Designing density fluctuation spectra in inflation

D. S. Salopek
Department of Physics, University of Toronto, Toronto, Canada M5S 1A7

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

J. M. Bardeen
Department of Physics, University of Washington, Seattle, Washington 98195
(Received 3 November 1988)

Scale-invariant (flat) fluctuation spectra are the most natural outcomes of inflation. Nonetheless current large-scale-structure observations seem to indicate more fluctuation power on large scales than flat spectra give. We consider a wide variety of models based on the chaotic inflation paradigm and sketch the effects that varying the expansion rate, structure of the potential surface, and the curvature coupling constants have on the quantum fluctuation spectra. We calculate in detail the quantum generation of fluctuation spectra by numerically solving the linearized perturbation equations for multiple scalar fields, the metric, and the radiation into which the scalars dissipate, following the evolution from inside the horizon through reheating. We conclude that (1) useful extended nonflat power laws are very difficult to realize in inflation, (2) double inflation leading to a mountain leveling off at a high-amplitude plateau at long wavelengths is generic, but to tune the cliff rising up to the plateau to lie in an interesting wavelength range, a special choice of initial conditions and/or scalar field potentials is required, and (3) small mountains (moguls) on the potential surface lead to mountains of extra power in the fluctuations added on top of an underlying flat spectrum. For quadratic and quartic couplings, the mountain fluctuations may obey Gaussian statistics but the spectral form will be very sensitive to initial conditions as well as potential parameters; non-Gaussian mountain fluctuations which depend upon potential parameters but not on initial field conditions will be the more likely outcome. However, adding cubic couplings can give mountains obeying Gaussian statistics independently of initial conditions. Since observations only probe a narrow patch of the potential surface, it is possible that it is littered with moguls, leading to arbitrarily complex "mountain range" spectra that can only be determined phenomenologically. We also construct an inflation model which houses the chaotic inflation picture within the grand unified theory (GUT) framework. The standard chaotic picture requires an unnaturally flat scalar field potential, $\lambda \approx 5\times 10^{-14}$, and a strong curvature coupling parameter bound, $\xi < 0.002$. By allowing the Higgs field to be strongly coupled to gravity through a large negative curvature coupling strength, $\xi \sim -10^4$, so the Planck mass depends on the GUT Higgs field, the Higgs field can be strongly coupled to matter fields [with $\lambda \sim (\xi/10^5)^2$]. This leads to both a flat Zeldovich spectrum of the "observed" amplitude and a high reheating temperature ($\sim 10^{15}$ GeV), unlike the $\lambda \sim 10^{-13}$ standard case. The large $-\xi$ would be related to the ratio of the Planck scale to a typical GUT scale. Although a single dynamically important Higgs multiplet gives flat spectra, a richer Higgs sector could lead to broken scale invariance.

I. INTRODUCTION

One of the most promising inflationary models is Linde's chaotic scenario.[1] In this model, the Universe emerges from the Planck epoch in a chaotic state, where any value of the scalar field consistent with a density $\ll m_P^4$ is allowed. In particular, there exist small patches of space where derivatives of the field are tiny (e.g., near a peak), and where the energy density of the Universe is potential dominanted, $\dot\phi^2, (\nabla\phi)^2 \ll V(\phi)$. The criterion for inflation, that the distance between any two points accelerates, $\ddot a/a = -(4\pi G)(p+\rho/3) > 0$, rather than decelerates, is then satisfied. The comoving Hubble distance $(Ha)^{-1}$ thus decreases with time, sweeping inward to encompass ever smaller comoving length scales k^{-1}. Once $(Ha)^{-1}$ drops below k^{-1}, the "Hubble damping" of the quantum fluctuations spontaneously generated before this on scale k^{-1} is arrested. The result is that the ϕ fluctuations form a homogeneous and isotropic Gaussian random field which is completely characterized by its power spectrum $\mathcal{P}_\phi(k)$, the variance in ϕ per logarithmic interval of k. The natural outcome of chaotic inflation driven by a single scalar field is a flat scale-invariant spectrum $\mathcal{P}_\phi(k)$ for $k^{-1}>(Ha)^{-1}$. A region of size smaller than the Planck length is inflated by a factor $\sim e^{60+N_I}$ to a region (much) larger than the current Hubble length $H_0^{-1}=3000\ h^{-1}$Mpc provided the number of e-foldings during inflation is $N_I \gtrsim 70$. Without inflation, the region would have expanded by only a factor $\sim e^{60}$ to less than a micron across by the current time—an aspect of the "horizon" problem which inflation so successfully solves. In spite of its naturalness, researchers have been slow to accept chaotic inflation because it relies on Planck era physics since a large ($\sim 5m_P$) initial value for the scalar field is required to ensure enough inflationary expansion. However, perhaps we should adopt the view that the structure of the Universe is one of the few conceivable windows we have to constrain Planck-scale physics.

An inflationary cosmology with an initially scale-invariant spectrum is completely defined (1) by a "biasing factor" b_ρ parametrizing the amplitude of the fluctuations and (2) by specifying such global parameters as the Hubble constant

$$h \equiv H_0/(100\ \mathrm{km\,s^{-1}/Mpc}) \tag{1.1}$$

and the mass densities (relative to closure density) of the various constituents of the Universe, baryons Ω_B, dark matter of various types Ω_X, relativistic particles Ω_{er}, and vacuum energy Ω_{vac} (nonzero Λ). These relative densities must all sum to a total Ω which is necessarily very nearly unity in inflation models. The post-inflation evolution of the fluctuations prior to the onset of nonlinearity when the first objects collapse in the Universe is then completely determined, characterized by a power spectrum for density fluctuations, $\mathcal{P}_\rho(k)$. The most successful scenario to date is the cold-dark-matter (CDM) model, with $\Omega_X \sim 0.9$, with a baryon density $\Omega_B \sim 0.1$, with $h \approx 0.5$ and $b_\rho \sim 1.4$–2.5. The CDM model does very well at explaining the observed structure on scales $\lesssim 10\ h^{-1}$Mpc (Refs. 2 and 3), but seems to conflict with some observations probing larger scales, such as the clustering of Abell clusters, as measured by the cluster-cluster correlation function: at a separation of 25 h^{-1}Mpc, the Bahcall and Soneira[4] result gives an amplitude of unity, while the CDM prediction is about an order of magnitude less [Bardeen, Bond, and Efstathiou[5] (BBE)]. Furthermore, there are indications from the apparently large coherent streaming velocity[6] of galaxies within a distance of order 25 h^{-1}Mpc from us that the CDM biasing factor would have to be nearly 1 to accommodate the observations.[7] Low values of $b_\rho(\sim 1)$ are inconsistent with the inflation requirement that $\Omega \approx 1$ (Ref. 3). There are other observations which do not indicate a lack of power in the CDM model. For example, the angular clustering of galaxies about galaxies and of galaxies about clusters are both now apparently compatible with the CDM model.[8–12] Taken all together, the data as currently interpreted would impose severe restrictions on the form of the density fluctuation spectrum that could give rise to it. However, we are not yet at the state observationally to definitively determine the form of the fluctuation spectrum from observations, and should regard the data as indicating what spectral shapes it would be desirable to construct theoretically.

One could accommodate the observations by doing either of two things. First, one may consider alternative inflationary models that produce more exotic primordial ("initial" or post-inflation) density fluctuation spectra, $\mathcal{P}_\rho^{(i)}(k)$. Second, one may alter the transfer function $T(k)$ which describes how physical processes working within the horizon affect density fluctuations. The spectrum prior to the onset of nonlinearity is $\mathcal{P}_\rho(k)=T^2(k)\mathcal{P}_\rho^{(i)}(k)$. BBE have systematically varied the free parameters of $T(k)$, namely, Ω_B, Ω_X, Ω_{vac}, h, the lifetime of massive neutrinos, and the biasing factor b_ρ. They find no natural satisfactory combination, although CDM with large baryon density (violating nucleosynthesis constraints) or short-lived neutrinos (little motivation from particle physics) come close. The first possibility was then considered by BBE and Bond:[12] phenomenological primordial spectra with more power at large scales than the scale-invariant spectrum were constructed, independently of the requirements of consistency with inflationary models (CDM+plateau and CDM+mountain, see Sec. II). In this paper, we discuss the possibility of designing general density fluctuation spectra in the context of chaotic inflationary models with multiple scalar fields. Since the present large-scale structure observations are tentative, we will not limit ourselves to these phenomenological spectra.

The standard chaotic model is unsatisfactory from a theoretical standpoint as it requires fine-tuning of the cosmological constant $\Lambda < 10^{-120} m_P^2$, the scalar field self-interaction $\lambda \approx 10^{-13}$, and the curvature coupling parameter $\xi < 0.002$. [The last constraint follows from requiring that the effective value of Newton's constant, $G_{\mathrm{eff}}=(m_P^2-8\pi\xi\phi^2)^{-1}$ be greater than zero.[13,14]] The problem of how a scalar field which interacts only extremely weakly can reheat the Universe has also not been solved. We explore the density fluctuations produced by theories which strongly couple scalar fields to gravity in an attempt to explain these fine-tunings.

In Sec. II we transform the phenomenological density fluctuation spectra suggested by the large-scale structure data into a form more suited for comparing with the output of our computations of post-inflationary spectra, namely, in terms of the gauge-invariant variable ζ which is a measure of the gravitational potential fluctuations. ζ must have an amplitude about $10^{-4}/b_\rho$ to explain the current level of structure in the Universe. In Sec. III we present our general Lagrangian for multiple scalar fields interacting among themselves through a potential and coupled nonminimally to gravity. With nonminimal couplings, Newton's constant varies with position. Using a conformal transformation, which is interpreted as a position-dependent change of units, it becomes constant, and the action transforms to that of standard Einsteinian gravity with minimally coupled scalar fields. To describe the development of fluctuations, we adopt a framework which is consistent within linear perturbation theory in which the background fields are treated classically and the fluctuations quantum mechanically. The high degree of isotropy of the microwave background justifies this approach. We derive the perturbation equations in Sec. IV and their initial conditions in Sec. V, many of which are set by quantum requirements.

In Sec. VI the heart of this paper, we describe the models we have considered. In Sec. VI A we present numerical calculations of fluctuations for a single scalar field interacting through a Coleman-Weinberg potential of "new" inflation and a chaotic ($\sim\lambda\phi^4$) potential. Scalar field quantum fluctuations are created inside the horizon and are transformed into density fluctuations after the Universe reheats. In Secs. VI B–VI D we generalize to multiple fields. Using a simple but illuminating model, we illustrate how features of various shapes may arise in the fluctuation spectrum. It is possible to generate fluctuation spectra with more power at large scales through double inflation,[15–18,13] though to tune the characteristic wave number to be around the scale associated rich clusters would require highly selected initial conditions for the scalar fields. We also explore the case of two scalar fields interacting through a general quartic potential as a means of breaking scale invariance in inflation. We show that extra power localized over a limited wavelength range (mountains) is the most likely way to break scale invariance.[16,18,13,19]

In Sec. VII we consider models in which Newton's constant is a function of a GUT Higgs field, determined by a $[m^2/(16\pi)-\xi\phi^2/2]R$ coupling of the Higgs field to gravity. The standard result for the amplitude of density fluctuations at horizon crossing (parametrized by the gauge-invariant variable ζ), $\zeta=400\lambda^{1/2}$ is now replaced by $\zeta=10\lambda^{1/2}/|\xi|$, valid for large $-\xi$. Thus, if, for example, $\lambda\approx5\times10^{-2}$, and $\xi=-2\times10^4$, density fluctuations are of the "observed" level. Particularly interesting is the induced gravity model[20–22] in which the bare value of the Planck mass m is set to zero and its present value is generated when the scalar field moves to the bottom of its potential. However, this model has difficulties with reheating since the scalar decouples from gauge fields (at least at the classical level). We construct a closely related model with $m\sim m_P$ which does not require the unnaturally small value of λ required by chaotic inflation, and has prompt reheating at the GUT scale. After this paper was written, we learned that the large $\xi<0$ fluctuation formula was obtained using different techniques by Spokoiny and that Fakir and Unruh have also independently derived it.[23] These papers do not address how this can aid in meshing chaotic inflation in a GUT scenario. These models still suffer from the cosmological constant problem and from an uncertain origin for the required large negative curvature coupling constant.

(In this paper, we loosely speak of fluctuations leaving the horizon and re-entering the horizon during inflation; what we mean is that $k<Ha$ and $k>Ha$. By the second horizon crossing, we mean the second time ak^{-1} equals the Hubble parameter. When we speak of the horizon scale now, we mean the comoving distance that light could have traveled since the end of inflation. Thus the horizon scale now is 6000 h^{-1}Mpc for $\Omega=1$ cosmologies. This imprecise terminology has by now become conventional.)

II. PHENOMENOLOGICAL POST-INFLATION FLUCTUATION SPECTRA

In this section, we normalize possible post-inflation Gaussian fluctuation spectra allowed by the data as currently interpreted. Bardeen, Bond, and Efstathiou,[5] Bond[24,12] and Bond and Couchman[10] have used the tests of large-scale streaming velocities, the cluster-cluster, cluster-galaxy, and galaxy-galaxy correlation functions, $\xi_{cc},\xi_{cg},\xi_{gg}$, and constraints on microwave background (CMB) anisotropies to construct a phenomenological linear density fluctuation spectrum $\mathcal{P}_\rho(k)$ which, when evolved, satisfies all of the current cosmic structure data. The straightforward interpretation of the data is that it is largely determined by different regions of the density fluctuation spectrum $\mathcal{P}_\rho$. However, we cannot prescribe a unique spectrum, or even upper and lower limits on allowed variations that hold over the entire spectral range of interest for structure formation since other processes such as fragmentation or explosions may have contributed to the formation and growth of structure. The best bet we have for a simple mapping of observations to the form of $\mathcal{P}_\rho$ is the large-scale structure data; in this regime, the density fluctuations are difficult to generate by nongravitational means (e.g., explosions), and the fluctuations are apparently still linear in amplitude. For example, limits on large-angle microwave background anisotropies determined for the Soviet RELICT experiment as reported by Strukov, Skulachev, and Klypin[25] place a stringent constraint on how large $\mathcal{P}_\rho$ may be in roughly the wave-number range 600 $h^{-1}\mathrm{Mpc}\lesssim k^{-1}\lesssim 6000$ h^{-1}Mpc, while the cluster-cluster correlation function data would restrict the range from ~5 h^{-1}Mpc to about ~50 h^{-1}Mpc. Within hierarchical models for structure formation, in which galaxies form before clusters, as the data would seem to indicate, ξ_{gg} and ξ_{cg} would constrain wave numbers between ~5 h^{-1}Mpc and ~20 h^{-1}Mpc; it is not completely clear at present that these spectral limitations are compatible with the ξ_{cc} spectral constraints. This is discussed more fully below and in Bond.[12]

The linear spectrum $\mathcal{P}_\rho(k)$ (ρ variance per $\ln k$) is fixed in shape once the Universe enters into a phase where it is dominated by pressureless nonrelativistic matter; the shape is retained even if it subsequently becomes vacuum dominated or curvature dominated. The transition from the initial post-inflation spectrum $\mathcal{P}_\rho^{(i)}(k)$ to $\mathcal{P}_\rho(k)$ is described by the transfer function $T(k)$: $\mathcal{P}_\rho(k,t)=D^2(t)T^2(k)\mathcal{P}_\rho^{(i)}(k)$. Here $D(t)$ is an overall growth factor (e.g., Peebles,[26] Sec. 2.3). For wave numbers which enter the horizon in the matter-dominated epoch, $T(k)=1$. For shorter scales it is necessarily less than one. It is conventional to define a local power-law index n to characterize the shape of $\mathcal{P}_\rho(k)$:

$$\mathcal{P}_\rho(k)=\frac{k^3}{2\pi^2}\left\langle\left|\frac{\delta\rho}{\rho}(k)\right|^2\right\rangle\propto k^{3+n}\ . \tag{2.1}$$

The average, defined more precisely in Sec. IV, is taken with respect to the random field ensemble. The $2\pi^2$ factor depends upon the Fourier transform convention; we use $\delta\rho(\mathbf{k})/\rho=\int d^3x\,[\delta\rho(\mathbf{x})/\rho]\exp(-i\mathbf{k}\cdot\mathbf{x})$. For initially adiabatic perturbations, the scale invariant (Zeldovich) spectrum has $n=1$. A clearer way of presenting the phenomenological spectra which emphasizes the scale invariance is to use the linear power spectra for the gravitational potential $\mathcal{P}_{\Phi_H}$, defined by

$$\mathcal{P}_{\Phi_H} \equiv \frac{k^3}{2\pi^2} \langle |\Phi_H(k)|^2 \rangle \ . \tag{2.2a}$$

The potential perturbations are related to the comoving energy density fluctuation $\delta\rho$ by[27]

$$\nabla^2 \Phi_H = -4\pi G a^2 \delta\rho \ , \tag{2.2b}$$

$$\mathcal{P}_{\Phi_H} = \frac{9}{4} \left[\frac{\Omega^{1/2} H a}{k} \right]^4 \mathcal{P}_\rho \ . \tag{2.2c}$$

Thus, for adiabatic spectra for which $n=1$ initially, $\mathcal{P}_{\Phi_H}$ is k independent on large scales, and just reflects the shape of $T^2(k)$ on shorter scales. Further, in the matter-dominated phase, $\mathcal{P}_{\Phi_H}$ is time independent (although the amplitude, but not the shape, does evolve in subsequent vacuum- or curvature-dominated eras).

Although Φ_H has the advantage of a simple physical interpretation as a gravitational potential obeying the Poisson-Newton equation (2.2b), it is also a gauge-invariant variable in general relativistic perturbation theory. We use it extensively in our numerical computations in Sec. VI. However, there is an even more useful gauge-invariant measure of the amplitude of the metric fluctuations arising in the post-inflation era, $\zeta_{\rm BST}$, introduced by Bardeen, Steinhardt, and Turner.[28] In this paper, following Bardeen, Bond, Kaiser, and Szalay,[3] we introduce a new variable ζ which is a factor of 3 larger than $\zeta_{\rm BST}$ so that ζ reduces to $\delta\rho/\rho$ well inside the horizon for nonrelativistic matter (see also Bardeen[29]):

$$\begin{aligned} \zeta &\equiv 3\zeta_{\rm BST} \\ &\equiv \frac{2}{1+p/\rho}(\Phi_H + H^{-1}\dot{\Phi}_H) \\ &\quad + \left[1 + \frac{2k^2}{9a^2H^2}\frac{1}{1+p/\rho} \right] 3\Phi_H \ . \end{aligned} \tag{2.3}$$

Here p and ρ refer to the total pressure and energy density of the background matter including that of any scalar fields present. ζ is also hypersurface invariant for perturbations outside of the horizon, and is only large ($|\zeta|>1$) if either the matter perturbation or the metric perturbation is large. Furthermore, ζ has the remarkable property of remaining constant in time for adiabatic perturbations $[\delta p/\delta\rho = (dp/dt)/(d\rho/dt)]$ outside of the horizon during the radiation- and matter-dominated phases, and even through inflation and reheating.

For scales just outside our present horizon, the value of $\mathcal{P}_\zeta$ is simply related through (2.3) (with $p=0$) to the value of $\mathcal{P}_{\Phi_H}$,

$$\mathcal{P}_\zeta^{1/2} = 5\mathcal{P}_{\Phi_H}^{1/2} = \left[\frac{15}{2}\frac{\delta\rho}{\rho} \right]_H \ . \tag{2.4}$$

The characterization of the amplitude of the fluctuations by the rms value on the horizon is the conventional (but imprecise) way of presenting the spectrum normalization.

In Fig. 1 we plot $\mathcal{P}_{\Phi_H}$ for two generic phenomenological spectra (CDM+pl and CDM+mt, where pl stands for plateau and mt stands for mountain); these are compared with the CDM model and some variants constructed from scale-invariant spectra ($\mathcal{P}_{\Phi_H}$=const), but with different cosmological parameters. The phenomenological models preserve the success of the CDM model on small (galaxy to cluster) scales by maintaining the same flat initial spectrum at high k with essentially the same amplitude as the CDM spectrum. To agree with the large-scale data, an *ad hoc* $n=-1$ rise from a scale k_u^{-1} of order 25 h^{-1}Mpc to $\sim$200 h^{-1}Mpc is added, followed by a return to the flat spectrum, either by leveling off to a plateau (CDM+pl) or by dropping down to the same lev-

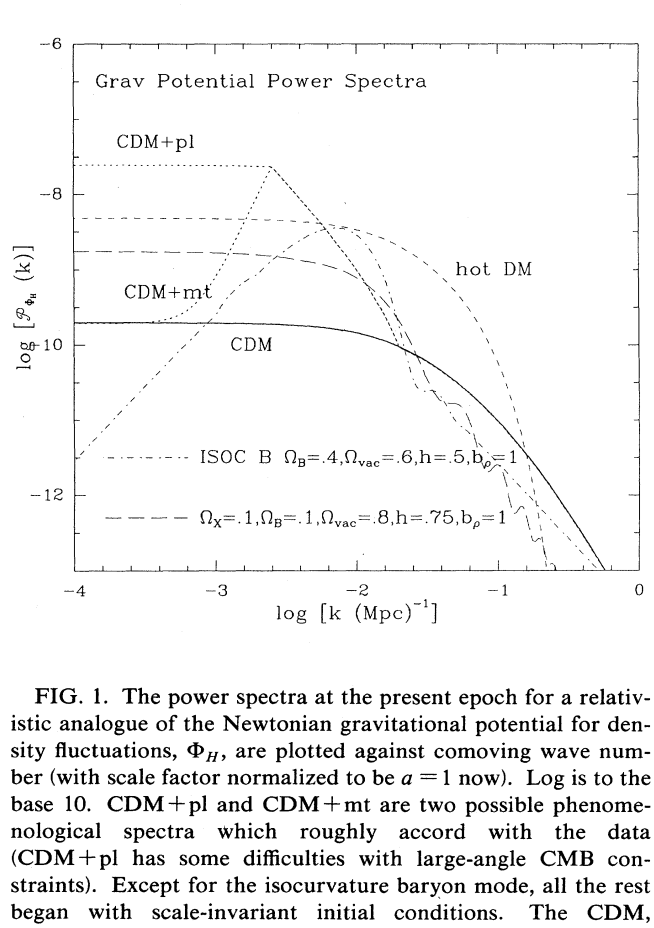

FIG. 1. The power spectra at the present epoch for a relativistic analogue of the Newtonian gravitational potential for density fluctuations, Φ_H, are plotted against comoving wave number (with scale factor normalized to be $a=1$ now). Log is to the base 10. CDM+pl and CDM+mt are two possible phenomenological spectra which roughly accord with the data (CDM+pl has some difficulties with large-angle CMB constraints). Except for the isocurvature baryon mode, all the rest began with scale-invariant initial conditions. The CDM, CDM+pl and CDM+mt models are all normalized with the (rather low) biasing factor $b_\rho=1.44$ (i.e., $b_\rho^2\mathcal{P}_{\Phi_H}$ would be the power spectrum if the galaxies are good tracers of the mass distribution). The hot-dark-matter model has $b_\rho=0.53$ corresponding to a redshift of 1 when rms density fluctuations become nonlinear. The low $\Omega_{\rm nr}$ model is normalized assuming galaxies are fair tracers of the mass. The isocurvature baryon model, with initial spectral index $n_i=-1$, is also normalized this way. To make $\Omega=1$ in this model, $\Omega_{\rm vac}=0.6$ is required. This does not change the normalization from that of the open model. These spectra are also proportional to $T^2(k)$.

el as at high k (mountain, CDM+mt). The starting point for the rise must be delayed to well beyond cluster scale to accommodate the observations of the drop from a power law in ξ_{gg} as inferred from the angular correlation function of galaxies.[8-10] The plateau or mountain cannot be too high or else intermediate angle $\Delta T/T$ constraints are violated.[5,24] The normalization of these models and of the standard biased CDM model is given by the biasing factor $b_\rho=1.44$, at the lower limit of the range required to make the CDM model compatible with $\Omega=1$.

