Non-Gaussianity of inflationary field perturbations from the field equation

David Seery¹, Karim A Malik² and David H Lyth³

¹ Centre for Theoretical Cosmology, Department of Applied Mathematics and Theoretical Physics, Wilberforce Road, Cambridge CB3 0WA, UK
² Astronomy Unit, School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
³ Cosmology and Astroparticle Physics Group, Department of Physics, University of Lancaster, Lancaster LA1 4YB, UK
E-mail: djs61@cam.ac.uk, k.malik@qmul.ac.uk and d.lyth@lancaster.ac.uk

Received 6 February 2008
Accepted 27 February 2008
Published 17 March 2008

Abstract. We calculate the tree-level bispectrum of the inflaton field perturbation directly from the field equations, and construct the corresponding $f_{NL}$ parameter. Our results agree with previous ones derived from the Lagrangian. We argue that quantum theory should only be used to calculate the correlators when they first become classical a few Hubble times after horizon exit, the classical evolution taking over thereafter.

Keywords: cosmological perturbation theory, inflation, quantum field theory on curved space, physics of the early universe

ArXiv ePrint: 0802.0588
1. Introduction

Recent advances in observational astronomy have allowed maps of the cosmic microwave background to be constructed in more detail than ever before [1]. The availability of such maps, with good noise properties and controlled foreground subtraction, offers the exciting prospect of studying the earliest ages in the evolution of our universe in a relatively direct way [2, 3].

In order to connect any theory of early universe physics with the cosmic microwave sky, one must make predictions for properties of the curvature perturbation, $\zeta$ [4], which sets the initial condition for those fluctuations in the matter and radiation densities which we can now observe as temperature anisotropies. To carry out this programme effectively, one needs both a model for the relevant physics and an efficient calculational tool with which to make predictions.

One of the most attractive models which has been proposed for describing the evolution of the early universe is inflation (for a review, see [5]). In an inflationary scenario the universe is supposed to have undergone a phase of accelerated expansion in the very distant past. During this phase each light scalar field acquires a fluctuation, which is close to scale invariance when the rate at which the universe expands is almost constant. These fields each contribute a proportion of the total energy density of the universe, and the relative importance of these contributions is sufficient for determining how $\zeta$ is composed of the separate fluctuations in each light scalar [6]–[10]. The fluctuations themselves can be computed, for example, by using the methods of quantum field theory. Inflation thus provides a framework within which the properties of $\zeta$ can readily be calculated, and for a large class of models these predictions are in extremely good agreement with the observational data [11].

Which properties of $\zeta$ should we compute? Any measurement of the CMB anisotropy can be framed in terms of the $n$-point expectation values of the curvature perturbation, and current data are primarily sensitive to the simplest such expectation value, which is
Non-Gaussianity of inflationary field perturbations from the field equation

given by the two-point function or power spectrum. Although this observable has sufficed for obtaining a good deal of precise information about the very early universe, there are limits to how much we can learn from it. Therefore, in order to discriminate effectively between the different models of inflation, it will be necessary to broaden our observational methods so that higher \( n \)-point functions of \( \zeta \) become experimentally accessible, and at the same time develop theoretical predictions for these functions in all relevant models. On the experimental side this process has been under way since the Cosmic Background Observer (COBE) satellite made the first all-sky map of the CMB in the mid-1990s [12,13]. Such efforts have borne fruit with the availability of more sensitive measurements taken by the NASA Wilkinson Microwave Anisotropy Probe with high angular resolution [14,11], and will be improved still further by ESA’s Planck satellite, due for launch in mid-2008. The first-year WMAP data release placed significant constraints on the level of non-Gaussianity which could be present in the CMB (often quantified in terms of a so-called ‘non-linearity parameter’, \( f_{NL} \) [15]), and this quantitative estimate has subsequently been refined [16,17], leading to a recent high confidence exclusion of \( f_{NL} = 0 \) [18]. In the long term, it may be possible to obtain even better estimates from observations, perhaps based on maps of the 21 cm emission of neutral hydrogen [19,20]. On the theoretical side, predictions for the non-Gaussianity generated using a large collection of relevant models have become available over the last few years [21,22,10], [23]–[30], [8], [31]–[33], following Maldacena’s successful calculation of the bispectrum produced in single-field, slow-roll inflation [34].

These developments mean that non-Gaussianity is now a standard cosmological observable, comparable to the spectral index \( n \) of the scalar power spectrum, and is a powerful discriminant between competing models. It is therefore extremely desirable to have at our disposal a simple calculational scheme—analogueous to the familiar formula which allows the scalar spectral index to be estimated [5]—which can be used to predict the non-Gaussianity generated in any model of our choice.

Most computations of the primordial non-Gaussianity have been carried out using the tools of quantum field theory, and have been explicitly framed in the context of the Lagrangian formalism, which leads to a method of computation closely related to many calculations in particle physics [34,29,35]. For example, when expanded perturbatively, the Lagrangian formalism naturally gives rise to a variant of the Feynman diagrams which are familiar from the calculation of scattering amplitudes. There is no doubt that this is a useful development, which immediately permits methods and intuition developed in the context of quantum field theory to be imported into cosmology. Nevertheless, there is considerable merit in exploring alternative calculational strategies, either to enlarge the class of theories in which the non-Gaussianity can be computed, or to take advantage of technical simplifications in the computation.

In the present paper we return to the more familiar method of employing the field equations, which does not invoke a Lagrangian. This is how the tree-level spectrum \((H/2\pi)^2\) of a scalar field in the de Sitter spacetime was first calculated [36], and also how the tree-level spectrum of the curvature perturbation was first calculated using the Mukhanov–Sasaki field equation [37,38]. More recently it has been shown how the

\[4\] ‘Tree-level’ here means leading order in the perturbation, which does indeed correspond to the tree-level Feynman graph [39,40].
method allows one to calculate all expectation values of field perturbation to all orders in de Sitter space \cite{40}. Here we use the method to calculate the tree-level bispectrum of the inflaton field perturbation, to second order in the field perturbation and including the associated metric perturbation.

