Nonlinear evolution of long-wavelength metric fluctuations in inflationary models

D. S. Salopek
NASA/Fermilab Astrophysics Center, P.O. Box 500 MS-209, Batavia, Illinois, 60510

J. R. Bond
Canadian Institute for Advanced Research Program, Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Canada M5S 1A1

(Received 2 July 1990)

Stochastic inflation can be viewed as a sequence of two-step processes. In the first step a stochastic impulse from short-distance quantum fluctuations acts on long waves—the interaction. In the second step the long waves evolve semiclassically—the propagation. Both steps must be developed to address whether fluctuations for cosmic structure formation may be non-Gaussian. We describe a formalism for following the nonlinear propagation of long-wavelength metric and scalar-field fluctuations. We perform an expansion in spatial gradients of the Arnowitt-Deser-Misner equations and we retain only terms up to first order. At each point the fields obey evolution equations like those in a homogeneous universe, but now described by a local scale factor e^{α} and Hubble expansion rate H. However, the different points are joined together through the momentum constraint equation. The gradient expansion is appropriate for inflation if the long-wave fields are smoothed over scales below $e^{-\alpha}H^{-1}$. Our equations are naturally described in the Einstein-Hamilton-Jacobi framework, which governs an ensemble of inhomogeneous universes, and which may be interpreted as a semiclassical approximation to the quantum theory. We find that the Hubble parameter, which is a function of the local values of the scalar field, obeys a separated Hamilton-Jacobi equation that also governs the semiclassical phase of the wave functional. In our approximation, time hypersurface changes leave the equations invariant. However, the stochastic impulses that change the field initial conditions are most simply given on uniform expansion factor hypersurfaces whereas propagation is most easily solved on uniform Hubble hypersurfaces, in terms of $\alpha(x^j,H)$, the nonlinear analog of ζ of linear perturbation theory; we therefore pay special attention to hypersurface shifting. In particular, we describe the transformation process for the fluctuation probability functional. Exact general solutions are found for the case of a single scalar field interacting through an exponential potential. For example, we show that quantum corrections to long-wavelength evolution of the metric are characteristically small using exact Green's-function solutions of the Wheeler-DeWitt equation for this potential. Approximate analytic solutions to our classical system for slowly evolving multiple scalar fields are also easy to obtain in this formalism, contrasting with previous numerical approaches.

I. INTRODUCTION

Current cosmological data have cast doubt on whether Gaussian scale-invariant density fluctuations, generated by quantum noise during inflation, are sufficient to explain the large-scale structure of the Universe. Alternative paradigms are perhaps required. The observations include the patterns in galaxy redshift surveys,[1,2] the angular clustering of galaxies on the sky,[3] the clustering of clusters,[4] and the magnitude and coherence of the large-scale flow of galaxies.[5] Although the interpretation of each of these is controversial, there is a strong hint that there is more large-scale structure than the most successful version of the Gaussian scale-invariant models, the adiabatic cold-dark-matter scenario, can accommodate. The alternative theories include various types of topological defects such as strings, domain walls and "cosmic textures," late-time phase transitions, and various hydrodynamical and radiative processes in the later Universe. They all predict non-Gaussian density perturbations of one form or another. Many theorists have been reexamining the inflationary paradigm to ascertain under what conditions either scale invariant or Gaussian statistics would break down. At the moment the consensus is that inflation driven by a single scalar field does lead to Gaussian fluctuations over the observable length scales and is most likely to give a nearly scale-invariant perturbation spectrum—although certain exotic potentials can modify this somewhat. However, with more than one dynamically important scalar field in inflation, the question is not whether scale invariance can be broken, but at what scale the breaking appears.[6,7] (It is usually argued that it is unnatural for the scale to be in our observable range.) Broken scale invariance, but with Gaussian statistics maintained, can be addressed entirely within linear perturbation theory. Although its implementation can be technically tricky, it requires no new conceptual ideas. It is even self-consistent at the quantum level.

The non-Gaussian issue requires a qualitatively different approach, capable of treating nonlinearities in field theory on inhomogeneous spacetimes, for both the gravitational field and scalar fields. The beginnings[8] of such a framework have been emerging over the past 5 years under the name "stochastic inflation," in which a separation is made between short-distance quantum fluctuations which oscillate on scales below the instantaneous Hubble radius and large distance fluctuations which are treated as classical fields. The short-distance components communicate with the long-wavelength classical fields by stochastic noise terms. We believe that this approach should be vigorously developed both at the level of fundamental field theory to address the conceptual foundations[9] and at the operational level to provide a calculational tool for quantitatively addressing those issues of nonlinear inhomogeneities of relevance for cosmic structure formation.

This is the first in a series of papers formulating a stochastic inflation approach to fluctuation generation and evolution and leading to computation of observable field configurations such as microwave-background anisotropy patterns and large-scale structures. In this paper we show how one can describe the long-wavelength fields self-consistently and evolve them from their initial conditions using a Hamilton-Jacobi approach. We do not explicitly include any short-distance communication. In the second paper[10] in this series (which will be referred to as SB2), we consider the stochastic evolution of the long-wavelength fields as a sequence of two-stage processes, a stochastic kick from short-distance effects (diffusion) followed by free propagation (drift). The stochastic impulse effectively resets the fields to a new set of initial conditions for the next round of propagation. Thus the current paper treats the propagation phase, given the initial conditions.

We find that the natural variables for describing the dynamical state of the metric are the local expansion factor $a(x^j,t)$ [or its logarithm, $\alpha=\ln a(x^j,t)$, which proves more useful] and the local Hubble expansion rate $H(x^j,t)$. These physical fields are useful for clarifying various aspects of the linear perturbation theory of inflation. Although, for a single scalar field, much analytic perturbation work[11-14] has been undertaken, a decisive step occurred when Bardeen, Steinhardt, and Turner[14] (BST) introduced a gauge-invariant metric variable ζ, which has the virtue of being a constant of motion once the wavelength of the perturbation exceeds the Hubble radius. There is a natural nonlinear generalization of ζ which we propose here which can be used to simplify the treatment of nonlinear interaction of many scalar fields among themselves and with the gravitational field during inflation, in spite of the complex interplay of scalar, vector, and tensor modes of the metric. Our version of ζ is 3α evaluated on hypersurfaces with H constant; for nonzero wave numbers, this agrees with the perturbation theory definition. [This ζ is $3\zeta_{\rm BST}$, where $\zeta_{\rm BST}$ is the original definition given by Bardeen, Steinhardt, and Turner; see Salopek, Bond, and Bardeen[7] (SBB).] In the variables α and H, we can solve three-dimensional nonlinear gravitational problems during the inflation epoch for fluctuations whose length scales are greater than the Hubble radius. Our approach also proves useful in semiclassical treatments. Historically, the long-wavelength evolution for a universe consisting of dust was first solved by Lifshitz and Khalatnikov,[15] and then extended by Eardley, Lang, and Sachs.[16] They restricted their attention to comoving hypersurfaces upon which Einstein's equations could be solved. Our work may be viewed as an extension to inflation models where the pressure is not zero and there may be many scalar fields present. Our formalism applies to arbitrary time choices, thus making it a useful calculational tool for stochastic inflation.

We consider the evolution of the metric and scalar fields from some initial classical configuration defined on a spacelike hypersurface assuming fluctuations have been smoothed out on scales smaller than the comoving Hubble radius $(Ha)^{-1}$. The evolution of regions separated by more than $(Ha)^{-1}$ will roughly evolve like independent universes because there is no causal contact between the points. This separation is not exact because the various regions are connected by large-scale gradients appearing in the gravitational and scalar-field equations. We are thus motivated to perform a spatial gradient expansion of Einstein's equations. In the zeroth-order approximation, all spatial gradients are set to zero, yielding the homogeneous Friedmann equations, which must be satisfied point by point. The next-order equations, given in Sec. II, contain first-order spatial gradient terms as well. The new ingredient is the momentum constraint equation of general relativity. To this order it can be integrated exactly, showing that the Hubble parameter $H(\phi_j)$ can only be a function of the scalar field ϕ_j and that there is no explicit dependence on the (arbitrary) time variable. The energy constraint equation (the usual Friedmann equation for H^2) is then cast into a partial differential equation for H, which we call the separated Hamilton-Jacobi equation.

In Sec. III we illustrate our formalism using analytic solutions for a single scalar field interacting with an exponential potential. We find that our equations conveniently isolate the "growing mode" of a single scalar field of arbitrary potentials. In this case the simplest choice of time is ϕ because the remaining equations are then easy to integrate; for multiple fields the local scale factor of the Universe is the natural time choice. In Sec. III we also exhibit approximate solutions of the Hamilton-Jacobi system assuming slow-rolling of fields. One can use this, for example, to derive analytically, yet quantitatively, fluctuation spectra for multiple-scalar-field models such as double inflation, which previously required a complicated numerical code (SBB).

In Sec. IV we show how the long-wavelength equations may be directly derived from the Hamilton-Jacobi formalism for inhomogeneous scalar fields and metric. To be consistent to first order in spatial gradients, one must solve both the energy and momentum constraints. A redundancy theorem,[17] which states that the momentum constraint is satisfied everywhere if the energy constraint is satisfied everywhere, does not apply in this situation. However, we can satisfy the functional momentum constraint equation; indeed, we obtain a general class of solutions to it. We also apply the Hamilton-Jacobi formalism to the evolution of the classical probability function for inhomogeneous fields, which is of relevance for cosmic structure formation studies. We develop a transformation theory relating the probability on different time hypersurfaces, which will prove useful for our later treatment of stochastic inflation. In Sec. V we discuss the quantum theory of long-wavelength evolution, which is the natural arena for combining the Hamilton-Jacobi approach with the probability evolution equation. This connects our analysis with the vast literature on quantum cosmology (see Halliwell[18] for a bibliography). In particular, we must solve the minisuperspace Wheeler-DeWitt equation at each point. Pilati[19] and Teitelboim[20] were the first to develop the quantum theory of long-wavelength fields, which they referred to as the strong gravitational coupling limit $G \equiv m_{\mathcal{P}}^{-2} \to \infty$. We exhibit analytic Green's functions of the Wheeler-DeWitt equation for the case of an exponential potential. We use this to show that quantum effects on the long-wavelength-field propagation are small because we smooth all field configurations on scales smaller than the Hubble radius; for the same reason macroscopic objects behave classically. Once again, the momentum constraint equation relates different spatial points, but we encounter difficulties satisfying it and the Wheeler-DeWitt equation beyond the semiclassical level.

II. SPATIAL GRADIENT EXPANSION FOR INFLATING COSMOLOGIES

A. Arnowitt-Deser-Misner formulation of Einstein and scalar-field equations

In the Arnowitt-Deser-Misner (ADM) formalism, the metric

$$g_{00} = -N^2 + \gamma^{ij} N_i N_j , \quad g_{0i} = g_{i0} = N_i , \quad g_{ij} = \gamma_{ij}$$

is parametrized by the three-metric γ_{ij} and the lapse and shift functions N and N^i, which describe the evolution of the timelike hypersurfaces. The ADM form of the action for n minimally coupled scalar fields $\phi_1, \ldots, \phi_n$ coupled to Einstein's gravity and self-interacting through a potential $V(\boldsymbol{\phi})$ is[21]

$$\mathcal{I} = \int d^4x \sqrt{-g} \left[\frac{m_{\mathcal{P}}^2}{16\pi} R - \frac{1}{2} g^{\mu\nu} \partial_\mu \phi_k \partial_\nu \phi_k - V(\phi_k) \right]$$

$$= \int d^4x \, N \sqrt{\gamma} \left[\frac{m_{\mathcal{P}}^2}{16\pi} ({}^{(3)}R + K_{ij} K^{ij} - K^2) + \frac{1}{2} [(\dot\phi_k - N^i \phi_{k|i})^2 / N^2 - \phi_{k|i} \phi_k{}^{|i}] - V(\phi_k) \right] .$$

We have adopted the summation convention. Vertical bars denote three-space-covariant derivatives with connection coefficients determined from the γ_{ij}. The three-space curvature associated with the metric γ_{ij} is ${}^{(3)}R$ and the extrinsic curvature three-tensor is

$$K_{ij} = \left[N_{i|j} + N_{j|i} - \frac{\partial \gamma_{ij}}{\partial t} \right] \Big/ (2N) . \tag{2.1a}$$

The traceless part of a tensor is denoted by an overbar. In particular,

$$\bar K_{ij} = K_{ij} - \tfrac{1}{3} K \gamma_{ij} , \quad K = K_i^i . \tag{2.1b}$$

The trace K is a generalization of the Hubble parameter that appears in isotropic cosmologies.

Variation of the action with respect to N and N^i yields the energy and momentum constraint equations (see, e.g., Ref. 22)

$$\bar K_{ij} \bar K^{ij} - \frac{2}{3} K^2 - {}^{(3)}R + \frac{16\pi}{m_{\mathcal{P}}^2} \epsilon = 0 , \tag{2.2}$$

$$\bar K^j_{i|j} - \frac{2}{3} K_{|i} + \frac{8\pi}{m_{\mathcal{P}}^2} \Pi^{\phi_k} \phi_{k|i} = 0 . \tag{2.3}$$

Variation of $\mathcal{T}$ with respect to γ_{ij} yields the dynamical gravitational-field equations

$$\frac{\partial K}{\partial t} - N^i K_{|i} = -N^{|i}{}_{|i} + N \left[\frac{3}{4} \bar K_{ij} \bar K^{ij} + \frac{1}{2} K^2 + \frac{1}{4} {}^{(3)}R + \frac{4\pi}{m_{\mathcal{P}}^2} S \right] , \tag{2.4}$$

$$\frac{\partial \bar K^i_k}{\partial t} + N^i{}_{|l} \bar K^l_k - N^l{}_{|k} \bar K^i_l - N^l \bar K^i_{k|l} = -N^{|i}{}_{|k} + \frac{1}{3} N^{|l}{}_{|l} \delta^i_k + N \left[K \bar K^i_k + {}^{(3)}\bar R^i_k - \frac{8\pi}{m_{\mathcal{P}}^2} \bar S^i_k \right] , \tag{2.5}$$

Variation of $\mathcal{I}$ with respect to ϕ_k gives the scalar-field equations of motion:

$$\left[\frac{\partial \Pi^{\phi_k}}{\partial t} - N^i \Pi^{\phi_k}{}_{|i} \right] \Big/ N - K \Pi^{\phi_k} - N_{|i} \phi_k{}^{|i} / N - \phi_{k|i}{}^{|i} + \frac{\partial V}{\partial \phi_k} = 0 , \tag{2.6}$$

where the scalar-field momentum is

$$\Pi^{\phi_k} = (\dot\phi_k - N^i \phi_{k|i}) / N . \tag{2.7}$$

The energy density on a constant-time surface is

$$\epsilon = \tfrac{1}{2} [(\Pi^{\phi_k})^2 + \phi_{k|i} \phi_k{}^{|i}] + V(\phi_k) , \tag{2.8a}$$

and the stress three-tensor is

$$S_{ij}=T_{ij}$$
$$=\phi_{k|i}\phi_{k|j}+\gamma_{ij}[\tfrac{1}{2}(\Pi^{\phi_k})^2-\tfrac{1}{2}\phi_{k|i}\phi_k{}^{|i}-V(\phi_k)] \ . \quad (2.8\text{b})$$

Although the goal is of course to solve these highly nonlinear coupled equations in a cosmological setting, its realization is a long way off. Rather drastic approximations are required to make progress at the current state of the art in computational relativity. A highly successful approach is to assume homogeneity of the fields to give a background solution and linearization of the equations to describe deviations from spatial uniformity (e.g., SBB). The smallness of cosmic microwave-background anisotropies and of large-scale galaxy density contrasts provides some justification that the Universe we see may indeed be accurately described within this perturbation framework. Even if this criterion does hold for our local observable patch of the Universe, there is no reason to suppose that it will be valid on much larger scales. Further, it is also possible that inherent nonlinearities may have played a role even in the patch we observe. Indeed, such inflationary models are the motivation for this current work. The attack on nonlinear aspects of the inhomogeneous ADM equations has been very limited so far. Progress has been made if some symmetry has been adopted to restrict the spatial dependence to at most one variable, both numerically for spherical and planar systems, and analytically for some very restrictive classes of metrics. If the Universe is inflating, another class of approximations is suggested, the stochastic inflation picture pioneered by Vilenkin, Starobinski, and others,[8] in which the short-distance behavior of the fields communicates with the large-distance structure through stochastic forces. In this paper the equations for the long-wavelength fields are obtained by systematically neglecting large-scale gradients, which leads to a self-consistent set of equations as we now show.

B. Spatial gradient expansion of the ADM and scalar-field equations

It is reasonable to expand in spatial gradients whenever the forces arising from temporal changes in the fields sufficiently exceed the forces from the spatial gradients. A standard example of this occurs in linear perturbation theory, when one solves the perturbation equations for evolution "outside of the horizon:" A typical time scale for evolutionary changes is the Hubble time H^{-1}, which is assumed to exceed the gradient scale ak^{-1}, where k is the comoving wave number of the perturbation, suggesting we expand in powers of k/Ha. Once nonlinear terms are included, the expansion parameter is not so straightforward. However, provided we are interested in structure on scales larger than the horizon, it is reasonable to expand in the nonlinear analog of $(Ha)^{-1}\nabla$. This is particularly appropriate for inflation models. For example, in linear perturbation theory, spatial gradients become exponentially negligible after a few e-foldings of expansion beyond $k=Ha$.

It is therefore a useful approach to split the fields [e.g., $\phi_k(t,x^j)=\phi_{bk}+\phi_{fk}$] into smoothed long-wavelength "background" fields $\phi_{bk}(t,x^j)$ and residual short-wavelength fluctuating fields $\phi_{fk}(t,x^j)$. It is not clear how to come up with a gauge-invariant form for this split which is useful. In this paper we shall assume that there is a gauge with coordinates t,x^i in which ϕ_{bk} has the form

$$\phi_{bk}(t,x)=\int S(t,x-x')\phi_k(t,x')d^3x' \ , \quad (2.9)$$

where S is a smoothing function whose Fourier transform falls off at high spatial momentum. Below we argue that there is a preferred timelike hypersurface within the stochastic inflation framework in which this is a reasonable definition. In order to ensure that ϕ_{bk} remains the same in another gauge, the relation between ϕ_{bk} and ϕ_k is found by transforming this convolution form. The relation in the new gauge will then not be as simple. The rigorous definition of ϕ_{bk} is problematic since it depends upon the specific choice of timelike hypersurface; i.e., the smoothing is not gauge invariant. For stochastic inflation the natural smoothing scale is the comoving Hubble length $(Ha)^{-1}$ and the natural hypersurfaces are those on which Ha is constant, at least within linear perturbation theory.[10] In that case a fundamental difference between ϕ_{fk} and ϕ_{bk} is that the short-wavelength components are essentially uncorrelated at different times, while longer-wavelength components are deterministically correlated.

By convolving the ADM equations with S, we get equations for the background fields; subtracting these from the original equations gives the equations for the fluctuating fields. The nonlinearity makes these two sets of equations very complicated indeed unless suitable approximations are made. The philosophy here is to expand in the spatial gradients of, e.g., ϕ_{bk} which operate on the background fields and to treat the terms explicitly, depending upon the fluctuating fields (e.g., ϕ_{fk}) that appear in the background-field equation as stochastic forces describing the connection of the short wavelengths to the long ones. In this paper we focus on the gradient expansion and set the fluctuating field ϕ_{fk} to zero. We do include its influence in the form of initial conditions for the background field. We retain only those terms which are at most first order in spatial gradients, neglecting such terms as $\phi_{\mathbf{bk}|i}^{|i}$, $\phi_{bk}{}^{|i}\phi_{bk|i}$, ${}^{(3)}R$, ${}^{(3)}R^i_j$, and $\bar{S}^i_j$. The terms involving second derivatives of ϕ_{fk} and the second spatial derivatives of the fluctuating part of the metric variables give the stochastic forces, which are treated in SB2. Since we are neglecting the fluctuating fields, we shall also drop the subscript bk in the following.

