]. From momentum conservation, its dependence on the three momenta defines a triangular shape. Because different inflationary models predict a peak for different triangular shapes, the shape of non-Gaussianities is an interesting tool to distinguish between models [45, 46, 54].

The simplest inflationary scenario, single field slow-roll inflation with canonical kinetic energy in a Bunch-Davies vacuum, predicts only non-Gaussianities that are too small to be observable [55–57], cf. (2.21). Therefore, current literature is focussed on any possible observation of non-Gaussianities, as it may indicate a variety of violations of the assumptions, favoring e.g. multi-field inflation [58, 59], non-canonical kinetic terms [60–62], non-standard initial states [62–66] or a different scenario for inflation altogether [67, 68]. Developments into this direction are very exciting, especially with the preliminary indication that such non-Gaussianities may be present in the upcoming release of data by the Planck mission [10, 69], but are beyond the scope of this work. We will limit ourselves to the three-point function of primordial curvature perturbations in single field slow-roll inflation with canonical kinetic terms. Even though these non-Gaussianities are beyond the observable level in any near future experiment, from the structure of the correlation functions of even this simplest inflationary model, we can learn a lot about the nature of the early universe.

2.3.2 The power spectrum

Let us now explain the origin of the primordial spectrum of density perturbations from inflation. As in other places in physics, we can calculate the small fluctuations to the inflationary evolution using perturbation theory,

$$\phi(t, \boldsymbol{x}) = \overline{\phi}(t) + \delta\phi(t, \boldsymbol{x}), \qquad g_{\mu\nu}(t, \boldsymbol{x}) = \overline{g}_{\mu\nu}(t) + \delta g_{\mu\nu}(t, \boldsymbol{x}).$$

In cosmological perturbation theory, such a split is not well-defined, since a coordinate transformation may redefine the background fields $\overline{\phi}$ and $\overline{g}_{\mu\nu}$. Therefore one has to be careful to consider only true perturbations and to distinguish these from induced perturbations caused by coordinate redefinitions. Before, in the FLRW ansatz with a homogeneous field $\phi(t)$, this dependence on the coordinate choice did not pose a problem, since we had a clear preferred choice in which the metric looks homogeneous and isotropic. Once perturbations are allowed, such a preferred choice no longer exists, leaving only gauge-invariant statements meaningful. From the scalar perturbations $\delta\phi$ and the curvature perturbation Ψ, defined by $R^{(3)} = \frac{4}{a^2}\nabla^2\Psi$ with $R^{(3)}$ the curvature of the spatial slices, we can construct the gauge-invariant object

$$\zeta = \Psi + \frac{H}{\dot{\overline{\phi}}}\delta\phi. \tag{2.10}$$

Not surprisingly, two popular gauge choices exist in which to calculate the scalar curvature perturbations produced during inflation: the spatially flat gauge $\Psi = 0$ and the comoving gauge $\delta\phi = 0$ [28, 37–43].

In the spatially flat gauge, $\Psi = 0$, one can first simply consider the perturbations of a scalar field in a (flat) de Sitter background. The gauge invariant perturbations ζ are directly obtained from the fluctuations in the field, via $\zeta = (H/\dot{\overline{\phi}})\delta\phi$. At the end of the calculation, the generalization to quasi-de Sitter backgrounds is straightforward. To compute the power spectrum of $\delta\phi$, the two field expectation value, we need to solve its equations of motion, quantize the system and compute the expectation value. Let us choose a massless field for simplicity. With some rewriting, $v_{\boldsymbol{k}} = a\delta\phi_{\boldsymbol{k}}$, of the Fourier modes of the fluctuations $\delta\phi(t, \boldsymbol{x}) = (2\pi)^{-3/2}\int \mathrm{d}^3\boldsymbol{k}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\delta\phi_{\boldsymbol{k}}(t)$, the scalar field equation for the fluctuations,

$$\delta\ddot{\phi} - \frac{1}{a^2}\nabla^2\delta\phi + 3H\delta\dot{\phi} = 0,$$

reduces to

$$v_k'' + \left(k^2 - \frac{2}{\tau^2}\right) v_k = 0 \tag{2.11}$$

in Fourier-space, where a prime $'$ denotes differentiation with respect to conformal time $\tau = -1/(aH)$ and where $k = |\boldsymbol{k}|$. A solution to this equation is

$$v_k(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}} \left(1 - \frac{i}{k\tau}\right). \tag{2.12}$$

On subhorizon scales, $k \gg aH$ or equivalently $k|\tau| \gg 1$, the modes oscillate, while on superhorizon scales, $k \ll aH$ or $k|\tau| \ll 1$, the fluctuations $\delta\phi_k = v_k/a$ are frozen out at a constant value $|\delta\phi_k| = H/\sqrt{2k^3}$. The conservation on superhorizon scales is very convenient. It enables one to calculate the fluctuations at horizon exit, knowing that they will not change until the modes re-enter the horizon. After horizon re-entry, the transfer function $\Delta_l(k)$ relates the primordial fluctuations with the temperature anisotropies.