The method of field equations permits us to make contact with the cosmological literature, which has traditionally approached perturbation theory from the standpoint of the Einstein equations. A second advantage is that it allows an immediate calculation of the non-Gaussianity generated in any model which can be cast in the form of an effective Klein–Gordon equation, whether or not the model is descended from an effective action principle. We do not invoke the Einstein–Hilbert action, only the Einstein field equation together with the inflaton field equation which is obtained from the flat-space Klein–Gordon equation by the replacement $\eta_{\mu\nu} \to g_{\mu\nu}$ (which is the content of the equivalence principle). We are able to give a simple, essentially ‘mechanical’ process for computing the bispectrum in such a model.

The outline of this paper is as follows. In section 2 we construct an explicit solution for the Heisenberg picture field directly from the equations of motion. In section 3 we use this solution to construct the three-point correlator of the field fluctuations. This is the essential ingredient in a calculation of the bispectrum of $\zeta$, since it serves as the required initial condition for the well-known $\delta N$ formula. In section 4 we write down an explicit formula for the non-linearity parameter $f_{NL}$ before giving a brief summary of the calculation in section 5, which concludes with a discussion.

Throughout this paper, we work in natural units where the Planck mass is set equal to unity, $M_P \equiv (8\pi G)^{-1/2} = 1$. The background metric is taken to be of Robertson–Walker form,

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

(1)

where the scale factor $a(t)$ obeys the usual Friedmann constraint $3H^2 = \rho$, with $\rho$ being the energy density. Throughout, $H \equiv \dot{a}/a$ is the Hubble parameter and an overdot denotes a derivative with respect to $t$. It is sometimes more convenient to work in terms of a conformal time variable, $\eta$, which is defined by $\eta \equiv \int_{t_0}^{t} dt'/a(t')$. When $\eta$ is used, it is useful to define the conformal Hubble rate $\mathcal{H}$, which satisfies $\mathcal{H} \equiv a'/a = aH$.

2. Constructing the Heisenberg picture field

For simplicity we will assume that there is only a single scalar degree of freedom, although our method generalizes easily to the case of many scalar fields. The perturbed metric can be written as

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

(2)

We work in the uniform curvature gauge which corresponds to setting $h_{ij}$ equal to its unperturbed value (given by equation (1)), where we assume explicitly that all tensor modes are absent\footnote{A background of gravitational waves would not contribute to the bispectrum to leading order in the perturbations; its contribution would therefore be strictly subleading. For this reason, the truncation to zero tensor modes should be acceptable for the purpose of computing $f_{NL}$.}. The quantities $N$ and $N^i$ are, respectively, the lapse function and
the shift vector [41]. These are unambiguously determined by the Einstein constraint equations once a gauge has been chosen.

Let \( \phi \) be a canonically normalized field, whose background value sources the metric (1). We assume that \( \phi \) is driven by a potential \( V \), which supports inflation in some region of field space. One then allows \( \phi \) to be perturbed by a small amount \( \delta \phi \), which obeys an effective Klein–Gordon equation. Perturbation theory is an expansion in powers of \( \delta \phi \).

In the cosmological literature a further decomposition is sometimes made (see, for example, [42]), by separating some part of \( \delta \phi \) which can be called \( \delta \phi_1 \). One then writes all quantities in the perturbation theory, including \( \delta \phi \), as an expansion in powers of \( \delta \phi_1 \),

\[
\delta \phi \equiv \delta \phi_1 + \frac{1}{2} \delta \phi_2 + \cdots + \frac{1}{n!} \delta \phi_n + \cdots.
\]

It is necessary to impose some arbitrary auxiliary condition to determine \( \delta \phi_1 \), which is usually done implicitly by demanding that \( \delta \phi_1 \) obey a linear equation of motion. It follows that \( \delta \phi_1 \) exhibits precisely Gaussian statistics. This formulation of perturbation theory is most useful when \( \delta \phi \) becomes increasingly Gaussian at early times, which implies that the \( \delta \phi_n \) for \( n \geq 2 \) vanish at past infinity. These two pictures of perturbation theory have complementary merits, and can be used interchangeably.

We will work in the second picture and choose \( \delta \phi_1 \) to be that part of \( \delta \phi \) which obeys precisely Gaussian statistics, as discussed above, and formulate perturbation theory as an expansion in powers of \( \delta \phi_1 \). As we have described, this implies that \( \delta \phi_1 \) should be chosen to obey a linear equation of motion, so that each Fourier mode evolves according to the equation

\[
\delta \phi''_1 + 2H \delta \phi'_1 + k^2 \delta \phi_1 = 0,
\]

where a prime ‘ denotes a derivative with respect to \( \eta \). Eventually we are aiming to calculate only to leading order in the slow-roll expansion and therefore higher order slow-roll contributions in equation (4) have been omitted. It is useful to introduce the notation \( \theta_k \) for solutions of (4), so that \( \delta \phi_1(k, \eta) \equiv \theta_k(\eta) \).

So far, this theory is entirely classical. Quantization of \( \delta \phi_2 \) is straightforward, and proceeds by the canonical method (see, e.g., [43]). The quantum field corresponding to a conjugate pair of solutions \( \{\theta_k, \bar{\theta}_k\} \) is

\[
\hat{\delta} \phi_1(x, \eta) = \int \frac{d^3k}{(2\pi)^3} e^{ik \cdot x} \left\{ a_k^* \theta_k(\eta) + a_{-k} \bar{\theta}_k(\eta) \right\},
\]

where the normalization of the \( \{\theta_k, \bar{\theta}_k\} \) has been adjusted so that the canonical commutation relation for \( \delta \hat{\phi} \) and its conjugate momentum is obtained when \( \{a_k, a_{k'}^*\} \) obey the usual creation–annihilation algebra,

\[
[a_k, a_{k'}^*] = (2\pi)^3 \delta(k - k').
\]

On the other hand, \( \delta \phi_2 \) obeys a non-linear equation with quadratic source terms. In this paper we are going to use the slow-roll approximation to control these sources. As
usual we take slow-roll to be defined by the following set of conditions:

$$\epsilon \equiv -\frac{\dot{H}}{H^2} \approx \frac{1}{2} \left( \frac{V'}{V} \right)^2 \approx \frac{1}{2} \frac{\dot{\phi}^2}{H^2} \ll 1,$$

$$\eta \equiv \frac{V''}{V} \approx -\frac{\ddot{\phi}}{H \dot{\phi}} + \epsilon \ll 1.$$  

(In particular, one should take care that $\eta$ is not confused with the conformal time.)