The equations simplify considerably if we set the shift N^i to zero. The evolution equation (2.5) for the traceless part of the extrinsic curvature is then $\partial\bar{K}^i_k/(N\partial t)=K\bar{K}^i_k$. Using $K=-\partial\ln(\sqrt{\gamma})/(N\partial t)$ from Eq. (2.1), where γ is the determinant of γ_{ij}, the solution is $\bar{K}^i_k\propto\gamma^{-1/2}$. Since during inflation $\gamma^{-1/2}\equiv a^{-3}$, where a is the overall expansion factor, $\bar{K}^i_k$ decays extremely rapidly. We can therefore set $\bar{K}^i_k$ to zero in the following—although the $\bar{K}^i_k\neq 0$ case where the long-wavelength gravitational radiation evolves in time also turns out to be tractable.[23] The most general form of the three-metric with vanishing $\bar{K}^i_k$ is

$$\gamma_{ij}=a^2(t,x^j)h_{ij}(x^j),\quad a(t,x^j)\equiv\exp[\alpha(t,x^j)]\ , \tag{2.10}$$

where the time-dependent conformal factor $a(t,x^j)$ is interpreted as a spatially dependent expansion factor. It is more convenient for us to work in α. The time-independent three metric $h_{ij}(x)$, which we assume has unit determinant, describes the three-geometry of the conformally transformed space. Within linear perturbation theory, $\alpha(t,x^j)$ would contribute only to the scalar perturbation modes, whereas the vector and tensor modes are time independent, reflecting the constancy of $h_{ij}(x^j)$. For the scalar models, if the longitudinal gauge is chosen, then $\alpha(t,x^j)$ can be written as the sum of a homogeneous $\bar{\alpha}(t)$ and Bardeen's gauge-invariant metric variable $\Phi_H(t,x^j)$. Based on the work of Lifshitz and Khalatnikov[15] for a dust-dominanted universe, Starobinski[8] suggested that an equation of the form (2.10) is valid in synchronous and comoving gauge. From our development this form follows for an arbitrary choice of the lapse function.

Since $a(t,x^j)$ is interpreted as a scale factor, we now use the Hubble parameter

$$H(t,x^j)\equiv\dot{\alpha}(t,x^j)/N(t,x^j)=-K(t,x^j)/3\ ,$$

in place of the trace K of the extrinsic curvature. The momentum constraint equation (2.3) can now be written

$$H_{|i}=-\frac{4\pi}{m_{\mathcal{P}}^2}\Pi^{\phi_k}\phi_{k|i}\ . \tag{2.11}$$

In the general solution, H is a function of the scalar-field values and of time:

$$H(t,x^j)\equiv H(\phi_k(t,x^j),t)\ . \tag{2.12a}$$

The scalar-field momenta $\Pi^{\phi_k}=N^{-1}\dot{\phi}_k$ must then obey

$$\Pi^{\phi_k}=-\frac{m_{\mathcal{P}}^2}{4\pi}\left[\frac{\partial H}{\partial\phi_k}\right]_t\ . \tag{2.12b}$$

We now show that the time dependence of H arises only through its dependence upon ϕ_k. Comparing the evolution equation (2.4) with the time derivative of Eq. (2.12a),

$$\begin{aligned}\frac{1}{N}\left[\frac{\partial H}{\partial t}\right]_{x^j}&=\frac{1}{N}\left[\frac{\partial\phi_k}{\partial t}\right]_{x^j}\left[\frac{\partial H}{\partial\phi_k}\right]_t+\frac{1}{N}\left[\frac{\partial t}{\partial t}\right]_{x^j}\left[\frac{\partial H}{\partial t}\right]_{\phi_k}\\&=-\frac{m_{\mathcal{P}}^2}{4\pi}\left[\frac{\partial H}{\partial\phi_k}\right]_t^2+\frac{1}{N}\left[\frac{\partial H}{\partial t}\right]_{\phi_k}\ ,\end{aligned}$$

we see that

$$\left[\frac{\partial H}{\partial t}\right]_{\phi_k}=0\ ,$$

and hence

$$H(t,x^j)=H(\phi_k(t,x^j))\ .$$

Thus, if one can neglect second-order spatial gradients, then the comoving gauge for a scalar field $\phi=$const is identical to the uniform Hubble gauge. This equivalence breaks down when the stochastic forces are included which do depend upon second-order spatial gradients.[10]

From Eq. (2.12b) we also see that the scalar-field momenta at a point are functions only of the scalar-field values at that point: $\Pi^{\phi_k}(t,x^j)=\Pi^{\phi_k}(\phi_{k'}(t,x^j))$. We need to show that these results are consistent with the rest of our equations. By differentiating the energy constraint equation, now in the form $H^2=(8\pi/3)\epsilon/m_{\mathcal{P}}^2$, with respect to ϕ_k, while holding t fixed, we obtain

$$\Pi^{\phi_k}\left[\frac{\partial\Pi^{\phi_j}}{\partial\phi_k}\right]_t+3H\Pi^{\phi_j}+\frac{\partial V}{\partial\phi_j}=0\ .$$

By comparing with the scalar-field evolution equation (2.6) and using the identity

$$\left[\frac{\partial\Pi^{\phi_j}}{\partial t}\right]_t=\Pi^{\phi_k}\left[\frac{\partial\Pi^{\phi_j}}{\partial\phi_k}\right]_t+\left[\frac{\partial\Pi^{\phi_j}}{\partial t}\right]_{\phi_k}\ ,$$

we verify that the functional form $\Pi^{\phi_k}(\phi_{k'})$ has no explicit time dependence.

With the neglect of second-order spatial gradients, the complicated ADM and scalar-field equations of Sec. II A reduce to the simple collection of background-field equations:

$$H\equiv H(\phi_j)\ , \tag{2.13a}$$

$$H^2=\frac{m_{\mathcal{P}}^2}{12\pi}\sum_k\left[\frac{\partial H}{\partial\phi_k}\right]^2+\frac{8\pi}{3m_{\mathcal{P}}^2}V(\phi_j)\ , \tag{2.13b}$$

$$\frac{\dot{\phi}_j}{N}=-\frac{m_{\mathcal{P}}^2}{4\pi}\frac{\partial H}{\partial\phi_j}\ , \tag{2.14a}$$

$$\frac{\dot{\alpha}}{N}=H\ . \tag{2.14b}$$

These equations describe point by point the *nonlinear* evolution of the scalar *and* gravitational fields smoothed over the horizon in an inflating patch of the Universe. These equations look similar to those of the Hamilton-Jacobi formalism of classical mechanics,[24] a connection we explore in detail in Sec. IV. In anticipation of our findings there, we refer to (2.13b) as the separated Hamilton-Jacobi equation (SHJE).

Instead of explicitly finding the function $H(\phi_j)$, we could instead work directly with the equations of motion, which reduce to

$$\frac{\partial}{N\partial t}\frac{\partial}{N\partial t}\phi_k+3H\frac{\partial}{N\partial t}\phi_k+\frac{\partial V}{\partial\phi_k}=0\ , \tag{2.15a}$$

$$H^2=\frac{8\pi}{3m_{\mathcal{P}}^2}\left[\frac{1}{2}\sum_k\left[\frac{\partial}{N\partial t}\phi_k\right]^2+V(\phi_k)\right]\ , \tag{2.15b}$$

$$\frac{\partial}{N\partial t}\alpha=H\ , \tag{2.15c}$$

with the neglect of second-order spatial gradients. In particular, with the synchronous gauge choice $N=1$ (together with the $N^i=0$ condition), the field equations look just like the familiar homogeneous equations for inflation models, point by point. Fluctuations are described through spatial variation in ϕ_k, α, and H. However, we would have to ensure that the initial inhomogeneous data satisfy the momentum constraint equation $\nabla H = -(4\pi/m_{\mathcal{P}}^2)\Pi^{\phi_k}\nabla\phi_k$. The momentum constraint would, then, hold at later times because of the equations of motion, whereas in the Hamilton-Jacobi formalism it is automatically satisfied at all times.

If there are many scalar fields, the SHJE is a partial differential equation in ϕ_k space. One way to solve it is by the method of characteristics, which involves integrating trajectories using Eqs. (2.15). The initial-value problem is specified by a surface $f(\phi_k)=0$ upon which the Hubble parameter is constant. A trajectory emanates from each point on this surface, with the direction of each initial velocity given by ∇f and with the magnitude given by the SHJE. Each path may then be integrated independently using Eqs. (2.15), and $H(\phi_k)$ may then be found from a catalog of trajectories.[24] However, in cases in which the functions $H(\phi_k)$ can be found by other means, the Hamilton-Jacobi approach is more straightforward to implement. In particular, this is the case for the specific analytic solutions given in Sec. III.

Transformations of the timelike hypersurfaces do not upset the solutions for the metric or the form of Eqs. (2.13) and (2.14) in a first-order gradient expansion. That is, given a new time variable $T=T(t,x^j)$, we can find coordinates $X^j(t,x^j)$ such that the shift remains zero and the form of the metric, $\gamma_{ij}=\exp[2\tilde{\alpha}(T,X^j)]\tilde{h}_{ij}(X^j)$, is retained. This is demonstrated in Appendix A. Further, derivatives of a quantity $Q(t,x^j)$ with respect to the two different time variables are equal up to second-order spatial gradients:

$$\frac{1}{N_T}\left[\frac{\partial Q}{\partial T}\right]_{X^j}=\frac{1}{N_t}\left[\frac{\partial Q}{\partial t}\right]_{x^j}, \tag{2.16}$$

where $N_t=N_T(\partial T/\partial t)$ relates the lapse functions on the different hypersurfaces. This relation explicitly demonstrates that Eqs. (2.14) are invariant under arbitrary changes of the time hypersurfaces. Of course, we have implicitly assumed that the function T is nonsingular. For some choices of physical interest, this is not true, for example, for $T=\phi$, if the scalar field ϕ undergoes oscillations. However, on restricted intervals it could still serve as a valid local clock.

III. SOLVING THE HAMILTON-JACOBI EQUATIONS

In this section we illustrate the usefulness of the Hamilton-Jacobi form of the long-wavelength fluctuation equations equations, (2.13) and (2.14), by explicitly solving them in several situations of cosmological importance. In Sec. III A an analytic example for a single scalar field interacting through an exponential potential is given. In Sec. III B we analyze the general solution for a single scalar field and extend the Bardeen, Steinhardt, and Turner gauge-invariant metric variable ζ, which has proved so useful in linear perturbation theory to the nonlinear case. We extend the analysis to n scalar fields in Sec. III C, showing that differentiation of the Hubble parameter solution of the SHJE with respect to its n integration constants solves the field equations. We also argue that α is the most natural choice of time variable for this case. In Sec. III D we show how approximate solutions may be obtained for multiple fields using the slow-roll approximation and extensions based upon it. For example, one may apply these results to determine analytically fluctuation spectra arising in double inflation as well as other multiple-scalar-field models. In Sec. III E we solve exactly the Hamilton-Jacobi equations for two scalar fields interacting with a specific separable solution to illustrate how the Hubble parameter may be a multivalued function of the scalar fields.

A. Exact solution for a single inflaton in an exponential potential

An exponential potential of the form

$$V(\phi)=V_0\exp\left[-\left[\frac{16\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right]$$

has proved very useful for generating analytic results (with $p>1$ required for inflation). For example, Lucchin and Matarrese[25] gave an exact solution of the cosmological background equations, and Ratra[6] has given some exact results for linear perturbation theory. The particular solution to the energy constraint equation (2.13b),

$$H_{\rm att}(\phi)=\left[\frac{8\pi V_0}{3m_{\mathcal{P}}^2}\frac{1}{1-1/(3p)}\right]^{1/2}\times\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right], \tag{3.1}$$

defines the attractor of Halliwell,[27] toward which all trajectories tend. A more complex parametric solution $H(\phi)$ is required to describe the motion for arbitrary initial conditions. We obtain this by defining a new dependent variable f,

$$H(\phi)=\left[\frac{8\pi V_0}{3m_{\mathcal{P}}^2}\right]^{1/2}\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right]f(\phi),$$

which transforms (2.13b) to

$$\frac{m_{\mathcal{P}}}{\sqrt{12\pi}}\frac{df}{d\phi}=\frac{1}{\sqrt{3p}}f\pm(f^2-1)^{1/2}.$$

A change of variables, $f=\cosh(u)$, suggested by this equation, yields the following parametric solution, $H\equiv H(u)$, $\phi\equiv\phi(u)$:

$$H(u)=\left[\frac{8\pi V_0}{3m_{\mathcal{P}}^2}\right]^{1/2}\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi(u)}{m_{\mathcal{P}}}\right]\cosh(u), \tag{3.2a}$$

$$\phi(u)=\phi_m-\frac{m_{\mathcal{P}}}{\sqrt{12\pi}}\left[1-\frac{1}{3p}\right]^{-1}\times[u+(3p)^{-1/2}\ln|\cosh(u)-\sqrt{3p}\,\sinh(u)|], \tag{3.2b}$$

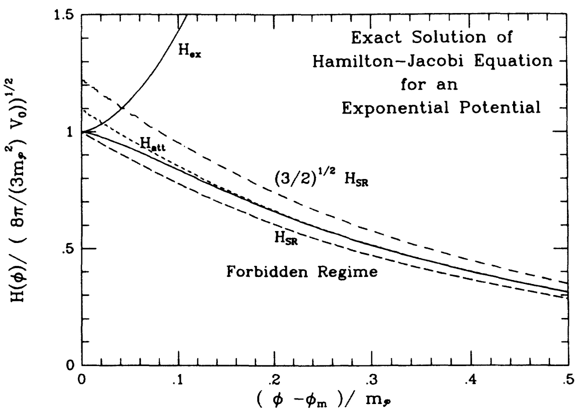


FIG. 1. The analytic Hubble function H_{ex} given by Eq. (3.2) for an exponential potential $V(\phi)=V_0\exp(-(16\pi/p)^{1/2}\phi/m_{\mathcal{P}})$ with $p=2.0$ is compared with the attractor solution H_{att} and the slow-rollover function $H_{SR}(\phi)$. The scalar field begins with a large value of H_{ex}, moves rapidly to the left, and reaches its minimum at ϕ_m, where H_{ex} has a cusp. The field then rolls down the potential, quickly evolving to the attractor H_{att}. At all times $H(\phi)$ must be greater than the $H_{SR}(\phi)$, which is typically a good approximation to H_{att} if the potential is very flat, $p>2$. Above the curve $(3/2)^{1/2}H_{SR}$, the Universe is not inflating. Even when inhomogeneities are incorporated, ϕ_m must be spatially independent, and remarkably, the function $H(\phi)$ is the same at all comoving spatial points in the long-wavelength approximation. Thus, if two spatial points have the same scalar-field value, then their Hubble function must also agree.

where

$$-\infty < u < \operatorname{arctanh}(1/\sqrt{3p}\,)\ , \tag{3.2c}$$

or

$$\operatorname{arctanh}(1/\sqrt{3p}\,) < u < \infty\ . \tag{3.2d}$$

The constant of integration ϕ_m must be spatially independent in order to satisfy the momentum constraint equation (2.11). There are two solutions corresponding to the case where the initial value of the parameter u is in the range (3.2c) or in the range (3.2d). The first case is in general double valued, as shown in Fig. 1. For these, ϕ_m is the minimum value that ϕ obtains. As u increases from $-\infty$ to 0, ϕ and H decrease, until a cusp in $H(\phi)$ is reached at the minimum ϕ_m. Near the turning point at ϕ_m,

$$H(\phi)\approx\left[\frac{8\pi V_0}{3m_{\mathcal{P}}^2}\right]^{1/2}\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi_m}{m_{\mathcal{P}}}\right]\times\left\{1\pm\left[\frac{p}{3}\right]^{1/2}\left[\left[\frac{16\pi}{p}\right]^{1/2}\left[\frac{(\phi-\phi_m)}{m_{\mathcal{P}}}\right]\right]^{3/2}\right\}\ ,\quad \phi\approx\phi_m\ . \tag{3.3}$$

The $+$ $(-)$ sign corresponds to motion to the left (right) in the figure. For positive u, ϕ increases from ϕ_m at $u=0$ to ∞ at $u=\operatorname{arctanh}(1/\sqrt{3p}\,)$, while H continues decreasing, rapidly approaching the attractor solution which is marked by the broken line in Fig. 1. The second range of u [Eq. (3.2d)] applies to the case where the scalar field would be initially moving to the right in Fig. 1 with a $H(\phi)>H_{att}(\phi)$. As u decreases from positive values to $\operatorname{arctanh}(1/\sqrt{3p}\,)$, H quickly approaches H_{att}.

Along the attractor solution the relation between the scalar factor at each point x^j and ϕ is found by dividing (2.14b) by (2.14a) and integrating, giving the linear law

$$\alpha(\phi,x^j)-\alpha(\phi_0,x^j)=(4\pi p)^{1/2}(\phi-\phi_0)/m_{\mathcal{P}}\ , \tag{3.4a}$$

which prescribes the point-by-point evolution of the expansion factor between the comoving hypersurfaces at times ϕ and ϕ_0, as shown in Fig. 2. This figure also illustrates how to transform from spatial fluctuations in α on constant ϕ hypersurfaces to spatial fluctuations in ϕ on constant α hypersurfaces, which is relevant for the treatment of perturbations in stochastic inflation as we describe in Sec. III B.

Remarkably, one can solve for the $\alpha(\phi,x^j)$ trajectories

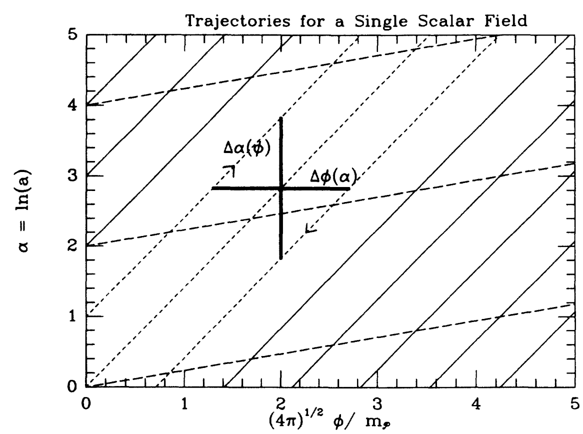


FIG. 2. Trajectories in the ϕ-α plane are displayed for cosmologies driven by an exponential potential. The solid and short-dashed lines represent trajectories of the attractor solution [Eqs. (3.1) and (3.4a)]. The long-dashed lines are surfaces of constant phase S of the Hamilton-Jacobi solution (4.8); they are orthogonal to the trajectories, as measured by the supermetric (4.17a). The thick solid lines illustrate the mechanics of a hypersurface transformation. The length of the horizontal line depicts a variation $\Delta\phi(\alpha)$ on a surface of constant α, which is analogous to the way the initial conditions for galaxy formation are set when a scale crosses the Hubble radius during inflation. The length of the vertical line represents the variation $\Delta\alpha(\phi)$ on a constant ϕ surface, which gives the fluctuations in the nonlinear analog of the metric fluctuation $\zeta/3$. Long-wavelength evolution in the time variable ϕ is simple to calculate. To effect the hypersurface transformation, the left end point is projected upward along a trajectory, whereas the right is projected down, until their scalar-field values coincide.

analytically in the general case as well, using the methods described in Sec. III C:

$$\alpha(\phi,x^j)-\alpha(\phi_0,x^j)=b(\phi)-b(\phi_0)\ , \tag{3.4b}$$

$$\begin{aligned} b(\phi)&\equiv-\frac{1}{3}\ln\left|\frac{\partial H}{\partial\phi_m}\right| \\ &=(4\pi p)^{1/2}\frac{\phi}{m_{\mathcal{P}}}-(4\pi p)^{1/2}\frac{\phi_m}{m_{\mathcal{P}}}\left[1-\frac{1}{3p}\right] \\ &\quad+\left[\frac{p}{3}\right]^{1/2}u(\phi)-\frac{1}{6}\ln\left[\frac{32\pi^2}{3p}\frac{V_0}{m_{\mathcal{P}}^4}\right]. \end{aligned} \tag{3.4c}$$

The linear law (3.4a) is recovered if ϕ or $-\phi_m$ is large.

We are primarily interested in the regime where inflation occurs, in which case we must have $\rho+3p<0$, i.e., $(\Pi^\phi)^2<V$. In Fig. 1 this region is given by $H(\phi)<\sqrt{3/2}H_{\rm SR}(\phi)$, where the "slow-roll-down" Hubble parameter is

$$H_{\rm SR}\equiv\left[\frac{8\pi}{3m_{\mathcal{P}}^2}\right]^{1/2}V^{1/2}(\phi)\ .$$

The solution (3.4) is still valid in the upper noninflating region above this line, whereas the region below $H_{\rm SR}(\phi)$ is forbidden.