The classical dynamics can be quantized by promoting the solution (2.12) to a quantum operator

$$\hat{v}_{\boldsymbol{k}} = v_k(\tau)\hat{a}_{\boldsymbol{k}} + v^*_{-k}(\tau)\hat{a}^\dagger_{-\boldsymbol{k}}, \tag{2.13}$$

where $\hat{a}_{\boldsymbol{k}}$ and $\hat{a}^\dagger_{-\boldsymbol{k}}$ are the usual creation and annihilation operators of the set of harmonic oscillators described by (2.11), with commutation relation $[\hat{a}_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k}'}] = (2\pi)^3\delta(\boldsymbol{k} - \boldsymbol{k}')$. The commutation relation imposes a normalization of the modes v_k. Together with the choice for a Bunch-Davies vacuum, $\hat{a}_{\boldsymbol{k}}|0\rangle$, —defined by the requirement that it is equal to the Minkowski vacuum in the far past [70]— this imposes sufficient boundary conditions to uniquely determine (2.12) as the solution of the second order differential equation (2.11). Using (2.12) we can compute the power spectrum of the $\delta\phi_k$ perturbations in the superhorizon limit, $P_{\delta\phi}(k) = \left(\frac{H}{2\pi}\right)^2$, which is equal to the value at horizon crossing, $k \approx aH$. As a result, we can easily generalize the de Sitter calculation to the slow-roll situation in which the Hubble parameter varies slightly or when the field is massive. In that case, different modes exit the horizon at slightly different times $k = a(t)H(t)$. Using the relation between $\delta\phi$ and ζ, the power spectrum of the gauge invariant curvature perturbations generated during slow-roll inflation is, in units $M_{pl} = 1$,

$$P_\zeta(k) = \frac{1}{8\pi^2} \frac{H^2}{\epsilon}\bigg|_{k=aH}, \tag{2.14}$$

which has to be evaluated at horizon crossing. In the slow-roll regime, this is completely controlled by the effective value of H at horizon crossing. The power spectrum is scale invariant in the de Sitter limit. Departure from scale invariance is defined in terms of the spectral index $n_s - 1 = \mathrm{d}\log P/\mathrm{d}\log k$. Using the relations $\mathrm{d}\log H/\mathrm{d}t = -H\epsilon$ and $\mathrm{d}\log\epsilon/\mathrm{d}t = 2H(\epsilon - \eta)$ and the relation between $k = a(t)H(t)$ and t at horizon exit, $\mathrm{d}\log k/\mathrm{d}t = H - H\epsilon$, the spectral index n_s is given by

$$n_s - 1 = 2\eta - 4\epsilon \tag{2.15}$$

to first order in slow-roll.

The same result can be obtained by a calculation in the comoving gauge, $\delta\phi = 0$, as is explained in [57]. It is convenient to work in the ADM formalism, in which the metric is parameterized via a lapse function N and shift vector N^i [71],

$$\mathrm{d}s^2 = -N^2\mathrm{d}t^2 + h_{ij}(\mathrm{d}x^i + N^i\mathrm{d}t)(\mathrm{d}x^j + N^j\mathrm{d}t).$$

The slow-roll action (2.3) is then given by

$$S = \frac{1}{2}\int \mathrm{d}^4x\,\sqrt{h}\,[NR^{(3)} - 2NV + N^{-1}(E_{ij}E^{ij} - E^2 + (\dot\phi - N^i\partial_i\phi)^2) - Nh^{ij}\partial_i\phi\partial_j\phi], \tag{2.16}$$

where $E_{ij} = \frac{1}{2}(\dot h_{ij} - \nabla_i N_j - \nabla_j N_i)$ and $E = E^i{}_i$. Spatial indices can be raised and lowered by h_{ij} and ∇_i is the covariant derivative of this spatial metric. In the comoving gauge, the scalar perturbations to the metric are given by writing

$$h_{ij} = a^2 e^{2\zeta}\delta_{ij} \approx a^2(1 + 2\zeta)\delta_{ij} \tag{2.17}$$