The non-linear equation which describes $\delta \phi_2$ can now be written in the form

$$\delta \phi''_2 + 2H \delta \phi'_2 + k^2 \delta \phi_2 = -a^2 V''' \delta \phi'_2 + F_2(\delta \phi_1) + G_2(\delta \phi'_1),$$

where $V'''$ is the third derivative of the potential for $\phi$, and $F_2$ and $G_2$ are source terms, which are respectively quadratic in $\delta \phi_1$ and $\delta \phi'_1$. It has recently been shown that \cite{44,45}

$$F_2 \equiv -\sqrt{2} \epsilon \left\{ \frac{1}{2} \delta \phi_1 \partial \delta \phi_1 - 2 \delta \phi_1 \partial^2 \delta \phi_1 - \partial^{-2} \left[ \partial^2 \delta \phi_1 \partial^2 \delta \phi_1 + \partial \delta \phi_1 \partial \partial^2 \delta \phi_1 \right] \right\},$$

and

$$G_2 \equiv -\sqrt{2} \epsilon \left\{ \frac{1}{2} \delta \phi'_1 \partial \delta \phi'_1 + 2 \partial \delta \phi'_1 \partial \partial^2 \delta \phi'_1 - \partial^{-2} \left[ \delta \phi'_1 \partial^2 \delta \phi'_1 + \partial \delta \phi'_1 \partial \partial^2 \delta \phi'_1 \right] \right\},$$

where we have dropped terms that are suppressed by factors of $\epsilon$ and $\eta$. It is useful to separate the $a^2 V'''$ term from the remainder, because its different dependence on the scale factor endows it with a distinctive time dependence which we shall see below leads to a logarithmic divergence at late times.

Among the small terms—subleading in powers of $\epsilon$ and $\eta$—which have been dropped in equations (9)–(11) are source terms proportional to the cross-product $\delta \phi_1 \delta \phi'_1$. Such terms also lead to logarithmic divergences which have the rough form $\sim \epsilon^{3/2}$. For any mode of wavelength $k$ the quantity $N \equiv \ln |k\eta|$ measures by how many $e$-foldings it is outside the horizon at some time of observation, $\eta$, and a subscript ‘*’ denotes evaluation at horizon crossing. Since we are assuming that slow-roll applies at that time these terms are small when $N \approx 0$, but they grow in magnitude with $N$ and when $N \sim 1/\epsilon_*$ they can no longer be ignored. Indeed, such terms form part of a tower of divergences which have the rough form $\sim \epsilon_*^{m+1/2} N^m$ for $m \geq 1$. It follows that the slow-roll approximation will be close to breaking down if we wish to evaluate expectation values at the end of inflation, where the scales of interest are $N$ $e$-foldings outside the horizon (with values of order $N \sim 60$ being typical), but in many theories $\epsilon_* \sim 1/N$. This would lead to a nonsensical perturbation theory in powers of unity.

This argument shows that we cannot use an expansion in powers of slow-roll parameters to predict expectation values at the end of inflation. One should instead think of these divergences as terms in a Taylor series expansion around the time of horizon exit, where $N = 0$ \cite{46}. In principle they could be resummed by the method of the renormalization group or an equivalent technique, after which they would merely correspond to the classical time evolution \cite{47,49}. However, there is no need to perform such a complicated resummation. Indeed, we believe that it is most accurate to evaluate all expectation values almost immediately after horizon exit, where slow-roll is an excellent approximation and all strictly positive powers of $N$ are entirely negligible. One must then use some other method, which does not rely on an expansion in terms of $e$-foldings since horizon exit, to evolve these expectation values until the desired time of evaluation.
is equivalent to the argument of the renormalization group, but makes use of the known simplification that evolution outside the horizon is simply classical. For the curvature perturbation the $\delta \dot{N}$ formula is very convenient [6, 36, 7, 47, 8]. More generally one can use the separate universe approach [51, 7, 8, 25, 52], or an equivalent gradient expansion [9, 10]. It may also be possible to use some formulations of conventional perturbation theory, provided they avoid the appearance of powers of $N$ [44]. The crucial point is that if the correlation functions are evaluated immediately after horizon exit, it is only necessary to compute the leading term in the Taylor series, which will be the constant term in any model giving rise to a Klein–Gordon equation of the form (9). This term is given by the lowest order slow-roll approximation.

In theories which are more general than Einstein gravity coupled to a scalar field, one might encounter examples where the constant term in the Taylor expansion is absent, or where the linear term has a comparable magnitude. An example of such a model is the case of non-local inflation studied by Barnaby and Cline [53], and we shall see below that it can occur even in single-field, slow-roll inflation if the third derivative of the potential is exceptionally large in comparison with the first and second ones. In such cases one can compute enough terms in the Taylor series to obtain a satisfactory approximation before using the separate universe picture or some equivalent technique to obtain the superhorizon evolution. If any $\delta \phi_1 \delta \phi'_1$ terms participate in this process, then they can be computed using a small modification of the argument outlined below.