B. General metric fluctuation formula for a single scalar field in an arbitrary potential

For a single scalar field with a general potential, one can view the separated Hamilton-Jacobi equation as a first-order ordinary differential equation in the time variable ϕ. Each solution $H(\phi,I)$ is characterized by a single parameter I, uniquely determined after one has specified H_0 at $\phi=\phi_0$. For example, for the exponential potential, I can be taken to be ϕ_m. As in Fig. 1 we must always have $H(\phi)>H_{\rm SR}$ and also have $H(\phi)<\sqrt{3/2}H_{\rm SR}$ for inflation to occur. Two different solutions $H(\phi,I_1)$ and $H(\phi,I_2)$ will approach each other exponentially rapidly, at least if I_1 and I_2 are close to one another. Letting $\Delta I=I_2-I_1$, we have, on comoving (constant ϕ) hypersurfaces, $H(\phi,I_2)-H(\phi,I_1)\approx(\partial H/\partial I)_\phi\Delta I$ for linear perturbations. Taking the derivative of the SHJE with respect to I at fixed ϕ, we have

$$\begin{aligned}\left[\frac{\partial\ln(\partial H/\partial I)_\phi}{\partial\phi}\right]_I&=\frac{12\pi}{m_{\mathcal{P}}^2}\frac{H}{(\partial H/\partial\phi)_I} \\ &=-3\left[\frac{\partial\alpha}{\partial\phi}\right]_I .\end{aligned}$$

In deriving the last equality, we have applied (2.14a) and (2.14b). The solution of this equation may be written as

$$\begin{aligned}\alpha(\phi,I)-\alpha(\phi_0,I)=&-\tfrac{1}{3}\ln|(\partial H/\partial I)_\phi(\phi,I)| \\ &+\tfrac{1}{3}\ln|(\partial H/\partial I)_\phi(\phi_0,I)|\ .\end{aligned}$$

[With $I=\phi_m$ that is what was used to obtain Eq. (3.4c) for the exponential potential.] The conclusion

$$H(\phi,I_2)-H(\phi,I_1)\approx\left[\frac{\partial H}{\partial I}\right]_\phi\Delta I\propto\exp[-3\alpha(\phi,I_1)]\Delta I\ , \tag{3.5}$$

is that all solutions rapidly approach one another in the inflationary regime. The transient $\propto\exp(-3\alpha)$ corresponds to the decaying mode which always appears in cosmological perturbation theory.[22] In this sense, during inflation, the solution of the separated Hamilton-Jacobi equation is unique up to small perturbations. [This result is not valid generally for multiple fields because of the multivalued nature of $H(\phi_k)$ as is shown in Sec. III E.]

Given a solution of the SHJE, we now wish to integrate the trajectory equations [Eqs. (2.14a) and (2.14b)]. Since $H\equiv H(\phi)$, ϕ is the simplest choice for our time coordinate if there is a single inflaton field. With the lapse function

$$N(\phi,x^j)=-\left[\frac{4\pi}{m_{\mathcal{P}}^2}\right]\Big/\left[\frac{\partial H}{\partial\phi}\right]_x$$

given by Eq. (2.14a), substituted into Eq. (2.14b), the trajectories $\alpha(\phi,x^j)$ can be integrated, yielding

$$\alpha(\phi,x^j)-\alpha(\phi_0,x^j)=b(\phi,I)-b(\phi_0,I)\ , \tag{3.6}$$

where the function

$$b(\phi,I)\equiv-\frac{4\pi}{m_{\mathcal{P}}^2}\int_{\phi_0}^{\phi}\left[H(\phi',I)\Big/\frac{\partial H(\phi',I)}{\partial\phi'}\right]d\phi' \tag{3.7}$$

is independent of spatial coordinates, but, of course, depends upon the parameter I needed to specify the specific $H(\phi,I)$. As a consequence of the momentum constraint equation, I is a global constant independent of space and time. (This is not necessarily true for multiple fields.[23]) Consider a fiducial spatial point x_0^j. The (nonlinear) metric fluctuation on constant ϕ surfaces taken relative to this point is

$$\Delta\alpha(\phi,x^j)\equiv\alpha(\phi,x^j)-\alpha(\phi,x_0^j)\ , \tag{3.8}$$

which is constant in time:

$$\Delta\alpha(\phi,x^j)=\Delta\alpha(\phi_0,x^j)\ . \tag{3.9}$$

Bardeen, Steinhardt, and Turner[14] introduced an extremely valuable gauge-invariant variable ζ defined in linear perturbation theory. The definition which is appropriate for connection with our work here is that $\zeta/3$ is the fluctuation in α on uniform Hubble parameter hypersurfaces:

$$\zeta(H,x^j)\equiv3[\alpha(H,x^j)-\bar{\alpha}]\ , \tag{3.10}$$

where $\bar{\alpha}$ is a suitable spatial average of α. [The best choice seems to be $3\bar{\alpha}=\ln\langle\exp(3\alpha)\rangle_V$, with the average over a comoving volume V.] Since for one field $H=$const surfaces coincide with $\phi=$const surfaces, the quantity $3\Delta\alpha(\phi,x^j)$ is the nonlinear generation of ζ. The importance (and beauty) of ζ for a single scalar field is that it remains constant outside of the horizon even as the field passes from a vacuum-energy-dominated regime through an oscillating regime about its potential minimum, even though Π^ϕ can vanish and $d\alpha/d\phi$ can become singular. Equation (3.9) expresses the conservation of ζ in the nonlinear framework. Once we have set the initial value $\Delta\alpha(\phi_0,x^j)$, we have the complete solution.

In SBB we showed that $\Delta\alpha(\phi_0)$ is constrained to be very small in popular models of galaxy formation. For example, in the adiabatic scale invariant cold-dark-matter model, $\Delta\alpha(\phi_0)\approx(1-3)\times10^{-5}$, to explain the two-point correlation function of galaxies. Determining the amplitude of the initial fluctuations $\Delta\alpha(\phi_0,x^j)$ in inflation models can be a very complicated problem, even if we linearize in the fluctuations occurring inside the horizon. In this paper we adopt a crude model for the fluctuations which contains the essence of the more detailed treatment of SB2. In stochastic inflation the long-wavelength fields are modified by the action of short-wavelength fluctuations as they cross the horizon, become time coherent, and contribute to the amplitude of the long-wavelength fields. This initial value is set on a surface of constant Ha. Because H is changing slowly during inflation compared with the changes in the expansion factor, the Ha=const surfaces are approximately those of constant α. In this paper we take the fluctuation $(\delta\phi)_\alpha(\alpha_0,x)=\phi(\alpha_0,x)-\phi(\alpha_0,x_0)$ as a function of position relative to our fiducial position x_0 to be given at time α_0. The stochastic addition to the scalar-field fluctuation at time α_0 when the Hubble parameter is H_0 is proportional to the Hawking temperature $H_0/(2\pi)$ at that time, $(\delta\phi)_\alpha(\alpha_0,x)\propto H_0/(2\pi)$; the stochastic addition to the metric approximately vanishes (except for the small differences between Ha and H). A useful conceptual picture is to think of the background fields as receiving a series of impulsive kicks from the fluctuating stochastic forces, while in between kicks they evolve according to the background equations (2.13) and (2.14). For simplicity we here assume there is one stochastic impulse at α_0 and follow the subsequent evolution. In SB2 we present a more general formalism for dealing with the fluctuations in which stochastic and drift forces are treated together and demonstrate that the approximations made here are reasonable ones provided the fields are evolving slowly. We also demonstrate that although stochastic kicks may change I, it remains spatially independent.

We therefore need to connect the spatial fluctuation $(\delta\phi)_\alpha(\alpha_0,x)$ in ϕ on a constant α hypersurface to one in α, $(\delta\alpha)_\phi(\phi_0,x)$ on a constant ϕ surface, so that we can make use of the solution (3.9). A geometric representation of the transition from a constant α surface to one of constant ϕ is shown in Fig. 2 for the case of an exponential point. The solid and short-dashed diagonal lines represent trajectories. (The long-dashed are lines of constant phase; see Sec. IV.) Given $(\delta\phi)_\alpha$, a fluctuation spread over the thick horizontal line, one must evolve each spatial point independently to obtain $(\delta\alpha)_\phi$ to obtain the spread over the thick vertical line. Thus the metric fluctuation on constant ϕ surfaces is

$$\Delta\alpha(\phi,x^j)=(\delta\alpha)_\phi(\phi_0,x^j)$$
$$=+\frac{4\pi}{m_{\mathcal{P}}^2}\int_{\phi_0}^{\phi_0+(\delta\phi)_\alpha(\alpha_0,x^j)}\left[H(\phi')\Big/\frac{\partial H(\phi')}{\partial\phi'}\right]d\phi' . \tag{3.11}$$

When Eq. (3.11) is linearized and Fourier transformed, it describes the mode-by-mode evolution of the amplitude in linear perturbation theory. In that case the interpretation of $(\delta\phi)_\alpha$ is as Gaussian random perturbation with power spectrum $\mathcal{P}_\phi=k^3/(2\pi^2)|\delta\phi(k)|^2$, where $\delta\phi(k)$ is the Fourier amplitude, and α_0 is the time at which the wave number k equals Ha. (See SBB.) For general potentials and small fluctuations $(\delta\phi)_\alpha\approx H_0/(2\pi)$, we recover the usual result[11-14]

$$\Delta\alpha(\phi)=2\left[\frac{H^2/m_{\mathcal{P}}^2}{\partial H/\partial\phi}\right]_0=-\left[\frac{H^2}{2\pi\dot\phi/N}\right]_0 . \tag{3.12}$$

In particular, this is usually used with the slow-roll approximation $3H\dot\phi/N=-\partial V/\partial\phi$.

For the exponential potential, an exact result can be given. For example, for the attractor (steady-state) solution $b(\phi)=(4\pi p)^{1/2}(\phi-\phi_0)/m_{\mathcal{P}}$, we have

$$\Delta\alpha(\phi)=-\left[\frac{p}{\pi}\right]^{1/2}\frac{H_0}{m_{\mathcal{P}}} , \tag{3.13}$$

if we take $(\delta\phi)_\alpha\approx H_0/(2\pi)$. Thus, to have $p\sim10$, one would require that $H_0/m_{\mathcal{P}}\approx10^{-5}$ to agree with observations of cosmic structures.[7]

C. Multiple fields and the integration of trajectories

The separated Hamilton-Jacobi equation (2.13b) is a first-order nonlinear partial differential equation for n scalar fields whose complete solution depends on n constant parameters I_j. The general solution of the cosmological background-field equations follows from this, as in the classical mechanics analog.[24] In this subsection we use $\alpha=\ln(a)$ as the natural choice for the time parameter since in general for many fields there will be no preferred ϕ_k hypersurfaces.

Using a proof similar to that leading to Eq. (3.5), we now show that the solutions $H\equiv H(\phi_k,I_k)$ of (2.13b) with differing I_k also approach each other rapidly. Differentiating (2.13b) with respect to I_k gives

$$H=\frac{m_{\mathcal{P}}^2}{12\pi}\sum_c\frac{\partial H}{\partial\phi_c}\frac{\partial}{\partial\phi_c}\ln\left|\frac{\partial H}{\partial I_k}\right|$$
$$=-\frac{1}{3N}\frac{\partial}{\partial t}\ln\left|\frac{\partial H}{\partial I_k}\right| .$$

The time derivative $\partial/\partial t$ is evaluated along a physical trajectory $\dot\phi_j/N=-[m_{\mathcal{P}}^2/(4\pi)](\partial H/\partial\phi_j)$. This equation may be integrated exactly with the help of (2.14b):

$$e^{-3\alpha}J_k=\frac{m_{\mathcal{P}}^2}{4\pi}\left[\frac{\partial H}{\partial I_k}\right]_\phi , \tag{3.14}$$

where the J_k are integration constants independent of time. Imitating (2.14a), we have introduced the numerical factor $m_{\mathcal{P}}^2/(4\pi)$. Thus the difference in H for different I's decreases as $e^{-3\alpha}$, as in Sec. III B. Solving (3.14) for $\phi_j = \phi_j(I_k, J_k, \alpha)$ gives the trajectories as functions of α. This is the complete solution: Given arbitrary initial conditions at $\alpha = \alpha_0$, one may choose the parameters I_k, J_k so that

$$e^{-3\alpha_0} J_k = \frac{m_{\mathcal{P}}^2}{4\pi} \frac{\partial H(\phi_k(\alpha_0), I_k)}{\partial I_k}$$

and

$$\Pi^{\phi_k}(\phi_k(\alpha_0)) = -\frac{m_{\mathcal{P}}^2}{4\pi} \frac{\partial H(\phi_k(\alpha_0), I_k)}{\partial \phi_k}$$

are satisfied. More generally one may view the Hubble function as the generator of a canonical transformation from (ϕ_k, Π^{ϕ_k}) to the new canonical coordinates (I_k, J_k), which are constant parameters. In most cases it is difficult to find an n-parameter solution of the separated Hamilton-Jacobi equation. An explicit solution for $n=1$ was given in Sec. III A, where α was integrated as a function of ϕ by differentiation of ϕ_m through (3.14). Analytic cases for many fields including gravitational radiation are given elsewhere.[23]

The new canonical momenta J_k are functions of position, but not of time. For one scalar field, $J(x)/J(x_0)$ is just the conserved quantity $\exp(3\Delta\alpha)$ [Eq. (3.9)], i.e., $\exp(\zeta)$, where x_0 is the fiducial point relative to which spatial fluctuations are measured. With many fields there are n constants $J_k(x)/J_k(x_0)$ which describe the nonlinear state of the system just as the single constant ζ does in the single-inflaton case. In the many-field case, the nonlinear version of ζ, defined by (3.10) in terms of uniform Hubble parameter surfaces, is still a useful concept, but it approaches a constant only when one of the fields dominates the energy density of the Universe. Hence the $\{J_k\}$ provide a more powerful description than the single variable ζ. (In general, the Hamilton-Jacobi equation for multiple fields is valid point by point, and the parameters I_k may also depend upon x. The consequences of this point are explored by Salopek.[23])

Although we showed in Sec. II B that the choice of time surface is essentially arbitrary in Eq. (2.14), the variable α is the best motivated choice since it does not appear in the SHJE, which leads to the natural expression of the trajectories as functions of α according to Eq. (3.14). It is also monotonically increasing in inflationary models and, hence, is a viable clock at all times unlike the scalar fields. (For a more general discussion of the problem of time in general relativity, see Unruh and Wald.[28])

D. Numerical and approximate solutions of the separated Hamilton-Jacobi equation

Since the Hamilton-Jacobi equations are analytically solvable only for special cases, it is worthwhile to develop approximate methods, the most important of which is the slow-roll approximation[29,30] in which the momentum terms are ignored in the Hubble function.

1. *Slow-roll approximation for a single scalar field*

For one field, neglecting the $\partial H/\partial\phi$ term in the SHJE gives the zeroth-order approximation

$$H_{(0)}^2 = H_{SR}^2 \equiv \frac{8\pi}{3m_{\mathcal{P}}^2} V(\phi) \ . \quad (3.15a)$$

Although higher accuracy may be obtained by substituting $H_{(0)}$ into the right side of the SHJE,

$$H_{(1)}^2 = \frac{8\pi}{3m_{\mathcal{P}}^2} V(\phi) \left[1 + \frac{m_{\mathcal{P}}^2}{48\pi} \left[\frac{\partial \ln V}{\partial \phi} \right]^2 \right] , \quad (3.15b)$$

higher-order terms in this series only slowly improve the accuracy over the zeroth-order approximation. We have tried a number of other expansions, but none have proved generally useful.

As a concrete example, we use the quadratic potential $V(\phi) = m^2\phi^2/2$. Even for this simple V, the SHJE cannot be solved analytically since, as we have seen in Sec. III A, $H(\phi)$ is singular whenever ϕ changes direction. In Fig. 3 the solid curve shows a numerical integration of the background scalar-field equations of motion (2.15), beginning at $\phi = 10m_{\mathcal{P}}$, well into the slow-roll regime. Note the number of cusps in H in this case, associated with the oscillations near the bottom of the potential well. The other curves show the zeroth- and first-order approximations:

$$\frac{H_{(0)}}{m} = \left[\frac{4\pi}{3} \right]^{1/2} \frac{\phi}{m_{\mathcal{P}}} ,$$
$$\frac{H_{(1)}}{m} = \frac{1}{3} \left[1 + 12\pi \frac{\phi^2}{m_{\mathcal{P}}^2} \right]^{1/2} . \quad (3.16)$$

A nice feature of this potential is that the solution scales in m according to $H = mf(\phi)$; one only needs a solution for a single m to obtain all other solutions that begin in the slow-roll regime. As mentioned in Sec. III B, $H_{(0)} = H_{SR}$ defines the envelope for allowed solutions. Figure 3 shows that $H_{(0)}$ is accurate for $\phi > 1m_{\mathcal{P}}$ and $H_{(1)}$ is accurate for $\phi > 0.3m_{\mathcal{P}}$. Higher-order approximations than $H_{(1)}$ do not aid matters much.

To estimate the degree of nonlinearity produced in models with a power-law potential $V(\phi) = V_0\phi^n$, we consider the nonlinear response of the background fields to one e-folding's stochastic impulse $(\delta\phi)_\alpha$ applied 60 e-foldings prior to the end of inflation, corresponding to the scale of the current Hubble length today. Using H_{SR} in Eq. (3.11) for the metric fluctuations on constant ϕ surfaces, we have

$$\Delta\alpha(\phi, x) = \frac{1}{n} \frac{4\pi}{m_{\mathcal{P}}^2} [2\phi_0(\delta\phi)_\alpha + (\delta\phi)_\alpha^2] \ , \quad (3.17)$$

which is constant in time and depends only on the initial value of the scalar field ϕ_0. The relation between the number of e-foldings from the hypersurface ϕ_0 to the hypersurface ϕ at a given position is the solution of (3.6):

$$\alpha(\phi, x^j) - \alpha(\phi_0, x^j) = \frac{1}{n} \frac{4\pi}{m_{\mathcal{P}}^2} (\phi_0^2 - \phi^2) \ . \quad (3.18)$$

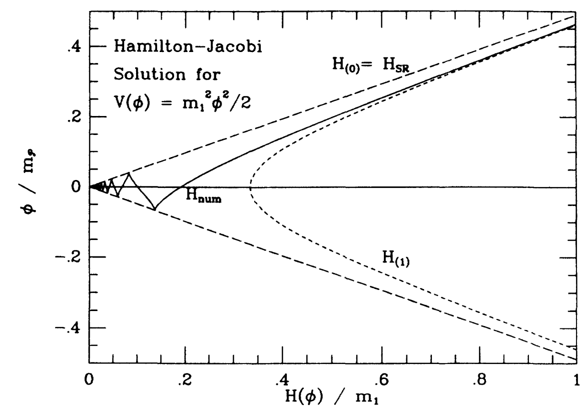


FIG. 3. Solid curve shows a numerical solution $H_{\rm num}$ of the Hamilton-Jacobi equation for a quadratic potential. At large values the scalar field follows the slow-roll solution $H_{\rm SR}$ (long-dashed curve) and, even more closely, the next-order improvement $H_{(1)}$ [Eq. (3.15b)] (short-dashed curve), and then begins to oscillate when $\phi \approx 0.35 m_{\mathcal{P}}$. As in the solution of Fig. 1, there is cusp behavior when the scalar field changes direction at zero field momentum and thus along the $H_{\rm SR}$ curve.

For $n=2$, $\phi_0 \approx 3.1 m_{\mathcal{P}}$ gives 60 e-foldings of expansion. Recall that the rms value for the fluctuations crossing the horizon in one e-folding is $(\delta\phi)_\alpha \approx H_0/(2\pi)$. Thus the ratio of the quadratic nonlinear term in $(\delta\phi)_\alpha$ to the linear one is $(12\pi)^{-1/2} m/m_{\mathcal{P}}$. We require $m \approx 5\times 10^{-7} m_{\mathcal{P}}$ to give fluctuations $\Delta\alpha$ at the observational level $(\sim 2\times 10^{-5})$; hence, as expected, the corrections to linearity are very small indeed, a conclusion which remains valid even with the inclusion of the fluctuations from all 60 e-foldings.[10]

For a quartic potential $V(\phi)=\lambda\phi^4/4$, $\phi_0 \approx 4.4 m_{\mathcal{P}}$ is required to get 60 e-foldings and $\lambda \approx 5\times 10^{-14}$ is required to get the $\Delta\alpha$ amplitude at the observationally inferred value. The quadratic correction is then at the 10^{-7} level for one e-folding's worth of stochastic impulse.