The spectra have one overall normalization constant. It is conventionally parametrized by the biasing factor b_ρ, determined by relating mass fluctuations $\Delta M/M$ evaluated assuming linear theory to fluctuations in the number of galaxies $\Delta N_g/N_g$, suitably averaged over some large scale r. More precisely, we use "J_3 normalization":[30]

$$\left\langle \frac{\Delta M}{M}(<r)\frac{\delta\rho}{\rho}(0)\right\rangle \equiv \frac{1}{b_\rho^2}\left\langle \frac{\Delta N_g}{N_g}(<r)\frac{\delta n_g}{n_g}(0)\right\rangle \approx \frac{0.81}{b_\rho^2}\ , \quad r=10\ \mathrm{h}^{-1}\mathrm{Mpc}\ .$$

The normalization for the scale-invariant spectrum is, for the CDM model,

$$\mathcal{P}_\zeta^{1/2}=10^{-4}/b_\rho\ . \tag{2.5}$$

As mentioned in the Introduction, b_ρ is currently not yet well determined by the model, and could be anywhere in the range ~ 1.4–2.5. The normalization for the hot dark-matter (massive neutrino-dominated) model shown in Fig. 1 is essentially identical, except that $b_\rho \lesssim 0.5$, the lower value being required to ensure galaxy formation occurs sufficiently early. (This model can be ruled out by small-angle CMB constraints.[31,32])

Large-angle CMB anisotropy experiments strongly constrain allowed extra power models. For example, in universes with $\Omega=\Omega_{nr}=1$, for any spectrum which is scale invariant for $k^{-1}>300\ \mathrm{h}^{-1}\mathrm{Mpc}$, the resolution scale of the RELICT experiment, the constraint on how large ζ can be is quite stringent:[25,24]

$$\mathcal{P}_\zeta^{1/2}<10^{-3.6},\quad 95\%\ \text{confidence limit}\ . \tag{2.6}$$

In the plateau case shown in Fig. 1, to which this limit applies, the $\mathcal{P}_\zeta(k)$ spectrum rises from $\mathcal{P}_\zeta^{1/2}=10^{-4}/b_\rho$ at short distances to $\mathcal{P}_\zeta^{1/2}\approx 10^{-3}/b_\rho$, therefore violating (2.6). Having a lower amplitude plateau could alleviate this problem, but at the expense of a drop in ξ_{cc} beyond $\sim 50\ \mathrm{h}^{-1}\mathrm{Mpc}$ (where the data cannot be believed anyway). These phenomenological spectra can be obtained by a finely tuned fluctuation generation mechanism which imprints the scale k_u^{-1} in the initial spectrum $\mathcal{P}_{\Phi_H}^{(i)}$, as discussed in Secs. VI B and VI C. For the mountain case, $\mathcal{P}_\zeta^{1/2}\approx 10^{-3}/b_\rho$ is approximately the peak value we would be shooting for. The large-angle CMB constraints are at present not as strong for mountain spectra, and indeed there is a reported observation[33] of anisotropy on intermediate angular scales, probing the spectral region around $k^{-1}\sim 300\ \mathrm{h}^{-1}\mathrm{Mpc}$.

A more conservative approach is to relate k_u^{-1} to known physical scales that are imprinted upon the transfer function $T(k)$ from evolutionary effects. Since $T(k)$ must fall below its small-k value of unity, such spectra must be of the plateau variety. Therefore we may expect it to be difficult to avoid excessive large-angle CMB anisotropies in such models. Further, it is unlikely[10] that any $T(k)$ modification can give both the spectral drop beyond cluster scale required to reproduce the ξ_{gg} data and the rise to push up the large-scale amplitude of ξ_{cc}.

To illustrate the problems, consider the effects of varying the baryon and vacuum energy contents of the Universe on $T(k)$. A reasonable variant of the CDM model would be one with the baryon abundance increased to the limit imposed by primordial nucleosynthesis constraints, $\Omega_B\sim 0.2$ (Yang *et al.*[34]). However, only larger variations which violate the nucleosynthesis constraint, $\Omega_B\sim 0.5$ give significant large-scale structure in $T(k)$. An extreme example of adding a large vacuum energy ($\Omega_{vac}=0.8$ and $\Omega_{nr}=0.2$) and also increasing the Ω_B/Ω_X ratio to 1 is shown in Fig. 1. Although the features in $\mathcal{P}_{\Phi_H}$ do give ample large-scale power, the price is violation of small angle microwave background anisotropies,[5] a low redshift of galaxy formation, and, of course, unpalatable assumptions. Further, although this model does give an adequate ξ_{cc}, it does not give[10] the break inferred observationally for ξ_{gg}, the problem mentioned above.

Another conservative approach is to assume the initial perturbations are isocurvature ($\mathcal{P}_{\Phi_H}^{(i)}=0$), with fluctuations in some component of the matter which eventually is an important contributor to the mass density. Examples are isocurvature baryon perturbations, in which fluctuations in the baryon density are compensated initially by opposite fluctuations in the rest of the relativistic plasma present, ensuring that the net comoving density perturbation vanishes. In this case, the initial perturbations are characterized by the power spectrum

$$\mathcal{P}_{n_B}^{(i)}=\frac{k^3}{2\pi^2}\left\langle \left|\frac{\delta n_B}{n_B}(k)\right|^2\right\rangle \propto k^{3+n_i}\ . \tag{2.7}$$

Similarly, in CDM models there can be isocurvature axion perturbations,[35] with a perturbed axion mass density initially opposed by energy fluctuations in the quark-gluon plasma. The relevant initial power spectrum to be specified is $\mathcal{P}_{\rho,\mathcal{A}}^{(i)}$. Any pseudo-Goldstone boson candidate which attains a nonzero mass in the post-inflation era may give such perturbations. The transfer function relating initial to final spectra now has $T(k)=1$ on short scales, before the perturbed baryons or axions begin to dominate the density, dropping as k^2 on large scales once they do.

Initially scale-invariant isocurvature perturbations have an initial spectral index $n_i=-3$. In this case, both isocurvature axion and baryon perturbations can be convincingly ruled out since they violate large-angle CMB anisotropy constraints if their amplitude is large enough to generate structure on smaller scales.[36,37] For such models to be viable, scale invariance must therefore be broken. A plateau variation would not be viable since it would give worse CMB anisotropy problems. A mountain variation could avoid this difficulty only if its amplitude plunged below the short-distance value, somewhere beyond cluster scale. However, one difficulty with such models is that the redshift of galaxy formation is uncomfortably low.[3,5]

Another possibility is an initial power-law slope steeper than scale invariant over some range in k-space; with $n_i > -2$, the falloff in power at large scales is great enough to satisfy the CMB limits. Of course, power laws with $n_i < -3$ strongly violate CMB limits. For steeper n_i, the relative amplitude between galaxy and cluster scales is increased, alleviating the problem of the redshift of galaxy formation. Extra large-scale power can be obtained by including vacuum energy. For example, Peebles,[38–40] has recently advocated considering isocurvature baryon fluctuations with an arbitrary power-law initial spectrum of fluctuations in the entropy-per-baryon in open universe models with $\Omega_B \sim 0.2$, saturating the primordial nucleosynthesis limit. With $n \sim -1$, this was the preferred phenomenological spectrum of the 1970s (Gott and Rees[41]). The transfer function for such universes naturally imprints large-scale features of the desired form. This is illustrated by the $\Omega_B = 0.4$, $n_i = -1$ model in Fig. 1; the spectrum has many similarities to the phenomenological mountain model of Fig. 1. Provided the Universe remained ionized, the small-angle temperature anisotropies in these isocurvature models can be below current limits.[39,37] To be compatible with the flatness predicted by inflation, we must assume $\Omega_{\rm vac} = 1 - \Omega_B$, which lowers the anisotropies even further.

Nonscale-invariant power laws for adiabatic perturbations are more difficult to constrain. Spectra shallower than $n = 1$ have more large-scale power. Provided $n_i \gtrsim 0$, they cannot be ruled out by large-angle CMB constraints.[42] On the other hand, without violating these constraints the extra large-scale power required by $\xi_{\rm cc}$ is not possible. The plunge beyond cluster scale to give the $\xi_{\rm gg}$ break would also not occur.[10] Further, the redshift of galaxy formation becomes uncomfortably low.

III. LAGRANGIANS FOR INFLATIONARY MODELS

A. General Lagrangian

In this section, we analyze an action which describes a large variety of inflationary models, with N scalar fields, ϕ_k, self-interacting through an effective potential $V(\phi_k)$, and interacting with gravity via a $f(\phi_k)R$ term:

$$S = \int d^4x \sqrt{-g}\, [f(\phi_k)R - \tfrac{1}{2}T^{ij}(\phi_k) g^{\mu\nu}\partial_\mu\phi_i\partial_\nu\phi_j - V(\phi_k)] \; . \qquad (3.1)$$

For generality, we have also allowed the kinetic matrix $T^{ij}(\phi_k)$ to be arbitrary.

Fermions and gauge bosons are not included in Eq. (3.1) since they only have influence on the effective potential and during the reheating phase of inflation models. If fermions have a conserved current associated with their number, the number density will decline $\propto a^{-3}$ during inflation. Gauge fields can be produced classically only if charged sources are present, but the density of these will have exponentially declined. Further, gauge fields are conformally invariant in four dimensions; consequently, $\rho \propto F_{\mu\nu}F_{\alpha\beta}g^{\mu\alpha}g^{\nu\beta} \propto a^{-4}$, implying that any initial energy density in gauge fields will decay just like ordinary radiation. (Note that if conformal invariance is broken then the above conclusions are not valid. See Turner *et al.*,[43] who try to produce primordial magnetic fields by including nongauge invariant terms such as $R A_\mu A^\mu$.)

One-loop quantum corrections to gravity generate a term quadratic in the Ricci scalar bR^2 with b a dimensionless constant. If we naively choose $b \sim 1$, then the corrections to our models will be exceedingly small, of the order $H^2/m_P^2 \approx 10^{-10}$. For the opposite case, $b \gg 1$, see Kofman, Linde, and Starobinsky[44] and Sec. VII A 4.

The action (3.1) contains the essential elements necessary to produce viable models of density fluctuations. It encompasses:

(1) New and chaotic inflation. These require only a single scalar field, $N = 1$. The coefficient of the kinetic energy is normalized to unity, $T = 1$, and the interaction with gravity is minimal, i.e., $f(\phi) = m_P^2/(16\pi)$. The potential for new inflation is conventionally taken to be of the Coleman-Weinberg form (Sec. VI A)

$$V(\phi) = \frac{25}{256\pi^2} g_G^4 \{\sigma^4/2 + \phi^4[\ln(\phi^2/\sigma^2) - 0.5]\} \; , \qquad (3.2)$$

which arises from one-loop gauge-boson corrections to $\lambda\phi^4/4$. [In this case, the single scalar ϕ is interpreted as the component of the Higgs field in the SU(3) $\times$SU(2)$\times$SU(1) direction. The normalization convention is given in Appendix B.]

The chaotic inflation potential is unrestricted, except that it must be very flat for $\phi > m_P$; it is usually assumed to have a quartic and possibly quadratic piece:

$$V(\phi) = \frac{\lambda}{4}\phi^4 + \frac{m^2}{2}\phi^2 \; . \qquad (3.3)$$

The generalization to multiple scalar fields (Secs. VI B and VI C) is a natural one.

(2) The Brans-Dicke scalar, the dilaton, and induced gravity. If $N = 1$, $T^{ij}(\phi_k) = 1$, $V(\phi_k) = 0$, and $f(\phi_k) = -\xi\phi^2/2$, where ξ is the curvature coupling constant, then ϕ is related to the Brans-Dicke[45] scalar $\varphi_{\rm BD}$ by $\varphi_{\rm BD} \equiv -8\pi\xi\phi^2$; the Brans-Dicke coupling constant is just $\omega = -1/(4\xi)$. This theory has a history dating back to Jordan,[46] and plays an important role in supersymmetry and superstrings,[47] where ϕ is called the dilaton. If one assumes instead a potential of the form $V(\phi) = \lambda(\phi^2 - \sigma^2)^2/4$, then the Planck mass is generated through symmetry breaking,

$$\frac{m_P^2}{16\pi} = f(\sigma) \; , \qquad (3.4)$$

in the same way that the Higgs field generates the weak-boson mass. The extensive literature on this subject is reviewed by Adler.[48] We consider these models in Sec. VII.

(3) The nonlinear sigma model. An example of a

theory with a nontrivial kinetic matrix $T^{ij}(\phi_k)\neq\delta^{ij}$ is the nonlinear sigma model, in which the transformation of $T^{ij}(\phi_k)$ under redefinition of fields is

$$T^{i'j'}=T^{ij}(\phi_k)\frac{\partial\phi_i}{\partial\phi_{i'}}\frac{\partial\phi_j}{\partial\phi_{j'}}\ ,$$

just like a rank-2 tensor. To ensure parametrization invariance of the quantum fields, the potential $V(\phi_k)$ must vanish. By itself, the nonlinear sigma model is not of particular interest for inflation since the pressure equals the density. To obtain inflation, one must have a nonzero potential. It is possible that these may be generated by quantum corrections. A more interesting possibility for inflation models is the appearance of a nontrivial kinetic matrix, $T^{ij}(\phi_k)\neq\delta^{ij}$, which occurs for nonminimally coupled fields when the conformal transformation discussed in Sec. III B is performed. With $T^{ij}(\phi_k)\neq\delta^{ij}$, the normalization of scalar field quantum fluctuations is altered, as we demonstrate in Appendix A.

B. Conformal transformation of the Lagrangian

The form of the action (3.1) is very inconvenient for studying density fluctuations in theories with nonminimally coupled scalar fields. For example, the stress-energy tensor contains many additional terms beyond the $\xi=0$ case[49] and the analysis is difficult. More importantly, in any quantum treatment we wish our action to be a function only of first derivatives in order to define field momenta, e.g., $P=\delta S/\delta\dot{\phi}$, so we can impose the equal-time commutation relations, e.g., $[\phi(x,t),P(x',t)]=i\delta^3(x-x')$; similarly we would introduce momentum for the gravitational field. However, R contains *second derivatives* of metric terms, and consequently $f(\phi_k)R$ must be integrated by parts, producing derivatives in ϕ_k. The details are presented in Appendix A. The main point is that the standard commutator $[\phi(x,t),\partial\phi(x',t)/\partial t]=ia^{-3}\delta^3(x-x')$ used in the normalization of scalar field fluctuations (Sec. V) is no longer valid. With the commutator given by Eq. (A2) of Appendix A, we can treat both small- and large-ξ couplings unlike Acceta, Zoller, Turner,[22] and Lucchin, Mataresse, and Pollock,[50] whose results are valid only for small ξ, the regime they were primarily interested in.

A more elegant approach to deal with the troublesome $f(\phi_k)$ term is to perform a conformal transformation $g_{\mu\nu}=\Omega^2\tilde{g}_{\mu\nu}$, where $\Omega^2=(m_P^2/16\pi)f^{-1}$. Using the identity[49]

$$R(\Omega^2\tilde{g}_{\mu\nu})=\Omega^{-2}R(\tilde{g}_{\mu\nu})-6\Omega^{-3}\tilde{g}^{\mu\nu}\Omega_{;\mu\nu}\ , \tag{3.5}$$

where the covariant derivative is taken with respect to $\tilde{g}_{\mu\nu}$, and integrating by parts, the action (3.1) becomes

$$S=\int d^4x\sqrt{-\tilde{g}}\left[\left[\frac{m_P^2}{16\pi}\right]\tilde{R}-\tfrac{1}{2}K^{ij}(\phi_k)\tilde{g}^{\mu\nu}\partial_\mu\phi_i\partial_\nu\phi_j - U(\phi_k)\right]\ . \tag{3.6}$$

Here,

$$g_{\mu\nu}=\left[\frac{m_P^2}{16\pi}\right]f^{-1}\tilde{g}_{\mu\nu}\ , \tag{3.7a}$$

$$\tilde{R}\equiv R(\tilde{g}_{\mu\nu})\ , \tag{3.7b}$$

$$K^{ij}(\phi_k)=\left[\frac{m_P^2}{16\pi}\right]f^{-2}\left[3\frac{\partial f}{\partial\phi_i}\frac{\partial f}{\partial\phi_j}+fT^{ij}(\phi_k)\right]\ , \tag{3.7c}$$

$$U(\phi_k)=\left[\frac{m_P^2}{16\pi}\right]^2 f^{-2}V(\phi_k)\ . \tag{3.7d}$$

The result is therefore the normal Einstein action in the new metric $\tilde{g}_{\mu\nu}$, plus minimally coupled scalars interacting through a modified potential, which now have a more complicated kinetic matrix. The field equations following from the action (3.6) are much easier to deal with than those derived from (3.1).

The physical meaning of this transformation was given by Brans and Dicke.[45] In the Brans-Dicke theory, the Planck mass is taken to be a function of spacetime coordinates through the scalar fields, $m_{P\text{eff}}^2=16\pi f(\phi_k)$, as in Eq. (3.1). The conformal transformation (3.7a) is just a position-dependent change of units, arranged to make the coefficient of $\tilde{R}$ in (3.6) constant. The potential and kinetic energies also change units, and, because of the position dependence, an extra term also appears in the kinetic matrix (3.7c).

IV. EQUATIONS

In Sec. IV A we discuss the general features of our treatment of the action (3.6) to first-order perturbation theory. We write the scalar fields, ϕ_j, $j=1,\ldots,N$, as the sum of a spatially homogeneous background field $\bar{\phi}_j$ and a first-order perturbation, $\delta\phi_j$. The scalar fields are assumed to interact through the effective potential (3.7d), $U(\phi_1,\ldots,\phi_N)$ and a flat kinetic matrix, $K^{ij}(\phi_k)=\delta^{ij}$, dissipate into radiation through a phenomenological friction law, and couple to scalar perturbations of the transformed gravitational metric $\tilde{g}_{\mu\nu}$ which obeys regular Einsteinian gravity. In Sec. IV B we present the actual equations we solve and in Sec. V the initial conditions for the fluctuation generation problem. Most of these initial conditions are set by quantum conditions. We are free only to choose the initial background values $\bar{\phi}_j(t_i)$.

A. Perturbation theory of quantum fluctuations

Expanding to linear order in perturbation theory, we have $\phi_j(\mathbf{x},t)\equiv\bar{\phi}_j(t)+\delta\phi_j(\mathbf{x},t)$, $j=1,\ldots,N$, for the N scalar fields and $\tilde{g}_{\mu\nu}(\mathbf{x},t)=a^2(t)[\eta_{\mu\nu}+h_{\mu\nu}(\mathbf{x},t)]$, where $a(t)$ is the expansion factor of the Universe, $\eta_{\mu\nu}$ is the background metric, and $h_{\mu\nu}(\mathbf{x},t)$ is the metric perturbation. A flat background has been assumed. This will be adequate for our calculations of the fluctuation spectrum provided the scale of curvature is inflated far outside the current Hubble length, the conventional assumption in inflation models. We assume a perfect-fluid stress-energy tensor to describe the other matter present: $T^0_{r0}=\bar{\rho}_r+\delta\rho_r$, $T^i_{r0}=-(\rho_r+P_r)v^i$, $T^j_{ri}=\delta^j_i(P_r+\delta P_r)$, where we take $P_r=\rho_r/3$, appropriate for a tightly coupled relativistic fluid. The form of the dissipative coupling describing the damping of the scalar field into radiation is discussed in Sec. IV B. [We adopt the diag$(-,+,+,+)$ form for η and Misner, Thorne, and Wheeler[51] sign conventions. We also take $\hbar=c=1$, and define the Planck mass by $G=m_P^{-2}$, so $m_P=1.221\times10^{19}$ GeV.]

We assume that $\bar{\phi}_j$ and a are classical homogeneous "background" fields, ignoring their quantum nature (see below). The perturbations to the scalar fields and the metric coefficients *are* treated as quantum (Heisenberg) operators; they can be expanded in terms of normal modes labeled by comoving wave vectors $\mathbf{k}$:

$$\delta\phi_j(\mathbf{x},t)=\sum_s\int\frac{d^3k}{(2\pi)^3}[\delta\phi_j(\mathbf{k},s,t)e^{i\mathbf{k}\cdot\mathbf{x}}+\delta\phi_j^\dagger(\mathbf{k},s,t)e^{-i\mathbf{k}\cdot\mathbf{x}}]\ , \quad (4.1a)$$

$$h_{\mu\nu}(\mathbf{x},t)=\sum_s\int\frac{d^3k}{(2\pi)^3}[h_{\mu\nu}(\mathbf{k},s,t)e^{i\mathbf{k}\cdot\mathbf{x}}+h_{\mu\nu}^\dagger(\mathbf{k},s,t)e^{-i\mathbf{k}\cdot\mathbf{x}}]\ . \quad (4.1b)$$

Here $e^{i\mathbf{k}\cdot\mathbf{x}}$ are spatial eigenfunctions of ${}^{(3)}\nabla^2$ for the flat background which we adopt. The eigenfunctions are much more complicated if the background is taken to be closed or open (see Bardeen[27] and Halliwell and Hawking[52] for recent discussions). The additional index s is required to specify the modes, for example, labeling whether the perturbations are scalar, vector, or tensor. In this paper we only treat scalar modes. The other modes are independent of the scalar modes and the vector modes decay rapidly. The tensor modes (gravitational waves) are generated quantum mechanically, but the tensor part of h_{ab} is totally decoupled from the perturbations of the scalar fields.[53,54,52]

The scalar models are associated with dynamical oscillations of the scalar fields and the fluid. The quantum generation of sound waves (phonons) has been considered by Lukash[55] and by Chibisov and Mukhanov[56] in the context of perfect-fluid cosmological models, but here any radiation present when Fourier modes of astrophysical interest leave the horizon is diluted by the inflationary expansion. The reheating of the Universe following inflation completely swamps any primordial phonons, and the fluid perturbations in the Friedmann epoch evolve out of primordial scalar field fluctuations.