In the classical theory, equation (9) can be solved by the use of a Green’s function. The correct choice is the so-called causal or retarded Green’s function, which in momentum space satisfies

$$G_k (\eta, \tau) = i a(\tau)^2 \left\{ \begin{array}{ll} 0 & \eta < \tau, \\ \theta_k(\tau) \bar{\theta}_k(\eta) - \theta_k(\tau) \theta_k(\eta) & \eta > \tau, \end{array} \right.$$ (12)

where $\{\theta_k, \bar{\theta}_k\}$ are a complex conjugate pair of solutions to the non-interacting Klein–Gordon equation (4) for comoving wavenumber $k$, normalized as in equations (5) and (6), and a bar denotes complex conjugation. To use $G_k$ to solve for $\delta \phi_2$, one must translate the source terms $F_2$ and $G_2$ into Fourier modes. The behaviour of each interacting mode can then be computed, and the result reassembled, giving

$$\delta \phi_2 (\eta, \mathbf{x}) = \int \frac{d^3 q}{(2\pi)^3} e^{i q \cdot \mathbf{x}} \left\{ \int_{-\infty}^{\eta} d\tau \int \frac{d^3 k_1 d^3 k_2}{(2\pi)^6} G_q (\eta, \tau) \delta (q - k_1 - k_2) S_2 \right\},$$ (13)

where $S_2$ is defined by

$$S_2 \equiv \left[ -a^2 (\tau) V'' + \mathcal{F}_2 (k_1, k_2) \right] \delta \phi_1 (k_1, \tau) \delta \phi_1 (k_2, \tau) + \mathcal{G}_2 (k_1, k_2) \frac{d \delta \phi_1 (k_1, \tau)}{d \tau} \frac{d \delta \phi_1 (k_2, \tau)}{d \tau},$$ (14)

and the momentum space source functions $\mathcal{F}_2$ and $\mathcal{G}_2$ satisfy

$$\mathcal{F}_2 \equiv - \frac{\dot{\phi}'}{\mathcal{H}} \left( -\frac{1}{2} k_1 \cdot k_2 + 2 k_2^2 + \frac{1}{(k_1 + k_2)^2} \left[ k_1^2 k_2^2 + k_2^4 k_1 \cdot k_2 \right] \right),$$ (15)

$$\mathcal{G}_2 \equiv - \frac{\dot{\phi}'}{\mathcal{H}} \left( \frac{1}{2} + \frac{2}{k_2^2} k_1 \cdot k_2 - \frac{1}{(k_1 + k_2)^2} \left[ k_2^2 + k_1 \cdot k_2 \right] \right).$$ (16)
In the quantum theory the same construction applies, with the understanding that
we are to substitute the quantum field \( \delta \phi_1(q) = q_0 a_q^\dagger \theta_q(\eta) + a_q \theta_q(\eta) \) for the non-interacting solution \( \delta \phi_1(q) \equiv \theta_q \) in the source term \( S_2 \). Thus, once the quantization of \( \delta \phi_1 \) has been
determined, this is sufficient to fix the quantization of the \( \delta \phi_n \) for \( n \geq 2 \).

3. The three-point correlator

The final step is to use our solution for the interacting Heisenberg field to compute the
three-point expectation value, \( \langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \). Since \( \delta \phi_1 \) obeys precisely Gaussian
statistics, it is forbidden to have a non-trivial three-point correlator. The leading
contribution therefore comes from a correlation between the second-order part of one of the
fields with the first-order part of the remaining two, so that the correlation is schematically
of the form \( \langle \delta \phi_1 \delta \phi_2 \delta \phi_3 \rangle \sim \frac{1}{2} \langle \delta \phi_1 \delta \phi_1 \delta \phi_1 \rangle \), where ‘\(*\)’ denotes a convolution. The resulting four-\( \delta \phi_1 \) expectation can be evaluated by use of Wick’s theorem, together with the
Wightman function for \( \delta \phi_1 \),

\[
\langle \delta \phi_1(k, \eta) \delta \phi_1(k', \eta') \rangle = (2\pi)^3 \delta(k + k') \bar{\theta}_k(\eta) \theta_k(\eta').
\]  

Notice that there is no time ordering in equation (17).

3.1. The \( V'''' \) term

Consider first the \( V'''' \) term, leading to a logarithmic divergence as described above. The
three-point expectation value sourced by such a term was considered by Zaldarriaga, who
calculated it in the context of the inhomogeneous decay rate model [47].

This term gives a contribution to the three-point function which can be written as

\[
\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \supseteq -i(2\pi)^3 \delta(k_1 + k_2 + k_3) \int_{-\infty}^{\infty} d\tau a(\tau)^4 V''''
\]

\[
\times \left\{ \left[ \theta_{k_3}(\tau) \bar{\theta}_{k_3}(\eta) - \bar{\theta}_{k_3}(\tau) \theta_{k_3}(\eta) \right] \bar{\theta}_{k_1}(\eta) \theta_{k_2}(\eta) \theta_{k_1}(\eta) \theta_{k_2}(\eta) \theta_{k_3}(\tau) \right. \\
\left. + \left[ \theta_{k_2}(\tau) \bar{\theta}_{k_2}(\eta) - \bar{\theta}_{k_2}(\tau) \theta_{k_2}(\eta) \right] \bar{\theta}_{k_1}(\eta) \theta_{k_1}(\tau) \theta_{k_1}(\eta) \theta_{k_2}(\eta) \theta_{k_3}(\tau) \right. \\
\left. + \left[ \theta_{k_1}(\tau) \bar{\theta}_{k_1}(\eta) - \bar{\theta}_{k_1}(\tau) \theta_{k_1}(\eta) \right] \bar{\theta}_{k_2}(\eta) \theta_{k_2}(\tau) \theta_{k_2}(\eta) \theta_{k_3}(\tau) \right\}.
\]  

(18)

It is easy to verify that four of the six terms in (18) cancel against each other, leaving the
sum of a quantity and its complex conjugate,

\[
\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \supseteq -i(2\pi)^3 \delta(k_1 + k_2 + k_3) \int_{-\infty}^{\infty} d\tau a(\tau)^4 V'''' \prod_i \theta_{k_i}(\tau) \bar{\theta}_{k_i}(\eta) + \text{complex conjugate},
\]  

(19)

where \( i \in \{1, 2, 3\} \). Such cancellation always occurs in the Lagrangian version of the
in–in formalism, leading to expectation values which are manifestly real. In the present
formalism any cancellation is contingent on the possibility of factoring the \( \theta, \bar{\theta} \) terms. As
we shall see below, this does not always occur, leading to calculations which are somewhat
longer than the Lagrangian equivalent and to the loss of manifest reality. At the end of the
calculation, of course, the two methods must produce equivalent results.
The integral in (19) is badly divergent, but almost all the divergent terms are imaginary and cancel out of the final result. This is easiest to see after integration by parts, which isolates the divergent pieces. One finds \(^6\)

\[
\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \supseteq \frac{(2\pi)^3}{4!} \delta(k_1 + k_2 + k_3) \frac{H_2^2 V''}{2 \sum k_i^2},
\]

\[
\times \left( -\frac{4}{9} k_i^2 + \frac{k_i}{\prod_{i<j} k_i k_j} + \frac{1}{3} \left( \frac{1}{3} + \gamma + \ln |k_i \eta_*| \right) \sum_i k_i^2 \right),
\]

where \(i \in \{1, 2, 3\}, k_i = \sum_i k_i, \) and for the purposes of obtaining an analytic solution we have assumed that the \(k_i\) have some approximately equal magnitude, which defines a more or less unique time of horizon exit \(\eta_*.\) A subscript ‘\(*\)‘ denotes evaluation at this time, and \(\gamma \approx 0.577 \ 216\) is Euler’s constant.