to first order in ζ. The field fluctuations $\delta\phi$ are zero, which means that all spatial derivatives on $\phi(t, \boldsymbol{x})$ vanish. The power of the ADM formalism is that the equations of motion for the Lagrange multipliers N and N^i are simply constraint equations, the hamiltonian and momentum constraints. Solving these constraints perturbatively in terms of ζ,

$$N = 1 + \frac{\dot\zeta}{H} + \ldots, \qquad N_i = \partial_i\left(-\frac{\zeta}{H} + \epsilon\frac{a^2}{H}\partial^{-2}\dot\zeta\right) + \ldots, \tag{2.18}$$

and substituting the result back into the action, then gives the action solely in terms of ζ. In order to find the quadratic action for ζ, it is sufficient to solve N and N^i only to first order in ζ, as the quadratic piece of N and N^i multiplies the zeroth order constraint equation which vanishes for a background solution satisfying the equations of motion [57]. Performing this procedure up to quadratic order gives

$$S^{(2)} = \int \mathrm{d}t\mathrm{d}^3\boldsymbol{x}\, a^3\epsilon\left[\dot\zeta^2 - a^{-2}(\partial_i\zeta)^2\right] = \frac{1}{2}\int \mathrm{d}\tau\mathrm{d}^3\boldsymbol{x}\left[w'^2 + \frac{z''}{z}w^2 - (\partial_i w)^2\right],$$

where $w = z\zeta$ and $z = a\sqrt{2\epsilon}$, which has as equation of motion in Fourier-space,

$$w_k'' + \left(k^2 - \frac{z''}{z}\right) w_k = 0.$$

To lowest order in slow-roll $\frac{z''}{z} \approx \frac{a''}{a} \approx \frac{2}{\tau^2}$ and we find exactly the same differential equation as (2.11). Hence, from (2.12) we read off that the power spectrum of w_k in the superhorizon limit is $P_w(k) = a^2H^2/4\pi^2$ and again we find $P_\zeta(k) = \frac{1}{z^2}P_w(k)\Big|_{k=aH} = \frac{1}{8\pi^2}\frac{H^2}{\epsilon}\Big|_{k=aH}$. Although the calculation is technically more involved in the comoving gauge, the advantage is that one directly uses the variable of interest ζ. It is this gauge invariant object that is conserved on superhorizon scales [40, 57].

2.3.3 Non-Gaussianities

The procedure to find the bispectrum B_ζ of primordial curvature perturbations in slow-roll inflation was laid down in [57]. In the comoving gauge, it is a direct generalization of the calculation of the two-point function, expanding (2.16) up to third order in ζ. Again it suffices to solve the hamiltonian and momentum constraints up to first order in ζ, cf. (2.18). The third order terms again multiply the constraint equations at zeroth order, while the second order terms multiply the constraint equations to first order, which vanish by the first order solution (2.18) [57, 62]. Substituting (2.18) into (2.16) and keeping cubic contributions, gives

$$S^{(3)} = \int \mathrm{d}t\mathrm{d}^3\boldsymbol{x} \left(a^3\epsilon^2 \left[\dot{\zeta}^2\zeta + a^{-2}(\partial_i\zeta)^2\zeta - 2\dot{\zeta}\partial_i\zeta\partial_i\partial^{-2}\dot{\zeta}\right] + f(\zeta)\left.\frac{\delta L}{\delta\zeta}\right|_{(1)} + \ldots\right). \tag{2.19}$$

The ellipsis contain terms that are of higher order in the slow-roll approximation. They are omitted to keep the calculation simple, with the justification that only —if any— the leading order contributions are likely to be observable [61]. The term prior to the ellipsis is proportional to the first order equations of motion and can therefore be removed by field redefinitions.

Once the third order action is known, the three-point function can be calculated. As was emphasized in [57, 61], the three-point function is an expectation value, defined with respect to the vacuum $|\mathrm{in}\rangle$ of the interacting theory at a given time,

$$\langle \mathrm{in}|\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}|\mathrm{in}\rangle.$$

As in ordinary quantum field theory [72], the vacuum of the interacting theory can be obtained from an evolution of the free vacuum $|0\rangle$ using the interaction hamiltonian

$S^{(3)} = -\int d\tau H_{\rm int}(\zeta^{(2)})$. The interaction hamiltonian depends on the quantum operator $\zeta^{(2)}$ corresponding to the solution of the free theory (2.12). In a cosmological context, this procedure is summarized in the "in-in"-formalism [57, 61, 73–75], which results into

$$\langle \zeta_{\boldsymbol{k}_1}(\tau)\zeta_{\boldsymbol{k}_2}(\tau)\zeta_{\boldsymbol{k}_3}(\tau)\rangle = -i\int d\tau' \langle 0| \left[\zeta^{(2)}_{\boldsymbol{k}_1}(\tau)\zeta^{(2)}_{\boldsymbol{k}_2}(\tau)\zeta^{(2)}_{\boldsymbol{k}_3}(\tau), H_{\rm int}(\tau')\right] |0\rangle. \tag{2.20}$$