2. *Slow-roll approximation for multiple scalar fields*

The slow-roll approximation also gives an accurate treatment of fluctuations when there is more than one scalar field for a large class of potentials $V(\phi_k)$. The great advantage of this approximation is that it does not depend upon any constants I_k and can quickly and elegantly yield quantitative results with much more simplicity than, e.g., our previous linear perturbation theory calculations.[7] We illustrate this for the double inflation potential $V(\phi_k)=V_1(\phi_1)+V_2(\phi_2)$, whose $H_{\rm SR}$ is

$$H_{\rm SR} = \left[\frac{8\pi}{3m_{\mathcal{P}}^2} [V_1(\phi_1)+V_2(\phi_2)] \right]^{1/2} . \tag{3.19a}$$

The exact solution $H^2(\phi_k)$ does not maintain the separability of $H_{\rm SR}^2$, which can be seen at the level of the next-order approximation. The equations of motion derived using (2.14a) are the usual slow-roll ones:

$$\frac{\dot\phi_1}{N} = -\frac{m_{\mathcal{P}}^2}{4\pi}\frac{\partial H_{\rm SR}}{\partial\phi_1} = -\frac{\partial V_1/\partial\phi_1}{3H_{\rm SR}} , \tag{3.19b}$$

$$\frac{\dot\phi_2}{N} = -\frac{m_{\mathcal{P}}^2}{4\pi}\frac{\partial H_{\rm SR}}{\partial\phi_2} = -\frac{\partial V_2/\partial\phi_2}{3H_{\rm SR}} . \tag{3.19c}$$

Assume that the field ϕ_2 is the one that dominates the energy density of the Universe at late times. It is convenient to take it as the time coordinate, and so the lapse is

$$N = -\frac{3H_{\rm SR}}{\partial V_2/\partial\phi_2} \tag{3.20}$$

and the remaining equation of motion is

$$\frac{d\phi_1}{d\phi_2} = \frac{\partial V_1/\partial\phi_1}{\partial V_2/\partial\phi_2} , \tag{3.21}$$

which is solved by two independent integrations:

$$\int_{\phi_{10}}^{\phi_1} \frac{d\phi_1}{\partial V_1/\partial\phi_1} = \int_{\phi_{20}}^{\phi_2} \frac{d\phi_2}{\partial V_2/\partial\phi_2} . \tag{3.22}$$

To obtain the metric fluctuations, we integrate (2.14b):

$$\alpha(\phi_2, x^j) - \alpha(\phi_{20}, x^j) = b(\phi_k) - b(\phi_{k0}) , \tag{3.23a}$$

where

$$b(\phi_k) = -\frac{8\pi}{m_{\mathcal{P}}^2}\int_{\phi_{10}}^{\phi_1} \frac{V_1(\phi_1')d\phi_1'}{(\partial V_1/\partial\phi_1')(\phi_1')} - \frac{8\pi}{m_{\mathcal{P}}^2}\int_{\phi_{20}}^{\phi_2} \frac{V_2(\phi_2')d\phi_2'}{(\partial V_2/\partial\phi_2')(\phi_2')} . \tag{3.23b}$$

The total expansion factor of the Universe is therefore the product of the scale factors for two independent inflationary epochs, a result derived by Starobinski.[31] To generally obtain the metric fluctuations at some late time ϕ_{2e}, we first form $\alpha(\phi_{2e}, x) - \alpha_0$, with the ϕ_2 integration evaluated between $\phi_2(\alpha_0, x)$ and ϕ_{2e}, and the ϕ_1 integration evaluated between $\phi_1(\alpha_0, x)$ and $\phi_{1e}(x)$. We have defined $\phi_{1e}(x) \equiv \phi_1(\phi_{2e}, x)$. In Sec. III B we adopted the notation $(\delta\phi_k)\alpha(x) \equiv \phi_k(\alpha_0, x) - \phi_{k0}$, for the fluctuations from the stochastic impulse at α_0, and took it to vanish at our fiducial point x_0. The fluctuation between the values at x and our fiducial point x_0 is then

$$\Delta\alpha(\phi_{2e}, x) \equiv \alpha(\phi_{2e}, x) - \alpha(\phi_{2e}, x_0)$$

$$= -\frac{8\pi}{m_{\mathcal{P}}^2}\int_{\phi_{20}+(\delta\phi_2)_\alpha(x)}^{\phi_{20}+(\delta\phi_2)_\alpha(x_0)} \frac{V_2(\phi_2')d\phi_2'}{(\partial V_2/\partial\phi_2')(\phi_2')} - \frac{8\pi}{m_{\mathcal{P}}^2}\int_{\phi_{10}+(\delta\phi_1)_\alpha(x)}^{\phi_{10}+(\delta\phi_1)_\alpha(x_0)} \frac{V_1(\phi_1')d\phi_1'}{(\partial V_1/\partial\phi_1')(\phi_1')} - \frac{8\pi}{m_{\mathcal{P}}^2}\int_{\phi_{1e}(x_0)}^{\phi_{1e}(x)} \frac{V_1(\phi_1')d\phi_1'}{(\partial V_1/\partial\phi_1')(\phi_1')} . \tag{3.24}$$

Recall that we have adopted the convention that $(\delta\phi_2)_\alpha(x_0)=(\delta\phi_1)_\alpha(x_0)=0$ for the fiducial point at x_0. To find the last term in (3.24), we would have to make use of Eq. (3.22):

$$\int_{\phi_{10}+(\delta\phi_1)_\alpha(x)}^{\phi_{1e}(x)} \frac{d\phi_1}{\partial V_1/\partial\phi_1}=\int_{\phi_{20}+(\delta\phi_2)_\alpha(x)}^{\phi_{2e}} \frac{d\phi_2}{\partial V_2/\partial\phi_2}\ , \tag{3.25}$$

at x and x_0, which can get messy. However, in most cases of interest the first field will have settled into the trough of its potential and damped away, so that ϕ_{1e} will be spatially independent, and the last integral in (3.24) vanishes. In this case the result is very simple, just the sum of the fluctuations form the two phases of inflation.[31] If the stochastic impulse occurs in the first phase, both $(\delta\phi_1)_\alpha$ and $(\delta\phi_2)_\alpha$ contribute fluctuations of amplitude $H_{\rm SR}/(2\pi)$, giving a large amplitude response. If the impulse occurs in the second phase, after ϕ_1 has settled down to its potential minimum, then $(\delta\phi_1)_\alpha$ vanishes, and only ϕ_2 fluctuations contribute, but at a diminished level since $H_{\rm SR}/(2\pi)$ has dropped so much.

If we are explicitly interested in the time development of $\Delta\alpha$, we would need to evaluate (3.25), which is messy for all but the simplest potentials. To illustrate the main features, we adopt quadratic potentials $V_1(\phi_1)=m_1^2\phi_1^2/2$ and $V_2(\phi_2)=m_2^2\phi_2^2/2$. In Fig. 4 we show the structure of the trajectories (solid lines) and lines of constant Hubble parameter (short-dashed lines) for $m_1^2=2m_2^2$. Trajectories begin near the top of the plot where the energy density of the Universe is dominated by ϕ_1. They then descend toward the origin, and as they cross the long-dashed line, the Universe becomes ϕ_2 dominated. When the Hubble decreases below m_1, shown as the dark dashed curve, the slow-roll approximation breaks down because ϕ_1 begins to oscillate. Integration of (3.25) yields

$$\phi_{1e}(x)=[\phi_{10}+(\delta\phi_1)_\alpha(x)]\left[\frac{\phi_{2e}}{[\phi_{20}+(\delta\phi_2)_\alpha(x)]}\right]^{m_1^2/m_2^2}\ , \tag{3.26}$$

and hence the fully time-dependent amplitude of the metric fluctuations as a function of ϕ_{2e} is

$$\Delta\alpha(\phi_{2e},x)=\frac{2\pi}{m_P^2}[2\phi_{20}(\delta\phi_2)_\alpha+(\delta\phi_2)_\alpha^2]+\frac{2\pi}{m_P^2}[2\phi_{10}(\delta\phi_1)_\alpha+(\delta\phi_1)_\alpha^2]$$
$$-\frac{2\pi}{m_P^2}\left[\frac{\phi_{2e}}{\phi_{20}}\right]^{2m_1^2/m_2^2}\left[[\phi_{10}+(\delta\phi_1)_\alpha]^2\left[\frac{1}{[1+(\delta\phi_2)_\alpha/\phi_{20}]}\right]^{2m_1^2/m_2^2}-\phi_{10}^2\right]\ . \tag{3.27}$$

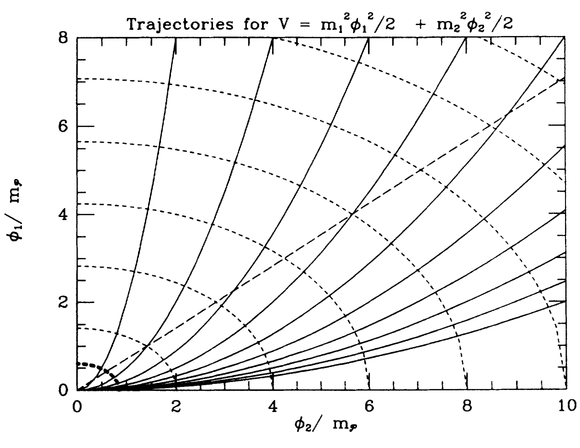


FIG. 4. Surfaces of constant Hubble parameter calculated using the slow-rollover approximation (short-dashed curves) are shown for the two-dimensional potential $V(\phi_1,\phi_2)=m_1^2\phi_1^2/2+m_2^2\phi_2^2/2$, with $m_1^2=2m_2^2$. The trajectories (solid curves) fall towards the origin and are orthogonal to constant Hubble parameter curves. Above the long-dashed curve, ϕ_1 dominates the energy density of the Universe, whereas below it ϕ_2 dominates. This approximate solution breaks down when the Hubble parameter decreases below m_1 (heavy short-dashed curve) because ϕ_1 begins to oscillate (see Fig. 5). One may use these results to analytically calculate primordial fluctuation spectra arising from double-inflation models (Sec III D 2).

The last term rapidly decays for $\phi_{2e}\ll\phi_{20}$, giving an asymptotic formula for the fluctuations which is the sum of two single-field [Eq. (3.17)] contributions. These semianalytic results give a simple qualitative understanding of the time evolution of ζ revealed in the detailed numerical calculations of SBB as well as providing a rapid method for quantitative calculations.

E. Separable example for two scalar fields

In this subsection we use a two-field analytic solution of the SHJE to illustrate the complexities that can arise in practice. In particular, we wish to address the behavior of the Hubble function when one of the scalar fields is oscillating in its trough, a point we could not deal with in Sec. III D 2. In this case we find that at a given point in the $\phi_1-\phi_2$ plane, the Hubble parameter is multivalued, which makes the Hamilton-Jacobi approach cumbersome to use for oscillating fields.

In classical mechanics the Hamilton-Jacobi equation is not particularly useful unless there are separable solutions. In this cosmological setting we would like to find a separable solution which has a potential with a trough that the field settles into. The potential

$$V(\phi_1,\phi_2)=\exp\left[-\left[\frac{16\pi}{p}\right]^{1/2}\frac{\phi_2}{m_P}\right]U(\phi_1) \tag{3.28}$$

has this property. A class of solutions to the SHJE can be obtained by assuming H is separable:

$$H(\phi)=\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi_2}{m_{\mathcal{P}}}\right]\tilde{H}(\phi_1) . \tag{3.29}$$

The two-dimensional SHJE is satisfied provided $\tilde{H}$ obeys

$$\tilde{H}^2=\frac{m_{\mathcal{P}}^2}{12\pi}\left[1-\frac{1}{3p}\right]^{-1}\left[\frac{\partial\tilde{H}}{\partial\phi_1}\right]^2 +\frac{8\pi}{3m_{\mathcal{P}}^2}\frac{U(\phi_1)}{1-1/(3p)} . \tag{3.30}$$

This looks like a one-dimensional SHJE equation, a similarity we can make more explicit by rewriting (3.30) as

$$\tilde{H}^2=\frac{m_{\mathcal{P}}^2}{12\pi}\left[\frac{\partial\tilde{H}}{\partial\tilde{\phi}_1}\right]^2+\frac{8\pi}{3m_{\mathcal{P}}^2}\tilde{V}(\phi_1) , \tag{3.31a}$$

in terms of the variables

$$\tilde{\phi}_1=\sqrt{1-1/(3p)}\,\phi_1 ,\qquad \tilde{V}(\tilde{\phi}_1)=U\left[\frac{\tilde{\phi}_1}{\sqrt{1-1/(3p)}}\right]\Big/\left[1-\frac{1}{3p}\right] . \tag{3.31b}$$

If we can solve for the single scalar-field trajectories $\tilde{\phi}_1(\tilde{\alpha})$ and $\tilde{H}(\tilde{\alpha})$ associated with Eq. (3.31), then we can solve the two-dimensional problem as a function of $\tilde{\alpha}$. The equations of motion become

$$\frac{\dot{\phi}_1}{N}=-\frac{m_{\mathcal{P}}^2}{4\pi}\sqrt{1-1/(3p)} \times\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi_2}{m_{\mathcal{P}}}\right]\frac{\partial\tilde{H}}{\partial\tilde{\phi}_1} , \tag{3.32a}$$

$$\frac{\dot{\phi}_2}{N}=\frac{m_{\mathcal{P}}}{\sqrt{4\pi p}}\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi_2}{m_{\mathcal{P}}}\right]\tilde{H} , \tag{3.32b}$$

$$\frac{\dot{\alpha}}{N}=\exp\left[-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi_2}{m_{\mathcal{P}}}\right]\tilde{H} . \tag{3.32c}$$

It is advantageous to choose α as time so that $\tilde{H}$ disappears from the ϕ_2 evolution equation, giving

$$\phi_2(\alpha)-\phi_2(\alpha_0)=\frac{m_{\mathcal{P}}}{\sqrt{4\pi p}}(\alpha-\alpha_0) . \tag{3.33}$$

If we reinterpret (3.33) to be an expression for $\alpha(\phi_2)$, we see that in the ϕ_2 time variable the remarkable feature that the evolution of the metric is independent of the first scalar field, whether it is inflating or oscillating, although the initial value $\alpha(\phi_{20})$ will depend on the initial value $\phi_1(\phi_{20})$, just as in double inflation (SBB).

By comparing $d\phi_1/d\alpha$ with $d\tilde{\phi}_1/d\tilde{\alpha}$, we see that α is linear in $\tilde{\alpha}$:

$$\tilde{\alpha}=[1-1/(3p)]\alpha . \tag{3.34}$$

Therefore, the ϕ_1 trajectory as a function of α is given in terms of the parametrized one-dimensional trajectory $\tilde{\phi}_1(\tilde{\alpha})$ by

$$\phi_1(\alpha)=[1-1/(3p)]^{-1/2}\tilde{\phi}_1([1-1/(3p)]\alpha) . \tag{3.35}$$

As a concrete example, consider $U=m_1^2\phi_1^2/2$. In general, orbitals will cross as shown in Fig. 5, where we have taken $p=1.01$ (and the magnitude of m_1 is unimportant for this illustration). Given one trajectory, one may obtain all others by translating in the ϕ_2 direction; hence $H(\phi_1,\phi_2)$ will depend upon a continuous index, and a given point (ϕ_1,ϕ_2) may take on many values for the Hubble parameter. The lines of uniform Hubble parameter H shown in Fig. 5 are given by (3.29):

$$\phi_2/m_{\mathcal{P}}=(p/4\pi)^{1/2}\ln[\tilde{H}(\sqrt{1-1/(3p)}\,\phi_1)/H] . \tag{3.36}$$

These curves each have cusps which would confuse the figure, and so only small segments of the uniform Hubble lines are shown.

IV. HAMILTON-JACOBI EQUATIONS FOR INFLATING COSMOLOGIES

In this section we demonstrate explicitly how the separated Hamilton-Jacobi (HJ) equation [Eq. (2.13b)] may in fact be derived from the Einstein-Hamilton-Jacobi equation and the functional momentum constraint for inhomogeneous long-wavelength fields. We discuss the limitations inherent in following a truncated form of the full equations. We also relate the form of the probability functional for ensembles of universes on various hypersurfaces which plays an important role in stochastic inflation.

A. Overview of Hamilton-Jacobi theory in general relativity

Although Misner, Thorne, and Wheeler,[21] for example, extol the virtues of the Einstein-Hamilton-Jacobi equation as providing the fastest route to quantum theory, it

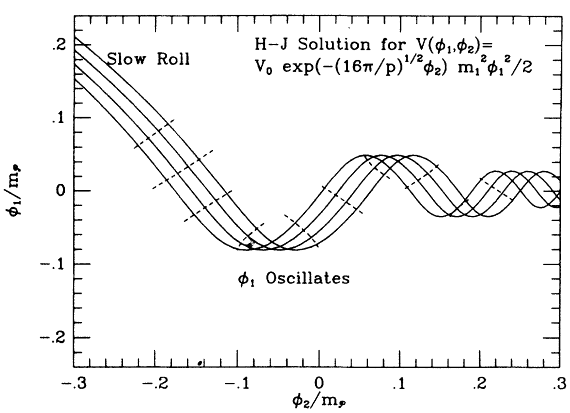


FIG. 5. Separable Hamilton-Jacobi solution is shown for two scalar fields rolling down a trough in the potential. Shown are surfaces of constant H (short-dashed) and their orthogonal trajectories (solid) for a potential which is a product of an exponential in ϕ_2 and a term quadratic in ϕ_1 [Eq. (3.28)]. The trajectories start in the slow-rollover regime, but as ϕ_2 increases, ϕ_1 begins to oscillate and damp. The trajectories eventually cross; hence the Hubble function may take on many values at the same point in the ϕ_1-ϕ_2 plane.

has not been exploited in the study of inhomogeneous systems. Let us recall the basic framework. The action $\mathcal{I}$ of Sec. II A in the ADM formulation is computed only for those $\gamma_{ij}(t,x^j)$ and $\phi_k(t,x^j)$ histories which extremize it. In the integration the lower end point is taken to be on a hypersurface with three-geometry specified by γ_{ij}^I upon which the initial scalar-field values ϕ_k^I are specified. This is considered fixed. The upper end point of the integration is taken to be a hypersurface with three-geometry that of $\gamma_{ij}(x^j)$ with scalar-field values $\phi_k(x^j)$. Hamilton's principal functional $S[\gamma_{ij}(x^j),\phi_k(x^j)]$ is just I computed this way, which can also be written as

$$S[\gamma_{ij}(x^j),\phi_k(x^j)]=\int_{\gamma_{ij}^I,\phi_k^I}^{\gamma_{ij},\phi_k}\left[\pi^{\gamma_{ij}}\frac{\partial\gamma_{ij}}{\partial t}+\pi^{\phi_k}\frac{\partial\phi_k}{\partial t}\right]d^4x\ . \tag{4.1}$$

The momenta conjugate to γ_{ij} and ϕ_k are

$$\pi^{\gamma_{ij}}(x)\equiv\frac{\delta S}{\delta\gamma_{ij}(x)}=\frac{m_P^2}{16\pi}\gamma^{1/2}(\gamma^{ij}K-K^{ij})\ , \tag{4.2a}$$

$$\pi^{\phi_k}(x)=\frac{\delta S}{\delta\phi_k(x)}=\gamma^{1/2}(\dot\phi_k-N^i\phi_{k,i})\ . \tag{4.2b}$$

In terms of these momenta, the energy and momentum constraints (2.2) and (2.3) are

$$0=\mathcal{H}(x)\equiv\frac{16\pi}{m_P^2}\gamma^{-1/2}\pi^{\gamma_{ij}}\pi^{\gamma_{ml}}\left[\gamma_{jm}\gamma_{il}-\frac{1}{2}\gamma_{ij}\gamma_{ml}\right]+\frac{1}{2}(\pi^{\phi_k})^2+\gamma^{1/2}V(\phi_k) \tag{4.3a}$$

$$+\left[-\frac{m_P^2}{16\pi}\gamma^{1/2}({}^{(3)}R)+\frac{1}{2}\gamma^{1/2}\gamma^{ij}\phi_{k,i}\phi_{k,j}\right]\ , \tag{4.3b}$$

$$0=\mathcal{H}_i(x)\equiv-2(\gamma_{il}\pi^{lm})_{,m}+\pi^{\gamma_{lm}}\gamma_{lm,i}+\pi^{\phi_k}\phi_{k,i}\ . \tag{4.4}$$

The Einstein-Hamilton-Jacobi equation is (4.2) with the momenta expressed in terms of the functional derivatives of $\mathcal{S}$ given in (4.2). One can interpret $\mathcal{S}$ as a generator of a canonical transformation in terms of which the new Hamiltonian functional vanishes strongly. In this approach the solutions of the HJ equation depend on an infinite number of parameters which are interpreted as new canonical coordinates.[23,32] The functional $\mathcal{S}$ can also be viewed as the phase of the wave function $\Psi[\gamma_{ij},\phi_k]$ that appears in quantum treatments of the gravitational field in the semiclassical approximation $\Psi=\mathcal{P}^{1/2}e^{i\mathcal{S}}$, $\mathcal{P}[\gamma_{ij},\phi_k]$ is the probability functional for the three-geometry $\gamma_{ij}(x)$ and the scalar-field configuration $\phi_k(x)$, which is of importance in the theory of fluctuations for cosmic structure, a connection we explore in Sec. IV D.