Therefore, we suppress the index s and consider as the independent annihilation operators the $\delta\phi_j(\mathbf{k},t)$ for the Fourier modes of the scalar field. If there is just one scalar field, the quantum operators $\delta\phi(\mathbf{k},t)$ and $\delta\phi^\dagger(\mathbf{k},t)$ are related to the conventionally normalized annihilation and creation operators for the modes of wave vector $\mathbf{k}$, $a(\mathbf{k})$, and $a^\dagger(\mathbf{k})$, by

$$\delta\phi(\mathbf{k},t)=\psi(\mathbf{k},t)a(\mathbf{k})\ \text{ and }\ \delta\phi^\dagger=\psi^*(\mathbf{k},t)a^\dagger(\mathbf{k})\ , \quad (4.1c)$$

where $\psi(\mathbf{k},t)$ is the mode function, a complex solution of the classical mode evolution equations. For multiple fields the picture is slightly more complicated. Since there are N scalar fields, there are an equal number of annihilation operators, $a_j(\mathbf{k})$, $j=1,N$, which satisfy the commutation relations

$$[a_j(\mathbf{k}),a_l^\dagger(\mathbf{k}')]=(2\pi)^3\delta_{jl}\delta^3(\mathbf{k}-\mathbf{k}')\ , \quad (4.1d)$$

with all others vanishing. In general,

$$\delta\phi_j(\mathbf{k},t)=\sum_l\psi_{jl}(\mathbf{k},t)a_l(\mathbf{k}) \quad (4.1e)$$

is a sum over the annihilation operators for all of the fields and not just the jth one; the single mode function is now replaced by ψ_{jl}, the mode function matrix. The reason for this complication is that the mass-squared matrix

$$m_{jl}^2\equiv\partial^2U/\partial\phi_j\partial\phi_l \quad (4.1f)$$

need not be diagonal, leading to mixing of the various annihilation operators within the perturbation equations [(4.7c)–(4.7f) below]. Even if it were diagonal, metric perturbations would cause a small amount of mixing.

In a classical treatment, $a_l(\mathbf{k})$ would be a complex random variable having a probability distribution to describe the initial occupation of the mode, and $a_l^\dagger(\mathbf{k})$ would be its complex conjugate. Once the $\mathbf{k}$-mode content of the Universe is specified, it is fixed for all time until mode-mode coupling occurs.

To determine the evolution of the operators (random fields) $\delta\phi_j(\mathbf{k},t)$, we need only solve the classical mode evolution equations. Our calculations are self-consistent quantum mechanically only if $\delta\phi_j(\mathbf{k},t)$ and $h_{\mu\nu}(\mathbf{k},t)$ are *both* treated as quantum operators (random fields), since they are coupled to one another. (The gravitational potentials we actually solve for involve combinations of the metric perturbations and their derivatives as described in Sec. IV B.) Ignoring the quantum aspects of the background fields $\bar{\phi}_j$ and a does not introduce a mismatch between classical and quantum variables. This approach of taking $\delta\phi_j$ and $h_{\mu\nu}$ as quantum fields defined upon a background manifold is similar to treating the graviton as a spin-2 field upon a Minkowski background.

The initial quantum state $|\Psi(t_i)\rangle$ is all that is required within our formalism to determine the full evolution of the state. For pure states, it can be expressed as a superposition of states $|\{n(\mathbf{k})\}\rangle$ having different occupation numbers for the modes. In some cases, it would be more appropriate to consider a mixed state described by a density matrix with a distribution (e.g., thermal) characterized by an energy scale T_i (e.g., a temperature). The usual assumption in inflation calculations is that $|\Psi(t_i)\rangle$ is the Bunch-Davies vacuum of de Sitter space which has none of the modes occupied: $a(\mathbf{k})|\Psi\rangle=0$. If the distribution is similar to the thermal form, modes within our current Hubble length have $k/a\gg T_i$ and therefore essentially zero occupation number. Lower k scales with nonzero occupation would be inflated outside of our current horizon. Since the length scales within our horizon were smaller than the Planck length before inflation and the linear perturbation approximation breaks down when the wavelengths are very small compared with the horizon, such arguments can only be regarded as suggestive. In this paper, we assume the initial state has zero occupation number. The fluctuations are then purely zero-point oscillations of the perturbation fields in the flat background spacetime. The wave functional $\Psi[\delta\phi_j(\mathbf{x},t),h_{\mu\nu}(\mathbf{x},t)]$ is then Gaussian, and can be written as a product of independent multivariate Gaussians, one for each of the k-modes. For example, for one scalar, the probability amplitude for the $\mathbf{k}$-mode field is Gaussian with dispersion equal to the classical value $|\psi(\mathbf{k},t)|^2$, the power spectrum of the field. As in Sec. II, we use the power per logarithm of wave number $\mathcal{P}_\phi$ to characterize the power spectrum. It is related to the equal-time two-point function of the field by

$$\mathcal{P}_\phi \equiv \frac{k^3}{2\pi^2}\int d^3x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\langle\Psi|\delta\phi(\mathbf{x})\delta\phi(\mathbf{0})|\Psi\rangle = \frac{k^3}{2\pi^2}|\psi|^2 . \tag{4.2a}$$

For multiple fields, the generalization is

$$\mathcal{P}_{\phi jl} \equiv \frac{k^3}{2\pi^2}\int d^3x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\langle\Psi|\delta\phi_j(\mathbf{x})\delta\phi_l(\mathbf{0})|\Psi\rangle = \frac{k^3}{2\pi^2}\sum_m \psi_{jm}\psi^*_{lm} , \tag{4.2b}$$

which is not necessarily diagonal.

To solve the full quantum gravity plus scalar field problem quantum mechanically is, of course, beyond the scope of current theory. The formal treatment within the context of quantum gravity is to solve the Wheeler-DeWitt equation for the wave functional $\Psi[\bar{\phi}_i,a,\delta\phi_j,h_{ij}]$, treating $\bar{\phi}_j$ and a as quantum fields. At best this can be done with the equations linearized, with no back action of the fluctuations on the background fields included (Fischler, Ratra, and Susskind,[57] Halliwell and Hawking.[52]) Halliwell and Hawking consider a massive scalar field and the initial quantum conditions are specified by the Hartle-Hawking[58] ansatz for the form of the wave function of the Universe. The equations for the fluctuations when cast into the Heisenberg framework are similar to ours for chaotic inflation with a massive noninteracting scalar field potential, given that the Universe is at a specific time fixed by requiring the volume (or a) to have some definite value. With their ansatz, inflation is shown to be a natural outcome of the quantum behavior in the $\bar{\phi}$-a sector. The mass in their case must still be chosen small enough to ensure enough e-foldings of expansion.

The inclusion of dissipation of the scalars into matter destroys the simplicity of the quantum analysis, since even the background scalar field with no momentum content will decay into particles with high momenta which depend upon $\bar{\phi}$'s temporal frequency content. However, these momenta are large compared with the momenta k of the modes we are following, so it is an adequate approximation to treat the creation of high-momentum modes in a phenomenological manner, and yet still treat the low-momentum modes we are concerned with quantum mechanically. The $\mathbf{k}$-modes for the density fluctuations in the radiation are then analogous to long-wavelength phonons, while the high-frequency fluctuations form a "thermal" bath. We will therefore expand the radiation energy density fluctuation and the radiation momentum density potential $(\rho+P)_r\mathbf{v}_r$ in terms of the annihilation and creation operators (IV B). It proves more tractable numerically to use instead the total energy density fluctuation $\delta\rho$ and the total momentum density potential Ψ, defined by $\nabla\Psi\equiv\sum_i\nabla\Psi_i\equiv -a\sum_i(\rho+P)_i\mathbf{v}_i$, where the sum is over the scalar fields and the radiation; expanding in terms of a Fourier integral, we have

$$\delta\rho = \int\frac{d^3k}{(2\pi)^3}[\delta\rho(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{x}}+\delta\rho^\dagger(\mathbf{k},t)e^{-i\mathbf{k}\cdot\mathbf{x}}] , \tag{4.3a}$$

$$\Psi = \int\frac{d^3k}{(2\pi)^3}[\Psi(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{x}}+\Psi^\dagger(\mathbf{k},t)e^{-i\mathbf{k}\cdot\mathbf{x}}] , \tag{4.3b}$$

where once again $\delta\rho$ and ψ are quantum operators.

B. Perturbation equations

We must choose a gauge, the spacelike hypersurfaces upon which to measure the fluctuations, and the specific combination of variables to solve for. The primary criterion used for this choice should be suitable for accurate and simple calculations, whether the computations are to be done numerically, as in our case, or analytically, the approach taken by most authors to fluctuation generation. We present equations for two gauge choices, one the venerable synchronous gauge that has been so widely used in perturbation theory in the expanding universe, beginning with the classic work of Lifshitz,[59] and the other, the longitudinal gauge. The choice of metric variables used in the synchronous gauge is rather novel.[29]

Although the general kinetic matrix $K^{ij}(\phi_k)$ introduced in Eq. (3.7c) need not be the identity, in this paper we assume it is $K^{ij}(\phi_k)=\delta^{ij}$. If $N=1$, the fields may always be redefined so this is true (Sec. VII). Varying the action (3.6), we find the following equation of motion for the scalar fields ϕ_j:

$$-\frac{1}{\sqrt{-\bar{g}}}\partial_\mu(\bar{g}^{\mu\nu}\sqrt{-\bar{g}}\,\partial_\nu\phi_j)+\partial U(\boldsymbol{\phi})/\partial\phi_j+J_j=0 , \tag{4.4a}$$

$$J_j=\Gamma_j\dot{\phi}_j,\quad \dot{\phi}_j\equiv\mathrm{sgn}(\dot{\phi}_j)|\partial_\mu\phi_j\partial^\mu\phi_j|^{1/2} , \tag{4.4b}$$

$$\Gamma_j=f_jM_j . \tag{4.4c}$$

The interaction term J_j in Eq. (4.4a) representing the coupling of the scalars to matter is generally extremely complicated. We adopt the phenomenological friction form Eq. (4.4b) for J_j. The time derivative here must be defined covariantly as indicated; $\mathrm{sgn}(\dot{\phi}_j)$ denotes the sign of the time derivative which is an invariant for any coordinate time t (as long as $\partial_\mu\phi_j$ is timelike). In working in gauges with $h_{00}\neq 0$, there are linear-order perturbative corrections to $\dot{\phi}_j$ as defined here which cannot be ignored (cf. Den and Tomita[60] who do neglect the extra term). The dissipation time scale Γ_j^{-1} is parametrized in Eq. (4.4c) by a dimensionless friction coefficient f_j and an energy M_j, which we take to be ϕ_j itself for a GUT inflation model and the scalars' masses for chaotic inflation models.

The equations of motion for the matter must guarantee that the energy and momentum dissipated from the scalar field end up in the radiation; i.e., the total stress-energy tensor must be covariantly conserved. It can be numerically advantageous to avoid following the details of this energy-momentum transfer, and use as variables in place of the radiation energy density and radiation momentum density the total energy density perturbation $\delta\rho$ and the total momentum current potential Ψ. In fact, we have found it desirable to go further and replace $\delta\rho$ by $\delta\rho_{\rm com}$,

$$\delta\rho_{\rm com} \equiv \delta\rho - 3H\Psi, \quad H \equiv \dot{a}/a \ . \tag{4.5}$$

Note that $\delta\rho_{\rm com}$ is a gauge-invariant variable which becomes the energy density perturbation in the comoving gauge, defined by $\Psi \equiv 0$.

The complete set of equations we must solve consists of the zeroth- and first-order versions of Eq. (4.4a) for the scalar field, the equations $\nabla \cdot T = 0$, and the Einstein equations for the gravitational field. The zeroth-order background equations are

$$da/dt = (8\pi G a^2 \rho/3)^{1/2}, \quad \rho \equiv \sum_j \rho_{\phi_j} + \rho_r \ , \tag{4.6a}$$

$$\rho_{\phi_j} = \tfrac{1}{2}\dot{\bar{\phi}}^2 + U(\phi), \quad p_{\phi_j} = \tfrac{1}{2}\dot{\bar{\phi}}^2 - U(\phi) \ , \tag{4.6b}$$

$$d^2\bar{\phi}_j/dt^2 + 3H d\bar{\phi}_j/dt + \partial U(\phi)/\partial\bar{\phi}_j + \Gamma_j \dot{\bar{\phi}}_j = 0 \ , \tag{4.6c}$$

$$\frac{1}{a^4}\frac{da^4\rho_r}{dt} = \sum_j \Gamma_j \dot{\bar{\phi}}_j^2 \ . \tag{4.6d}$$

In the synchronous gauge, $h_{00} = h_{0i} = 0$. This does not fix the gauge until the initial hypersurface and spatial coordinates in the initial hypersurface are specified. For scalar perturbations, the perturbation in the spatial part of the metric can be written in terms of two functions φ and γ: $h_{ij} = 2\varphi\delta_{ij} + 2\gamma_{,ij}$. While γ changes under a spatial gauge transformation, $\dot{\gamma}$ and φ are spatially gauge invariant. The nontrivial gauge transformations are time gauge transformations, which affect both φ and $\dot{\gamma}$. If we decompose the perturbation into independent modes characterized by the wave number $\mathbf{k}$, and choose the $\hat{\mathbf{k}}$ axis to be the three-axis, then $\varphi = (h - h_{33})/4$, $\gamma = k^{-2}(h - 3h_{33})/4$, and the scalar three-curvature of the constant time hypersurfaces is $a^2\mathcal{R} \equiv 4k^2\varphi$. The extrinsic curvature is $\kappa_{ij} = -\dot{h}_{ij}/2$. For scalar perturbations there is no intrinsic dynamics of the gravitational field and the δG_0^0 and δG_i^0 constraint equations suffice to determine the metric perturbations. Only $\dot{\varphi}$ and $\kappa \equiv -\dot{h}/2$ enter the matter equations of motion and are given by

$$\dot{\varphi} = 4\pi G \Psi \ , \tag{4.7a}$$

$$\kappa \equiv -\dot{h}/2 = [(k^2/a^2)\varphi - 4\pi G \delta\rho]/H \ . \tag{4.7b}$$

In most previous numerical work utilizing the synchronous gauge, the "dynamical" δR_0^0 equation was integrated to find κ. For the particular case of scalar fields plus tightly coupled radiation, φ does not enter the equations at all. However, the conventional approach can lead to numerical difficulties. The advantages of using Eq. (4.7b) for κ are discussed more fully in Bardeen.[29]

The equations for overall conservation of energy and momentum, after eliminating κ with Eq. (4.7b) and $\delta\rho$ in favor of $\delta\rho_{\rm com}$, are

$$\delta\dot{\rho}_{\rm com} = -3H[1 + \tfrac{1}{2}(1 + p/\rho)]\delta\rho_{\rm com} + (k^2/a^2)[H^{-1}(\rho + p)\varphi + \Psi] \ , \tag{4.7c}$$

$$\dot{\Psi} = -3H\Psi - \delta p \ , \tag{4.7d}$$

$$\delta p = \tfrac{1}{3}\delta\rho_{\rm com} + H\Psi + \tfrac{2}{3}\sum_j \left[\dot{\phi}_j \delta\dot{\phi}_j - 2\frac{\partial U}{\partial \phi_j}\delta\phi_j\right] \ . \tag{4.7e}$$

The advantage of using $\delta\rho_{\rm com}$ is that over much of the evolution $\delta\rho_{\rm com} \ll \delta\rho, 3H\Psi$. The perturbation in the scalar field obeys

$$\frac{d^2\delta\phi_j}{dt^2} + (3H + \Gamma_j)\frac{d\delta\phi_j}{dt} + (k^2/a^2)\delta\phi_j + \sum_i \frac{\partial^2 U}{\partial\phi_j \partial\phi_i}\delta\phi_i - \kappa\dot{\bar{\phi}}_j + \dot{\bar{\phi}}_j \delta\Gamma_j = 0 \ . \tag{4.7f}$$

Equations (4.7) are the full set of synchronous gauge equations cast in the form in which we solve them.

The stress-energy perturbations for the scalar fields and the explicit expression for the comoving density perturbation are

$$\delta\rho_{\phi_j} = \dot{\bar{\phi}}_j \delta\dot{\phi}_j + \frac{\partial U(\phi)}{\partial \phi_j}\delta\phi_j \ ,$$

$$\Psi_{\phi_j} = -\dot{\bar{\phi}}_j \delta\phi_j \ ,$$

$$\delta p_{\phi_j} = \dot{\bar{\phi}}_j \delta\dot{\phi}_j - \frac{\partial U(\phi)}{\partial \phi_j}\delta\phi_j \ ,$$

$$\delta\rho_{\rm com} = \sum_j \left[\dot{\bar{\phi}}_j \delta\dot{\phi}_j + 3H\dot{\bar{\phi}}_j \delta\phi_j + \frac{\partial U}{\partial\phi_j}\delta\phi_j\right] + (\delta\rho_r - 3H\Psi_r) \ .$$

The synchronous gauge provides a valid description until coordinate singularities arise, that is, until pancake formation first occurs, a phenomenon that will be delayed until the fluctuations enter the horizon for the second time, just before the epoch of galaxy formation.

The gauge-invariant variable Φ_H of Sec. II was introduced by Bardeen[27] to give a useful generalization in the cosmological setting of a perturbed Newtonian gravitational potential. Φ_H satisfies a Poisson-Newton equation with source the comoving density perturbation:

$$\Phi_H = k^{-2} 4\pi G a^2 \delta\rho_{\rm com} \ . \tag{4.8a}$$

Φ_H remains constant outside the horizon in the radiation-dominated phase. In the longitudinal gauge,[27] the metric perturbation is diagonal: $h_{ij} = 2\Phi_H \delta_{ij}$, $h_{00}/a^2 = -2\Phi_A$; if there is no anisotropic stress as in the case treated here, $\Phi_A = -\Phi_H$. The perturbations in this gauge take the simple form (see also Sasaki[61] and cf. Den and Tomita[60])

$$\delta\ddot{\phi}_{Lj}+(3H+\Gamma_j)\delta\dot{\phi}_{Lj}+(k^2/a^2)\delta\phi_{Lj}+\sum_i \frac{\partial^2 U}{\partial\phi_i\partial\phi_j}\delta\phi_{Li}$$
$$+\dot{\bar{\phi}}_j\delta\Gamma_{Lj}=\left[2\frac{\partial U}{\partial\phi_j}+\dot{\bar{\phi}}_j\Gamma_j\right]\Phi_H-4\dot{\bar{\phi}}_j\dot{\Phi}_H\ , \quad (4.8b)$$

$$\ddot{\Phi}_H+5H\dot{\Phi}_H+\left[\frac{k^2}{3a^2}+\frac{32\pi U}{3m_P^2}\right]\Phi_H$$
$$=\frac{8\pi}{3m_P^2}\sum_j\left[2\frac{\partial U}{\partial\phi_j}\delta\phi_{Lj}-\dot{\bar{\phi}}_j\delta\dot{\phi}_{Lj}\right]. \quad (4.8c)$$

The transformation of the scalar field perturbation from the longitudinal gauge to the synchronous one is

$$\delta\phi_{jL}=\delta\phi_j-2a^2k^{-2}(\kappa+3\dot{\varphi})\dot{\bar{\phi}}_j\ . \quad (4.8d)$$

Equations (4.7) and (4.8) are quantum operator equations. To reduce to the *complex*-valued equations which we solve numerically, we expand $\delta\phi_i(\mathbf{k},t)$ in terms of the N annihilation operators according to Eq. (4.1e), and similarly for the remaining variables:

$$\delta\rho_{\rm com}(\mathbf{k},t)=\sum_l r_l(\mathbf{k},t)a_l(\mathbf{k})\ ,$$
$$\Psi(\mathbf{k},t)=\sum_l y_l(\mathbf{k},t)a_l(\mathbf{k})\ , \quad (4.9a)$$

$$\kappa=\sum_l K_l(\mathbf{k},t)a_l(\mathbf{k}),\quad \varphi=\sum_l f_l(\mathbf{k},t)a_l(\mathbf{k})\ , \quad (4.9b)$$

$$\delta\Gamma_j=\sum_l G_{jl}(\mathbf{k},t)a_l(\mathbf{k})\ . \quad (4.9c)$$

Note that extra creation operators for the metric perturbations and the matter perturbation are not required because these fields are given directly in terms of the N scalar fields through Eqs. (4.7).

We have solved these equations numerically both in both synchronous and longitudinal gauges; both are satisfactory for accurate computations of fluctuation generation. As an explicit illustration of the equations we solve, the N operator equations (4.7f) now become N^2 complex equations or $2N^2$ real equations:

$$\frac{d^2\delta\psi_{jl}}{dt^2}+(3H+\Gamma_j)\frac{d\delta\psi_{jl}}{dt}+(k^2/a^2)\delta\psi_{jl}$$
$$+\sum_i\frac{\partial^2 U}{\partial\phi_j\partial\phi_i}\delta\psi_{il}-\dot{\bar{\phi}}_jK_l+\dot{\bar{\phi}}_j\delta G_{jl}=0\ . \quad (4.10)$$

If we fix l, we may think of ψ_{jl} and G_{jl}, $j=1,\ldots,N$ as vectors, r_l,K_l,γ_l, as scalars; these variables satisfy Eqs. (4.7a)–(4.7f), now interpreted as complex-valued equations, subject to the initial conditions given in Sec. V. Numerically, this set of equations must be solved N times, once for each l, in order to describe the full dynamics of the quantum system.