3.2. Zero-derivative terms

Now let us focus on those quadratic source terms in equation (9) when, within no time derivatives, described by the function \(F_2.\) Such terms make a contribution to the three-point expectation value of the form

\[
\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \supseteq i(2\pi)^3 \delta(k_1 + k_2 + k_3) \int_{\eta_*}^{\eta_*} d\tau a(\tau)^2
\]

\[
\times \left\{ \mathcal{F}_2(k_1; k_2; k_3) \left[ \theta_{k_1}(\tau) \theta_{k_2}(\eta) - \theta_{k_1}(\eta) \theta_{k_2}(\tau) \right] \theta_{k_3}(\eta) \theta_{k_2}(\tau) \theta_{k_3}(\tau) \right. \]
\[
+ \mathcal{F}_2(k_1; k_3; k_2) \left[ \theta_{k_1}(\tau) \theta_{k_3}(\eta) - \theta_{k_1}(\eta) \theta_{k_3}(\tau) \right] \theta_{k_2}(\eta) \theta_{k_3}(\eta) \theta_{k_2}(\tau) \theta_{k_3}(\tau) \]
\[
+ \mathcal{F}_2(k_2; k_3; k_1) \left[ \theta_{k_1}(\tau) \theta_{k_2}(\eta) - \theta_{k_1}(\eta) \theta_{k_2}(\tau) \right] \theta_{k_2}(\eta) \theta_{k_3}(\eta) \theta_{k_2}(\tau) \theta_{k_3}(\tau) \right\},
\]

where the fields which participate in the expectation value are each evaluated at time \(\eta_*\). The quantity \(\mathcal{F}_2(a; b; c)\) is to be obtained by symmetrizing \(\mathcal{F}_2(a; b; c)\) between \(a\) and \(b\) with weight unity. (This follows from the factor of \(1/2\) which is used in the definition of \(\delta \phi_2;\) if this factor is absent, the relative weighting of \(\mathcal{F}_2\) must be adjusted to match.) However, the result is subject to the constraint \(a + b + c = 0\) and therefore can be written in several equivalent forms. For this reason it is useful to fix the representation unambiguously, following Maldacena [34], by using the relation between \(\{a, b, c\}\) to eliminate all dot products of distinct vectors, leaving a result which depends only on the magnitudes \(\{a, b, c\}\). We find

\[
\mathcal{F}_2(k_1; k_2; k_3) = \frac{1}{H} \left( \frac{3}{2} (k_1^2 + k_2^2) - \frac{(k_1^2 - k_2^2)^2}{4k_3^4} - \frac{k_3^2}{4} \right).
\]

If \(\mathcal{F}_2(a; b; c)\) were symmetric over permutations of \(\{a, b, c\}\), then as above the six terms in equation (21) would partially cancel among themselves. Owing to its construction, however, \(\mathcal{F}_2\) is in general symmetric only under interchange of \(a\) and \(b\) and fails to exhibit any symmetry under exchange of \(c\) with \(\{a, b\}\). Therefore, it is not obvious

\(^6\) Note that this result differs from that given by Zaldarriaga by a sign and the replacement \(\gamma \mapsto \gamma + 1/3.\) We thank Filippo Vernizzi for pointing this out.

Journal of Cosmology and Astroparticle Physics 03 (2008) 014 (stacks.iop.org/JCAP/2008/i=03/a=014)
that equation (21) is real. To go further and demonstrate equality with the Lagrangian formulation, it is necessary to compute each of the six integrals.

Once this is done, we find that the zero-derivative contribution to the expectation value becomes

$$
\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \supset (2\pi)^3 \delta(k_1 + k_2 + k_3)
$$

$$
\times \frac{H^4}{8 \prod_i k_i^4} \left\{ f_1 F_2(k_1, k_2; k_3) + f_2 F_2(k_1, k_3; k_2) + f_3 F_2(k_2, k_3; k_1) \right\},
$$

(23)

where the momentum factors $f_i$ obey

$$
f_1 = \frac{2k_3^3(k_1^2 + 4k_1k_2 + k_2^2 - k_3^2)}{(k_1 + k_2 - k_3)^2 k_1^2},
$$

$$
f_2 = \frac{2k_2^3(k_1^2 - 4k_1k_3 + k_3^2 - k_2^2)}{(k_3 - (k_1 - k_3)^2)^2},
$$

$$
f_3 = \frac{2k_1^3(k_2^2 + 4k_2k_3 + k_3^2 - k_1^2)}{(k_1 + k_2 - k_3)^2 k_1^2},
$$

(24)

and we have again assumed that the $k_i$ have an approximate common value, $k$. The expectation value is to be evaluated at a time just after horizon exit of the mode with wavenumber $k$. As above, $H_*$ denotes the value of the Hubble parameter at this time. Note that the $f_i$ do not vary from theory to theory; they are fixed by the structure of the interaction, and once calculated do not need to be calculated again. The different theories which may govern the interactions of $\delta \phi$ influence the three-point expectation only through the term $F_2(a, b; c)$. Similar remarks apply for the $n$-point correlation functions for all $n$.

One might have expected that the $f_i$ should be related by cyclic permutations among the various $k_i$. In fact there are extra signs which are introduced, because the $f_i$ arise from integrating over combinations of the mode functions $\theta_{k_i}$ and their conjugates $\bar{\theta}_{k_i}$. This explains the apparent lack of symmetry in equation (24).