Although significant progress with the full nonlinear system does not seem feasible, a reasonably large subclass of problems in which the fields only contain long-wavelength contributions is tractable. The key approximation is to neglect the (4.3b) terms in the $\mathcal{H}=0$ equation. In the case with no scalar fields, Pilati[19] and Teitelboim[20] have considered this same situation, which they labeled the strong gravitational coupling limit $G\equiv m_P^{-2}\to\infty$. Their application was to quantum gravity (see Sec. V). In addition to incorporating matter fields, our long-wavelength limit interpretation has the flexibility of allowing for interactions with short-distance effects through the stochastic formalism. As a warmup, we remind the reader how HJ works in minisuperspace model before turning to the long-wavelength inhomogeneous model of interest to us.

B. Homogeneous and isotropic minisuperspace model

The simplest derivation of the SHJE from the full HJ equation occurs in a minisuperspace model of homogeneous scalar fields $\phi_k(t)$, evolving in a homogeneous universe with metric $ds^2=-N^2(t)dt^2+e^{2\alpha(t)}(dx^2+dy^2+dz^2)$ analogous to (2.10). The ADM action is

$$I=\int dt\ UNe^{3\alpha}\left[-\frac{3m_P^2}{8\pi}\frac{\dot\alpha^2}{N^2}+\frac{1}{2}\frac{\dot\phi_k^2}{N^2}-V(\phi_k)\right]\ , \tag{4.5}$$

where we denote $U=\int d^3x$ as the comoving volume of the Universe. The momenta are given by

$$p_\alpha=\frac{\partial S}{\partial\alpha}=-\frac{3m_P^2}{4\pi}\frac{Ue^{3\alpha}\dot\alpha}{N}\ ,\qquad p_{\phi_k}=\frac{\partial S}{\partial\phi_k}=\frac{Ue^{3\alpha}\dot\phi_k}{N}\ , \tag{4.6}$$

in terms of which the Hamiltonian is

$$H=-\frac{2\pi}{3m_P^2}U^{-1}e^{-3\alpha}p_\alpha^2+\frac{1}{2}U^{-1}e^{-3\alpha}p_{\phi_k}^2+Ue^{3\alpha}V(\phi_k)=0\ . \tag{4.7}$$

The Hamilton-Jacobi equation (HJE) is obtained by substituting the derivatives of Hamilton's principal function S for the momenta. The HJE is most transparently written in terms of the combination $\exp(-3\alpha)S$. If we assume this quantity is independent of α,

$$S=-\frac{m_P^2}{4\pi}Ue^{3\alpha}H(\phi_k)\ , \tag{4.8}$$

then H must satisfy the SHJE. Furthermore, if we substitute (4.8) into (4.6), we find that the scalar field velocity is just the gradient of the Hubble parameter, and we thus recover Eqs. (2.14). However, it is also clear that solutions to (4.7) exist for which $\exp(-3\alpha)S$ is not α independent. In the inhomogeneous case the momentum constraint prohibits these solutions. Lines of constant S are shown in Fig. 2 for the attractor solution (3.1) of Sec. III A.

C. Inhomogeneous long-wavelength fields

We now demonstrate that the solutions of the long-wavelength version of the HJE and momentum constraint equations lead to a separation of variables in the Hamilton principal function analogous to that in Eq. (4.8), point by point. In this section we restrict ourselves to a single scalar field.

In the long-wavelength approximation, we neglect the (4.3b) second-order gradient terms in the energy constraint functional $\mathcal{H}$. Motivated by the results of Sec. II B, we shall also assume that Hamilton's principal function $\mathcal{S}=\mathcal{S}[\alpha,\phi]$ is only a functional of

$$\alpha\equiv\tfrac{1}{6}\ln\gamma,\quad \gamma=\det(\gamma_{ij})\ , \tag{4.9a}$$

and of the scalar field. The gravitational momenta are then proportional to the three-metric,

$$\pi^{\gamma_{ij}}(x)=\frac{\delta\mathcal{S}}{\delta\gamma_{ij}(x)}=\frac{1}{6}\gamma^{ij}\frac{\delta\mathcal{S}}{\delta\alpha(x)}\ , \tag{4.9b}$$

which leads to the following constraint equations:

$$-\frac{2\pi}{3m_{\mathcal{P}}^2}e^{-3\alpha(x^j)}\left[\frac{\delta\mathcal{S}}{\delta\alpha(x^j)}\right]^2+\frac{1}{2}e^{-3\alpha(x^j)}\left[\frac{\delta\mathcal{S}}{\delta\phi(x^j)}\right]^2+e^{3\alpha(x^j)}V(\phi(x^j))=0\ , \tag{4.10a}$$

$$\frac{1}{3}\left[\frac{\delta\mathcal{S}}{\delta\alpha(x^j)}\right]_{,i}=\frac{\delta\mathcal{S}}{\delta\alpha(x^j)}\alpha_{,i}(x^j)+\frac{\delta\mathcal{S}}{\delta\phi(x^j)}\phi_{,i}(x^j)\ . \tag{4.10b}$$

The momentum constraint, which may be rewritten in the suggestive form

$$\frac{1}{3}(e^{-3\alpha}\pi^{\alpha})_{,i}=e^{-3\alpha}\pi^{\phi}\phi_{,i}\ , \tag{4.11}$$

where $\pi^{\alpha}(x)\equiv\delta\mathcal{S}/\delta\alpha(x)$, can be solved following the same line of argument as in Sec. II B: For given functions ϕ and α, we conclude that $e^{-3\alpha(x^j)}\pi^{\alpha}(x^j)$ is only a function $3F(\phi,\mathcal{J})$ of the local value of the scalar field as well as some integration constant $\mathcal{J}$,

$$e^{-3\alpha(x^j)}\pi^{\alpha}(x^j)=3F[\phi(x^j),\mathcal{J}]\ . \tag{4.12a}$$

The scalar-field momentum is then necessarily

$$e^{-3\alpha(x^j)}\pi^{\phi}(x^j)=\frac{\partial F}{\partial\phi}[\phi(x^j),\mathcal{J}]\ . \tag{4.12b}$$

On the right-hand side, no local dependence on $\alpha(x^j)$ is allowed. The integration constant may actually be a functional $\mathcal{J}[\alpha,\phi]$ of the fields ϕ and α because $\partial\mathcal{J}/\partial x^i=0$.

The volume factor $\gamma^{1/2}=e^{3\alpha(x^j)}$ enters only in a separable way to preserve the density character of π^{α} and π^{ϕ}. Substitution of Eqs. (4.12) into the energy constraint $\mathcal{H}=0$ leads directly to the SHJE [Eq. (2.13b)], provided we set

$$F[\phi(x^j),\mathcal{J}]=-\frac{m_{\mathcal{P}}^2}{4\pi}H[\phi(x^j),\mathcal{J}]\ . \tag{4.13}$$

The momentum constraint therefore serves to reduce the full class of solutions of the HJE to those which are separable and satisfy the SHJE.

Although we have now completed our promised demonstration that Eq. (2.13b) should rightly be called the separated Hamilton-Jacobi equation, it is of interest to integrate (4.12) to obtain the explicit form of the phase functional $\mathcal{S}$. Consider the case when the functional $\mathcal{J}$ is independent of ϕ and α; $\mathcal{S}$ is then essentially the same as the homogeneous result (4.8):

$$\mathcal{S}=-\frac{m_{\mathcal{P}}^2}{4\pi}\int d^3x\, e^{3\alpha(x^j)}H[\phi(x^j),\mathcal{J}]\ , \tag{4.14}$$

but it is valid point by point. The integral is convergent if one considers a finite patch of comoving space, corresponding, for example, to our observable Universe.

There are, however, more complicated solutions if $\mathcal{J}[\alpha,\phi]$ is nontrivial:

$$\mathcal{S}=-\frac{m_{\mathcal{P}}^2}{4\pi}\int d^3x\, e^{3\alpha(x^j)}H[\phi(x^j),\mathcal{J}(\phi,\alpha)]-\int^{\mathcal{J}}g(\mathcal{J}')d\mathcal{J}'\ . \tag{4.15a}$$

This phase function describes the evolution of many universes, each with differing values of $\mathcal{J}$, which are defined implicitly through the relation

$$g(\mathcal{J})=-\frac{m_{\mathcal{P}}^2}{4\pi}\int d^3x\, e^{3\alpha(x^j)}\frac{\partial H}{\partial\mathcal{J}}[\phi(x^j),\mathcal{J}]\ , \tag{4.15b}$$

where g is an arbitrary function, as is shown in Appendix B. This solution is used in the semiclassical treatment of Sec. V describing the wave functional of many universes. However, each distinct universe *is* characterized by a single value of $\mathcal{J}$, as we now show. Let us assume that g is not constant. Take a sequence of field configurations $\alpha(x),\phi(x)$, corresponding to one evolving universe. If Eq. (4.15b) is satisfied at one time for a given $\mathcal{J}$, then it will be satisfied at all subsequent times with the same $\mathcal{J}$, since by Eq. (3.14), $\exp[3\alpha(x^j)](\partial H/\partial\mathcal{J})[\phi(x^j),\mathcal{J}]$ is conserved in time.

Imposition of the momentum constraint, which has played such a crucial role in our analysis, has effectively been neglected in the literature because a theorem proved by Moncrief and Teitelboim[17] showed that if Hamilton's principal function satisfies the energy constraint everywhere in space, then it also satisfies the momentum constraint. (See also Kuchar[33] for a more general discussion of constraints.) The reason we cannot apply this theorem is that, by neglecting second-order spatial gradients, we have lost terms whose presence were crucial for the proof. For example, the Poisson brackets between our approximate Hamiltonian densities at two different points,

$$\{\mathcal{H}(x^j),\mathcal{H}(x^{j'})\}\equiv\int d^3y\left[\frac{\delta\mathcal{H}(x^j)}{\delta\alpha(y^j)}\frac{\delta\mathcal{H}(x^{j'})}{\delta\pi^{\alpha}(y^j)}+\frac{\delta\mathcal{H}(x^j)}{\delta\phi(y^j)}\frac{\delta\mathcal{H}(x^{j'})}{\delta\pi^{\phi}(y^j)}-(x^j\leftrightarrow x^{j'})\right] \tag{4.16a}$$

vanishes, whereas the exact Poisson brackets gives the momentum constraint

$$\{\mathcal{H}(x^j),\mathcal{H}(x^{j'})\}=[\gamma^{ij}(x^k)\mathcal{H}_j(x^k)+\gamma^{ij}(x^{k'})\mathcal{H}_j(x^{k'})]\times\delta^3_{,i}(x^k-x^{k'}) . \quad (4.16b)$$

The Moncrief and Teitelboim (MT) redundancy theorem depends crucially on (4.16b). If one were to include second-order gradient terms such as $\phi_{|i}\phi^{|i}$ in (4.16a), one would recover the exact Poisson brackets but this would take us back to the full nonlinear gravitational problem. (MT actually forsaw that their theorem would break down if one neglected second-order terms. See also Ref. 20.) We conclude that within a spatial gradient perturbation framework, both momentum and energy constraints must be explicitly satisfied.

A concrete example of a functional which satisfies the energy constraint for a scalar field with an exponential potential (Sec. III A), but not the momentum constraint, is

$$\mathcal{S}=-\left[\frac{V_0 m_{\mathcal{P}}^2}{6\pi(1-1/3p)}\right]^{1/2}\int d^3x\,\exp\left[3\alpha(x^j)-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi(x^j)}{m_{\mathcal{P}}}\right]\cosh\left[\left[\frac{3}{p}\right]^{1/2}\alpha(x^j)-\frac{\sqrt{12\pi}\phi(x^j)}{m_{\mathcal{P}}}-I(x^j)\right] ,$$

where $I(x)$ is independent of either α or ϕ. Unless $I(x)$ is suitably restricted, $\mathcal{S}$ is not even invariant under reparametrizations of the three-metric, whereas the momentum constraint guarantees this property for the solutions (4.14) and (4.15) (Ref. 21, p. 1187).

Recall that Hamilton's principal function is the phase in semiclassical theory, a nice reinterpretation of the momentum constraint, implicit in (4.10b), is to recast it in terms of the phase of each cell volume, $S(x^j) =-(m_{\mathcal{P}}^2/4\pi)Ue^{3\alpha}H(\phi_k)$, as in (4.8): $\nabla S=p_\alpha\nabla\alpha +p_{\phi_k}\nabla_{\phi_k}$. We also have $\dot{S}=p_\alpha\dot{\alpha}+p_{\phi_k}\dot{\phi}_k$, and so the total spatial and temporal variation obeys $dS(x^j,t) =p_\alpha(x^j,t)d\alpha(x^j,t)+p_{\phi_k}(x^j,t)d\phi_k(x^j,t)$. That is, the local minisuperspace criterion $dS=p_\alpha d\alpha+p_{\phi_k}d\phi_k$ must hold point by point and moment by moment. Within this language the nonduality between α and ϕ_k that seemed apparent in (4.10b) is therefore easily understood.

D. Probability functionals

The probability functional $\mathcal{P}[\gamma_{ij},\phi_k]$ for field configurations $\gamma_{ij}(x^j)$ and $\phi_k(x^j)$ has encoded in it the full history of the ensemble of universes. Time does not explicitly appear. It is an intrinsic quantity to be defined in terms of the fields.[33] A typical choice for us is the measure of the local volume expansion factor $\det(\gamma_{ij})$, i.e., of α. To find the probability that the fields ϕ_k have configuration $\phi_k(x^j)$ at time α_0, we must determine the constrained probability functional $\mathcal{P}[h_{ij},\phi_k|\alpha(x^j)=\alpha_0]$, where $h_{ij}\equiv\exp(-2\alpha)\gamma_{ij}$. To evaluate this we need to make use of the conservation law that $\mathcal{P}[\gamma_{ij},\phi_k]$ obeys (in the absence of stochastic forces), which expresses the vanishing of the divergence of a probability current in superspace. This follows from the Wheeler-DeWitt equation (see Sec. V).

Instead of presenting the general case, we confine ourselves to the long-wavelength approximation where $\mathcal{S}$ (and $\mathcal{P}$) are functionals only of the three-metric determinant and scalar fields. We also only treat a single scalar field. We first discuss the minisuperspace model of Sec. IV B. The conservation law for $\mathcal{P}$ is most conveniently expressed in terms of flows on superspace. Superspace has a geometry described by a supermetric

$$ds^2=G_{AB}dX^AdX^B , \quad (4.17a)$$

$$X^0=\alpha,\quad X^1=\phi ,$$
$$G_{00}=-\frac{3m_{\mathcal{P}}^2}{4\pi}\frac{Ue^{3\alpha}}{N},\quad G_{11}=\frac{Ue^{3\alpha}}{N} , \quad (4.17b)$$

designed to make the kinetic terms in the action $\propto ds^2$. The flow velocity is just the contravariant momentum p^A, given by $G^{AB}\partial S/\partial X^B$ [Eq. (4.6)]; the covariant momentum is $p_A=\partial S/\partial X^A$. The conservation equation for the probability current is

$$(p^AP)_{;A}=\frac{1}{\sqrt{-G}}\partial_A(P\sqrt{-G}G^{AB}\partial_B S)=0 , \quad (4.18)$$

where the semicolon denotes covariant derivative in minisuperspace.

Let us choose α as our time measure. The constrained probability of observing a scalar field value ϕ, given α, is defined by

$$P(\phi|\alpha)d\phi\equiv\frac{P(\alpha,\phi)\sqrt{-G}\,d\phi}{\int P(\alpha,\phi)\sqrt{-G}\,d\phi} . \quad (4.19)$$

By integrating the equation of continuity over all ϕ, we find that

$$\int P(\alpha,\phi)\sqrt{-G}G^{00}\left[\frac{\partial S}{\partial\alpha}\right]d\phi$$

is constant. However, from the definition of the momentum [Eq. (4.6)], $G^{00}(\partial S/\partial\alpha)=\dot{\alpha}$ is unity for the α time choice, and thus

$$P(\phi|\alpha)d\phi=K_\alpha\sqrt{-G}P(\alpha,\phi)d\phi =K_\alpha\left[\frac{3m_{\mathcal{P}}^2}{4\pi}\right]^{1/2}[Ue^{3\alpha}H(\phi_k)]P(\alpha,\phi)d\phi , \quad (4.20a)$$

where K_α is some normalization constant. More simply, note that $P(\phi|\alpha)$ is just proportional to the $A=0$ term appearing under the derivative in (4.18):

$$P(\phi|\alpha)\propto P\sqrt{-G}G^{00}\partial_\alpha S . \quad (4.20b)$$

The notion of constrained probability was first introduced by DeWitt.[34] For a more recent discussion, see Kandrup.[35]

One may generalize the above argument to other choices of time surface. For example, if ϕ is the inflaton (as in Sec. III D 2), we would wish to form

$$P(\alpha|\phi)d\alpha = K_\phi \sqrt{-G}\, P(\alpha,\phi)d\alpha = K_\phi \left[\frac{3m_{\mathcal{P}}^2}{4\pi}\right]^{1/2} \left[U e^{3\alpha} \left[-\frac{m_{\mathcal{P}}^2}{4\pi}\frac{\partial H}{\partial \phi}\right]\right] \times P(\alpha,\phi)d\alpha \ . \tag{4.21}$$

where $K_\phi^{-1} = \int P(\alpha,\phi)\sqrt{-G}\, d\alpha$ is constant. [Again, note that $P(\alpha|\phi)$ is just proportional to the $A=1$ term appearing under the derivative in (4.18).] Since $\sqrt{-G} \propto N^{-1}$ changes when different hypersurfaces are chosen, and so the normalization K will also depend upon this choice. Thus the probabilities measured on different hypersurfaces may differ considerably because of the different $\sqrt{-G}$. The ability to transform from one to another is useful for stochastic inflation calculations.[10]

The continuity equation for the probability current implies that P is conserved along classical trajectories. Choosing α as the time coordinate, we find the following general solution for the constrained probability:

$$P(\phi(\alpha,\phi(\alpha_0)),|\alpha) = J^{-1} P(\phi(\alpha_0)|\alpha_0) \ , \quad J = \frac{\partial(\phi(\alpha,\phi(\alpha_0)))}{\partial(\phi(\alpha_0))} \ . \tag{4.22a}$$

Here $\phi \equiv \phi(\alpha,\phi(\alpha_0))$ describes the evolution of the scalar field as a function of time α and the initial field values $\phi(\alpha_0)$. J is the Jacobian of the transformation linking the Eulerian coordinates $\phi(\alpha,\phi(\alpha_0))$ to the Lagrangian ("initial") coordinates $\phi(\alpha_0)$. Applying Eq. (3.7), we find that $J = [db(\phi_0)/d\phi_0]/[db(\phi)/d\phi]$; we can therefore write the solution in terms of an arbitrary function f:

$$P(\phi|\alpha) = f(\alpha - b(\phi))\frac{db}{d\phi} \ . \tag{4.22b}$$

For example, in the case of an exponential potential $b = \sqrt{4\pi p}\, \phi/m_{\mathcal{P}}$, Eq. (4.22b) describes a form-invariant probability. However, every time the system receives a stochastic kick, we would generally expect the form to change.