As we mentioned in Sec. II, ζ defined in terms of Φ_H by Eq. (2.2) is an extremely useful variable for us to monitor in our calculations. It is related to the synchronous gauge variables φ, the scalar three-curvature, and $\delta\rho$, the density fluctuation, by

$$\zeta=3\varphi+\delta\rho/(\rho+p)\ . \quad (4.11)$$

C. Connecting with the present length scale

The integration of the equations of motion is terminated as soon as the Universe becomes radiation dominated. In practice, we stop the calculations when $\rho_\phi<10^{-3}\rho_{\rm tot}$, and we then determine the end temperature T_e from the simple thermodynamics of relativistic particles:

$$\rho_e=g_{\rm eff}\frac{\pi^2}{30}T_e^4\ . \quad (4.12)$$

The effective number of degrees of freedom at temperature T expressed in terms of the number of bosonic and fermionic degrees of freedom g_B and g_F is

$$g_{\rm eff}(T)=g_B(T)+\tfrac{7}{8}g_F(T)\ . \quad (4.13)$$

In this paper, we use the minimal SU(5) value, $g_{\rm eff}=160.75$, although other GUT models would give higher values. T_e should not be confused with the reheat temperature, $T_{\rm reh}$, which is defined as the maximum temperature attained during reheating; both, of course, are of the same order of magnitude. In order to link scales of the inflationary phase with those of the present epoch, we must determine the expansion factor, a_0/a_e, in the radiation- and matter-dominated eras. Assuming conservation of the entropy in relativistic particles per comoving volume,

$$s_*=\tfrac{4}{3}a^3\rho/T\ ,$$

we have

$$a_e^3g_{\rm eff}(T_e)T_e^3=10.75T_{\nu0}^3a_0^3$$
$$=(2T_{\gamma0}^3+\tfrac{21}{4}T_{\nu0}^3)a_0^3\ . \quad (4.14)$$

We have assumed that there is no entropy generation between neutrino decoupling at $T\sim1$ MeV, when $g_{\rm eff}=10.75$ for 6 neutrino degrees of freedom, and the present, so the relation between the neutrino temperature and the observed photon temperature is the conventional $T_{\nu0}=(\frac{4}{11})^{1/3}T_{\gamma0}$. We take $T_{\gamma0}=2.7K$. The number of e-foldings from the end of our computation to the present is then

$$N_e\equiv\ln(a_0/a_e)$$
$$=72.583\,28+\ln(T_e/m_P)+\tfrac{1}{3}\ln(g_{\rm eff})\ . \quad (4.15)$$

For convenience we set the present value of the scale factor a_0 to unity. Thus physical and comoving length scales coincide at the present epoch: $k_{\rm phys}^{-1}=a_0k^{-1}=k^{-1}$. The value of the scale factor at the end of inflation is then e^{-N_e}. To solve the horizon problem, the initial values of the background fields, $\phi_j(t_i)$, must be chosen so that the physical scale starts well within the Hubble radius, $k/a(t_i)\gg H(t_i)$ or equivalently the number of e-foldings during inflation, N_I must exceed $\ln(H/k)-N_e$. One can then prescribe the initial values of the perturbation equations (Sec. V). Typically, $N_I\approx60$ for $k^{-1}=5000$ h^{-1}Mpc.

V. INITIAL CONDITIONS INSIDE THE HORIZON

The calculations of the mode functions are begun when (1) the background scalar fields are in their slow-rolldown phase and (2) the waves are well inside the horizon $[k/(Ha)\sim 50]$. Given the initial values $\bar{\phi}_j(t_i)$, $\dot{\bar{\phi}}_j(t_i)$ is uniquely determined by (1), $\dot{\bar{\phi}}_j(t_i)\approx-(3H+\Gamma_j)^{-1}\partial U/\partial\phi_j$.

We first consider the case of a *single scalar field.* Using the commutation relation $[\phi(\mathbf{x},t),\partial\phi(\mathbf{x}',t)/\partial t]=ia^{-3}(t)\delta^3(\mathbf{x}-\mathbf{x}')$ to set the overall amplitude we find that Eq. (4.7f) or (4.8b) has the WKB solution

$$\delta\phi(\mathbf{k},t)=\psi(\mathbf{k},t)a(\mathbf{k})\ , \tag{5.1a}$$

$$\psi(\mathbf{k},t)=\frac{\exp\left[-ik\int dt/a\right]}{(2k)^{1/2}a}$$
$$\text{for } k/Ha\gg 1, dH^{-1}/dt, |m_{11}|/H\ , \tag{5.1b}$$

$$\mathcal{P}_\phi=\left[\frac{H}{2\pi}\right]^2\left[\frac{k}{Ha}\right]^2,\quad (\mathcal{P}_\phi^{1/2})_{\rm hor}\approx\frac{H}{2\pi}\ . \tag{5.1c}$$

The influence of metric perturbations is only felt at higher order in Ha/k in the ψ solution. The WKB solution for the metric perturbations f and K of Eq. (4.9b) are given by

$$f(\mathbf{k},t)=-4\pi iG\frac{a}{k}\dot{\phi}\psi(\mathbf{k},t)\ , \tag{5.2a}$$

$$K(\mathbf{k},t)=12\pi iG\dot{\phi}\psi(\mathbf{k},t)\ . \tag{5.2b}$$

Note that f and K are then purely oscillatory and that $K\approx 0$. Equation (5.2b) explicitly illustrates that the metric terms have no influence on ψ [since $K(\mathbf{k},t)\dot{\psi}\ll(k/a)^2\psi(\mathbf{k},t)$].

The initial conditions for our longitudinal gauge computations are ψ_L is still given by (5.1b), with the difference between ψ_L and ψ given by Eq. (4.8d) being of higher order; Φ_H is given by Eq. (4.8a) with the radiation contribution neglected; $\dot{\Phi}_H$ is given by

$$\dot{\Phi}_H=-H\Phi_H-4\pi G\dot{\phi}\delta\phi\ . \tag{5.2c}$$

For all the calculations shown here, $\dot{\rho}_r$ was also set equal to zero initially. Whether we use Eqs. (5.2a) and (5.2b), or Eq. (5.2c) or even assume that all metric perturbations initially vanish, the results are essentially identical, differing by less than a part in a thousand.

In addition to the "positive energy" WKB solutions (5.1b), there are negative-energy solutions. We choose the positive energy ones. The state with no mode occupation $(a(\mathbf{k})|\Psi\rangle)$ is then the ground state. However, if the expansion is highly dynamic, an initially positive-frequency mode will generate negative-frequency components; hence, different results could be obtained depending upon precisely how far within the horizon we begin the evolution. For the models treated here, this is not a problem. Since (5.1b) gives complex initial conditions, we must solve complex equations; the real and imaginary parts are decoupled, but we must solve both sets of perturbation equations.

The modifications of Eqs. (5.1) and (5.2) required to treat *multiple scalar fields* are straightforward. The overall amplitude of the mode matrix ψ_{jl} is determined by the equal-time commutation relations

$$[\phi_j(\mathbf{x},t),\partial\phi_l(\mathbf{x}',t)/\partial t]=ia^{-3}(t)\delta_{jl}\delta^3(\mathbf{x}-\mathbf{x}')\ .$$

Substituting into Eq. (4.1e) gives the normalization criteria

$$2a^3\sum_m \mathrm{Im}\left[\psi_{jm}(\mathbf{k},t)\frac{\partial}{\partial t}\psi^*_{lm}(\mathbf{k},t)\right]=\delta_{jl}\ . \tag{5.3}$$

Far within the horizon, the positive-energy WKB solutions of (4.7f) or (4.8b) which obey (5.3) are

$$\psi_{jl}=\frac{\exp\left[-ik\int dt/a\right]}{(2k)^{1/2}a}\delta_{jl}$$
$$\text{for } k/Ha\gg 1, dH^{-1}/dt, |m_{ij}|/H\ . \tag{5.4a}$$

The cross-correlation power spectrum (4.2b),

$$\mathcal{P}_{\phi jl}=\left[\frac{H}{2\pi}\right]^2\left[\frac{k}{Ha}\right]^2\delta_{jl}\ , \tag{5.4b}$$

demonstrates that the various scalar fields are initially uncorrelated. Rotation to a different combination of the fields will not change this. However, once horizon crossing is approached, nondiagonal terms in the mass matrix, $m^2_{jl}\neq 0$, will lead to nonzero ψ_{jl}; when the metric terms cease being oscillatory, we also have non-negligible ψ_{jl}.

The initial conditions for the synchronous metric mode functions f_l and κ_l defined by Eqs. (4.9) generalize Eqs. (5.2a) and (5.2b) to

$$f_l(\mathbf{k},t)=-4\pi iG\frac{a}{k}\sum_j\dot{\phi}_j\psi_{jl}(\mathbf{k},t)\ , \tag{5.5a}$$

$$K_l(\mathbf{k},t)=12\pi iG\sum_j\dot{\phi}_j\psi_{jl}(\mathbf{k},t)\ . \tag{5.5b}$$

Longitudinal initial metric perturbations are given by Eq. (4.8a) and

$$\dot{\Phi}_H=-H\Phi_H-4\pi G\sum_j\dot{\phi}_j\delta\phi_j\ . \tag{5.5c}$$

Translation to appropriate mode functions is obvious.

VI. CALCULATIONS OF INFLATION SPECTRA

In this section, we present numerical calculations of fluctuation evolution for various inflationary models. Since there is no definitive particle-physics model for the form of the potential V of Eq. (3.1) and the dissipation term, Eq. (4.4b), they are at our disposal. We primarily choose examples with quartic potentials—formally required to ensure renormalizability of the scalar field theory. More general phenomenological potentials are discussed in Sec. VI D. For illustration, we consider two cases for $V(\phi)$ in Sec. VI A: a Coleman-Weinberg potential and a typical chaotic inflation potential. The parameters defining the potential, the dissipation term and $\phi_j(t_i)$ are chosen to satisfy the basic requirements for "successful" inflation: (1) At least 60 e-foldings of expansion in addition to the ~ 60 e-foldings that occur in the standard radiation-dominated phase of the big bang to ensure that flatness, isotropy, and homogeneity are obtained. (2) Fluctuations in the gravitational potential of order 10^{-5} arise to ensure that galaxy scale structure can form without overgenerating anisotropies in the microwave background. As is usual in inflation, obtaining (1) and (2) impose severe restrictions on the flatness of the potential V. To create features in the fluctuation spectrum at an astrophysically interesting length scale imposes further restrictions on our choice of parameters. Mechanisms which can drive such features are explored in Sec. VI B in the context of a simple single scalar field model by allowing the Hubble parameter and effective mass to be time dependent, which isolates the important physics of more complex multiple scalar field models. Detailed numerical calculations of chaotic inflation models with two scalar fields are given in Sec. IV C. We show that, in addition to the fine-tuning of the potential parameters required to solve (1) and (2), the initial conditions of the background fields must often be carefully chosen to place the features in the fluctuation spectrum near the scale of rich clusters, as in the phenomenological spectrum of Sec. II.

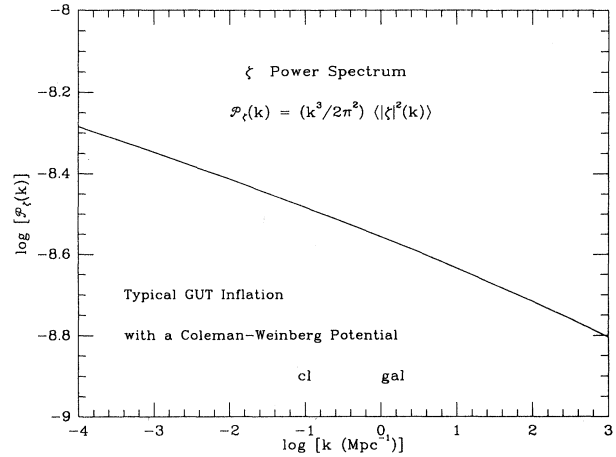

FIG. 2. The $\mathcal{P}_\zeta(k)$ fluctuation spectrum that arises in a GUT inflation scenario with a Coleman-Weinberg potential is plotted against k in Mpc^{-1} for an $h=0.5$ model. The extremely weak breaking of scale invariance (a factor of 3 drop over seven decades) has a negligible effect on structure formation. The amplitude of the spectrum depends mainly on the coupling strength, which is chosen to be $g_G=5.4\times10^{-4}$, whereas the natural value is close to $\frac{1}{2}$. The relevant regime for galaxy formation and clustering is indicated by the scales for clusters (cl) and for galaxies (gal).

A. Standard inflation with one scalar field

1. New inflation

Before turning to the promising chaotic scenario, we illustrate the use of our numerical techniques on the well-studied standard new inflation model utilizing a GUT-inspired Coleman-Weinberg potential, Eq. (3.2), for a massless and minimally coupled scalar field. For the purposes of this study, however, we may take the parameters σ, where the potential has its minimum, and g_G, a measure of the interaction strength, as tunable, with values guided by cosmological rather than particle-physics requirements. $V(\phi)$ has a relatively flat "slow rollover" region near the origin, followed by a fairly abrupt drop to a well centered on $\phi=\sigma$.

In Fig. 2 we show the $\mathcal{P}_\zeta(k)$ fluctuation spectrum that emerges in a GUT inflation scenario with a Coleman-Weinberg (CW) potential. The small downward slope of the CW potential during slow-rollover results in H decreasing slightly with expansion, giving the slightly smaller fluctuations at high k than at low k. Some of the parameter choices are, of course, entirely nonstandard in order that the density perturbations remain small; here, the coupling of the scalar was taken to be $g_G=5.4\times10^{-4}$, whereas the natural value is closer to $\frac{1}{2}$, which would give $\mathcal{P}_\zeta^{1/2}\approx 75$ instead of the $\sim 7\times10^{-5}$ value obtained in Fig. 2. The other parameter choices are more conventional: $\sigma=1.2\times10^{15}$ GeV, $\Gamma=0.05g_G^2\sigma$, and $g_{\rm eff}=160.75$. The tiny coupling effectively requires that ϕ be a gauge singlet. Some models have been proposed to fit into a GUT framework a gauge singlet with tiny g_G^4 utilizing supersymmetry, but none are very attractive.[62] A difficulty with such weak coupling is that Γ should also be tiny, necessitating low-temperature baryon synthesis. In Sec. VII we are able to drop the small g_G condition.

The background field ϕ shown in Fig. 3 begins at $3\times10^{-6}\sigma$ to ensure sufficient e-foldings, and changes very slowly until about 50 e-foldings after our current Hubble length, $\sim 10^4$ Mpc, left the horizon. Once the scalar nears the bottom of its potential, it only takes about 10^{-3} of a Hubble time for its vigorous oscillations to have decayed into radiation. Of course, the tiny choice of Γ gives a small reheating temperature, $T_{\rm reh}\sim 10^{10.8}$ GeV. Also shown in Fig. 3 are the fluctuations, $\delta\phi$, Φ_H, and ζ associated with the scale $k^{-1}=5000$ h^{-1}Mpc. The amplitude $\delta\phi$ Hubble damps as $1/a(t)$, Eq. (5.1b), within the horizon, and, upon crossing, freezes out at the Hawking temperature, $H/2\pi$. The negative curvature of the Coleman-Weinberg potential causes $\delta\phi$ to grow by 11 orders of magnitude during the slow-rollover regime. During this time, Φ_H is essentially zero and only during reheating does it grow, reaching a value of $\zeta/4.5$ after a few e-foldings.

If only the magnitude of the adiabatic fluctuations is desired for single-field inflation, the precise behavior during reheating is unimportant, since the perturbation ζ remains constant outside the horizon, even through reheating, as Fig. 3 indicates. An accurate analytic result[28] utilizing this constancy gives the amplitude of $\zeta(k)$ for single scalar field inflation: $\zeta=-3H\delta t$, where $\delta t=\delta\phi/\dot{\bar\phi}$ is the lag of the perturbation relative to the background value and $H=H(k)$ is the Hubble parameter when the wave crosses outside of the horizon, $Ha=k$. With Eq. (5.1c) giving the amplitude of the fluctuations and slow rolldown giving $\dot{\bar{\phi}}$, we have

$$\mathcal{P}_\zeta^{1/2}(k) = 3H^2\dot{\bar{\phi}}^{-1}(\mathcal{P}_\phi^{1/2}/H)_{\rm hor} \approx \frac{3}{2\pi} H^2 \dot{\bar{\phi}}^{-1} . \quad (6.1)$$

Substituting the values used in Fig. 2, we obtain excellent agreement with our numerical results. By including the

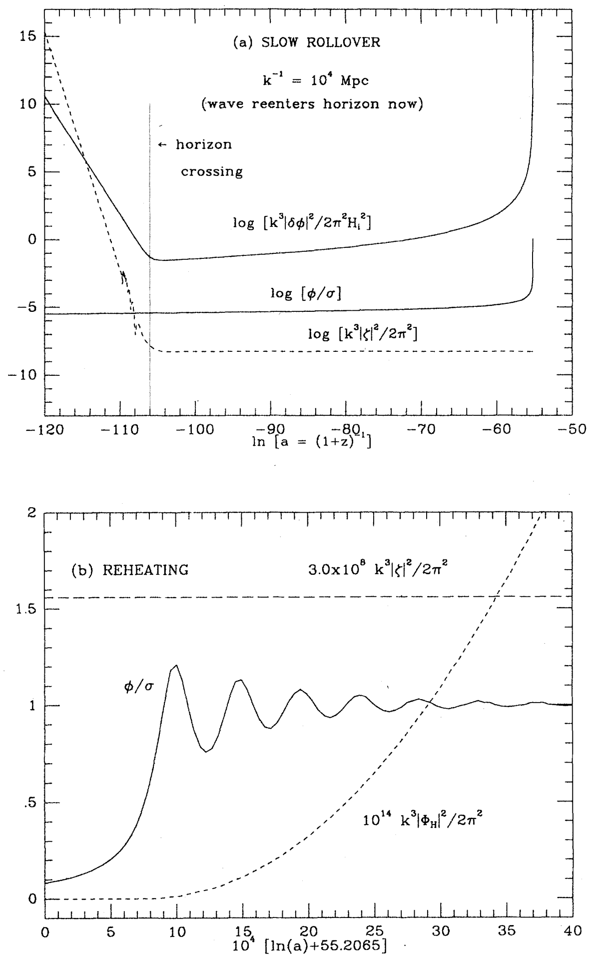

FIG. 3. The evolution of the background field during inflation demonstrates the gradual increase during slow rollover, while ϕ follows the flat portion of the CW potential, and the rapid rise towards high-frequency oscillations about σ as ϕ falls into the steep-sided potential well centered on $\phi=\sigma$, which damp as the field dissipates its energy into radiation. The number of particle degrees of freedom was taken to be 160.75, and the reheat temperature was $10^{10.8}$ GeV$\approx g_G\sigma$. Log denotes $\log_{10}$ and ln denotes $\log_e$. Note that (b) represents only about 10^{-3} of a Hubble time. For a wave with $k^{-1}=5000\ h^{-1}$Mpc, which is just now reentering the current "horizon," the amplitude of the scalar field fluctuations $\mathcal{P}_\phi^{1/2}$ drops as a^{-1} [Eq. (5.1b)] while inside the horizon during inflation, reaches the Hawking temperature $H/(2\pi)$ at horizon crossing, and then grows slowly (but significantly) outside the horizon in response to the gradually downcurving CW potential. ζ is extremely constant outside the horizon, whereas the gravitational potential grows only during reheating, reaching a final value $\zeta/4.5$ several Hubble times later.

modification with k of $\dot{\bar{\phi}}$ as the field rolls down the potential, the rise from small to large distances can also be obtained quantitatively.

Our results disagree with the numerical computations with CW potentials of Den and Tomita,[60] who solved equations similar to (4.8), but started far outside the horizon just prior to the onset of fast rolldown into the well, while assuming fluctuation levels appropriate to horizon crossing. As we have seen, a significant amount of growth in $\mathcal{P}_\phi$ occurs between horizon crossing and fast rolldown. Den and Tomita therefore seriously underestimated the amplitude of the density fluctuations.

2. *Chaotic inflation*

In this scenario, a simple potential of the form of Eq. (3.3) is assumed, with $\lambda=5\times10^{-14}$ chosen so the final adiabatic fluctuations come out at the right level, and m chosen to be very small so that the $\lambda\phi^4$ piece dominates. This value of λ and the parameters $m=10^{-8}m_P$, $\Gamma=0.1m$, and $\phi(t_i)=5m_P$ were chosen to generate the fluctuation spectrum for chaotic inflation with a single field given in Fig. 8 below (curve nearest the bottom). Again there is a small rise from short to large distances, but the spectrum is basically scale invariant.

Inflation is possible with such a potential provided the field is initially large enough and homogeneous enough that the potential energy $V(\phi)$ dominates the kinetic piece, $\dot{\phi}^2/2+(\nabla\phi)^2/2$. For homogeneous fields to give $N_I \gtrsim 60$ e-foldings of expansion during inflation, the field must start far out on the potential, $\phi(t_i) \approx (N_I/\pi)^{1/2}m_P \gtrsim 4.4m_P$. $\phi(x,t_i)$ must attain such values coherently over a scale $a_ik^{-1} \gtrsim N_I^{1/2}H_i^{-1}$, $\sim 15H_i^{-1}$ for $\phi(t_i)=5m_P$ as in Fig. 8 below. [$H\sim4\times10^{12}(\phi/m_P)^2$ GeV.]

For chaotic inflation, the analytic result equation (6.1) gives $\mathcal{P}_\zeta^{1/2}(k) \approx (6\pi)^{1/2}\lambda^{1/2}(\phi/m_P)^3$, which, for the wave which is just crossing the horizon at t_i, gives $10^{-4.1}$, dropping slowly as k increases to values which agree very well with our computations.

B. Breaking scale invariance with one driven field

In standard chaotic inflation,[1] initial values of the background fields as large as $\phi\sim\lambda^{-1/4}m_P\approx3000m_P$ are possible without requiring a quantum gravity treatment. Yet the structure we observe that could derive directly from primordial perturbations, from $k^{-1}\sim1$ kpc to the current horizon scale $k^{-1}\sim10^4$ Mpc, only corresponds to scalar field values in the narrow between $3.7m_P$ and $4.4m_P$ [using $N_I\approx\pi(\phi/m_P)^2$ for a $\lambda\phi^4$ chaotic inflation average potential]. Reheating occurs a little later, usually when $\phi\approx1m_P$, corresponding to a scale $k^{-1}\sim1$ m. In CDM models, we could possibly have dark-matter condensations down to scales $k^{-1}\sim1$ m; with a very large amplitude for the fluctuations these could be stable enough to ultimately influence gas or be observable in dark-matter halo searches. We could gain information on the 1 m to 1 kpc scales only if such exotic phenomena were to occur, but the limits we can currently place are very crude. Even over the more directly observable range the evidence for a flat spectrum in any one regime is not very compelling. (The best evidence comes from galaxy clustering, covering the region from $4.0m_P$ to $4.15m_P$, and even this can be argued as being largely due to dynamics rather than primordial spectral shape.) Thus from an observational point of view models with even severely broken scale invariance cannot be excluded. The smallness of this observational window on the potential surface is emphasized by Fig. 4.