### 3.3. Two-derivative terms

We can carry out a calculation similar to that in the previous section for the terms in the $\delta \phi_2$ equation which are quadratic in $\delta \phi_1$, given in (11). One arrives at a contribution to the three-point expectation which can be written as

$$
\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \supset i(2\pi)^3 \delta(k_1 + k_2 + k_3) \int_{-\infty}^{\eta_*} d\tau a(\tau)^2
$$

$$
\times \left\{ G_2(k_1, k_2; k_3) \left[ \theta_{k_3}(\tau) \bar{\theta}_{k_3}(\eta) - \bar{\theta}_{k_3}(\tau) \theta_{k_3}(\eta) \right] \theta_{k_1}(\eta) \theta_{k_2}(\tau) \frac{d\theta_{k_1}(\tau)}{d\tau} \frac{d\theta_{k_2}(\tau)}{d\tau} \right\}
$$

$$
+ G_2(k_1, k_3; k_2) \left[ \theta_{k_3}(\tau) \bar{\theta}_{k_3}(\eta) - \bar{\theta}_{k_3}(\tau) \theta_{k_3}(\eta) \right] \bar{\theta}_{k_1}(\eta) \theta_{k_1}(\tau) \frac{d\theta_{k_1}(\tau)}{d\tau} \frac{d\bar{\theta}_{k_3}(\tau)}{d\tau}
$$

$$
+ G_2(k_2, k_3; k_1) \left[ \theta_{k_3}(\tau) \bar{\theta}_{k_3}(\eta) - \bar{\theta}_{k_3}(\tau) \theta_{k_3}(\eta) \right] \theta_{k_2}(\eta) \theta_{k_1}(\tau) \frac{d\bar{\theta}_{k_2}(\tau)}{d\tau} \frac{d\theta_{k_1}(\tau)}{d\tau} \right\},
$$

(25)
where $G_2(a, b; c)$ is obtained from $G_2(a, b)$ by a rule analogous to that relating $F_2$ and $F_2$: one symmetrizes over $a$ and $b$, and then uses the relation $a + b + c = 0$ to eliminate dot products of distinct vectors. This gives

$$G_2(k_1, k_2; k_3) = \frac{\delta'(k_1^2 + k_2^2)(k_1^2 + k_2^2 - k_3^2)}{2k_1^2k_2^2}.$$ (26)

In equation (25) there is no cancellation between terms even in the case where $G_2(a, b, c)$ is totally symmetric. However, we can follow a line of argument similar to that used in the case of the zero-derivative terms and find

$$\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \geq (2\pi)^3 \delta(k_1 + k_2 + k_3)$$
$$\times \frac{H^4}{8 \prod_i k_i^3} \left\{ g_1 G_2(k_1, k_2; k_3) + g_2 G_2(k_1, k_3; k_2) + g_3 G_2(k_2, k_3; k_1) \right\}. \quad (27)$$

The two-derivative momentum factors $g_i$ are defined by

$$g_1 = \frac{4 \prod_i k_i^2}{(k_1 + k_2 - k_3)^2 k_1^2};$$
$$g_2 = \frac{4 \prod_i k_i^2}{(k_2^2 - (k_1 - k_3)^2)^2};$$
$$g_3 = \frac{4 \prod_i k_i^2}{(k_1^2 - (k_2 + k_3)^2)^2}. \quad (28)$$

We can now assemble the zero- and two-derivative terms to obtain an overall contribution to the three-point expectation value which corresponds to

$$\langle \delta \phi(k_1) \delta \phi(k_2) \delta \phi(k_3) \rangle \geq (2\pi)^3 \delta(k_1 + k_2 + k_3)$$
$$\times \frac{H^4}{8 \prod_i k_i^3} \frac{\dot{\phi}_i}{H_*} \left\{ \frac{1}{2} \sum_i k_i^3 - \frac{4}{k_i} \sum_{i < j} k_i^2 k_j^2 - \frac{1}{2} \sum_{i \neq j} k_i k_j \right\}, \quad (29)$$

in complete agreement with previous calculations using the action-based approach [29, 54].

4. The non-linearity parameter $f_{NL}$

The three-point correlator of $\delta \phi$ is not itself the object of primary interest, since it is not directly observable in the cosmic microwave background. Instead, experiments conventionally set limits on a parameter $f_{NL}$ which is related to the three-point correlator of the curvature perturbation, $\zeta$.

We define $f_{NL}$ by

$$\langle \zeta(k_1) \zeta(k_2) \zeta(k_3) \rangle = (2\pi)^3 \delta(k_1 + k_2 + k_3) \frac{6}{5} f_{NL} \sum_{i < j} P_\zeta(k_i) P_\zeta(k_j). \quad (30)$$

With this choice of sign and prefactor, the parameter $f_{NL}$ corresponds to the one conventionally used in the analysis of CMB observations [55], although other sign conventions exist [34].
Non-Gaussianity of inflationary field perturbations from the field equation

For a general theory one can use the \( \delta N \) formula to show that

\[
\zeta(x) \equiv \delta N = \frac{\partial}{\partial \phi^*} \delta \phi^*(x) + \frac{1}{2} \frac{\partial^2}{\partial \phi^*} \delta \phi^*(x)^2 + \cdots,
\]

which holds in coordinate space and is easily generalized to the case of multiple fields. In this equation, the \( \delta \phi \) are evaluated on a flat hypersurface at time \( \eta_* \), and \( N \) is a function of both \( \eta_* \) and the time of observation. To second order in \( \delta \phi \), \( \zeta \) can be calculated in terms of the so-called ‘separate universe approach’, and takes the form

\[
\zeta(k) = \frac{1}{\sqrt{2} \epsilon_*} \delta \phi_*(k) - \frac{1}{2} \left( 1 - \frac{\eta_*}{2 \epsilon_*} \right) \int \frac{d^3q}{(2\pi)^3} \delta \phi_*(k_1 - q) \delta \phi_*(q) + \cdots,
\]

where ‘\( \cdots \)’ denotes terms of higher order in \( \delta \phi_*(k) \) which have been omitted. In this expression, \( \ast \) denotes evaluation at some time \( t \) after the mode with wavenumber \( k \) has left the horizon, after which \( \zeta(k) \) maintains a constant value.