Our analysis can be generalized to inhomogeneous fields. The supermetric $ds^2(x)$ has the same form as (4.17), but varies from point to point, with U now interpreted as a cell volume of the scale we have smoothed over. We define the probability functional on $\alpha(x)$ hypersurfaces to be

$$\mathcal{P}[\phi(x)|\alpha(x)] = \mathcal{P}[\alpha(x),\phi(x)]\mathcal{K}_\alpha \times \prod_x \sqrt{-G(x)}\, G^{00}(x)\frac{\delta S}{\delta\alpha(x)} \ , \tag{4.23}$$

where $\mathcal{K}_\alpha$ is a normalization factor. It is conserved because it is the first term appearing under the derivative in the functional equation of continuity for $\mathcal{P}$:

$$\frac{\delta}{\delta X^A(x)}\left[\mathcal{P}\sqrt{-G(x)}\, G^{AB}(x)\frac{\delta S}{\delta X^B(x)}\right] = 0 \ . \tag{4.24}$$

Since the configurations allowed in the probability functional should be required to satisfy the energy and momentum constraints, we expect that $\mathcal{P}$ must obey corresponding functional equations expressing this. The energy constraint on $\mathcal{P}$ is (4.24). The momentum constraint on $\mathcal{P}$ can be derived within the context of canonical quantum theory, as we describe the next section, and takes the same form as the momentum constraint equation obeyed by $\mathcal{S}$:

$$\frac{1}{3}\left[\frac{\delta\mathcal{P}}{\delta\alpha(x^j)}\right]_{,i} = \frac{\delta\mathcal{P}}{\delta\alpha(x^j)}\alpha_{,i}(x^j) + \frac{\delta\mathcal{P}}{\delta\phi(x^j)}\phi_{,i}(x^j) \ . \tag{4.25}$$

A consequence of this is that $\mathcal{P}[\alpha,\phi]$ is reparametrization invariant under spatial coordinate transformations. Just as we showed for the phase at the end of Sec. IV C, a transparent way to rewrite (4.25) which makes the invariance manifest is to use the density $\ln P(x^j,t)$ for a cell volume of size U [defined by $\frac{1}{3}\delta \ln\mathcal{P}/\delta\alpha(x^j)$]. The momentum constraint is just $\nabla \ln P = \nabla\alpha\, \partial \ln P/\partial\alpha + \nabla\phi\, \partial \ln P/\partial\phi$, requiring that coordinate changes respect the functional form of $\ln P$.

An explicit solution of (4.24), for a single scalar field with Hamilton's principal function given by the general class of solutions (4.15a), is

$$\mathcal{P}[\phi(x)|\alpha(x)] = f[\alpha(1) - b[\phi(1)], \alpha(2) - b[\phi(2)], \ldots] \times \frac{db}{d\phi}[\phi(1)]\frac{db}{d\phi}[\phi(2)]\cdots \ , \tag{4.26}$$

where f is an arbitrary function. This result may be verified using the identity (B3a) in Appendix B. Usually, one chooses all of the $\alpha(x)$ to be identical, in which case we interpret this functional to be the probability on a uniform α surface.

A sample functional form that arises in most single-inflaton models is the Gaussian-like form

$$\mathcal{P}[\phi(x)|\alpha(x)] = \exp\left[-\frac{1}{2}\int\int d^3x\, d^3y\, e^{3\alpha(x^j)}\frac{\partial H}{\partial I}(\phi(x^j),I)\xi^{-1}(x^j,y^j)e^{3\alpha(y^j)}\frac{\partial H}{\partial I}(\phi(y^j),I)\right] \times \frac{1}{\sqrt{\det(\xi)}}\prod_x\left[\frac{e^{3\alpha(x^j)}}{\sqrt{2\pi}}\frac{\partial^2 H}{\partial I\,\partial\phi}(\phi(x^j),I)\right] \ . \tag{4.27}$$

Here $\xi^{-1}(x^j,y^j)$ is the inverse of a two-point correlation function. We have assumed that $e^{3\alpha}\partial H/\partial I \propto \exp\{3[\alpha - b(\phi)]\}$, which follows from (3.7) and (3.14). Such a form could describe the spatial correlations that give rise to galaxy formation. Note that the connection between different spatial points expressed through ξ^{-1} in (4.27) cannot come from our truncated equations. For this reason, (4.27) does not obey (4.25). Within stochastic inflation these connections arise from spatial correlations in the quantum fluctuations cross the horizon and feeding the long wavelengths. Thus a self-consistent framework requires the explicit connection to the short-distance forces.[10,36]

V. QUANTUM EVOLUTION OF LONG-WAVELENGTH FLUCTUATIONS

Only within the quantum theory can the intimate relation between Hamilton's principal functional $\mathcal{S}$ and the probability functional $\mathcal{P}$ introduced in Sec. IV be properly appreciated. The wave functional $\Psi[\gamma_{ij},\phi_k]$ can be written in terms of a phase $\mathcal{S}$ and modulus $\mathcal{P}$: $\Psi=\mathcal{P}^{1/2}e^{i\mathcal{S}}$. Canonical quantization using Hamiltonian methods yields functional differential equations for Ψ expressing the constraint equations, which, when written in terms of $\mathcal{S}$ and $\mathcal{P}$, bear much similarity to the equations for $\mathcal{S}$ and $\mathcal{P}$ given in the last section; indeed they are identical in the limit in which $\hbar \rightarrow 0$: in Sec. IV we are dealing with the WKB limit of the quantum gravity theory.[34] Of course, the canonical quantization method has its unresolved controversies,[18] but it is at least straightforward for the well-studied homogeneous minisuperspace models for which the energy constraint (Wheeler-DeWitt) equation is all that is required. Our long-wavelength metric with its restricted degrees of freedom generalizes these homogeneous models.[19,20] Indeed, the Hamiltonian constraint commutes at different spatial points, so that at one point the Wheeler-DeWitt equation is the same as for homogeneous minisuperspace models. We solve exactly the Wheeler-DeWitt equation with finite $\hbar$ for the case of an exponential potential, the principal analytic model of this paper (Sec. V B). (We restrict the discussion in this section to one scalar field only. We also set $\hbar=1$ in the subsequent equations.)

As in the last section, the issue of the averaging volume U arises. In minisuperspace it is the comoving volume of the Universe. We view U as the comoving horizon volume when fluctuations in our observable Universe expanded beyond the Hubble radius during inflation, as motivated by our long-wavelength framework in which we spatially average the full theory over the horizon size. Since U is much larger than the Planck volume, quantum corrections to long-wavelength evolution are necessarily small. We conclude that quantum corrections to the WKB solutions of Sec. IV are unimportant for galaxy formation and do not result in significant non-Gaussian modifications on observable scales.

In Sec. V C we discuss the functional momentum constraint, which must also be explicitly satisfied along with the functional Wheeler-DeWitt equation. We find that the long-wavelength Wheeler-DeWitt equations do not, in general, satisfy the momentum constraint except in the WKB limit. We have not been able to construct a self-consistent quantum theory of long-wavelength fields. Since the short-wavelength components are integrated over to yield the long-wavelength quantum theory, our construction based on a wave functional of the long-wavelength fields may be fundamentally flawed.

A. Review of canonical formalism

Through this section we consider the case of a single scalar field only. In the canonical approach to the quantum theory of the gravitational field, the Hamiltonian and momentum constraints [Eqs. (4.3) and (4.4)] are expressed as functional operator equations acting on the wave functional Ψ, with the momenta replaced by the functional derivatives $\pi^{\gamma_{ij}}(x)=-i\delta/\delta\gamma_{ij}(x)$ and $\pi^{\phi}(x) = -i\delta/\delta\phi(x)$:

$$\mathcal{H}(x)\Psi=0, \quad \mathcal{H}_i(x)\Psi=0 . \tag{5.1}$$

In the long-wavelength approximation we drop all second-order spatial gradients. We further assume that $\Psi[\alpha,\phi]$ depends on the metric only through the volume factor $\alpha(x)=\ln[\gamma(x)]/6$, in which case the constraints become

$$\left[\frac{2\pi}{3m_{\mathcal{P}}^2}e^{-3\alpha(x)}\frac{\delta^2}{\delta\alpha(x)^2}-\frac{1}{2}e^{-3\alpha(x)}\frac{\delta^2}{\delta\phi(x)^2} + e^{3\alpha(x)}V(\phi(x))\right]\Psi=0 , \tag{5.2a}$$

$$\left[\frac{1}{3}\left[\frac{\delta}{\delta\alpha(x)}\right]_{,i}-\alpha_{,i}(x)\frac{\delta}{\delta\alpha(x)}-\phi_{,i}(x)\frac{\delta}{\delta\phi(x)}\right]\Psi=0 . \tag{5.2b}$$

The above expressions must be regularized to be meaningful. Here we split the spatial space into cells of equal comoving volume U and replace the functional derivative by an ordinary partial derivative $\delta/\delta\alpha(x)\rightarrow U^{-1}\partial/\partial\alpha(x)$. With this substitution the Wheeler-DeWitt equation is the same as that for homogeneous minisuperspace, except that it holds point by point (over cells of volume U). Although it generally has operator-ordering ambiguities for our choice of variables (5.2a) gives the correct ordering. It can be recast in a manifestly field-reparametrization-invariant form.[37]

$$-\tfrac{1}{2}\Psi_{;A}{}^{;A}+NUV(\phi_k)\Psi=0 . \tag{5.3}$$

All covariant derivatives are taken with respect to the metric (4.17a). Note that for a single scalar field the Wheeler-DeWitt equation is independent of N; i.e., it is conformally invariant in the supermetric. However, this is not necessarily true for multiple fields unless one explicitly introduces a term $\propto R\Psi$ coupled to the Ricci curvature of the supermetric. Otherwise, the time surface issue would have to be addressed.

To make contact with the classical analysis of Sec. IV, we express the complex wave function in polar form, $\Psi=P^{1/2}e^{iS}$, and split (5.3) into real and imaginary parts:

$$-\tfrac{1}{2}P^{-1/2}P^{1/2}{}_{;A}{}^{;A}+\tfrac{1}{2}S^{,A}S_{,A}+NUV(\phi_k)=0\ , \tag{5.4a}$$

$$(\sqrt{-G}PS^{,A})_{,A}=0\ . \tag{5.4b}$$

The phase and modulus have a quantum nature since the first term in (5.4a) is actually proportional to $\hbar^2$, and, hence, is of a quantum origin. The remainder is the classical Hamilton-Jacobi equation (4.7). Further, Eq. (5.4b) is just the equation of continuity (5.6). We now recognize these as the WKB or semiclassical limit of the quantum equations in which $\hbar\to 0$. In these equations $\hbar$ always appears in the combination $\hbar/U$. Hence the semiclassical limit $\hbar\to 0$ is essentially equivalent to the infinite volume limit $U\to\infty$. In these cases $S\to S_{\rm cl}$, Hamilton's principal function of Sec. IV, and $P\to P_{\rm cl}$, the classical probability function of Sec. IV.

However, it is not P we want, but the constrained probability on a specific time hypersurfaces. As in Sec. IV, we do this (up to a constant normalization factor) on a surface of constant α through the p_α operator:

$$P(\phi|\alpha)=i\left[\frac{\partial\Psi^*}{\partial\alpha}\Psi-\Psi^*\frac{\partial\Psi}{\partial\alpha}\right]=-2|\Psi|^2\frac{\partial S}{\partial\alpha}\ . \tag{5.5}$$

It is conserved, as one may readily verify through the equation of continuity

$$i[\sqrt{-G}\,(\Psi^*\Psi^{,A}-\Psi\Psi^{*,A})]_{,A}=0\ , \tag{5.6}$$

which follows from (5.3). In the general inhomogeneous case, one replaces the partial derivative by a functional derivative at point x.

The probability function Eq. (5.5), which coincides with our classical definition (4.20b), is positive for an expanding universe. However, one of the problems with the quantum theory is that $P(\phi|\alpha)$ may become negative, depending on the sign of $-\partial S/\partial\alpha$. We do not consider this a major problem, at least in the semiclassical limit, since α would have ceased to be a viable time coordinate. For if $-\partial S/\partial\alpha$ changes sign, then, by continuity, it must at some time vanish. Within our long-wavelength equations, this would imply that the positive-definite quantity $\frac{1}{2}U^{-2}e^{-6\alpha}(\partial S/\partial\phi)^2+V(\phi_k)$ in Eq. (4.7) would vanish as well, describing a universe which is static and with no matter energy. Of course, if we consider universes with positive three-curvature ${}^{(3)}R$, $-\partial S/\partial\alpha$ does change sign, and α can be a valid time coordinate only for either the expanding or contracting portion of the spacetime. With quantum fluctuations it is no longer clear that $P(\phi|\alpha)$ will always remain positive. In fact, our exact quantum solutions to the Wheeler-DeWitt equation can have apparently negative values of the probability.

B. Analytic long-wavelength Wheeler-DeWitt solutions in an exponential potential

Even for the quantum theory, the exponential potential leads to analytic solutions, which we display here. By making a change of variables, one may reduce the long-wavelength Wheeler-DeWitt equation describing a single scalar field with an exponential potential into a Klein-Gordon equation. By constructing analytic Green's functions, we can explore the entire solution space. Although there are numerous solutions depending upon boundary conditions, the Feynman Green's function, which includes the Vilenkin wave function[38,39] as a special case, is particularly useful because it describes an expanding universe with a scalar field rolling down the potential. By contrast, other Green's functions such as the retarded one, an example being the Hartle-Hawking wave function, have components which describe contraction. We show that quantum corrections to the long-wavelength evolution of the probability function derived from the Feynman Green's function, which is of relevance for galaxy formation, are usually small.

The Green's function is a solution of Eq. (5.3) with a δ-function term $\frac{1}{2}\delta(\alpha-\alpha_0)\delta(\phi-\phi_0)/\sqrt{-G}$ added to the right-hand side. (The factor of $\frac{1}{2}$ has been inserted for later convenience.) To transform the Wheeler-DeWitt equation into a Klein-Gordon form for a massive field, we change variables:

$$e=\exp\left[3\alpha-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal P}}\right]\times\cosh\left[\left[\frac{3}{p}\right]^{1/2}\alpha-\frac{\sqrt{12\pi}\phi}{m_{\mathcal P}}\right]\ , \tag{5.7a}$$

$$f=\exp\left[3\alpha-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal P}}\right]\times\sinh\left[\left[\frac{3}{p}\right]^{1/2}\alpha-\frac{\sqrt{12\pi}\phi}{m_{\mathcal P}}\right]\ . \tag{5.7b}$$

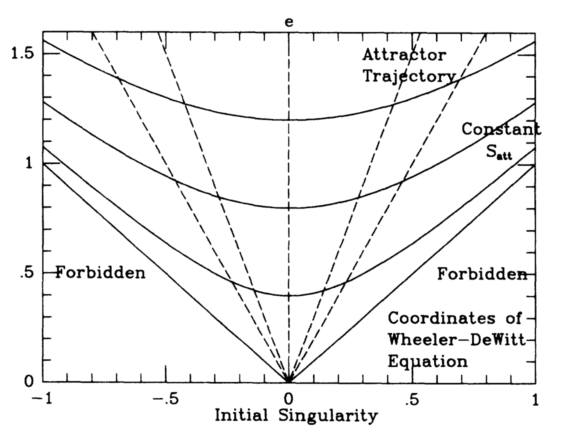


FIG. 6. The Wheeler-DeWitt equation for a scalar field with an exponential potential is transformed into the analytically tractable two-dimensional wave equation for a massive scalar field if one defines new fields $e(\alpha,\phi)$ and $f(\alpha,\phi)$ [Eq. (5.7)]. These new coordinates must lie with the region defined by $e>0$, $|f|<e$. Trajectories of the classical attractor $\alpha=\sqrt{4\pi p}\,\phi+\text{const}$ are straight lines. The hyperbolas are surfaces where Hamilton's principal function $S_{\rm att}$ is constant, i.e., where $\alpha=\sqrt{4\pi/p}\,\phi/(3m_{\mathcal P})+\text{const}$. The Universe has zero volume at the initial singularity $e=f=0$.

The inverse transformation is given by

$$\alpha=\frac{1}{3(1-1/3p)}\left[\frac{\ln(e^2-f^2)}{2}-\frac{\operatorname{arctanh}(f/e)}{\sqrt{3p}}\right], \tag{5.8a}$$

$$\phi=\frac{1}{1-1/3p}\frac{m_{\mathcal{P}}}{\sqrt{12\pi}}\left[\frac{\ln(e^2-f^2)}{2\sqrt{3p}}-\operatorname{arctanh}\left[\frac{f}{e}\right]\right]. \tag{5.8b}$$

Note that e must be positive and that $|f|<e$. The exponential term in (5.7) is proportional to the attractor solution of the classical Hamilton-Jacobi equation [see Eqs. (4.8) and (3.1)]. The argument of the cosh and sinh is a constant along the attractor trajectory [Eq. (3.4a)]. The new coordinates are plotted in Fig. 6. The supermetric line element now becomes

$$Ds^2=U\frac{m_{\mathcal{P}}^2}{12\pi}\left[1-\frac{1}{3p}\right]^{-1}(-de^2+df^2)$$
$$\times\exp\left[-3\alpha+\left[\frac{16\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right]\Big/N, \tag{5.9}$$

and the Green's function equation simplifies to

$$\frac{\partial^2\Psi}{\partial e^2}-\frac{\partial^2\Psi}{\partial f^2}+\kappa^2\Psi=\delta(e-e_0)\delta(f-f_0),$$

with (5.10)

$$\kappa^2=U^2V_0m_{\mathcal{P}}^2/6\pi[1-1/(3p)].$$

Changes of variables in the Wheeler-DeWitt equation have also proven to be useful in numerical[39,40] and analytic solutions[41] of homogeneous closed-universe models.

We wish to describe an inflating universe with a scalar field moving down the potential. The appropriate boundary conditions for the Green's function are the Feynman boundary conditions. The Feynman Green's function Ψ_F can be expressed[42] in terms of a Hankel function of the second kind, $H_0^{(2)}$ (consult Abramowitz and Stegun[43] for notation):

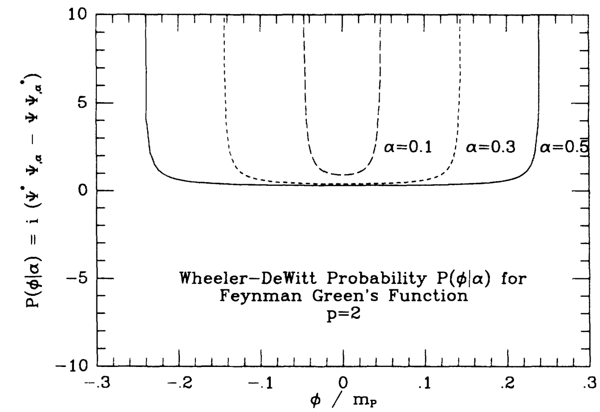


FIG. 7. The Feynman function [Eq. (5.11a)], which includes the Vilenkin wave function as a special case, is the most useful Green's function of the Wheeler-DeWitt equation if one wishes to evolve a probability function giving the fluctuations for galaxy formation. Originating at $\alpha_0=0$, $\phi_0=0$, the conserved probability $P(\phi|\alpha)$ [Eq. (5.12)] is plotted at the three times $\alpha=0.1$, 0.3, and 0.5 for a scalar field moving in an exponential potential with $p=2$ (see Sec. III A). It is positive for $|\phi|\le\sqrt{3/(4\pi)}\alpha m_{\mathcal{P}}$ and vanishes elsewhere. The singular behavior at the leading edges is removed if one smears over initial field values ϕ_0, giving a solution which agrees well with the classical probability function if the initial comoving spatial volume is large, corresponding to $\kappa\approx10^9$.

$$\Psi_F(e,f;e_0,f_0)=\tfrac{1}{4}H_0^{(2)}(\kappa[(e-e_0)^2-(f-f_0)^2-i\epsilon]^{1/2}), \tag{5.11a}$$

where ϵ is a positive infinitesimal. If the expression under the square root is positive, $(e-e_0)^2-(f-f_0)^2>0$, the function is complex,

$$\Psi_F(e,f;e_0,f_0)=\tfrac{1}{4}\{J_0(\kappa[(e-e_0)^2-(f-f_0)^2]^{1/2})-iN_0(\kappa[(e-e_0)^2-(f-f_0)^2]^{1/2})\}, \tag{5.11b}$$

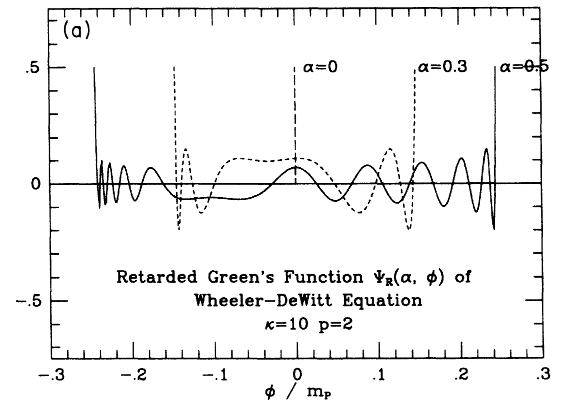
whereas it is purely imaginary for $(e-e_0)^2-(f-f_0)^2<0$:

$$\Psi_F(e,f;e_0,f_0)=\frac{i}{2\pi}K_0(\kappa[-(e-e_0)^2+(f-f_0)^2]^{1/2}). \tag{5.11c}$$