We now seek to modify the flat spectrum by changing the structure of the potential surface over this limited observational window. Since the shape of $\mathcal{P}_\zeta$ is largely determined by the shape of $\mathcal{P}_\phi$ outside of the horizon, it is natural to concentrate on how to change $\mathcal{P}_\phi$ for one scalar field. To isolate the most important terms we consider $N=1$ version of Eq. (4.7f) with the damping and metric terms omitted. We are still free to modify (1) the occupation number $\bar{n}(\mathbf{k})$, (2) the Hubble parameter $H(t)=\dot{a}/a$, which enters in the Hubble drag term $-3H\delta\dot{\phi}_j$, (3) the diagonal component of the effective mass matrix $m_{11}^2 \equiv \partial^2 V/\partial\phi_1^2$, or (4) the curvature coupling constant ξ, which enters in $f=[m_P^2/(16\pi)-\xi\phi^2/2]$, the coefficient of the Ricci scalar in (3.1). $H(t)$, and $m_{ij}(t)$ vary in time according to the behavior of the background fields, which we can control by a suitable choice of potential. In the action (3.1), the curvature coupling term enters as an effective term in the

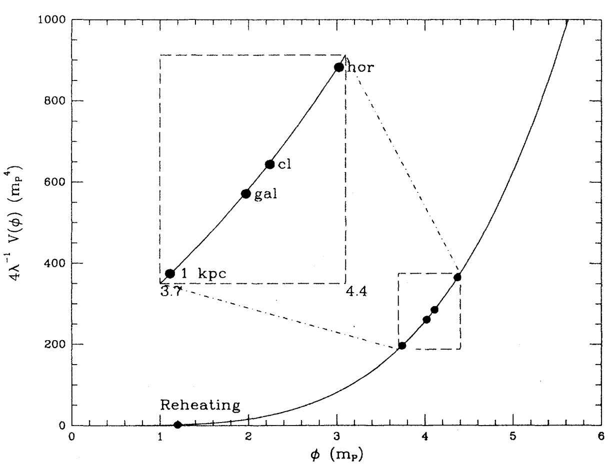

FIG. 4. In standard chaotic inflation, all observable structure in the Universe originates during the slow-rollover regime when ϕ is between $3.7m_P$ and $4.4m_P$, although inflation can presumably start with any initial value consistent with classical gravity, $\phi<\lambda^{-1/4}m_P\approx 3000m_P$. At $4.37m_P$, the present "horizon" scale, $\sim 10^4$ Mpc, would have left the horizon during inflation, assuming 60 e-foldings of inflation are required and $h=0.5$. (Depending upon the details of reheating this could be 60 ± 5.) Cluster scales (~ 10 Mpc) cross at $4.1m_P$, galaxy scales (~ 10 Mpc) at $4.0m_P$, and the post-recombination baryon Jeans length (~ 1 kpc), the smallest scale of gas condensations in the standard CDM model, at $3.74m_P$. Reheating occurs when $\phi\sim m_P$. Thus observations of cosmic structure probe only a small part of the potential surface and exotic fluctuation spectra are quite conceivable.

mass matrix; ξ may be positive or negative; it is 0 for minimally coupled fields, $\frac{1}{6}$ for conformally coupled fields, and is subject to renormalization so the value could evolve. Unfortunately, we cannot probe the role that the off-diagonal driving terms $m_{1j}^2\delta\phi_j$, $j\neq 1$, have on the behavior of $\delta\phi_1$ without doing a numerical integration. The role of such terms is considered in Sec. VI C.

As we discussed in Sec. IV A, it is usually assumed that the modes which are currently within our horizon would have had such large values of k/a relative to any characteristic energy scale describing mode occupation that $\bar{n}(\mathbf{k})$ can be taken to be zero, so that only zero-point fluctuations contribute to the spectrum. Since we are dealing with sub-Planck-scale physics it is by no means clear that this assumption is valid. If $\bar{n}(\mathbf{k})$ is nonzero then the fluctuation spectrum would be unlikely to be scale invariant, reflecting instead whatever physics would determine the primordial occupation of modes, and the fluctuations would be non-Gaussian.

An approximate solution to Eq. (4.7f) can be used to illustrate the effect of (slowly) varying the $H(t)$ and $m^2(t)$ histories: outside the horizon $k^{-1}<(Ha)^{-1}$, but before growth due to the perturbed metric occurs, the power spectrum is

$$\mathcal{P}_\phi(k)\approx[H(t_k)/(2\pi)]^2 \times\exp\left[-\int_{t_k}^{t}dt\,H(t)[3+n(t)]\right], \tag{6.2a}$$

$$n(t)\equiv \mathrm{Re}(-3\{1-4m^2(t)/[9H^2(t)]\}^{1/2}), \tag{6.2b}$$

$$\rho_\phi\propto\exp\left[-\int_{t_k}^{k}dt\,H(t)[3+n(t)]\right], \tag{6.2c}$$

$$a_k\equiv a(t_k)\equiv k/H . \tag{6.2d}$$

The time t_k and expansion factor a_k when a wave number k crosses the horizon are given by Eq. (6.2d). Equation (6.2a) is valid provided $n>0$ at t_k, and remains valid even if n subsequently vanishes $[2m/(3H)>1]$. If n is zero at t_k, then the prefactor $H^2(t_k)$ should be replaced by $H^3(t_k)/m(t_k)$. Note that if n is constant, then the exponential term in (6.2a) is simply $[k/(H(t_k)a)]^{3+n}$, a power law in k if H is constant. A number of interesting cases follow from this result:

1. Scale invariance

If $m^2\approx 0$, then the spectrum is scale invariant if H is independent of the expansion factor a, just as in Sec. VI A 1. The value of a when k first "crosses" the horizon is a_k. Since $H^2=(8\pi G/3)(\sum_j\dot{\phi}_j^2+V)$, judicious choice of V can lead to structure in $\mathcal{P}_\phi$. However, the most likely case over the ~ 6 orders of magnitude observable k range and the corresponding 14 e-foldings in a (which is relatively short compared to the total number of e-foldings possible during inflation) is that V will fall gently, leading to only slight deviations from scale invariance, with just a little more power on large scales than on small.

We can distinguish two types of behavior, depending upon whether the field is driving inflation (in which case it is the inflation) or its Hubble drag is driven by another field (in which case it is an isocon). The *inflation* leads to adiabatic fluctuations with the power spectrum $\mathcal{P}_\zeta \sim [3H(a_k)/\dot{\phi}]^2 \mathcal{P}_\phi(k,a_k)/H^2$. For an *isocon,* for which we take the axion as the generic case, the fluctuations in the axion mass density after the axion mass is generated is $\mathcal{P}_{\rho_{\mathcal{A}}} \propto \mathcal{P}_\phi$. In both cases, approximate scale invariance is the outcome. An isocon which preferentially dissipates into baryons rather than antibaryons can lead to isocurvature baryon perturbations. Again the δn_B spectrum will be scale invariant if the ϕ spectrum is. The constraints on adiabatic and isocurvature modes are discussed in Sec. II.

2. *Double inflation and plateaus*

Since $\mathcal{P}_\phi \propto H^2(a_k)$ for adiabatic fluctuations if $m=0$, one way to get more power on large scales is to let H drop, but still retain $m/H \ll 1$. The quantum fluctuation spectrum that arises if H drops suddenly is displayed in Fig. 5. It is easily computed using the exact solution of Eq. (4.7f) with metric terms neglected. The ringing would not appear if H were changed more gradually, which is more realistic if the change in H is attributed to the effects of two scalar fields with differing potentials (Sec. VI C). For the feature in $\mathcal{P}_\phi$ to appear at the preferred scale k_u (Sec. II) would require tuning the redshift of the drop quite precisely. This is the mechanism of double inflation. The generic outcome is a plateau spectrum.

Another way to get a plateau spectrum is to change the potential. If we allow for more than one degree of freedom for our scalar fields, then marvelous mountain ranges, valleys, and moguls can be envisaged for potential space. One's intuition regarding motion in this space can be utilized except that the rolling fields are subject to Hubble drag $(-3H\dot{\phi}_j)$ which can result in a terminal velocity (slow rollover), a phenomenon fundamental to the realization of inflation. Fluctuations $\delta\phi$ are small spreads in the field about the rolling background value $\bar{\phi}$. Fluctuations within the horizon at the specific epoch are still oscillating, while those outside have their shapes frozen in although the values may change with time. For some purposes it is useful to include those waves outside the horizon with the background field, to allow for a gentle variation from place to place.[63]

If there are a number of flat directions in potential space, inflation could be a complicated process, modifying the H profile with time and also the form of the mass matrix m_{ij}^2. This will certainly map onto structure in the fluctuation spectrum. However, to transform specific features in potential space to features in a particular range in k-space, we must arrange for the fields to pass through this V structure at a specific range of a values.

Consider first the case of two scalars having a potential forming a broad valley with the valley minimum line very gently sloping down towards the origin with somewhat steeper walls rising away from the minimum. This configuration leads to double inflation.[15-18] If the two-dimensional field begins high enough on the wall away from the origin, it will be potential dominated by the wall part, and experience inflation with a large value of H. The field will roll down towards the valley minimum, oscillating in one direction ($m^2>0$ so the power in the field in this direction damps away as $a^{-(3+n)}$ as above), while continuing to roll down towards the origin in the other direction with a lower value of H. The field in the second direction is all that is left after the end of inflation. Since it experiences first the high H for long waves as they leave the horizon, then the low H value for short waves, a plateau structure is generic. The ramp between the two levels will depend upon the specific form of the potential. To arrange for the location of the ramp to be tuned to an astrophysically interesting scale, the initial location of the field matters. If the field remains near the valley minimum for too long a period then the spectrum within the current Hubble length will only reflect the low value of H and be effectively scale invariant.

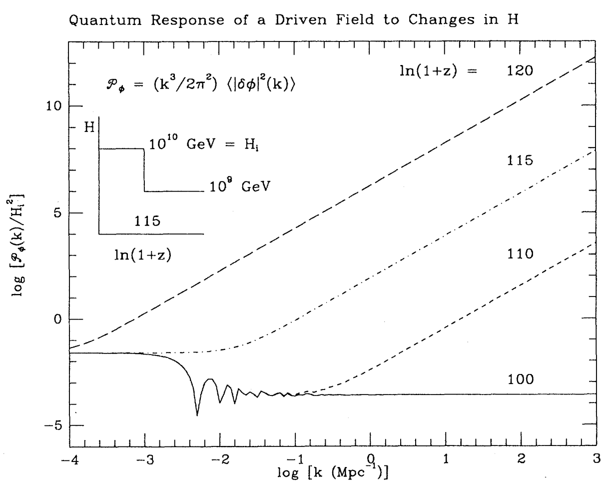

FIG. 5. Several snapshots in time of the $\mathcal{P}_\phi$ spectrum are shown as the Hubble constant drops abruptly from 10^{10} to 10^9 GeV in a simple model of double inflation. The earliest shows all scales within the horizon, where the modes are oscillating vigorously, and the spectrum is $\propto k^2$, Eq. (5.1b). As the scales leave the horizon, they freeze-out at a value equal to the Hawking temperature, $H/2\pi$, thus exhibiting a flat spectrum at large k and a plateau at small k. The transition region in between would be smoother if the Hubble constant varied continuously (see Fig. 8).

3. *Mountains*

Choosing m^2 positive or negative can give spectra rising or falling with increasing k, but at the expense of exponential decreases or increases in both the background field energy density and in the fluctuation power—assuming inflation is continuing.

If m^2 is fixed, the drop of H with time eventually leads to n reaching zero. The field oscillates coherently with an energy density averaged over an oscillation period decreasing as $\rho_\phi \sim a^{-3}$ like nonrelativistic matter, with fractional fluctuations $\mathcal{P}_\phi/\rho_\phi$ being constant. This is the mechanism by which the axion, once it attains its mass when the temperature of the Universe is about 200 MeV, behaves like cold nonrelativistic matter. However this occurs *after* inflation, so the a^{-3} law is not devastating.

To have a field whose fluctuations are still of interest we can only have m^2 positive or negative over a limited regime *during* inflation. To shape a mountain of power, we would want m^2 to begin positive, then become negative for about five e-foldings, and finally settle to zero. An example of this is shown in Fig. 6(b). To ensure the mountain is not too high or too short, too broad or too steep, a highly precise balance between positive and negative $m^2(t)$ behavior is required. This is in addition to the fine-tuning associated with fixing the position of the peak. If instead m^2 becomes negative first, the plateau plus ramp structure will arise, as is shown in Fig. 6(a).

If, in the example of Fig. 6(b), the driven (isocon) field had m^2 becoming positive rather then finally dropping to zero, its mountain spectrum would inflate away. This transient could still leave an imprint upon the observed fluctuations through its coupling to other fields, such as the inflaton. A mogul in potential space provides a realization of this mechanism, as discussed in Secs. VI C and VI D.

4. *Power laws*

One way to get extended power laws over some k range is to arrange for n to be constant over the associated range in a_k. This would be possible if m^2 scales with H^2. The natural way for this to occur is to make use of nonzero curvature coupling constants ξ, for then the local spectral index for isocons is $n=-3(1-16\xi/3)^{1/2}$. The severe price to pay is the $\sim a^{-(3+n)}$ fall of the energy density in the field throughout inflation. Only for $n\approx -2$ might we envisage getting the perturbation strength back by relative growth compared with the radiation, and this requires the very special choice $\xi\approx\frac{5}{48}$, which could prove disastrous if it causes Newton's constant to change sign (Sec. VI D). It is easier to contemplate ξ negative with $n<-3$, falling to short wavelengths, although this could be valid only for a limited time to avoid precipitous growth of the field. We have assumed that the isocon contribution to H is negligible and that $|\xi|$ is not very large. The $\xi \ll 0$ case is more complicated. (See the treatment for inflatons in Sec. VII.)

Another way to get power laws has been to invoke power-law inflation,[64,65] with the expansion going as $a\sim t^q$, with $q>1$ to ensure $\ddot{a}>0$. This necessitates an equation of state yielding fixed $p/\rho=-1+2/(3q)$. For the scalar field which drives power-law inflation, $\mathcal{P}_\phi\sim(H/2\pi)^2\sim k^{(3+n)}$, where $n=-(3q-1)/(q-1)$ $=3(1-p/\rho)/(1+3p/\rho)$. As shown in Sec. VII, this is also the k dependence of $\mathcal{P}_\zeta$, hence the density fluctuations subsequently generated would have the power-law index $n_\rho=n+4$. This always gives $n\le -3$, $n_\rho\le 1$, if inflation is realized, hence there is more power at large scales than the Zeldovich spectrum. A disadvantage of such spectra is that the redshift of cluster and galaxy formation will not be well separated for power laws with $n_\rho\lesssim 0$. Further, if n is too negative, large-angle CMB anisotropies become too large. With an interaction potential of the specific form

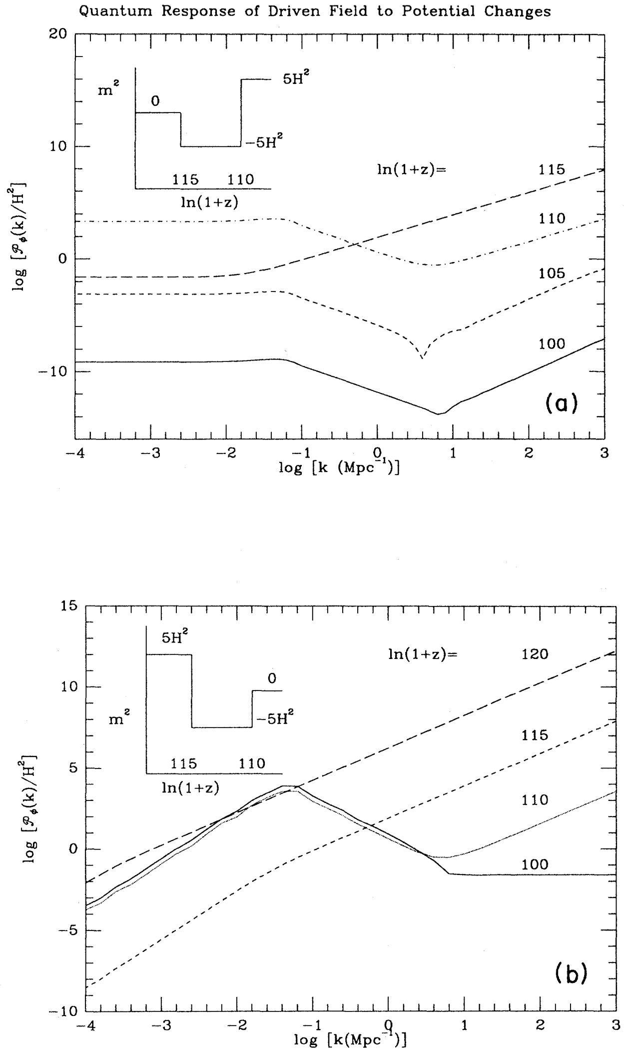

FIG. 6. Another way to modify the evolution of the mode functions of the scalar field (ϕ_1) is to change the effective masses m_{ij}^2. This figure illustrates the influence on $\mathcal{P}_\phi$ of sudden changes of the effective m_{11}^2, as indicated. In (a) m^2 passes from zero to a negative value then to a positive one; (b) is the reverse. During the negative phase, $\mathcal{P}_\phi$ grows for waves outside the horizon, while maintaining a spectrum of form Eq. (5.1b) inside provided $(k/a)^2 \ll |m_{11}^2|$. If m_{11}^2 were then to return to zero, a spectrum with more power on large scales would result. For positive m_{11}^2, fluctuations outside the horizon oscillate with frequency m_{11}, with an envelope decaying as a^{-3}, the law of decline of the density of nonrelativistic matter. If inflation continues, the ϕ fluctuations would diminish exponentially. The generic spectrum has plateau initial conditions. The reverse history for m_{11}^2 shown in (b) leads to mountain initial conditions. Small wave numbers damp since they are outside the horizon when m_{11}^2 is positive, intermediate wave numbers grow, since they leave the horizon when $m_{11}^2<0$, generating a ramp slowing downward toward higher k which leave the horizon later, and large wave numbers that leave the horizon when $m_{11}^2=0$ remain constant.

$V(\phi)=V_0\exp[-(16\pi/q)^{1/2}\phi/m_P]$ for the inflaton, power-law inflation could be realized, with n related to q as given above. Power-law inflation could also drive an isocon to develop power-law isocurvature perturbations. For example, a Goldstone boson such as the axion would have $\mathcal{P}_\phi \sim k^{3+n}$, with n also as given above. The axion density perturbations developed once the axion mass is generated would have the same power n; such isocurvature CDM spectra can be strongly ruled out by large-angle CMB anisotropies.[36]

From this discussion, there does not seem to be much hope that the $-1 \lesssim n \lesssim 0$ power-law isocurvature baryon spectra advocated by Peebles[66] will arise within the inflationary paradigm. In any case, these models would naturally have $\Omega_B=1$ yet smaller values $\Omega_B \sim 0.2-0.4$ are apparently preferred observationally. To agree with the $\Omega=1$ requirement of inflation, the remaining energy density would have to be made up with vacuum energy ($\Lambda \neq 0$), which has severe fine-tuning problems.

C. Specific realizations of broken scale invariance

Consider two minimally coupled scalar fields interacting through a chaotic inflation potential containing only quadratic and quartic terms:

$$V(\phi_1,\phi_2)=\frac{m_2^2\phi_2^2}{2}+\frac{\lambda_2\phi_2^4}{4}+\frac{m_1^2\phi_1^2}{2}+\frac{\lambda_1\phi_1^4}{4}-\frac{\nu\phi_1^2\phi_2^2}{2}+V_0 . \tag{6.3}$$

The constant V_0 is chosen so that the minimum of the potential is zero. We first consider the simplest case in which the scalar fields are decoupled, then add the extra complication of $\nu \neq 0$. Finally we add a cubic term $-\mu_1\phi_1\phi_2^2$ to show the effects of breaking the $\phi_1 \to -\phi_1$ symmetry. These cases illustrate the range of behavior expected for more general interaction potentials, as we discuss in Sec. VI D.

1. Decoupled scalars: $\nu=0$

If $\nu=0$, then inflation is likely to be double inflation: If $\lambda_1 \gg \lambda_2$, first ϕ_1 dominates H, then ϕ_2. We again assume m_i is small and that both fields couple to matter with a dissipative term of form Eq. (4.4b). The free parameters in the model are $\phi_1(t_i)$ and $\phi_2(t_i)$. We adopt the following standard set of potential and damping parameters: $\lambda_1=5\times10^{-10}$, $\lambda_2=5\times10^{-14}$, $m_1=10^{-6}m_P$, $m_2=10^{-8}m_P$, $\nu=0$, $\Gamma_1=10^{-3}m_1$, $\Gamma_2=10^{-1}m_2$; we also take $g_{\rm eff}=160.75$.

In Fig. 7(a) the evolution of the background fields $\bar{\phi}_j$ and the Hubble parameter H is given for the initial conditions $\phi_1(t_i)=2.3m_P$ and $\phi_2(t_i)=4.36m_P$. ϕ_1 dominates the energy density until redshift $\sim e^{123}$ when ϕ_2 takes over. Once the ϕ_1 energy is no longer potential dominated ($z\sim e^{70}$), ρ_{ϕ_1} behaves as a^{-3}, unless it damps into radiation, in which case the energy density falls as a^{-4}. Since ϕ_2 is rolling down its very flat potential so slowly, it will still be driving inflation at this time. It is therefore unavoidable that the density fluctuations in ϕ_1 or the radiation it decays into are unimportant.