To complete the calculation of \( f_{NL} \) we will use two more slow-roll parameters:

\[
\xi^2 \equiv \frac{V'''}{V'^2},
\]

\[
\sigma^3 \equiv \frac{(V')^2 (d^4V/d\phi^4)}{V'^3},
\]

and the slow-roll relations

\[
- \frac{d}{dN_*} \ln \epsilon_* \simeq 4 \epsilon_* - 2 \eta_* ,
\]

\[
- \frac{d}{dN_*} \eta_* \simeq 2 \epsilon_\ast \eta_* - \xi^2_* ,
\]

and

\[
- \frac{d}{dN_*} \xi^2_* \simeq 4 \epsilon_* \xi^2_* - \eta_* \xi^2_* - \sigma^3_* .
\]

The slow-roll approximation, which we defined through equations (7) and (8), does not in itself place conditions on \( \xi^2 \) and \( \sigma^3 \). Barring rapid oscillation of these quantities as functions of \( \phi \), or a narrow spike, we certainly need them to be much less than unity to preserve \( |\eta| \ll 1 \) over many e-folds. At a generic point it is reasonable to expect [56] \( |\sigma^3| \ll |\xi^2| \ll |\eta| \), but that will obviously fail if \( \eta \) goes through zero.

Using (32) to compute \( \langle \zeta \zeta \zeta \rangle \) and using equation (30) for \( f_{NL} \) gives

\[
\frac{6}{5} f_{NL} = \eta_* - \frac{3}{2} \epsilon_* + \xi^2_* \left( \frac{1}{9} + \gamma + N_* \right)
\]

\[
+ \frac{\epsilon_*}{\sum_i k_i^2} \left\{ 3 \xi^2_* \left( k_i \sum_{i<j} k_i k_j - \frac{4}{9} k_i^3 \right) - \frac{4}{k_i} \sum_{i<j} k_i^2 k_j^2 - \frac{1}{2} \sum_{i<j} k_i k_j^2 \right\},
\]

where we recall that \( N_* \) measures the number of e-folds which elapse between horizon exit of the mode \( k \) and the time of evaluation, \( \eta_* \).

If taken at face value, equation (38) may suggest that we can obtain an \( f_{NL} \) as large as we please by allowing a large number of e-foldings outside the horizon, making \( N_* \) very large. However, that cannot be the case because the separate universe approach shows
that $\zeta$ is conserved after horizon exit during single-field inflation. (More generally, $\zeta$ is conserved outside the horizon if isocurvature perturbations are negligible so that there is a unique relation between pressure and energy density \cite{57,8,58}. From (36) we see that the explicit $N_*$ dependence of $f_{NL}$ is cancelled by a contribution to the $N_*$ dependence of $\eta_*$. The remaining time dependence of $f_{NL}$ involves contributions proportional to $\epsilon_* \eta_*^2$ and $\sigma_*^3$, which would be cancelled if we took the calculation to a higher order in the field perturbation and abandoned the slow-roll approximation.

The present example is trivial because $\zeta$ is time independent. However, a similar caution applies to multiple-field inflation models where $\zeta$ is time dependent. There also, powers of $N_*$ may also appear in the expression for $f_{NL}$. If these terms are to source large non-Gaussianity (which occurs, for example, in the curvaton scenario \cite{59,52,31}), then they should be handled using a non-perturbative approach rather than relying on ‘naive’ expressions such as equation (38).

5. Conclusions

In this paper we have shown that the bispectrum of an interacting but canonically normalized scalar field can be calculated during an inflationary epoch using the details of its field equation directly, without constructing an effective action and using the rules of the in–in formalism.

Where a Lagrangian formulation exists, the action and the field equations to which it gives rise are manifestly equivalent. In this case, as a point of principle, our calculation is a straightforward rewriting of the usual one, although it may prove to be more convenient in examples where one works with the field equation from the outset. On the other hand, for models where no Lagrangian formulation exists, or none is known, our formula enables the non-Gaussianity to be computed easily and compared with experiment.

The recipe is straightforward. For reference, we summarize the steps here.

• Write the model in terms of a scalar field $\phi$, and determine the field equation for the perturbation in this field, $\delta \phi$, defined in the uniform curvature gauge.
• Separate $\delta \phi$ into a term $\delta \phi_1$ which obeys a linear equation of motion, and a part $\delta \phi_2$ which is quadratic in $\delta \phi_1$. Write down the evolution equation for $\delta \phi_2$.
• Determine the Fourier coefficients $F_2$ and $G_2$ (respectively) of the terms quadratic in $\delta \phi_1$ and $\delta \phi'_1$.
• Construct the quantities $F_2$ and $G_2$ from $F_2$ and $G_2$.
• Insert $F_2$ and $G_2$ in equations (23) and (27), respectively. Simplify the resulting expressions to give the complete three-point expectation value.

We have stressed that the computation of the $\delta \phi$ expectation value soon after horizon crossing is a simple exercise. However, we wish to emphasize that by itself this is insufficient for obtaining a prediction for the non-Gaussianity which is observed in the CMB. For that, one needs to make a prediction for the expectation values of the curvature perturbation, $\zeta$, using the classical evolution on superhorizon scales and the $\delta \phi$ expectation values evaluated soon after horizon crossing as an initial condition. This process avoids the appearance of large logarithms which would require resummation.
In the context of Einstein gravity coupled to a single scalar field this is simple to implement. The separate universe picture can be used in conjunction with rigorous results concerning the conservation of $\zeta$ on superhorizon scales to make robust predictions, especially in the case where isocurvature modes are absent. In more general theories it may be necessary to augment this procedure or find a replacement in order to make reliable estimates of the non-Gaussianity which can actually be observed in the CMB.

Acknowledgments

We would like to thank Jim Lidsey for useful discussions. DS would like to thank Misao Sasaki and the Yukawa Institute of Theoretical Physics, Kyoto, for their hospitality during the September 2007 programme Gravity and Cosmology, where part of this work was completed. DS is grateful to the Astronomy Unit at Queen Mary, University of London, for continued hospitality.