Here K_0 is a modified Bessel function, which decays exponentially for large values of its argument. The constrained probability (5.5) has a very simple form:

$$P(\phi|\alpha)=\frac{3}{8\pi}\left[\frac{1-1/\sqrt{3p}}{1-\exp-(1-1/\sqrt{3p})[3(\alpha-\alpha_0)+\sqrt{12\pi}(\phi-\phi_0)/m_{\mathcal{P}}]}\right.$$
$$\left.+\frac{1+1/\sqrt{3p}}{1-\exp-(1+1/\sqrt{3p})[3(\alpha-\alpha_0)-\sqrt{12\pi}(\phi-\phi_0)/m_{\mathcal{P}}]}\right], \tag{5.12}$$

for $\sqrt{12\pi}|\phi-\phi_0|/m_{\mathcal{P}}<3(\alpha-\alpha_0)$; and it vanishes elsewhere. This solution is shown in Fig. 7 for several times. It is positive everywhere and is singular whenever the fields approach "the null trajectory," $|\phi-\phi_0|/m_{\mathcal{P}}=\sqrt{3/(4\pi)}(\alpha-\alpha_0)$; hence it is not normalizable because its integral over ϕ is divergent. However, linear combinations of the various Feyn-



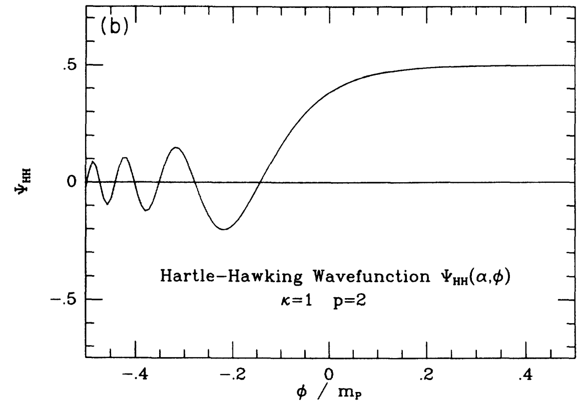


FIG. 8. (a) Retarded Green's function [Eq. (5.13a)] originating at $\alpha_0=0$, $\phi_0=0$, is shown at the times $\alpha=0.3$ and 0.5, for the same situation as Fig. 7 except that here the wave amplitude Ψ is plotted. It is oscillatory for $|\phi| \leq \sqrt{3/(4\pi)}\alpha m_{\mathcal{P}}$ and vanishes otherwise. Since Ψ is real, the conserved probability $P(\phi|\alpha)$ vanishes, implying these boundary conditions for Ψ are uninteresting for galaxy formation fluctuations. $\kappa=10$ is shown for plotting purposes, although $\kappa\sim 10^9$ is preferred. (b) The Hartle-Hawking wave function of the Universe [Eq. (5.13b)], shown here at one instance in time, is just the retarded Green's function originating at the singularity $\alpha_0=-\infty$. It oscillates vigorously for large negative ϕ, but approaches a constant in the opposite limit. It moves in time α as a wave of constant shape, with uniform phase velocity $d\phi/d\alpha=\sqrt{4\pi/p}\,m_{\mathcal{P}}/3$. For closed homogeneous universes this wave function follows from a specific choice of the ground state. However, we consider the inhomogeneous long-wavelength Wheeler-DeWitt equation as an evolution equation for an initial probability function which is generated by quantum effects within the Hubble radius, as suggested by stochastic inflation. Linear superpositions of the Feynman Green's functions are then more relevant. $\kappa=1$ is shown for plotting purposes.

man functions are not guaranteed to yield positive probabilities. The Hankel function phase is just S, which is negative, describing an expanding universe, as recommended by Vilenkin.[38,39]

Other Green's functions with differing boundary conditions are also of interest, but these are combinations of terms with positive as well as negative phase components. For example, the retarded Green's function is essentially the first term of Eq. (5.11b):

$$\Psi_R(e,f;e_0,f_0)=\tfrac{1}{2}\theta(|e-e_0|-|f-f_0|)J_0(\kappa[(e-e_0)^2-(f-f_0)^2]^{1/2}) . \tag{5.13a}$$

Unlike the Feynman function, it vanishes for $|e-e_0|<|f-f_0|$ because of the theta function: $\theta(x)=1$ for $x\geq 0$, $\theta(x)=0$ for $x<0$. It is displayed in Fig. 8(a). If the initial value α_0 is taken to be the zero volume limit $\alpha_0\to-\infty$, then e_0 and f_0 vanish, yielding the Hartle-Hawking[44] wave function for the exponential potential case:

$$\Psi_{HH}(\alpha,\phi)=\Psi_R(e,f;0,0)=\tfrac{1}{2}J_0\left[\kappa\exp\left[3\alpha-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right]\right] . \tag{5.13b}$$

As Fig. 8(b) shows, it describes a wave moving with uniform velocity and constant shape. Ψ approaches a constant in the zero volume limit with ϕ held fixed, as required by Hartle and Hawking. In a homogeneous closed model, which cannot be treated with our first-order gradient expansion, (5.13b) would follow from the ground-state wave function of the Universe. There has been much controversy in the literature over whether the Hartle-Hawking wave function or the Vilenkin wave function offers a better description of the ground state of the Universe. For us, however, the ground state for long-wavelength fields is irrelevant. By analogy with stochastic inflation, we consider the long-wavelength Wheeler-DeWitt equation as an evolution equation for some given initial wave function. This initial state is generated within the Hubble radius by short-scale physics which has not been included in our long-wavelength treatment. (There may not even be a ground state, as for the inverted harmonic oscillator in ordinary quantum mechanics for which the evolution of some initial probability distribution using the Schrödinger equation is still meaningful.)

An interesting question is whether quantum gravity corrections will affect the fluctuations for galaxy formation. For the long-wavelength component, the answer is probably no. Our version (5.2a) of the Wheeler-DeWitt equation is, of course, only an approximation to the full quantum system since all operators have been averaged on the comoving volume U, which is justified only if it exceeds the comoving horizon volume $H^{-3}e^{-3\alpha}$. In this case there is no causal contact between different spatial points, and they may be treated as independent universes, just as explained in Sec. II. However, when $\alpha_0\to-\infty$, the argument of the Hankel function in (5.11a),

$$\kappa\exp\left[3\alpha-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right]=\frac{1}{4\pi}\left[\frac{m_{\mathcal{P}}}{H}\right]^2 U e^{3\alpha}H^3>\frac{1}{4\pi}\left[\frac{m_{\mathcal{P}}}{H}\right]^2\approx 10^9 , \tag{5.14}$$

is exceedingly large, since the correct level of fluctuations are produced when $H/m_{\mathcal{P}} \approx 10^{-5}$ [see discussion after Eq. (3.12)]. One is justified in using the asymptotic expansion of the Hankel function, $H_0^{(2)}(x) \approx \sqrt{2/(\pi x)} e^{-i(x-\pi/4)}$: The analysis is essentially classical to 1 part in 10^9. However, it is quite possible that quantum gravity effects could affect the fluctuations for galaxy formation by altering the perturbations that arise within the horizon. One expects that something must happen here because scales less than the Planck length expand to encompass our present horizon size.

To see more clearly that long-wavelength quantum effects will not be of importance for the fluctuations useful for galaxy formation, consider a more general expansion solution, in which the Feynman function is smeared over a weight function $W(\phi_0)$:

$$\Psi = \int \Psi_F(e,f;e_0,f_0) W(\phi_0) d\phi_0 \ , \tag{5.15a}$$

Consider, at fixed α, scalar-field values which are far from the edges of the null trajectory, $|\phi - \bar{\phi}_0| < \sqrt{3/(4\pi)}(\alpha-\alpha_0) m_{\mathcal{P}}$; there, the argument of Ψ_F may be approximated as

$$\begin{aligned}
\kappa(e^2-f^2)^{1/2} & \left[\left[1 - \frac{e_0-f_0}{e-f} \right] \left[1 - \frac{e_0+f_0}{e+f} \right] \right]^{1/2} \\
& \approx \kappa(e^2-f^2)^{1/2} - \frac{\kappa}{2}(e_0^2-f_0^2)^{1/2} \left[\left[\frac{e+f}{e-f} \right]^{1/2} \left[\frac{e_0-f_0}{e_0+f_0} \right]^{1/2} + \left[\frac{e-f}{e+f} \right]^{1/2} \left[\frac{e_0+f_0}{e_0-f_0} \right]^{1/2} \right] \\
& = \kappa \exp \left[3\alpha - \left[\frac{4\pi}{p} \right]^{1/2} \frac{\phi}{m_{\mathcal{P}}} \right] - \kappa \exp \left[3\alpha_0 - \left[\frac{4\pi}{p} \right]^{1/2} \frac{\phi_0}{m_{\mathcal{P}}} \right] \cosh \left[\left[\frac{3}{p} \right]^{1/2} \left[\alpha - \alpha_0 - \frac{\sqrt{4\pi p}\,(\phi-\phi_0)}{m_{\mathcal{P}}} \right] \right] .
\end{aligned}$$

[Keeping away from the null trajectory is justified after the fact by noting that the wave function we obtain in (5.15b) has little weight on the null trajectory.] If the initial value ϕ_0 is chosen so that $\kappa e^{3\alpha_0 - \sqrt{4\pi/p}\,\phi_0/m_{\mathcal{P}}} \approx 10^9$, motivated by the maximum fluctuation level allowed by observations, then one can apply the asymptotic expansion of the Hankel function, yielding

$$\Psi(\alpha,\phi) \approx e^{-(3\alpha - \sqrt{4\pi/p}\,\phi/m_{\mathcal{P}})/2} e^{-i\kappa \exp(3\alpha - \sqrt{4\pi/p}\,\phi/m_{\mathcal{P}})} g(\alpha - \sqrt{4\pi p}\,\phi/m_{\mathcal{P}}) \ . \tag{5.15b}$$

The wave form g is

$$g(x) = \left[\frac{2}{\pi\kappa} \right]^{1/2} \int d\phi_0 W(\phi_0) \exp \left\{ i\kappa e^{3\alpha_0 - \sqrt{4\pi/p}\,\phi_0/m_{\mathcal{P}}} \cosh \left[\left[\frac{3}{p} \right]^{1/2} \left[x - \alpha_0 + \frac{\sqrt{4\pi p}\,\phi_0}{m_{\mathcal{P}}} \right] \right] \right\} . \tag{5.15c}$$

To make this tractable we assume that the weight function $W(\phi_0)$ is sharply peaked about $\phi_0 = \bar{\phi}_0$, the classical initial value. In this limit we can show that g is related to the Fourier transform of W. For real g the probability $P(\phi|\alpha) = 6\kappa g^2(\alpha - \sqrt{4\pi p}\,\phi/m_{\mathcal{P}})$, evolves in a similar way to the classical evolution [Eq. (4.22b)]. Thus we conclude that, provided the averaging volume is large compared to the Planck length and provided transients have decayed, long-wavelength quantum effects are tiny.

The Klein-Gordon equation admits a complete set of wave solutions $\Psi = \exp[-i(we - kf)]$, with $w = \pm(\kappa^2 + k^2)^{1/2}$, where k is the "wave number." However, for these, narrow wave packets describe collapsing probability distributions, as we now show. If one sums over a narrow band of positive-energy waves centered about $k = \bar{k}$, with weight given by $W(k)$, then it is a well-known result that the wave form moves at the group velocity $f/e = (d\omega/dk)(\bar{k})$, with a constant shape:

$$\Psi = \int dk\, W(k) e^{-i(we-kf)} \approx e^{-i[w(\bar{k})e - \bar{k}f]} g \left[\frac{d\omega}{dk}(\bar{k})e - f \right] ,$$

where

$$g[x] = \int dk\, W(k) e^{-ix(k-\bar{k})} ,$$

which, for convenience, we shall assume has its maximum at $x=0$. The constrained probability

$$P(\phi|\alpha) = \left[\frac{3}{\pi} \right]^{1/2} \omega(\bar{k}) m_{\mathcal{P}} g^2 \left[\frac{d\omega}{dk}(\bar{k})e - f \right] \frac{\partial}{\partial\phi} \left[\frac{d\omega}{dk}(\bar{k})e - f \right]$$

is explicitly conserved; for simplicity, g was assumed to be real. However, in the large $\alpha - \alpha_0$ limit, it approaches a δ function. For example, the argument of g^2 may be written as

$$-\kappa\omega^{-1}(\bar{k}) e^{3\alpha - \sqrt{4\pi/p}\,\phi/m_{\mathcal{P}}} \sinh \left[\left[\frac{3}{p} \right]^{1/2} \alpha - \frac{\sqrt{12\pi}\phi}{m_{\mathcal{P}}} - \operatorname{arcsinh} \left[\frac{\bar{k}}{\kappa} \right] \right] .$$

The point of maximum probability, $x=0$, evolves according to the attractor solution of (3.4a), and its width decays exponentially, $(\Delta\phi^2)^{1/2} \propto e^{-\alpha(3-1/p)/2}$. This does not mean that a typical long-wavelength wave function collapses, but rather that these solutions correspond to classical trajectories in the decaying mode. One must work to a higher order in $k-k_0$, as given, for example, in the Feynman functions, in order to produce probability functions of physical interest. Furthermore, in order to describe the generic collapse of a wave function, one must somehow introduce short-scale fluctuations.[45]

The change of variables [Eq. (5.7)] is the crucial step in obtaining our exact solutions. We end this section by showing how it was obtained. The homogeneous Wheeler-DeWitt equation is

$$\frac{4\pi}{3m_{\mathcal{P}}^3}\frac{\partial^2\Psi}{\partial\alpha^2}-\frac{\partial^2\Psi}{\partial\phi^2}+2U^2V_0\exp\left[6\alpha-\left[\frac{16\pi}{p}\right]^{1/2}\frac{\phi}{m_{\mathcal{P}}}\right]\Psi=0 \, .$$

When expressed in terms of null coordinates,

$$u=\phi+\left[\frac{3m_{\mathcal{P}}^2}{4\pi}\right]^{1/2}\alpha, \quad v=\phi-\left[\frac{3m_{\mathcal{P}}^2}{4\pi}\right]^{1/2}\alpha \, ,$$

this equation has a simpler structure:

$$4\exp\left[-\frac{\sqrt{12\pi}}{m_{\mathcal{P}}}\left[1-\frac{1}{\sqrt{3p}}\right]u\right]\frac{\partial}{\partial u}\left\{\exp\left[\frac{\sqrt{12\pi}}{m_{\mathcal{P}}}\left[1+\frac{1}{\sqrt{3p}}\right]v\right]\frac{\partial\Psi}{\partial v}\right\}+2U^2V_0\Psi=0 \, .$$

The obvious thing to do now is remove the exponentials through the change of variables

$$x=\exp\left[\frac{\sqrt{12\pi}}{m_{\mathcal{P}}}\left[1-\frac{1}{\sqrt{3p}}\right]u\right] \quad \text{and} \quad y=\exp\left[-\frac{\sqrt{12\pi}}{m_{\mathcal{P}}}\left[1+\frac{1}{\sqrt{3p}}\right]v\right] \, ,$$

and finally revert to a diagonal D'Alembertian, by introducing

$$x=e-f \quad \text{and} \quad y=e+f \, .$$

This yields the Klein-Gordon equation in the form (5.10).

C. Inhomogeneous fields

As in the semiclassical theory of Sec. IV, we might expect to have difficulties satisfying the quantum version of the functional momentum constraint. Just as for the Hamiltonian constraint, the momentum constraint for the wave functional $\Psi=\mathcal{P}^{1/2}\mathcal{S}$ has both real and imaginary parts. The real part of $\Psi^*\mathcal{H}\Psi$ gives the Hamilton-Jacobi equation for $\mathcal{S}$, and the imaginary part yields the equation of continuity for the probability $\mathcal{P}$ [Eq. (4.24)]. The real part of $\Psi^*\mathcal{H}_i\Psi$ is the familiar momentum constraint on $\mathcal{S}$ [Eq. (4.10b)], and the imaginary part is an identical equation for $\mathcal{P}$ [Eq. (4.25)]. Taken together, the two imaginary parts impose the requirement that the probability be both field reparametrization invariant and spatial coordinate reparametrization invariant.

Since the energy constraint (5.2a) holds point by point, a natural long-wavelength wave functional to adopt for an exponential potential is a product of Feynman Green's wave functionals, one for each cell of coordinate volume $U=d^3x$:

$$\Psi=\prod_{x^j}\frac{1}{4}H_0^{(2)}\left\{\kappa\exp\left[3\alpha(x^j)-\left[\frac{4\pi}{p}\right]^{1/2}\frac{\phi(x^j)}{m_{\mathcal{P}}}\right]\right\}=\prod_{x^j}\frac{1}{4}H_0^{(2)}\left[-\frac{S_{\mathrm{cl,att}}}{U}\right] \, . \tag{5.16}$$

We have taken $\alpha_0=-\infty$. Writing it in terms of the one-point classical Hamilton-Jacobi principal function attractor solution $S_{\mathrm{cl,att}}$ [see Eqs. (3.1) and (4.8)] makes the asymptotic $\sim\exp(iS_{\mathrm{cl,att}}/U)$ semiclassical phase behavior manifest. To be more general, we could smear the Hankel functions at each x over ϕ_0, as in Sec. V B. Direct substitution indeed shows that the trial solution does not satisfy (5.2b); indeed, even the real part of $\Psi^*\mathcal{H}_i\Psi$ does not vanish. Thus, in quantum theory, the redundancy theorem, which states that if the energy constraint is satisfied, then the momentum constraint is automatically true, breaks downs in the long-wavelength limit as expected from Sec. IV. At the semiclassical level, however, it was found that subsets of the Hamilton-Jacobi solution space do satisfy the momentum constraint.

We do not know how to proceed with this long-wavelength formulation beyond the semiclassical level. Indeed, it may be that the fluctuations must be semiclassical if one neglects second-order spatial gradients, as signaled by commutation of the long-wavelength Hamiltonian densities at different spatial points. Within a full quantum treatment it may be that the quantum communication between short and large scales invalidates the concept of a wave functional obeying long-wavelength functional equations.[9] Since many of the issues raised by the problem of cosmic structure formation ultimately require us to address the quantum nature of Ψ to get $\mathcal{P}$, even an approximate quantum theory based on the long-wavelength formulation would be worthwhile.

VI. DISCUSSION AND CONCLUSIONS

Misner, Thorne, and Wheeler[21] extol the virtues of the Einstein-Hamilton-Jacobi (EHJ) equation, for containing, as it does, "as much information as all ten components of Einstein's field equations"—provided its solutions are properly parametrized—and for providing the shortest "leap from quantum to classical dynamics." In spite of this, the EHJ equation has not been widely to solve inhomogeneous problems in general relativity. This is similar to the situation in nonrelativistic dynamics in which the Hamilton-Jacobi formulation provides valuable insight, but is not generally a good calculational tool. The surprising thing in our long-wavelength application to inflation is that Hamilton-Jacobi methods are extremely useful for rapidly providing quantitative results (Sec. III D).

At the level of the first-order spatial gradient expansion explored in this paper, the Hamilton-Jacobi equation does not contain as much information as advertised. Although there is a general theorem that the momentum constraint is redundant in the full theory, it is not so for our truncated system. Accordingly, a crucial ingredient in our analysis is a careful treatment of the momentum constraint which, if one neglects the evolution of gravitational radiation, gives the Hubble parameter as a function of the scalar-field values and some integration parameters $H \equiv H(\phi_k, I_k)$. Without the momentum constraint, the long-wavelength equations describe each point evolving like a separate universe. The momentum constraint must be explicitly satisfied in order to provide the first-order patching required to glue the points together. The momentum constraint was explicitly shown to restrict the solutions beyond the full set of solutions of the zeroth-order equations.

The momentum constraint provides the first level of patching required. The inclusion of higher-order spatial derivatives modifies the energy constraint and scalar-field equations as well. The ability to make a consistent approximation scheme at higher order is explicitly explored in a third paper in this series.[36] There, in a nonlinear extension of the longitudinal gauge (the same α parametrization of the metric but a specific lapse defining the time foliation), second-order spatial gradient terms are included. This gives both linear perturbation theory at one level and these nonlinear long-wavelength equations at another, and provides a self-consistent arena for matching the short- and long-distance fluctuations inherent in the stochastic approach.