The power spectrum of the ϕ_2 field follows the behav-

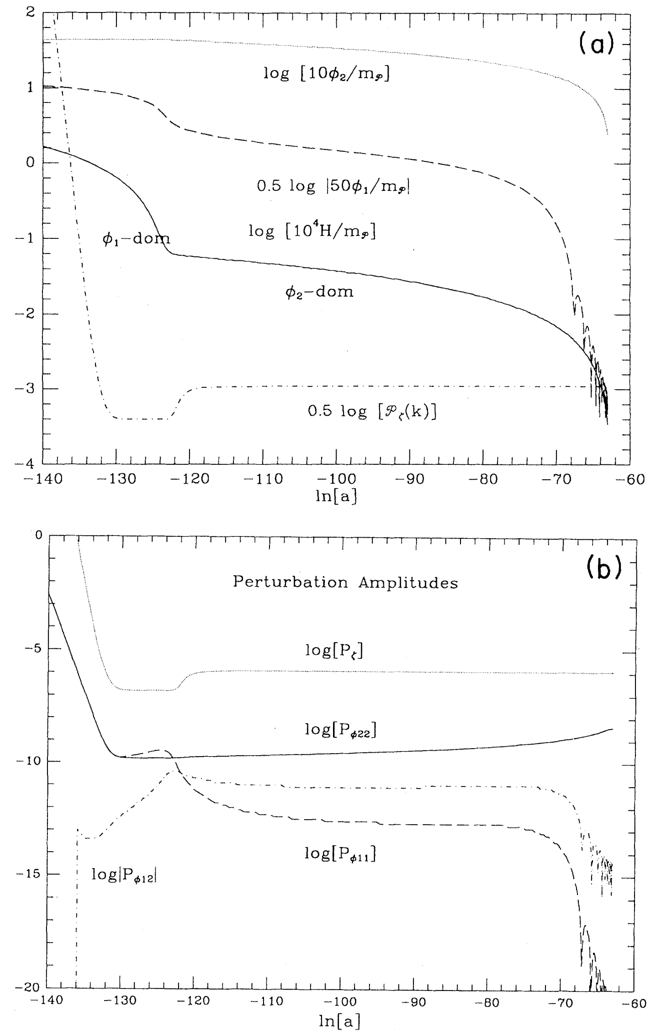

FIG. 7. (a) Evolution in a typical double inflationary model with two scalar fields interacting only through gravity. Since $\lambda_1 \gg \lambda_2$, ϕ_1 dominates the energy density of the Universe initially, but it quickly rolls down its potential, becoming energetically unimportant after redshift $z=e^{123}$, although it does not start to oscillate about the minimum of its potential until $z=e^{70}$. [The integration variable $\ln a=-\ln(1+z)$ in terms of the redshift.] The Hubble parameter reflects the behavior of the dominant field, becoming approximately constant once ϕ_2 dominates. When the scale $k^{-1}=5000$ h^{-1} Mpc leaves the horizon at $z=e^{132}$, $\mathcal{P}_\zeta$ acquires a value of $[k^3/(2\pi^2)](3H\psi_{11}/\dot{\phi}_1)^2$, where all quantities are evaluated at horizon crossing, but grows to the constant result of $[k^3/(2\pi^2)](3H\psi_{22}/\dot{\phi}_2)^2$, where now all values are determined at the time that ϕ_2 dominates. (b) Longitudinal gauge calculations of quantum fluctuations are shown for the scale $k^{-1}=5000$ h^{-1}Mpc in a double inflation scenario. The cross correlation $\mathcal{P}_{\phi12}$, which is set initially to zero, grows because of gravity, then freezes out at the start of the ϕ_2-dominated era, and finally decays when the mass of ϕ_1 exceeds the Hubble parameter. Both $\mathcal{P}_{\phi11}$ and $\mathcal{P}_{\phi22}$ cross the horizon with the same value $H_{\rm cr}/(2\pi)$. Whereas $\mathcal{P}_{\phi22}$ remains essentially constant, $\mathcal{P}_{\phi11}$ grows slightly because of the gravitational interaction, then quickly damps when m_{11} is larger than H. It freezes out when the reverse is true, and eventually decays in time much like $\mathcal{P}_{\phi12}$.

ior of H^2. The mechanism is identical to that of Fig. 5, although here the drop in H is gentler. Note that $\phi_1(t_i)$ is essentially irrelevant for the production of a break, since it just determines the number of e-foldings until ϕ_2 takes over. The precise value of $\phi_2(t_i)$ will determine the number of e-foldings until the end of inflation, thereby determining the location of the ramp in the spectrum.

Figure 7(b) displays the time evolution of the perturbation variables in the longitudinal gauge for a specific choice of k^{-1} (5000 h^{-1}Mpc). This figure illustrates how the computational procedure for more than one scalar field described in Secs. IV and V proceeds in practice. The calculations are based on two independent runs, the first using the quantum fluctuations ψ_{11} and ψ_{21} and the second using ψ_{12} and ψ_{22}. The power spectra of the correlation functions (4.2b) mix the output of the two runs:

$$\mathcal{P}_{\phi 11}=\frac{k^3}{2\pi^2}(|\psi_{11}|^2+|\psi_{12}|^2) , \tag{6.4a}$$

$$\mathcal{P}_{\phi 12}=\mathcal{P}^*_{\phi 21}=\frac{k^3}{2\pi^2}(\psi_{11}\psi^*_{21}+\psi_{12}\psi^*_{22}) , \tag{6.4b}$$

$$\mathcal{P}_{\phi 22}=\frac{k^3}{2\pi^2}(|\psi_{21}|^2+|\psi_{22}|^2) . \tag{6.4c}$$

The integration of the fluctuation equations (4.8) begins at $\ln a=-136$. At this time the cross-correlation power spectrum, $|\mathcal{P}_{\phi 12}|$ is set to zero. It grows in time because of the metric terms in Eq. (4.8b). (The small overshoot at the beginning is a consequence of starting at zero rather than at the true value.) It freezes out at the end of the ϕ_1-dominated era, when the Hubble parameter exceeds the effective mass of ϕ_1, then oscillates and damps away when the reverse is true at $\ln a=-70$. $\mathcal{P}_{\phi 22}(k)$ remains relatively near its horizon-crossing value $[H(a_k)/(2\pi)]^2$. $\mathcal{P}_{\phi 11}(k)$ also leaves with this value, grows slightly because of gravity, and then follows a path similar to $|\mathcal{P}_{\phi 12}|$. Both reflect the behavior of ψ_{12} since, in these latter phases, only the ψ_{22} and ψ_{12} terms contribute, so that $\ln|\mathcal{P}_{\phi 12}|$ is the average of $\ln\mathcal{P}_{\phi 22}(k)$ and $\ln\mathcal{P}_{\phi 11}(k)$.

The ζ spectrum also has two contributions, $\mathcal{P}_\zeta=k^3(|\zeta_1|^2+|\zeta_2|^2)/(2\pi^2)$, where $\zeta(k,t)=\zeta_1 a_1+\zeta_2 a_2$ in the language of Sec. IV. Outside the horizon and long before reheating, the dominant term for ζ in Eq. (2.3) is the one with the $1+p/\rho$ denominator. It may be evaluated using Eq. (5.5c) for $\dot{\Phi}_H$:

$$\zeta_1\approx-3H(\dot\phi_1\psi_{11}+\dot\phi_2\psi_{21})/(\dot\phi_1^2+\dot\phi_2^2) , \tag{6.5a}$$

$$\zeta_2\approx-3H(\dot\phi_1\psi_{12}+\dot\phi_2\psi_{22})/(\dot\phi_1^2+\dot\phi_2^2) . \tag{6.5b}$$

When the scale leaves the horizon during the ϕ_1-dominated regime, $\zeta_1\approx-3H\psi_{11}/\dot\phi_1$ and $\zeta_2\approx-3H\psi_{22}\dot\phi_2/\dot\phi_1^2$, and ζ_1 sets the value of $\mathcal{P}_\zeta$. This holds until $\ln a=-120$, when ϕ_2 dominates; then ζ_1 and ζ_2 quickly reach constant values of $-3H\psi_{21}/\dot\phi_2$ and $-3H\psi_{22}/\dot\phi_2$, and ζ_2 essentially determines the final value of $\mathcal{P}_\zeta$. This explains why $\mathcal{P}_\zeta$ rises from one constant value to another. We emphasize that $\mathcal{P}_\zeta^{1/2}$ does not reach its asymptotic value, $\propto\psi_{22}\approx H(a_k)/(2\pi)$, until ϕ_2 dominates. This is the naive value that one would have expected.

In Fig. 8 the ζ spectrum for the initial conditions of Fig. 7 (solid line) demonstrates that $\mathcal{P}_\zeta$ follows the evolution of H closely. These conditions are those required to produce an initial spectrum of the CDM+plateau form (long dashed). However, small variations in the initial value of the ϕ_2 background field move the position of the rise to the plateau: e.g., the dot-dashed curves with $\phi_1(t_i)=2.3m_P$, $\phi_2(t_i)=4.4m_P$ (ramp at low k), and with $\phi_1(t_i)=2.3m_P$, $\phi_2(t_i)=4.3m_P$ (ramp at high k). The short-dashed curve shows that it is also easy to get different shapes by flattening the ϕ_1 potential, $\lambda_1=10^{-10}$ instead of the standard $\lambda_1=5\times10^{-10}$; the initial conditions chosen were $\phi_1(t_i)=2.3m_P$, $\phi_2(t_i)=4.36m_P$. The spectrum for chaotic inflation with a single scalar field ϕ_2 with standard parameters and initial value $\bar\phi_2(t_i)=5m_P$ are shown for comparison. We conclude that two decoupled scalar fields give plateau initial fluctuation spectra. To reproduce the phenomenological spectrum involves a careful choice of the initial value $\bar\phi_2(t_i)$ to position the break at the required location and of the relative amplitude of the scalars' potential parameters to ensure the required shape.

2. *Interacting scalars: $v\neq0$*

We now consider the richer possibilities that can arise if $v\neq0$. In some circumstances, mountains can arise. We know from Sec. VI B 3 what the critical features are for the potential surface to lead to mountain spectra: one of

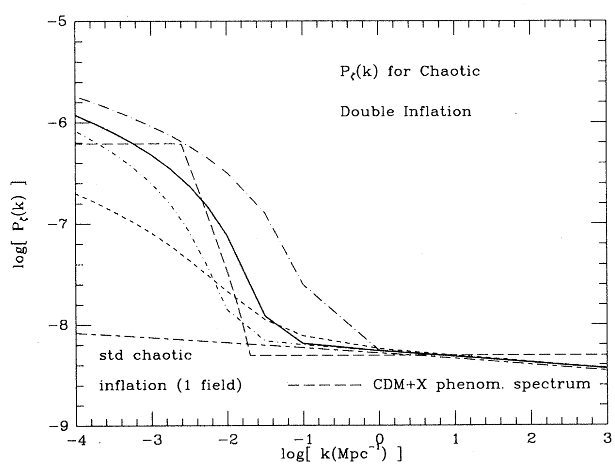

FIG. 8. Double inflation spectra for models with various potential parameters and background-field initial conditions. The solid curve has $\phi_1(t_i)=2.3m_P$ and $\phi_2(t_i)=4.36m_P$ in order to fit the CDM+plateau phenomenological spectrum (long dashed). If $\phi_2(t_i)$ is varied by $\approx1\%$, the dotted-dashed curves are the result. These spectra generate dramatically different large-scale structure than that for the solid curve. The short-dashed curve illustrates the result of decreasing λ_1 by a factor of 5 to $\lambda_1=10^{-10}$. All models assume $\lambda_2=5\times10^{-14}$, and the lowest-lying curve has $\phi_1(t_i)=0$, $\phi_2(t_i)=5m_P$, and thus is the standard inflationary model for a single scalar field.

the fields must enter into a region of negative mass squared. If the mountain is to survive in the isocon field, it is also essential to avoid a prolonged $m_{\rm eff}^2 > 0$ region for the isocon while inflation proceeds. A transient mountain in the isocon field spectrum can however imprint a mountain in the inflaton field spectrum, leading to adiabatic mountain spectra. It is now only necessary to work with the parameters of the potential (6.3) to demonstrate concrete examples.

It is not possible[18] to get mountains if $\nu > 0$. The potential surface has a valley where $\partial V/\partial\phi_1$ vanishes, given by a hyperbola centered about the ϕ_2 axis,

$$\phi_{1\rm tr}(t) = \pm\left[\frac{\nu\phi_2^2(t) - m_1^2}{\lambda_1}\right]^{1/2}, \quad \phi_2 > \frac{m_1}{\sqrt{\nu}}, \tag{6.6a}$$

whose two branches converge to a single trough along the ϕ_2 axis for smaller ϕ_2:

$$\phi_{1\rm tr}(t) = 0, \quad \phi_2 < \frac{m_1}{\sqrt{\nu}}. \tag{6.6b}$$

There would be a double inflation phase while ϕ_1 rolls down to its valley. After this the ϕ_2 field drives inflation. It always has $m_{22}^2 > 0$. However, the effective mass in the ϕ_1 direction,

$$m_{11}^2 \equiv \partial^2 V/\partial\phi_1^2 = 3\lambda_1\phi_1^2 - (\nu\phi_2^2 - m_1^2), \tag{6.7}$$

is negative within the lobes of the trough ($|\phi_1| < |\phi_{1\rm tr}|/\sqrt{3}$). However, it remains positive along the trough, except at $\phi_2 = m_1/\sqrt{\nu}$. Kofman and Linde[16] envisaged a model where the first scalar field would oscillate about $\phi_{1\rm tr}$. They assumed that ϕ_1 would find its way into the region of negative m_{11}^2 when $\phi_2 \approx m_1/\sqrt{\nu}$, leading to a mountain shape. However, numerical and analytical calculations show that m_{11}^2 is always positive about the true trajectory in (ϕ_1, ϕ_2) space. A typical run is shown in Fig. 9. The interaction terms were chosen to be $\lambda_1 = 10^{-9}$, $\lambda_2 = 10^{-14}$, $\nu = 10^{-12}$, the mass terms to be $m_1 = 4.5\times 10^{-6} m_P$, $m_2 = 10^{-8} m_P$, and the friction terms to be $\Gamma_1 = 10^{-3} m_1$, $\Gamma_2 = 10^{-1} m_2$; here $g_{\rm eff} = 1.0$. Although we started ϕ_1 in its trough, this figure demonstrates that ϕ_1 deviated from its trough value as ϕ_2 approached $m_1/\sqrt{\nu}$, and, as a result, m_{11}^2 never became negative.

This result is generally true: Hubble drag ensures that the ϕ_1 field will lag behind its trough value provided that the rough valley is curved. We can illustrate this in linear perturbation theory, assuming the background field ϕ_1 deviates from $\phi_{1\rm tr}$ by a small amount

$$f(t) \equiv \phi_1(t) - \phi_{1\rm tr}(t). \tag{6.8}$$

Substituting (6.8) into (4.6c), neglecting particle damping, and retaining only first-order terms, we find that f satisfies the equation

$$\ddot{f} + 3H\dot{f} + \left[\frac{\partial^2 V}{\partial\phi_1^2} f + \ddot{\phi}_{1\rm tr} + 3H\dot{\phi}_{1\rm tr}\right] = 0. \tag{6.9}$$

In about an e-folding, f will settle to the value

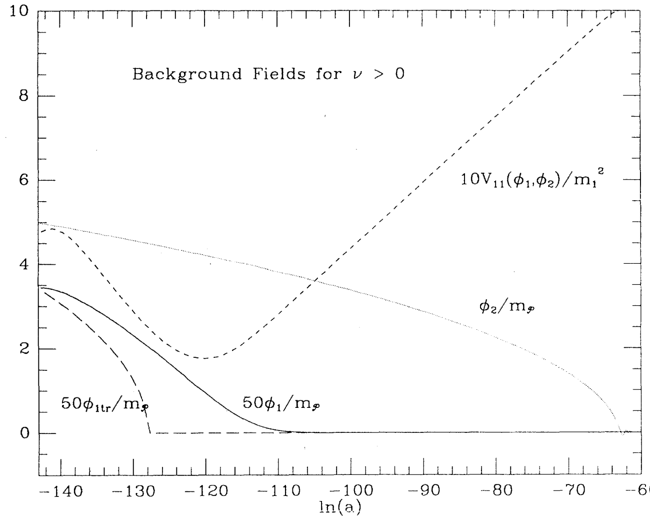

FIG. 9. Models with coupled scalars and $\nu > 0$ do not generate CDM + mountain spectra. While ϕ_2 is in the slow-rollover regime, ϕ_1 settles into its trough, but it deviates from its trough value as $\phi_2 \approx m_1/\sqrt{\nu}$. As a result ϕ_1 never reaches a region of negative mass squared.

$$f = -(\ddot{\phi}_{1\rm tr} + 3H\dot{\phi}_{1\rm tr})/m_{11}^2. \tag{6.10}$$

The second term is dominant. On the $\phi_1 > 0$ side, $\dot{\phi}_{1\rm tr} < 0$, hence f is positive and the lag from the trough is always in the opposite direction to the $m_{11}^2 < 0$ region: negative mass squared is therefore unattainable.

The $\nu > 0$, $m_1^2 < 0$ case has a minimum which is offset from $(\phi_1, \phi_2) = (0,0)$. In this situation, the region of negative mass squared is separated from the trough by a finite amount, and it seems unlikely that the scalar will deviate sufficiently from the trough value to reach this area.

A negative mass-squared region *can* be reached if $\nu < 0$, provided we also take $m_1^2 < 0$, as Kofman and Pogosyan[19] (KP) suggested. (If $m_1^2 > 0$, the trough lies along the ϕ_2 axis, with $m_{11}^2 > 0$, leading to double inflation with no extra ingredients.) The $\nu < 0$, $m_1^2 < 0$ potential surface is somewhat similar to the $\nu > 0$, $m_1^2 > 0$ surface, except that it is inverted:

$$\phi_{1\rm tr} = 0, \quad \phi_2 > (m_1^2/\nu)^{1/2}, \tag{6.11a}$$

and it bifurcates into an ellipse for smaller ϕ_2:

$$\phi_{1\rm tr} = \pm[(|m_1^2| - |\nu|\phi_2^2)/\lambda_1]^{1/2}, \quad \phi_2 < (m_1^2/\nu)^{1/2}. \tag{6.11b}$$

The $m_{11}^2 < 0$ region is shown by Eq. (6.7) to be an ellipse with the same semimajor axis, but a smaller semiminor axis. Isocurvature Gaussian mountain spectra do not survive at an interesting level in this model, since m_{11}^2 is positive in the trough. However, adiabatic Gaussian mountain spectra can arise. Nonetheless to achieve a prescribed location, width, and amplitude in k space, three parameters must be tuned. We illustrate this sensitivity in Figs. 10 and 11.

Figure 10(a) shows the evolution of the background fields and Fig. 10(b) of the perturbation variables for a specific wave number, $k^{-1}=158\ h^{-1}$ Mpc. The parameters chosen for Fig. 10, $\lambda_1=10^{-5}$, $\lambda_2=2\times10^{-13}$, $\nu=-7.935\times10^{-10}$, $(-m_1^2)^{1/2}=10^{-4}m_P$, $m_2=10^{-7}m_P$, $\Gamma_1=0$ and $\Gamma_2=0.1m_2$, give a mountain spectrum which is not violated by any observational data, and is similar to the phenomenological CDM+mt spectrum of Fig. 1. Of these, m_1^2 and ν are crucial. The spectrum is extremely sensitive to the initial value $\phi_2(t_i)$, which controls the

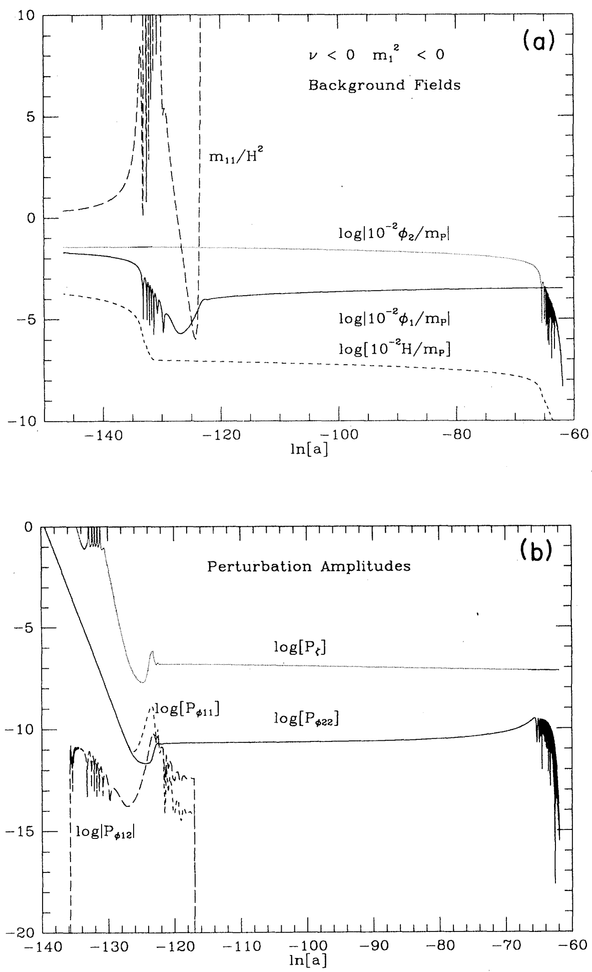

FIG. 10. Mountain spectra do arise when $\nu<0$ and $m_1^2<0$. ϕ_1 dominates the energy density of the Universe initially, but it quickly rolls down, oscillates, and damps to a small value, as (a) shows. Once ϕ_2 decreases below $(m_1^2/\nu)^{1/2}=3.55m_P$, m_{11}^2 becomes negative for about four e-foldings; ϕ_1 then grows very rapidly until it settles in its trough. Thereafter, ϕ_2 dominates, and at $z=e^{67}$ it oscillates because its effective mass exceeds the damping rate. In (b), the perturbation variables for $k^{-1}=158$ h^{-1}Mpc are shown. The rapid increase of $\mathcal{P}_{\phi 11}$ when $m_{11}^2<0$ is evident. Since ϕ_2 is coupled to ϕ_1 by the $-\nu\phi_1^2\phi_2^2/2$ term of the potential, both $\mathcal{P}_{\phi 22}$ and $\mathcal{P}_{\phi 12}$ also grow very rapidly. The net result is a peak in the $\mathcal{P}_{\phi 22}$ and $\mathcal{P}_\zeta$ spectra. After $z=e^{120}$, $\mathcal{P}_\zeta$ decreases by a factor of 2 because $\phi_{1\mathrm{tr}}$ is time dependent.