References

[1] Hinshaw G et al (WMAP Collaboration), Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: temperature analysis, 2007 Astrophys. J. Suppl. 170 288 [astro-ph/0603451]
[2] Martin J and Ringlev C, Inflation after WMAP3: confronting the slow-roll and exact power spectra to CMB data, 2006 J. Cosmol. Astropart. Phys. JCAP08(2006)009 [SPIRES] [astro-ph/0605367]
[3] Kinney W H, Kolb E W, Melchiorri A and Riotto A, Inflation model constraints from the Wilkinson Microwave Anisotropy Probe three-year data, 2006 Phys. Rev. D 74 023502 [SPIRES] [astro-ph/0605338]
[4] Bardeen J M, Steinhardt P J and Turner M S, Spontaneous creation of almost scale-free density perturbations in an inflationary universe, 1983 Phys. Rev. D 28 679 [SPIRES]
[5] Liddle A R and Lyth D H, 2000 Cosmological Inflation and Large-scale Structure (Cambridge: University Press) p 400
[6] Starobinsky A A, Stochastic de Sitter (inflationary) stage in the early universe, 1986 Field Theory, Quantum Gravity and Strings (Lecture Notes in Physics vol 246) ed H J De Vega and N Sanchez (Berlin: Springer) pp 107–26
[7] Sasaki M and Stewart E D, A general analytic formula for the spectral index of the density perturbations produced during inflation, 1996 Prog. Theor. Phys. 95 71 [SPIRES] [astro-ph/9507001]
[8] Lyth D H, Malik K A and Sasaki M, A general proof of the conservation of the curvature perturbation, 2005 J. Cosmol. Astropart. Phys. JCAP05(2005)004 [SPIRES] [astro-ph/0411220]
[9] Langlois D and Vernizzi F, Nonlinear perturbations of cosmological scalar fields, 2007 J. Cosmol. Astropart. Phys. JCAP02(2007)017 [SPIRES] [astro-ph/0610064]
[10] Rigopoulos G I and Shellard E P S, Non-linear inflationary perturbations, 2005 J. Cosmol. Astropart. Phys. JCAP10(2005)006 [SPIRES] [astro-ph/0405185]
[11] Peiris H V et al (WMAP Collaboration), First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: implications for inflation, 2003 Astrophys. J. Suppl. 148 213 [astro-ph/0302225]
[12] Bennett C L et al, 4-year COBE DMR cosmic microwave background observations: maps and basic results, 1996 Astrophys. J. 464 L1 [SPIRES] [astro-ph/9601067]
[13] Ferreira P G, Magueijo J and Gorski K M, Evidence for non-Gaussianity in the COBE DMR Four Year Sky Maps, 1998 Astrophys. J. 505 L1 [SPIRES] [astro-ph/9803256]
[14] Komatsu E et al (WMAP Collaboration), First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: tests of Gaussianity, 2003 Astrophys. J. Suppl. 148 119 [astro-ph/0302223]
[15] Komatsu E and Spergel D N, Acoustic signatures in the primary microwave background bispectrum, 2001 Phys. Rev. D 63 063002 [SPIRES] [astro-ph/0005036]
[16] Creminelli P, Senatore L, Zaldarriaga M and Tegmark M, Limits on $f_{NL}$ parameters from WMAP 3yr data, 2007 J. Cosmol. Astropart. Phys. JCAP03(2007)005 [SPIRES] [astro-ph/0610600]
[17] Yadav A P S et al, Fast estimator of primordial non-Gaussianity from temperature and polarization anisotropies in the cosmic microwave background II: partial sky coverage and inhomogeneous noise, 2007 Preprint 0711.4033 [astro-ph]
[18] Yadav A P S and Wandelt B D, Detection of primordial non-Gaussianity ($f_{NL}$) in the WMAP 3-year data at above 99.5% confidence, 2007 Preprint 0712.1148 [astro-ph]
Non-Gaussianity of inflationary field perturbations from the field equation

[49] Seery D, One-loop corrections to the curvature perturbation from inflation, 2007 Preprint 0707.3378 [astro-ph]

[50] Bartolo N, Matarrese S, Pietroni M, Riotto A and Seery D, On the physical significance of infra-red corrections to inflationary observables, 2007 Preprint 0711.4263 [astro-ph]

[51] Starobinsky A A, Multicomponent de Sitter (inflationary) stages and the generation of perturbations, 1985 JETP Lett. 42 152

[52] Lyth D H, Non-Gaussianity and cosmic uncertainty in curvaton-type models, 2006 J. Cosmol. Astropart. Phys. JCAP06(2006)015 [SPIRES] [astro-ph/0602285]

[53] Barnaby N and Cline J M, Large nongaussianity from nonlocal inflation, 2007 J. Cosmol. Astropart. Phys. JCAP07(2007)017 [SPIRES] [0704.3426] [hep-th]

[54] Vernizzi F and Wands D, Non-Gaussianities in two-field inflation, 2006 J. Cosmol. Astropart. Phys. JCAP05(2006)019 [SPIRES] [astro-ph/0603799]

[55] Komatsu E, Spergel D N and Wandelt B D, Measuring primordial non-Gaussianity in the cosmic microwave background, 2005 Astrophys. J. 634 14 [SPIRES] [astro-ph/0305189]

[56] Kohri K, Lyth D H and Melchiorri A, Black hole formation and slow-roll inflation, 2007 Preprint 0711.5006 [hep-ph]

[57] Wands D, Malik K A, Lyth D H and Liddle A R, A new approach to the evolution of cosmological perturbations on large scales, 2000 Phys. Rev. D 62 043527 [SPIRES] [astro-ph/0003278]

[58] Rigopoulos G I and Shellard E P S, The separate universe approach and the evolution of nonlinear superhorizon cosmological perturbations, 2003 Phys. Rev. D 68 123518 [SPIRES] [astro-ph/0306620]

[59] Lyth D H, Ungarelli C and Wands D, The primordial density perturbation in the curvaton scenario, 2003 Phys. Rev. D 67 023503 [SPIRES] [astro-ph/0208055]