At this order, gauge issues are not significant. It is possible to include gravitational radiation, and it is not necessary to assume that the shift function vanishes.[23] However, hypersurface shifts do play an important role in our formulation. Although we showed that the choice of time variables is arbitrary at first order in the spatial gradient expansion, α is a natural time choice as long as the Universe continues expanding. Nonetheless, other choices for time hypersurfaces are more useful in certain physical contexts. To determine matching with the short-distance physics, it is natural to use α hypersurfaces to evaluate the ϕ fluctuations crossing the horizon. In SB2 we argue that $\ln(Ha)$ hypersurfaces, which are not that different from α hypersurfaces in inflating models, is a somewhat better choice in the stochastic framework. Once the initial conditions are set at horizon crossing, however, the propagation phase is much more easily calculated on H hypersurfaces; when one scalar field ϕ dominates the energy density, these are almost identical with the ϕ hypersurfaces during inflation, but as a time variable, H does not break down during scalar-field oscillation at the end of inflation. Inflation ends when $H(\phi_k) > \sqrt{3/2}H_{SR}(\phi_k)$, which translates to a critical value of the Hubble function or of the inflaton field; thus H surfaces tell us when the Universe reheated. On these surfaces α fluctuations are precisely the nonlinear analog of the ζ fluctuations of linear perturbation theory. We also described the hypersurface shifting techniques that are required to follow through this picture of initial conditions for ϕ fluctuations specified on α surfaces, but for propagation of α fluctuations occurring on a foliation of ϕ surfaces.

We found that the Hubble function solutions of the separated Hamilton-Jacobi equation were often multivalued and characterized by different integration parameters, but that they rapidly approach each other as expansion proceeds. The loss of memory of detailed initial conditions signals the presence of an attractor solution. That transients are decaying and often ignorable lies at the heart of the calculational approach we take to stochastic inflation in SB2. Here we have noted that analytic solutions of the evolution equations can often be found if one knows the detailed form of the attractor. The attractor may be as simple as the slow-rollover Hubble parameter, which is approximately valid in a wide class of inflation models. For this case we showed that fluctuations can be obtained analytically even for apparently complex models with many scalar fields of dynamical importance, for which we previously resorted to numerical integration (SBB), such as double inflation. The attractor can also differ substantially from the slow-rollover form, as we saw for inflation with an exponential potential. As several other researchers of inflation have found, and as we have extensively shown in this paper and also show in SB2, the exponential potential serves as a nice proving ground for ideas since it is so amenable to analytic solutions.

Probability functionals are fundamental to any stochastic description. To be meaningful they must be referred to the time hypersurface they are measured on, described here by conditional probabilities. We explored the transformation from one time surface to another and related the probabilities to the modulus of the wave functional of the Universe. Although the full quantum theory is far from complete, we showed that the long-wavelength approach is self-consistent at the semiclassical WKB level where one considers only the phase of the wave function. During the propagation (drift) phase of evolution, quantum effects are small for the fluctuations of relevance to our observable patch of the Universe, within the local Hubble volume.

This paper is really just one step to the larger whole of building a consistent short-long split for inhomogeneous early Universe field theories. Although the approximate equations that one uses cannot do justice to the full range of nonlinear behavior expected, we believe that they can be developed sufficiently to capture the essence of fluctuation generation and propagation in inflationary cosmologies, and yet be amenable to numerical calculation.

ACKNOWLEDGMENTS

We would like to thank Jim Bardeen, Alexei Starobinski, and Bill Unruh for stimulating discussions on some of these issues. D.S.S. was funded by NSERC of Canada, the U.S. Department of Energy, and NASA at Fermilab (Grant No. NAGW-1340). J.R.B. was supported by NSERC and a Canadian Institute for Advanced Research Fellowship. We also thank John Bahcall for the hospitality of the Institute for Advanced Study where some of this paper was written. Many of the results contained in this paper originate in the doctoral dissertation of D.S.S.[46] The long-wavelength formalism was first presented at the Computing Cosmologies Workshop hosted by J. R. Bond and N. Kaiser of the Canadian Institute for Theoretical Astrophysics at The University of Toronto, 1989.

APPENDIX A: TIME HYPERSURFACE INVARIANCE IN A FIRST-ORDER GRADIENT EXPANSION

In deriving the long-wavelength equations [Eqs. (2.13) and (2.14)], we did not make any special assumption about the time coordinate. these equations were valid for any time choice. In this appendix we show explicitly that the long-wavelength equations are invariant under an arbitrary time hypersurface transformation. In particular, we demonstrate that choosing a new time hypersurface $T \equiv T(t,x^j)$ does not change the form of the three-metric [Eq. (2.10)], provided our new spatial coordinates $X^j \equiv X^j(t,x^j)$ are projected orthogonally to surfaces of constant T, and if, further, one neglects second-order spatial gradients.

Assume that the new time surface $T(t,x^j)$ is arbitrary and choose a set of spatial coordinates X^j on a $T=T_0$ hypersurface. Spatial coordinates on the other surfaces of constant T will be labeled by orthogonal projection of these initial spatial coordinates, giving us a complete set of new coordinates. This prescription ensures that the metric components $g_{(T,X^j)}$ vanish. Along lines of constant X^j the old coordinates change according to $dx^\mu = T^{,\mu}ds$, where s is some arbitrary parameter, which we eliminate in favor of the new time parameter T:

$$dT = T_{,\alpha}dx^\alpha = T_{,\alpha}T^{,\alpha}ds \ ,$$

so that

$$\left[\frac{\partial x^\mu}{\partial T}\right]_{X^j} = \frac{T^{,\mu}}{T_{,\alpha}T^{,\alpha}} \ . \tag{A1}$$

We have thus found 4 of the 16 transformation derivatives relating the old and new coordinates. The remaining 12 will depend on the initial choice of X^j coordinates, and so we do not expect to obtain an explicit expression, but one can obtain differential equations for them which we now derive. The transformation matrix

$$B_k^\mu = \left[\frac{\partial x^\mu}{\partial X^k}\right]_T \tag{A2}$$

evolves in T according to

$$\left[\frac{\partial B_k^\mu}{\partial T}\right]_{X^j} = \left[\frac{\partial}{\partial T}\left[\frac{\partial x^\mu}{\partial X^k}\right]_T\right]_{X^j} = \left[\frac{\partial}{\partial X^k}\left[\frac{\partial x^\mu}{\partial T}\right]_{X^j}\right]_T = \left[\frac{\partial}{\partial X^k}\left[\frac{T^{,\mu}}{T_{,\alpha}T^{,\alpha}}\right]\right]_T \ .$$

Changing the derivative in X^j to one in x^μ gives

$$\left[\frac{\partial B_k^\mu}{\partial T}\right]_{X^j} = B_k^\nu \left[\frac{T^{,\mu}}{T_{,\alpha}T^{,\alpha}}\right]_{,\nu} \ . \tag{A3}$$

This expression allows us to verify that the orthogonality relation

$$B_k^\mu T_{,\mu} = 0$$

is constant in time; that is, the columns of the transformation matrix are tangent to a surface of constant T. The first row of the matrix

$$B_k^0 = -\frac{B_k^i T_{,i}}{T_{,0}} \tag{A4}$$

can be substituted into Eq. (A3) to yield

$$\left[\frac{\partial B_k^l}{\partial T}\right]_{X^j} = \left[\left[\frac{T^{,l}}{T_{,\alpha}T^{,\alpha}}\right]_{,m} - \frac{T_{,m}}{T_{,0}}\left[\frac{T^{,l}}{T_{,\alpha}T^{,\alpha}}\right]_{,0}\right] B_k^m \ . \tag{A5}$$

The right-hand side vanishes if we drop second-order spatial gradients; hence, to this order, $B_k^l \equiv B_k^l(X^j)$ is independent of T.

By integrating x^j along a line of constant X^j using (A1), we obtain

$$x^j = f^j(X^j) + \int \frac{T^{,j}}{T_{,0}T^{,0}}dT \ , \tag{A6}$$

in terms of an arbitrary spatial function $f^j(X^j)$, which for simplicity we take to be $f^j = X^j$. Thus x^j and X^j differ by a term that is first order in spatial gradients. A function g evaluated at x^j or X^j will then be equal up to this order:

$$g(x^j) = g\left[X^j + \int \frac{T^{,j}}{T_{,0}T^{,0}}dT\right] = g(X^j) + g_{,j}\int \frac{T^{,j}}{T_{,0}T^{,0}}dT = g(X^j) + \text{second-order terms} \ .$$

If we write the metric as $\gamma_{ij} = \exp(2\alpha)h_{ij}$, then the transformed $h_{i'j'}$ in the T-X system is a function only of the X^j coordinates:

$$\gamma_{k',l'} = e^{2\alpha(t,x^j)} B^k_{k'}(X^j) B^l_{l'}(X^j) h_{kl}(x^j)$$
$$= e^{2\alpha(t,X^j)} B^k_{k'}(X^j) B^l_{l'}(X^j) h_{kl}(X^j) \ .$$

This completes the proof that the form of the three-metric given by (2.10) is invariant under the arbitrary time hypersurface transformation since it may be written as a conformal factor multiplied by some three-metric which is only a function of the spatial coordinates X^j.

Finally, it is useful to note that there is a simple relation between time derivatives of a quantity Q on different hypersurfaces. Applying the chain rule and Eq. (A1), we find that

$$\left[\frac{\partial Q}{\partial T}\right]_{X^j} = \left[\frac{\partial Q}{\partial t}\right]_{x^j} \Big/ T_{,0} + \left[\frac{\partial Q}{\partial x^k}\right]_t T^{,k} \Big/ (T_{,0} T^{,0}) \ .$$

Neglecting second-order spatial gradients, the second term drops out, leaving Eq. (2.16), which was used to show that the evolution equations (2.14) and the energy constraint (2.13b) are invariant under arbitrary time hypersurface changes if one neglects second-order spatial gradients.

APPENDIX B: FUNCTIONAL MOMENTUM CONSTRAINT SOLUTIONS

We show that a general class of solutions of the functional momentum constraint equation (4.10b) for long-wavelength classical fields is

$$\mathcal{S} = \int d^3x \, e^{3\alpha(x^j)} F[\phi(x^j), \mathcal{J}] - \int^{\mathcal{J}} g(\mathcal{J}') d\mathcal{J}' \ , \quad \text{(B1a)}$$

where g and F are arbitrary functions of one and two variables, respectively. Given F and g, the functional $\mathcal{J}$ is defined implicitly through

$$g(\mathcal{J}) = \int d^3x \, e^{3\alpha(x^j)} F_{,2}[\phi(x^j), \mathcal{J}] \ . \quad \text{(B1b)}$$

The notation $F_{,2}$ denotes the derivative with respect to the second variable $\mathcal{J}$. If we set $F = -[m_P^2/(4\pi)] H[\phi(x^j), \mathcal{J}]$, as in (4.13), then this leads directly to Hamilton's principal function (4.15a). The functional $\mathcal{J}$ can be determined by an iterative method: Given the functions $\phi(x)$ and $\alpha(x)$, choose a sequence of $\mathcal{J}$'s for which (B1b) is satisfied to progressively higher accuracy.

In Sec. IV C we showed quite generally that the solution to the momentum constraint equation is given by Eq. (4.12):

$$\frac{\delta \mathcal{S}}{\delta \alpha(x^j)} = 3 e^{3\alpha(x^j)} F[\phi(x^j), \mathcal{J}(\phi,\alpha)] \ ,$$
$$\frac{\delta \mathcal{S}}{\delta \phi(x^j)} = e^{3\alpha(x^j)} F_{,1}[\phi(x^j), \mathcal{J}(\phi,\alpha)] \ . \quad \text{(B2)}$$

We can integrate $\mathcal{S}$ if we know what forms of the functional $\mathcal{J}$ are allowed. Nontrivial constraints on $\mathcal{J}$ follow from the integrability condition that derivatives commute in the infinite-dimensional gradient appearing in (B2). For example, $[\delta/\delta\phi(y^j), \delta/\delta\alpha(x^j)]\mathcal{S} = 0$ implies that

$$\frac{1}{e^{3\alpha(y^j)} F_{,1,2}(y^j)} \frac{\delta \mathcal{J}}{\delta \phi(y^j)} = \frac{1}{3 e^{3\alpha(x^j)} F_{,2}(x^j)} \frac{\delta \mathcal{J}}{\delta \alpha(x^j)} = \mathcal{C}$$

(B3a)

[the abbreviated notation $F(x^j)$ denotes the full expression appearing in (B2)]. The remaining cross derivatives are $[\delta/\delta\phi(x^j), \delta/\delta\phi(y^j)]\mathcal{S} = 0$ and $[\delta/\alpha(x^j), \delta/\delta\alpha(y^j)]\mathcal{S} = 0$:

$$\frac{1}{e^{3\alpha(y^j)} F_{,1,2}(y^j)} \frac{\delta \mathcal{J}}{\delta \phi(y^j)} = \frac{1}{e^{3\alpha(x^j)} F_{,1,2,}(x^j)} \frac{\delta \mathcal{J}}{\delta \phi(x^j)} = \mathcal{C} \ ,$$

(B3b)

$$\frac{1}{3 e^{3\alpha(y^j)} F_{,2}(y^j)} \frac{\delta \mathcal{J}}{\delta \alpha(y^j)} = \frac{1}{3 e^{3\alpha(x^j)} F_{,2}(x^j)} \frac{\delta \mathcal{J}}{\delta \alpha(x^j)} = \mathcal{C} \ .$$

(B3c)

The left side of (B3a) is a function only of y^j and the right-hand side is a function only of x^j, implying that each is just a functional $\mathcal{C}$, which is independent of y^j and x^j. Because there an infinite number of variables appearing here, it proves illuminating to first consider a finite model with only three spatial points, denoted by 1, 2, and 3. Define $a_1 = \alpha(1)$, $a_2 = \alpha(2)$, and $a_3 = \alpha(3)$, as well as $f_1 = \phi(1)$, $f_2 = \phi(2)$, and $f_3 = \phi(3)$. Equation (B3c) then becomes

$$\frac{\partial I}{\partial a_1} = 3 C e^{3a_1} F_{,2}(f_1, I) \ , \quad \text{(B4a)}$$

$$\frac{\partial I}{\partial a_2} = 3 C e^{3a_2} F_{,2}(f_2, I) \ , \quad \text{(B4b)}$$

$$\frac{\partial I}{\partial a_3} = 3 C e^{3a_3} F_{,2}(f_3, I) \ , \quad \text{(B4c)}$$

where I and C are functions of six variables, i.e., $I \equiv I(a_1, a_2, a_3, f_1, f_2, f_3)$. We can eliminate C from these expressions by dividing (B4a), and (B4b) by (B4c) and applying the relation

$$\left[\frac{\partial I}{\partial a_1}\right] \Big/ \left[\frac{\partial I}{\partial a_3}\right] = -\left[\frac{\partial a_3}{\partial a_1}\right]_I \ ,$$

yielding

$$\left[\frac{\partial}{\partial(e^{3a_1})}\right]_I [e^{3a_3} F_{,2}(f_3, I)] = -F_{,2}(f_1, I) \ ,$$

$$\left[\frac{\partial}{\partial(e^{3a_2})}\right]_I [e^{3a_3} F_{,2}(f_3, I)] = -F_{,2}(f_2, I) \ .$$

Holding I constant, one may integrate this equation:

$$e^{3a_1} F_{,2}(f_1, I) + e^{3a_2} F_{,2}(f_2, I) + e^{3a_3} F_{,2}(f_3, I) = g(I) \ ,$$

with $g(I)$ an arbitrary function of I. Repeating this analysis for (B3a) and (B3b) leads to the same conclusion. In the infinite-dimensional case, the sum over all spatial points becomes an integral [Eq. (B1b)], which is the general form of the functional $\mathcal{J}$ consistent with (B2).

We now wish to integrate (B2) to determine the phase function $\mathcal{S}$. Define the functional $\mathcal{Q}$ by

$$\mathcal{S} = \int d^3x \, e^{3\alpha(x^j)} F[\phi(x^j), \mathcal{J}] - \mathcal{Q} \ . \quad \text{(B5)}$$

When substituted into (B2) this gives

$$\frac{\delta Q}{\delta\phi(y^j)} = \frac{\delta \mathcal{J}}{\delta\phi(y^j)} \int d^3x\, e^{3\alpha(x^j)} F_{,2}[\phi(x^j), \mathcal{J}] ,$$

$$\frac{\delta Q}{\delta\alpha(y^j)} = \frac{\delta \mathcal{J}}{\delta\alpha(y^j)} \int d^3x\, e^{3\alpha(x^j)} F_{,2}[\phi(x^j), \mathcal{J}] .$$

Since the integral appearing in these equations is just $g(\mathcal{J})$, the solution

$$Q = \int g(\mathcal{J}')d\mathcal{J}'$$

follows, which leads to (B1). Although we have thus shown that Eq. (B1) represents a general class of solutions, we do not know if all solutions are of this form.

[1]V. De Lapparent, M. J. Geller, and J. P. Huchra, Astrophys. J. **302**, L1 (1986).
[2]T. J. Broadhurst, R. S. Ellis, D. C. Koo, and A. S. Szalay, Nature **343**, 726 (1990).
[3]S. J. Maddox, G. Efstathiou, W. J. Sutherland, and J. Loveday, Mon. Not. R. Astron. Soc. **242**, 43P (1990).
[4]N. Bahcall and R. Soneira, Astrophys. J. **270**, 70 (1983).
[5]A. Dressler, S. M. Faber, D. Burstein, R. L. Davies, D. Lynden-Bell, R. J. Terlevich, and G. Wegner, Astrophys. J. Lett. **313**, L37 (1986).
[6]L. A. Kofman and A. D. Linde, Nucl. Phys. **B282**, 555 (1987).
[7]D. S. Salopek, J. R. Bond, and J. M. Bardeen, Phys. Rev. D **40**, 1753 (1989).
[8]A. Vilenkin, Phys. Rev. D **27**, 2848 (1983); A. A. Starobinski, in *Field Theory, Quantum Gravity, and Strings,* proceedings of the Seminar, Meudon and Paris, France, 1984–1985, edited by H. J. de Vega and N. Sánchez, Lecture Notes in Physics Vol. 246 (Springer, New York, 1986), p. 107; J. M. Bardeen and G. J. Bublik, Class. Quantum Grav. **5**, L113 (1988); A. S. Goncharov, A. D. Linde, and V. F. Mukhanov, Int. J. Mod. Phys. A **2**, 561 (1987).
[9]B. L. Hu, Maryland University report, 1990 (unpublished).
[10]D. S. Salopek and J. R. Bond, referred to as SB2 (unpublished).
[11]S. W. Hawking, Phys. Lett. **115B**, 295 (1982).
[12]A. A. Starobinski, Phys. Lett. **117B**, 175 (1982).
[13]A. Guth and S.-Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982).
[14]J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. D **28**, 679 (1983).
[15]E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz. Nauk **80**, 391 (1964) [Sov. Phys. Usp. **6**, 495 (1964)].
[16]D. Eardley, E. Lang, and R. Sachs, J. Math. Phys. **13**, 99 (1972).
[17]V. Moncrief and C. Teitelboim, Phys. Rev. D **6**, 966 (1972).
[18]J. J. Halliwell, Int. J. Mod. Phys. A **5**, 2473 (1990).
[19]M. Pilati, Phys. Rev. D **26**, 2645 (1982).
[20]C. Teitelboim, Phys. Rev. D **25**, 3159 (1982).
[21]C. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[22]J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).
[23]D. S. Salopek (unpublished).
[24]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1981).
[25]F. Lucchin and S. Matarrese, Phys. Rev. D **32**, 1316 (1985).
[26]B. Ratra, Princeton University Report No. 1102, 1989 (unpublished).
[27]J. J. Halliwell, Phys. Lett. B **185**, 341 (1987).
[28]W. G. Unruh and R. M. Wald, Phys. Rev. D **40**, 2958 (1989).
[29]A. D. Linde, Phys. Lett. **108B**, 389 (1982).
[30]A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).
[31]A. A. Starobinski, Pis'ma Zh. Eksp. Teor. Fiz. **42**, 124 (1985) [JETP Lett. **42**, 152 (1985)].
[32]U. Gerlach, Phys. Rev. **177**, 1929 (1969).
[33]K. Kuchar, in *Quantum Gravity 2,* edited by C. Isham, R. Penrose, and D. W. Sciama (Clarendon, Oxford, 1981).
[34]B. S. DeWitt, Phys. Rev. **160**, 1113 (1967).
[35]H. E. Kandrup, Class. Quantum Grav. **5**, 903 (1988).
[36]J. R. Bond and D. S. Salopek (unpublished).
[37]J. J. Halliwell and S. Hawking, Phys. Rev. D **31**, 1777 (1985).
[38]A. Vilenkin, Phys. Rev. D **33**, 3560 (1986).
[39]A. Vilenkin, Phys. Rev. D **37**, 888 (1988).
[40]U. Carrow and S. Watamura, Phys. Rev. D **32**, 1290 (1985).
[41]S. W. Hawking and D. N. Page, University of Alberta report, 1990 (unpublished).
[42]N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
[43]M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
[44]J. B. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983).
[45]J. J. Halliwell, Phys. Rev. D **39**, 2912 (1989).
[46]D. S. Salopek, Doctoral dissertation, University of Toronto, 1989.