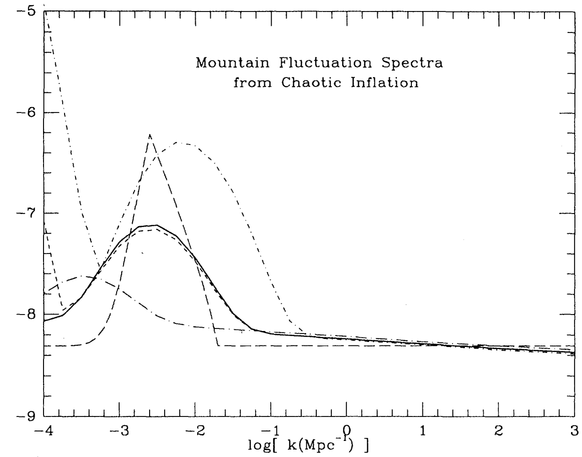

FIG. 11. Mountain spectra generated in a variety of models with $\nu<0$ and $m_1^2<0$ are shown. The short-dashed curve gives healthy cluster correlations and large-scale streaming velocities, and yet is within the RELICT CMB anisotropy limits. At small k, the amplitude is very high because these scales left the horizon during the first of the two inflation epochs. The parameters, m_1^2, ν, and $\phi_2(t_i)$ (listed in the text) must be finely adjusted to give the desired shape and position. For example, the dotted short-dashed and the dotted long-dashed-curves show the effects of varying $\phi_2(t_i)$ by less than a percent. The phenomenological CDM+mountain spectrum of Fig. 1 is shown for comparison (long-dashed). If a cubic term $-\mu\phi_1\phi_2^2$ is present in the potential, one can avoid the tuning of $\phi_2(t_i)$ but at the expense of small μ, $\mu=1.3\times10^{-15}$ (dark solid curve).

height and width of the mountain. In Fig. 10 $\phi_2(t_i)=3.71m_P$ and $\phi_1(t_i)\approx 2m_P$. Initially, ϕ_2 remains frozen at $\phi_2(t_i)$ while ϕ_1 dominates the potential before it rolls down to its trough where it oscillates and Hubble damps. Double inflation is the result, as indicated by the drop in H at $\ln(a)=-135$ in Fig. 10(a). Before the ϕ_1 oscillations have fully damped, the slow rolldown of ϕ_2 to below $(m_1^2/\nu)^{1/2}=3.55m_P$ causes ϕ_1 to enter the $m_{11}^2<0$ region, remaining there for about 4 e-foldings. The effective mass squared reaches a minimum of $-6H^2$. During this time ϕ_1 grows extremely rapidly until it reaches its new trough, given by Eq. (6.11b), where it oscillates and dies away. At this stage, we have two relevant time scales which are quite different, the ϕ_1 oscillation period and the ϕ_2 rolldown time. To follow the details of the ϕ_1 oscillations as it Hubble damps is computationally demanding and unimportant physically. Accordingly, we set $\phi_1=\phi_{1\mathrm{tr}}$ once ϕ_2 falls below $0.9(m_1^2/\nu)^{1/2}$ and ρ_{ϕ_1} falls below $<10^{-5}\rho_{\phi_2}$. Similarly, the fluctuations in the first scalar field are set to zero at this point. (See Appendix C for details.) With ϕ_1 in its trough at the end of inflation, the effective mass of ϕ_2 exceeds the damping rate Γ_2; hence, ϕ_2 oscillates for many periods before the Universe reheats, in contrast with the situation encountered in Fig. 7 even though the same Γ_2/m_2 was used.

Figure 10(b) shows the evolution of the perturbation variables. At $\ln a=-126$, the quantum fluctuation ψ_{11}

feels the effect of having a negative mass squared and grows exponentially. Because of the coupling term $-\nu\phi_1^2\phi_2^2/2$ in the potential, ψ_{22} behaves similarly. The abrupt drop of $\mathcal{P}_{\phi 11}$ and $\mathcal{P}_{\phi 12}$ to zero at $\ln a = -117$ occurs where we froze ϕ_1 into its trough to facilitate the numerical computation as mentioned above. Unlike our previous models, $\mathcal{P}_\zeta$ is not strictly constant after this time since the potential seen by ϕ_2 is time dependent—as is evident if one replaces ϕ_1 by $\phi_{1\mathrm{tr}}$ in Eq. (6.3). In this model, $\mathcal{P}_\zeta$ decreases by a factor of 2 from $\ln a = -120$ to $\ln a = -62$.

Our computed spectrum for the model of Fig. 10 is the short-dashed line in Fig. 11. The mountain peaks at $k^{-1} = 250\ h^{-1}$Mpc. There is sufficient large-scale power to produce large cluster correlations and large coherent streaming velocities. At the largest scales the power spectrum grows steeply, reflecting the epoch of ϕ_1 dominance, when the Hubble parameter was large. Figure 11 also illustrates how small variations in ν and $\phi_2(t_i)$ dramatically affect the position and width of the mountain. The dotted-long-dashed curve has $\nu = -7.506\times 10^{-10}$, $\phi_2(t_i) = 3.81 m_P$, and the dotted-short-dashed curve has $\nu = -8.16\times 10^{-10}$, $\phi_2(t_i) = 3.66 m_P$. All other parameters are the same. Although one must ultimately use trial and error to produce these curves, the order of magnitude of the parameters may be found through the following prescription: $\lambda_2 \approx 10^{-14}$ determines the large-k behavior; $\phi_2(t_i) \approx (m_1^2/\nu)^{1/2} \approx 4 m_P$ will produce a mountain near cluster scales; $(-m_1^2)^{1/2} \sim 10H$ yields a sufficiently large mogul; $[\phi_2(t_i) - (m_1^2/\nu)^{1/2}] \approx 0.2 m_P$ determines the width of the mountain; and finally $\lambda_1 > \nu^2/\lambda_2$ ensures that the effective coefficient of ϕ_2^4 is positive when $\phi_1 = \phi_{1\mathrm{Tr}}$. We argue in the next section that the more interesting case of $\phi_1(t_i)$ beginning exactly in the trough leads to non-Gaussian ϕ_2 fluctuations and that the perturbative treatment adopted here breaks down. The reason for this is that both the ridge and the trough have $\phi_1 = 0$.

One can, however, get Gaussian mountain fluctuations by ensuring that the mogul ridge is offset from the trough. To generally define ridges and troughs, we must consider the structure of a potential surface for N scalar fields, $V(\phi_1, \ldots, \phi_N)$. There will be one inflaton field, ϕ_N say, and $N-1$ isocon directions. The extrema $\partial V/\partial\phi_i = 0$, $i = 1, \ldots, N-1$ define curves parametrized by ϕ_N. For troughs m_{ij}^2 is positive definite while for ridges m_{ij}^2 is negative definite. At points where the $(N-1)\times(N-1)$ matrix m_{ij}^2 is singular, new curves appear. For $N > 2$, saddle lines as well as minimum and maximum lines arise, complicating the analysis. For the $N = 2$ case of interest here, ridges $\phi_{1\mathrm{ri}}(\phi_2)$ and troughs $\phi_{1\mathrm{tr}}(\phi_2)$ are the only possibilities. In the symmetric case discussed above, the $\phi_{1\mathrm{tr}} = 0$ curve goes smoothly over into the $\phi_{1\mathrm{ri}} = 0$ curve defining the backbone of the mogul at $\phi_2 = (m_1^2/\nu)^{1/2}$. Ensuring instead that $\phi_{1\mathrm{ri}} = 0$ is not a continuous extension of $\phi_{1\mathrm{tr}} = 0$ avoids the non-Gaussian behavior. With the natural starting condition $\bar\phi_1 = \bar\phi_{1\mathrm{tr}}$, appropriate if the ϕ_1-driven inflation epoch is over by the time the 10^4 Mpc wave leaves the horizon, the trajectory $\bar\phi_1(t)$ then passes along a specific side of the mogul. The final spectrum becomes sensitive to the level of offset, defined by the potential, rather than by the initial condition $\phi_2(t_i)$, which no longer requires such fine-tuning.

One realization of this is to add a cubic term, $-\mu_1\phi_1\phi_2^2$, to the potential (6.3). The trough and the ridge lines are solutions of

$$\phi_2^2 = (\lambda_1\phi_1^3 + m_1^2\phi_1)/(\nu\phi_1 + \mu_1)\ , \tag{6.12}$$

rather than of Eq. (6.11). For large values of ϕ_2, there is only one solution, a trough along the line $\phi_1 = -\mu_1/\nu$ rather than along the ϕ_2 axis; near $\phi_2 \approx (m_1^2/\nu)^{1/2}$, this trough bends smoothly toward the minimum of the potential which remains at $\phi_2 = 0$ and $\phi_1 = (-m_1^2/\lambda_1)^{1/2}$. Although the trough never actually crosses the $m_{11}^2 \le 0$ region [which is still given by Eq. (6.7)], the true trajectory $\bar\phi_1(t)$ deviates slightly from $\bar\phi_{1\mathrm{tr}}$, following Eq. (6.10), and slips into the $m_{11}^2 \le 0$ region for a short period of time. The subsequent evolution of the background fields then follow closely the $\mu_1 = 0$ case considered above. The ridge line of the $m_{11}^2 < 0$ region appears near $\phi_2 \approx (m_1^2/\nu)^{1/2}$ where Eq. (6.12) has three solutions for ϕ_1: it is connected to a trough with $\phi_{1\mathrm{tr}} < 0$, but is well disconnected from the global $\phi_{1\mathrm{tr}} > 0$ trough.

In Fig. 11 the solid curve is the fluctuation spectrum for a model with the same potential parameters as the short-dashed curve model discussed above except that $\mu_1 = 1.3\times 10^{-15}\ m_P$, a value chosen to give a suitable mountain. We started the scalar fields exactly in the trough, $\phi_1(t_i) = 2\times 10^{-5} m_P$, $\phi_2(t_i) = 3.8 m_P$. A similar mountain spectrum is therefore obtained by replacing the extremely sensitive tuning of $\phi_2(t_i)$ by a less sensitive tuning of μ_1. Nonetheless, a specific mountain shape still requires delicate adjustment of the potential parameters.

D. Prospects for Gaussian spectra with broken scale invariance

Based on our analytic and numerical computations, we feel that a strong case can be made against *useful* power-law spectra arising naturally in inflation. We have also shown that useful Gaussian plateau and mountain spectra are also improbable if we restrict ourselves to potentials involving quadratic and quartic terms. However, more general potentials could lead to interesting nonflat Gaussian spectra. We justify these assertions in the following.

For the double inflation calculations, the position of the ramp in k space is controlled by the initial value $\phi_2(t_i)$. In the spirit of chaotic inflation as originally proposed by Linde,[1] $\phi_2(t_i)$ should fluctuate in space on comoving scales similar to the comoving scales on which ϕ_1 fluctuates, with amplitudes which should range up to $\frac{1}{4}\lambda_2\phi_2^4 \sim m_P^4$. Given that the initial values of ϕ_1 allow the first stage of inflation to take place in some region, the inhomogeneities in ϕ_2 will inflate away (if ϕ_2 is not rough on arbitrarily small scales compared with the initial Planck scale). The value of ϕ_2 at a given location is frozen until the end of the first stage of inflation. Only if the frozen value of ϕ_2 is very precisely a certain value near $4.3 m_P$ will the ramp on the final perturbation spectrum be at an astrophysically interesting scale. If initial values of ϕ_2 do indeed range over $|\phi_2(t_i)| \lesssim \lambda_2^{-1/4} m_P$ and if, as is necessary for the amplitude of the density perturbations, $\lambda_2 \lesssim 10^{-14}$, then the desired range of $\phi_2(t_i)$ would be a tiny fraction of its possible range. Although the probability distribution of $\phi_2(t_i)$ over this range is unknown, it will undoubtedly be necessary to "fine-tune" the initial conditions—as well as the potential parameters—to place the ramp in the desired location, making this version of double inflation rather unattractive in our view. In any case, making the ramp interestingly large to generate structure will result in a high plateau giving a high-amplitude scale-invariant spectrum for microwave-background fluctuations which would violate the stringent bounds set by the Soviet RELICT experiment as discussed in Sec. II.

In new inflation, the initiation of inflation from "thermal" initial field configurations, an issue raised by Masenko, Unruh, and Wald, now appears feasible through Hubble damping, as Feldman and Brandenberger have demonstrated with numerical simulations.[67] However, although double inflation is also realizable in new inflation, the required tuning of ϕ_2 at the onset of the second stage of inflation to get the ramp in the right place does not follow from thermal initial conditions.

Double inflation also results if one adds a term quadratic in the Ricci scalar to the action (3.1) with a single scalar field.[44] Starobinsky[63] showed that the conformal anomaly of massless scalar fields (nonzero trace of the stress-energy tensor due to quantum effects) is of order R^2 and this term may drive inflation. However the conformal anomaly terms that were most likely to appear in typical theories were shown to be unlikely to lead to inflationary behavior. To avoid excessive density fluctuations, the R^2 terms in the gravitational Lagrangian was instead parametrized by a small mass scale M: $\mathcal{L}_G=(m_P^2/16\pi)[R+R^2/(6M^2)]$. To give P_ζ at the "observed" level, we require $M \approx 10^{-6} m_P$, for $b_\rho=1.44$ (see Sec. VII A 4). The Friedmann equation is now significantly modified over the form with no R^2 term in the Lagrangian. Provided $M>(\lambda/6\pi)^{1/2} m_P$, both the scalar field potential and the R^2 term drive the initial inflationary epoch, followed by an era where the scalar field dominates, leading to double inflation.[44] A ramp plus plateau spectrum remains the generic outcome. Again, the end of the first inflation phase must somehow coincide with the time that cluster scales cross the horizon.

Adding the coupling $\nu \neq 0$ in (6.3) does not aid matters appreciably. Models with both ν and m_1^2 negative can produce mountain fluctuation spectra which are consistent with cluster-cluster correlations and the large-scale streaming velocities and yet remain within the RELICT CMB constraints. However, they too suffer from the problem of unnaturalness since both ν and m_1^2 must be carefully tuned. In addition, the position and breadth of the mountain depend crucially on the initial value of ϕ_2; hence, the arguments against double inflation also apply here.

The more likely situation would be one in which the first scalar has settled into the $\phi_1=0$ trough before ϕ_2 hits the mogul, so the result is independent of $\phi_2(t_i)$. In this case, the quantum perturbations ψ_{11} would exceed the value of the background field $\bar{\phi}_1$, so it is these that drive the rolldown from the mogul, influencing the "background-field" value $\bar{\phi}_2$ through the ϕ_1-ϕ_2 coupling. The net result would be non-Gaussian fluctuations in ϕ_2. If the field ϕ_1 does not enter into a prolonged $m_{11}^2>0$ phase after the $m_{11}^2<0$ phase, significant non-Gaussian isocon fluctuations would survive. Thus the generic case for moguls centered about a $\phi_{1\mathrm{ri}}(\phi_2)$ ridge line which is continuous with the incoming $\phi_{1\mathrm{tr}}(\phi_2)$ trough line, as in the $\nu<0$, $\mu_1=0$ case of Sec. VI C, is a non-Gaussian "mountain," provided

$$\langle (\delta\phi_1)^2(\mathbf{x},t)\rangle^{1/2} \gtrsim |\bar{\phi}_1-\phi_{1\mathrm{ri}}| \;. \tag{6.13}$$

In this expression, the quantum fluctuations $\langle(\delta\phi_1)^2(\mathbf{x},t)\rangle$ are assumed to have the short-distance rapidly oscillating $k^{-1} \lesssim (Ha)^{-1}$ waves filtered out. Note that starting from $\phi_1=0$ with a symmetric mogul leads to bifurcation of the field, which can lead to domain walls. To avoid a domain-wall density problem, we must suppose that either the mogul is localized in ϕ_2 or another interaction is present at lower energies to destroy the walls. In either case, the large-scale non-Gaussian metric perturbations will survive intact.

So far in this subsection, we have considered only two scalars interacting via quadratic and quartic terms. The addition of cubic interaction terms to the potential (6.3) can also give $m_{11}^2<0$ over a short range. If they are symmetric about $\phi_{1\mathrm{tr}}$, the induced ϕ_2 fluctuations would again be non-Gaussian. Bardeen[68] has explored a model of this type, adding a $-\mu_2\phi_2\phi_1^2$ interaction to (6.3). The potential can be rewritten in the form of (6.3), with a $-\frac{1}{2}\nu(\phi_2-a)^2\phi_1^2$ term replacing the $-\frac{1}{2}\nu\phi_2^2\phi_1^2$ term, and with $\lambda_1 \gg -\nu \gg \lambda_2>0$, $0<-m_1^2 \ll -\nu a^2$. Since $m_{11}^2<0$ only temporarily, no isocon perturbations survive. The nonlinear interaction of $\delta\phi_1$ with $\delta\phi_2$ generates adiabatic perturbations whose primordial amplitude ζ is quadratic in the Gaussian field $\delta\phi_1$ as evaluated at the end of the $m_{11}^2<0$ period. The spectrum has a mountain centered on the comoving wave number leaving the horizon at the beginning of the $m_{11}^2<0$ period. The statistics and final amplitude of the density perturbations in this model depend only on the tuning of the potential parameters, not on the initial conditions. The probability distribution for the density perturbation field falls off with height as a simple exponential, rather than as a Gaussian. There is a patchy distribution of lower mass peaks of high amplitude on the scale of the mountain in the $\delta\phi_1$ power spectrum, which is what one might like to explain voids and other large-scale structure. However, rather elaborate tuning of the potential, particularly the parameters ν, a, and m_1^2, is required to get the scale and amplitude of the perturbation mountain.

If instead the added cubic interaction offsets the mogul peak from the trough position, so the background $\bar{\phi}_1$ trajectory determines the $\bar{\phi}_1+\delta\phi$ trajectories, a Gaussian mountain which is independent of the initial conditions for the fields can arise as we showed in Sec. VI C 2. In that model, the non-Gaussian criterion (6.13) is not met, for $|\bar{\phi}_1-\phi_{1ri}|$ exceeds $H/(2\pi)$ by $\sim 10^2$. However, to generate a viable mountain, the potential parameter μ_1 must be tuned precisely to be a tiny value, $\approx 10^{-15} m_P$.

It therefore seems quite unlikely that there would be just one mogul appropriately placed to solve our large-scale structure problems. A more likely variant would have a potential surface populated by many small moguls and valleys, perhaps of varying heights and width. As we emphasized in Sec. VI B and Fig. 4, such potential space structure over the narrow observable range is certainly not inconceivable. However, it would be necessary to control the bumpiness so that inflation still proceeds. On average, the ϕ_2 direction would still have the downward slope that drives slow rolldown. The ϕ_1 direction could still have on average $m_{11}^2>0$, as in the $\nu<0$, $m_1^2<0$ model of Sec. VI C, or be a Goldstone boson, with $m_{11}^2\approx 0$ in the regions without moguls. An example of the latter case would be a generalized axion model, with ϕ_1 and ϕ_2 related to the argument and modulus of a complex scalar, respectively. Isocon fluctuations could survive in such models. In both cases, one could get either non-Gaussian or Gaussian perturbations, depending upon whether or not Eq. (6.13) is satisfied. If Gaussian, the fluctuation spectrum would have a wavy appearance rather then a flat one, with peaks and troughs reflecting the topography in potential space that the $(\bar{\phi}_1,\bar{\phi}_2)$ trajectory passed over.

Power laws other than scale invariant can also be achieved in inflation. For an isocon field, this occurs quite naturally if it is nonminimally coupled. To avoid an excessively small energy density at the present epoch, one requires $\xi\approx\frac{5}{48}$ leading to a spectral index of $n=-2$. This index can be ruled out by RELICT observations. It also differs from the $0\gtrsim n\gtrsim -1$ law advocated by Peebles.[38–40] Power-law inflation produces power-law adiabatic fluctuations which generically give more power on large scales. However, the bound set by the RELICT experiment make, it difficult for these models to fit the cluster-cluster and large-scale streaming observations.

For single-inflaton models with the ultraweak couplings required by observations, although fluctuations can be quite non-Gaussian on scales very much larger than our current horizon size, they are very nearly Gaussian for the patch of the Universe corresponding to our observable Universe, even for such nonstandard cases as power-law inflation[71] (Bardeen and Bublik, Bond and Salopek, cf. Ortolan *et al.*). Some other possibilities for non-Gaussian fluctuations from inflation have been discussed in the literature. Allen, Grinstein, and Wise[69] considered a Universe with massive axions which produced non-Gaussian fluctuations through self-interactions. Kofman, Linde, and Einasto[70] suggested that the transition to a second phase of inflation might proceed by quantum tunneling through a potential barrier, as in Guth's original "old" inflation model, leading to bubbles being generated, superimposed upon a Universe smoothed during the first inflationary epoch. Such fluctuations would certainly be non-Gaussian, but it would be extremely difficult to arrange for their amplitude to be just right to be useful to explain the large-scale texture we observe now. We regard the addition of moguls in potential space, with trajectories obeying Eq. (6.13), as being a more promising realization of non-Gaussian fluctuations from inflation.

Finally, we discuss a general problem associated with all chaotic models. To obtain enough *e*-foldings of inflation with a potential for ϕ_2 of the form Eq. (6.3), it is necessary for the field to start at a value many times the Planck mass. If one were to assume nonminimal coupling in the action (3.1), $f=m_P^2/(16\pi)-\xi_1\phi_1^2/2-\xi_2\phi_2^2/2$, then Newton's constant, $G=(16\pi f)^{-1}$, may actually become negative. Once negative, it would remain so because it must cross a singularity before it changes sign. To avoid this catastrophe it is necessary that $\xi_2\lesssim 0.002$ (Refs. 13 and 14). The field must be effectively minimally coupled, or have $\xi_2<0$, a case we treat in detail in Sec. VII. A similar restriction would hold for ξ_1 if we were contemplating large initial values of ϕ_1 as well. Even if ξ_1 did not have to be zero, unless it is quite negative the potential gives a positive effective mass $m_{11}^2=m_1^2+12\xi_1 H^2+3\lambda_1\phi_1^2-\nu\phi_2^2$: the ϕ_1 fluctuations again inflate away, and only in special circumstances could they be important now (Sec. VI B 4).

VII. NONMINIMALLY COUPLED SCALAR FIELDS

In this section we show how to incorporate chaotic inflation in grand unified theories using nonminimally coupled fields, and apply it to an SU(5) GUT example in Sec. VII B. To illustrate the basic principles, in Sec. VII A we restrict our attention to a single scalar field, ϕ, with quadratic coupling to the Ricci scalar, $f=m^2/(16\pi)-\xi\phi^2/2$, where the free parameter m is the bare value of the Planck mass. Since a $-\xi R\phi^2$ coupling is required for the renormalizability of a $\lambda\phi^4/4$ potential in curved space-time,[49] it is natural to consider the cosmological significance of this coupling for arbitrary values of ξ.

We first provide an overview of the main results derived in Sec. VII A. We have already noted in Sec. VI D that if $m=m_P$ Newton's constant may become negative in chaotic inflation unless $\xi<0.002$. The constraint is an embarrassment for chaotic inflation if one insists that ξ be positive: one is simply tuning the curvature coupling constant in an *ad hoc* and unnatural manner. We shall therefore analyze the other possibility, $\xi<0$. One may object on the grounds that the field is now a tachyon, $m_\phi^2=\xi R<0$, but we will always choose potentials of form $V(\phi)=\lambda(\phi^2-\sigma^2)^2/4$ to stabilize the theory. The effect of a negative coupling constant is to flatten the potential for $\phi>m_P/(8\pi|\xi|)^{1/2}$. If, for example, $\lambda=0.05$, and ξ is chosen to be large and negative, $\xi=-2\times 10^4$, metric fluctuations of the correct level are produced. Gauge-boson radiative corrections, which proved disastrous for new inflation, do not destroy the desired flatness of the potential in this case. If $m\approx m_P$, the reheat temperature is quite high, $\approx 1.0\times 10^{15}$ GeV. To avoid monopole overproduction and yet have gauge bosons produce the baryon asymmetry of the Universe, we find the GUT symmetry-breaking scale must be found within the narrow limits, 1.5×10^{15} GeV $<\sigma<3.0\times10^{15}$ GeV, which is quite close to the value found from renormalization-group calculations for SU(5), $\sigma=1.2\times10^{15}$ GeV. However, if baryogenesis proceeds through the decay of the Higgs particle, then only the upper limit is necessary and the constraint is not as stringent.