With the interaction hamiltonian defined by $S^{(3)}$ and the free field solution $\zeta^{(2)}$ given by (2.12), the above prescription yields the bispectrum of primordial curvature perturbations produced during slow-roll inflation. To leading order in the slow-roll expansion, it is given by [57]

$$\begin{aligned} B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) &= (2\pi)^4 \left(P_\zeta\right)^2 \frac{1}{k_1^3 k_2^3 k_3^3} \left(A_\epsilon + A_\eta\right) + \ldots, \\ A_\epsilon &= \epsilon\left(\frac{1}{8}\sum_{j=1}^{3} k_j^3 + \frac{1}{8}\sum_{j\neq l} k_j k_l^2 + \frac{1}{k_t}\sum_{j<l} k_j^2 k_l^2\right), \\ A_\eta &= \eta\left(-\frac{1}{4}\sum_{j=1}^{3} k_j^3\right), \end{aligned} \tag{2.21}$$

where $k_t = k_1 + k_2 + k_3$ and where P_ζ is the power spectrum evaluated when the modes cross the horizon, under the assumption that this happens almost simultaneously for all modes.

Equation (2.21) depends only on the two leading order slow-roll parameters ϵ, η. It can be generalized to other inflationary scenarios with more parameters. For example, multi-field slow-roll inflation has a set of multi-dimensional slow-roll parameters [59] and the bispectrum of the most general single field scenario depends on a set of five parameters [62]. Equation (2.21) can also be calculated directly from the field equations, as was done in [76]. In that case, one directly solves the second order equation for $\delta\phi$, rather than using the "in-in"-formalism to calculate the three-point function in the interacting theory from the solutions of the free theory. The result can be written as

$$\begin{aligned} B_\zeta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) &= (2\pi)^4 \left(P_\zeta\right)^2 \frac{1}{k_1^3 k_2^3 k_3^3} \left(A_\epsilon + A_\eta + A_\xi\right) + \ldots, \\ A_\xi &= \xi_V^2 \frac{1}{4}\left((-1 + \gamma + \log[-k_t\tau_*])\sum_{j=1}^{3} k_j^3 + k_1 k_2 k_3 - \sum_{j\neq l} k_j k_l^2\right), \end{aligned} \tag{2.22}$$

where γ is Euler's constant and τ_* is the (conformal) time of horizon crossing. Compared to (2.21), this result includes a contribution proportional to the higher order (potential) slow-roll parameter

$$\xi_V^2 = \frac{V'V'''}{V^2}. \tag{2.23}$$

It is the contribution coming from an interaction term $V'''\delta\phi^3$ in the action, as calculated by [77, 78].

The calculation of [77, 78] is actually a much simpler calculation, because the field under consideration acts as a (massless) spectator field in an expanding de Sitter background, i.e. the field is not responsible for driving the accelerated expansion. The specific form A_ξ of (2.22) corresponds to this bispectrum of a massless scalar spectator field, as the gauge invariant density perturbation can in many aspects be thought of as a massless scalar field (i.e. its solution to the equation of motion, cf. (2.20) and (2.11)). However, the perturbations of the inflaton field are also coupled to gravity and the gauge invariant curvature perturbations obtain contributions both from the fluctuations of the inflaton field as well as from metric perturbations. As argued in [57] the $V'''\delta\phi^3$-contribution to the bispectrum resides within the . . . of (2.19), indicating higher order slow-roll contributions, and is neglected in that calculation. The result (2.22) confirms this expectation and explicitly shows that the contribution from a direct interaction between the scalar fields is second order in slow-roll.

Strictly speaking, by including the ξ_V^2-contribution, one should also include the other contributions that are second order in slow-roll, i.e. those proportional to ϵ^2, η^2 and $\epsilon\eta$. Since these effects will be beyond the observable threshold, the effort of correctly combining all higher order slow-roll contributions is not a relevant exercise at this time, although first results into this direction are known [79]. For our purposes, the appearance of the ξ_V^2-proportional term and in particular of the momentum structure given by $\log[-k_t\tau_*]$ in (2.22) is interesting from a more fundamental point of view, as will be further discussed in chapter 6.