A. Induced gravity and related models

1. ζ fluctuations from nonminimally coupled scalar fields

Here we demonstrate how to relate $\tilde{\mathcal{P}}_\zeta$ determined for the conformally transformed model of Sec. III B with $f(\phi_k)=m^2/(16\pi)-\xi\phi^2/2$ to that determined in the untransformed system $\mathcal{P}_\zeta$. For single scalar field models, it is possible to define a new field χ which has a standard kinetic energy term and for which $\tilde{\mathcal{P}}_\zeta$ can be calculated using techniques for minimally coupled fields:

$$\chi=\int[K^{11}(\phi)]^{1/2}d\phi\ , \tag{7.1a}$$

$$K^{11}=\frac{\dfrac{m^2}{m_P^2}+8\pi|\xi|(1+6|\xi|)\dfrac{\phi^2}{m_P^2}}{\left[\dfrac{m^2}{m_P^2}+8\pi|\xi|\dfrac{\phi^2}{m_P^2}\right]^2}\ . \tag{7.1b}$$

The kinetic term K^{11} therefore transforms to unity in the χ variable. Integrating (7.1a) yields

$$e^{\alpha\chi}=(8\pi|\xi|)^{1/2}\left[\frac{1+s}{1-s}\right]^{s/2}\frac{\phi+\sqrt{\phi^2+(\beta m)^2}}{2m_P}\times\left[\frac{\sqrt{\phi^2+(\beta m)^2}-s\phi}{\sqrt{\phi^2+(\beta m)^2}+s\phi}\right]^{s/2}, \tag{7.2a}$$

where α, β, and s are constants:

$$\alpha=\left[\frac{8\pi|\xi|}{1+6|\xi|}\right]^{1/2}m_P^{-1}\ ,$$
$$\beta=[8\pi|\xi|(1+6|\xi|)]^{-1/2}\ , \tag{7.2b}$$
$$s=\left[\frac{6|\xi|}{1+6|\xi|}\right]^{1/2}.$$

The effective potential is a function of χ only, and is given in (3.7d):

$$U(\chi)=\left[\frac{m^2}{m_P^2}+8\pi|\xi|\frac{\phi^2(\chi)}{m_P^2}\right]^{-2}V(\phi(\chi))\ . \tag{7.2c}$$

Similar formulas were also obtained by Futamase and Maeda,[14] who used them to discuss the background-field evolution. Here we are more interested in the fluctuations that are generated. The important point here is that although the effective potential $U(\chi)$ is identical to $V(\phi)$ for $\phi\ll m/(8\pi|\xi|)^{1/2}$, it approaches a constant, $U_\infty=\lambda[m_P^2/(16\pi|\xi|)]^2$, in the opposite limit, $\phi\gg m/(8\pi|\xi|)^{1/2}$.

For the transformed metric $\tilde{g}_{\mu\nu}$, the standard formula for metric fluctuations, $\tilde{\mathcal{P}}_\zeta^{1/2}=(3/2\pi)(H^2/\dot{\chi})$, Eq. (6.1), is still valid because only Einsteinian gravity appears in the transformed Lagrangian, (3.6). Since the original metric $g_{\mu\nu}$ is related to $\tilde{g}_{\mu\nu}$ through the conformal transformation (3.7a), their perturbations satisfy

$$\delta g_{\mu\nu}=\left[\frac{m_P^2}{16\pi f}\right]\left[-\frac{df}{f}\tilde{g}_{\mu\nu}+\delta\tilde{g}_{\mu\nu}\right].$$

The ratio of the first term to the second is of order $\tilde{\mathcal{P}}_\zeta^{-1/2}(df/f)$. Since df quickly decreases to zero once the scalar field settles down to the minimum of its potential, the second term dominates after the inflationary epoch. Thus after inflation has stopped the metric fluctuations are identical in both frames, and $\mathcal{P}_\zeta=\tilde{\mathcal{P}}_\zeta$.

2. Fluctuations in the induced gravity model

In induced gravity, the bare Planck mass is zero, its present value being generated when the Higgs field becomes trapped at the minimum of its potential which occurs at a value determined from Eq. (3.4), $\sigma=m_P/(8\pi|\xi|)^{1/2}$. In the $m=0$ limit, Eqs. (7.2) become

$$U(\chi)=\frac{\lambda}{4}\left[\frac{m_P^2}{8\pi|\xi|}\right]^2(1-e^{-2\alpha\chi})^2,\quad e^{\alpha\chi}=\phi/\sigma\ . \tag{7.3}$$

The potential U, shown in Fig. 12(a), grows exponentially for $\chi\ll-\alpha^{-1}$, attains its minimum at $\chi=0$, and then approaches its limiting value, U_∞ for $\chi\gg\alpha^{-1}$. In this case, although χ varies from $-\infty$ to $+\infty$, it only parametrizes half of the original potential and neglects $\phi<0$. Another field, for example, $\alpha^{-1}\ln(-\phi/\sigma)$, is required to describe the other half.

There are two inflationary models to consider here, depending on which side of the origin the scalar field originates.

Case 1. $\chi(t_i)\gg\alpha^{-1}$: During the slow-rollover epoch, all scales k leave the horizon with approximately the same value of H,

$$H_i\approx\left[\frac{\lambda}{96\pi\xi^2}\right]^{1/2}m_P\ ,$$

due to the flatness of U. The amplitude of the metric fluctuations (6.1),

$$\mathcal{P}_\zeta^{1/2}(k)=\frac{9}{2\pi}\frac{H_i^3}{\dfrac{\partial U}{\partial\chi}[\chi(t_k)]}\ , \tag{7.4a}$$

can be related to the number of e-foldings from the time t_k when the scale k crosses the horizon to the end of inflation at t_e,

$$N_i(t_k)\approx H_i(t_e-t_i)\approx\frac{3}{2\alpha}\frac{H_i^2}{\dfrac{\partial U}{\partial\chi}[\chi(t_k)]}\ , \tag{7.4b}$$

found by integrating the slow-rollover equation. The final elegant expression

$$\mathcal{P}_\zeta^{1/2}=\frac{1}{\sqrt{8\pi^2}}\frac{\sqrt{\lambda}}{|\xi|}N_I(t_k),\quad \xi\ll-\tfrac{1}{6}\ , \tag{7.5}$$

only involves the dimensionless variables of the theory. Of course, $N_I(t_k)$ is a reparametrization of $\chi(t_k)$.

For definiteness, we now choose a reasonable value of λ consistent with the radiative corrections of GUT models, $\lambda=5\times10^{-2}$. The scale where the fluctuation spectra are typically normalized, as discussed in Sec. II, is $k^{-1}\sim10$ Mpc. The number of inflationary e-foldings after horizon crossing would be $N_I(t_k)\sim50$. To get the "observed" fluctuation level $\mathcal{P}_\zeta^{1/2}\sim7\times10^{-5}$ would then require $\xi\approx-2\times10^4$. The number of e-foldings associated with our current horizon is only slightly larger $N_I(t_k)\sim60$: the spectrum is very nearly scale invariant. An extremely small λ is therefore not required to get the observed fluctuation level, but the price is the introduction of a large negative curvature coupling constant.

Some characteristic quantities for this successful fluctuation generation model are

$$\begin{aligned}
&\sigma=\sigma_I\equiv\frac{m_P}{\sqrt{8\pi|\xi|}}=1.7\times10^{16}\text{ GeV},\\
&m_A=g_G\sigma_I\approx8.6\times10^{15}\text{ GeV},\\
&\phi(t_k)=[4N_I(t_k)/3]^{1/2}\sigma_I=1.5\times10^{17}\text{ GeV},\\
&\chi(t_k)=\alpha^{-1}\ln[\phi(t_k)/\sigma_I]=1.3\times10^{19}\text{ GeV},\\
&H_i=\left[\frac{\lambda}{96\pi\xi^2}\right]^{1/2}m_P=7.9\times10^{12}\text{ GeV},
\end{aligned}\tag{7.6}$$

where m_A is the superheavy boson mass and $\phi(t_k)[\chi(t_k)]$ is the value of the scalar field required to give a sufficient number of e-foldings. Note that the induced gravity energy scale, σ_I, is close to that of grand unified theories, and hence it is plausible that the scalar field could actually be the Higgs particle for a GUT as well as for gravity.

Case 2: $\chi_0\ll-\alpha^{-1}$. In this limit, the potential is an exponential,

$$U(\chi)=\frac{\lambda}{4}\left[\frac{m_P^2}{8\pi\xi}\right]^2e^{-4\alpha\chi}\tag{7.7}$$

and the background field has the exact solution[65]

$$\chi(t)=(2\alpha)^{-1}\ln(bt)$$
$$\text{with } b^2=\frac{\lambda\alpha^4m_P^6}{96\pi^3\xi^2}\left[1-\frac{\alpha^2m_P^2}{3\pi}\right]^{-1},\tag{7.8}$$

describing power-law inflation, $a(t)\propto t^q$, with index $q=(1+6|\xi|)/(8|\xi|)$ which may vary from $\frac{3}{4}$ to ∞. Halliwell[72] has shown this solution to be a stable attractor for $q>1$ ($\xi>-\frac{1}{2}$). The potential is too steep for larger $|\xi|$ for inflation to be possible. Therefore, the only one we need consider is the $\xi>-\frac{1}{2}$ case.

Fluctuations follow from (6.1) which is valid even if the slow-rollover approximation does not apply:

$$\mathcal{P}_\zeta^{1/2}=\frac{3q}{\pi}H(t_k)\alpha\ .\tag{7.9}$$

The resulting metric fluctuation spectrum

$$\begin{aligned}
&\mathcal{P}_\zeta^{1/2}=\mathcal{P}_{\zeta_1}^{1/2}\left[\frac{k}{k_1}\right]^{1/(1-q)},\\
&\mathcal{P}_{\zeta_1}^{1/2}=3[(1+6|\xi|)/(8\pi|\xi|)]^{1/2}(H_1/m_P)
\end{aligned}\tag{7.10}$$

generalizes the flat Zeldovich spectrum to a power law with negative exponent (see also Sec. VI B). Here k_1 is some characteristic scale, e.g., the present horizon scale. $\mathcal{P}_{\zeta_1}^{1/2}$ is of the right order if $H_1/m_P\approx10^{-5}$. However, we must choose the parameters of the model to be very small, typically $\lambda=10^{-23}$ and $\xi=-0.029$, to ensure that the Universe inflates by at least a factor of e^{60}:

$$\begin{aligned}
e^{60}\ll\frac{a(t_e)}{a(t_1)}&=\left[\frac{t_e}{t_1}\right]^q\\
&=\left[\frac{8\pi^2}{9\lambda}\frac{3q-1}{q^2(4q-3)^2}\mathcal{P}_{\zeta_1}\right]^{q/2}.
\end{aligned}$$

Accetta, Zoller, and Turner[22] first analyzed this possibility, using a different technique, and also concluded that λ had to be extremely tiny. In this expression, the initial time, t_1, is determined by the level of fluctuations, $\mathcal{P}_{\zeta_i}^{1/2}\approx7\times10^{-5}$, and the final time, t_e, occurs when the background field reaches the bottom of its potential well, $\chi(t_e)=0$. In addition, this model is unnatural because the exponential behavior for large negative χ, (7.7), depends crucially on the vanishing of the bare Planck mass. Unless this is guaranteed by some symmetry, we have yet another example of a fine-tuning. For these reasons, we favor case (1) for which the potential is extremely flat provided only that $\phi(t_k)$ exceeds both $m/(8\pi|\xi|)^{1/2}$, and $m_P/(8\pi|\xi|)^{1/2}$.

Unfortunately, induced gravity suffers from a perhaps fatal flow: although the gauge bosons acquire a mass, given by $m_A^2=g_G^2m_P^2/(8\pi|\xi|)$, where g_G is some coupling constant, they are *totally* decoupled dynamically from the scalar field. The same holds true for fermions. Therefore there is no reheating.

To see this, let us assume for simplicity that ϕ is a complex scalar coupled to some U(1) gauge field A_μ which can be accommodated by adding terms quadratic in the field strength, $F_{\mu\nu}$, to the Lagrangian (3.1), and by replacing the kinetic term $\partial_\mu\phi\partial_\nu\phi$ by $(\partial_\mu\phi^*+ig_GA_\mu\phi^*)(\partial_\nu\phi-ig_GA_\nu\phi)$. (The non-Abelian case is treated in Appendix B.) Performing the now familiar conformal transformation, and then redefining fields, $\phi=Be^{iS}$, $A_\mu=\tilde{A}_\mu+S_{,\mu}/g_G$, where B and S are real fields, we find the action to be

$$S=\int d^4x\sqrt{-\tilde{g}}\left[\frac{m_P^2}{16\pi}\tilde{R}-\frac{1}{2}\frac{m_P^2}{16\pi}f^{-2}(3f_{,\mu}f^{,\mu}+fB_{,\mu}B^{,\mu})+\left[\frac{m_P^2}{16\pi f}\right]^2V(B)-\frac{1}{2}\frac{m_P^2}{16\pi}f^{-1}B^2g_G^2\tilde{A}_\mu\tilde{A}^\mu-\tfrac{1}{4}\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu}\right],\tag{7.11}$$

where $\tilde{F}_{\mu\nu}=\partial_\mu \tilde{A}_\nu-\partial_\nu \tilde{A}_\mu$. If $f=-\xi\phi^*\phi/2=-\xi B^2/2$ then the gauge field becomes a superheavy boson, $m_A^2=g_G^2 m_P^2/(8\pi|\xi|)$, and, remarkably, B totally decouples from $\tilde{A}_\mu$.

Models that couple the scalar field to fermions (which would become superheavy upon symmetry breaking) do not solve the reheating problem either. Fermions are also conformally coupled to gravity, and they behave much like gauge bosons. To illustrate this point, we consider a simple model containing a left-handed fermion field, ψ_L, carrying a U(1) charge $-g_G$, and a neutral right-handed field, ψ_R. The Lagrangian density is

$$\mathcal{L}=i\bar{\psi}_L\gamma^\alpha(\mathcal{D}_\alpha+ig_G V_\alpha{}^\mu A_\mu)\psi_L+i\bar{\psi}_R\gamma^\alpha\mathcal{D}_\alpha\psi_R -h\phi\bar{\psi}_R\psi_L-h\phi^*\bar{\psi}_L\psi_R\ . \tag{7.12}$$

The covariant derivative $\mathcal{D}_\alpha$ is defined in terms of the vierbein[73] $V_\alpha{}^\mu$:

$$\mathcal{D}_\alpha=V_\alpha{}^\mu(\partial_\mu+\tfrac{1}{2}\sigma_{\beta\gamma}V^{\beta\nu}V_{\nu;\mu}{}^\gamma)\ , \quad \sigma_{\beta\gamma}=-\tfrac{1}{4}[\gamma_\beta,\gamma_\gamma]\ , \tag{7.13}$$

and h is some dimensionless coupling constant. If we write $g_{\mu\nu}=\Omega^2(\mathbf{x},t)\tilde{g}_{\mu\nu}$, then the vierbein transforms as $V^\alpha{}_\mu=\Omega\tilde{V}^\alpha{}_\mu$, and the Lagrangian density becomes

$$\tilde{\mathcal{L}}=i\bar{\tilde{\psi}}_L\gamma^\alpha(\tilde{\mathcal{D}}_\alpha+ig_G\tilde{V}_\alpha{}^\mu A_\mu)\tilde{\psi}_L+i\bar{\tilde{\psi}}_R\gamma^\alpha\tilde{\mathcal{D}}_\alpha\tilde{\psi}_R -h\Omega\phi\bar{\tilde{\psi}}_R\tilde{\psi}_L-h\Omega\phi^*\bar{\tilde{\psi}}_L\tilde{\psi}_R\ . \tag{7.14}$$

The covariant derivative $\tilde{\mathcal{D}}_\alpha$ is just (7.13) with $V_\alpha{}^\mu$ replaced by $\tilde{V}_\alpha{}^\mu$, and the transformed fermion field is $\tilde{\psi}_{L,R}=\Omega^{3/2}\psi_{L,R}$, which produces a Dirac mass $h\Omega\phi$. In our case $\Omega=m_P/(8\pi|\xi|\phi^*\phi)^{1/2}$; hence, the transformed fermion field also totally decouples from the scalar field.

This situation is disastrous because the Universe would not reheat. ϕ would act like cold dark matter as it oscillates endlessly about the minimum of its potential. Radiation and baryonic matter would be produced only in trace amounts through gravitational particle creation. This problem may perhaps be resolved by calculating the quantum corrections to the action (3.1). Except for the scalar self-interaction, there is no scale in the theory. Quantum corrections will probably break this symmetry,[48] possibly in the same way asymptotic freedom is broken in low-energy quantum chromodynamics. More work is required to test whether this type of broken scale invariance can make induced gravity viable.

3. Fluctuations in the variable Planck mass model

To ensure that reheating will occur, we drop the condition that m_P be totally induced by the scalar field, and present a less ambitious proposal, which we call the variable Planck mass model, with

$$\sigma \ll \frac{m_P}{\sqrt{8\pi|\xi|}}\ . \tag{7.15}$$

The bare mass must then be $m\approx m_P$. In this case, ϕ acts only as the Higgs to the GUT bosons fields; as its minimum it provides only a nominal contribution to m_P.

The effective potential $U(\chi)$ is plotted in Fig. 12(b). Fortunately, the extreme flatness of the potential at large values of ϕ, $\phi \gg m_P/(8\pi|\xi|)^{1/2}$, is not destroyed by radiative corrections because gauge bosons, as in the case for

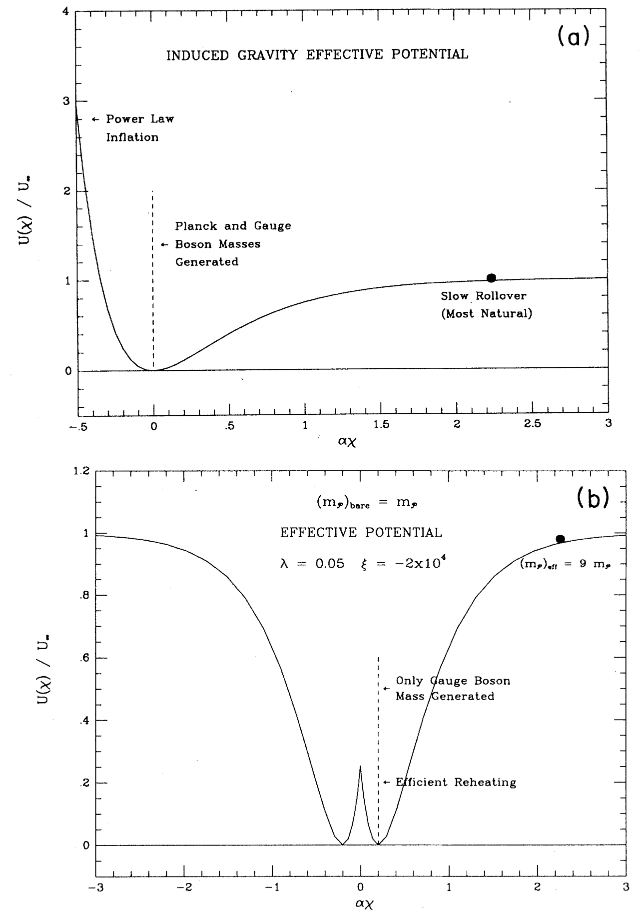

FIG. 12. The effective potentials [Eq. (7.2c)] for the induced gravity (a) and variable Planck mass (b) models are plotted as functions of the transformed GUT Higgs field χ. In (a), $m_{P_{\rm eff}}=m_P\exp(\alpha\chi)$ is generated entirely by the χ field, reaching its standard asymptotic value when $\chi=0$ at the bottom of its potential. Power-law inflation results if $\alpha\chi(t_i)\ll-1$, but to ensure sufficient e-foldings λ must be exceedingly small. In a chaotic setting, $\alpha\chi(t_i)\gg 1$ is more natural. The black dot marks the position where the scale 5000 h^{-1}Mpc left the horizon. If $\lambda=0.05$ and $\xi=-2.0\times10^4$, $\mathcal{P}_\zeta^{1/2}$ of the "observed" amplitude, 7×10^{-5}, is obtained, thus solving the small-λ problem in exchange for large $-\xi$. However, this model has severe reheating problems. The model of Fig. 11(b) has a bare value for m_P equal to its present value. At the 5000-h^{-1}Mpc horizon crossing point (black dot again), $m_{P_{\rm eff}}$ is an order of magnitude larger than m_P. For $\alpha\chi\gg1$, the model retains the asymptotic behavior of the induced gravity model, with successful $\mathcal{P}_\zeta$ generation using the same λ and ξ. Coupling to gauge bosons is small on the plateau, and radiative corrections do not destroy the flatness of the potential. At the minimum, coupling is strong and reheating is efficient. The GUT symmetry-breaking parameter σ was chosen to be $\sigma=10^{-3}m_P$ for plotting purposes, although the $\sigma\approx10^{-4}m_P$ is theoretically preferred.

induced gravity, essentially decouple from ϕ. To show this, consider the one-loop radiative correction to the effective potential given, for example, in Cheng and Li:[74]

$$V_{\rm rad}=\frac{3m_{\tilde A}^4}{64\pi^2}\ln\left[\frac{m_{\tilde A}^2}{t^2}\right], \qquad (7.16)$$

where, t is some fixed normalization scale, and $m_{\tilde A}^2$ is the effective value of the boson mass, which may be read from (7.11), $m_{\tilde A}^2=(m_P^2/16\pi)g_G^2B^2/f$. This correction term may be combined with the tree-level potential. Its effect is to produce a running coupling constant λ_R:

$$\lambda_R=\lambda+\frac{3g_G^4}{16\pi^2}\ln\left[\frac{m_{\tilde A}^2}{t^2}\right]. \qquad (7.17)$$

If $\lambda=5\times10^{-2}$ and $g_G=0.5$, then the radiative correction is of the same order as the tree value of λ, but does not overwhelm it, which was a fatal flaw for earlier models with $\lambda=5\times10^{-14}$ (Sec. VI A). The modification due to fermions is at a similar level. Although these arguments are crude because the potential $U(\chi)$ is not renormalizable, they are nonetheless extremely suggestive.

Using Eq. (7.15), one can easily show that the effective potential in this model has the same large ϕ asymptotic behavior as the induced gravity model. Therefore the $\mathcal{P}_\zeta$ formula (7.5) remains valid. As in Sec. VII A 2, there is no small λ problem in this model. Moreover, when the scalar field rolls down to the bottom of its potential, it becomes strongly coupled to gauge bosons via the term $g_G^2B^2\tilde A_\mu\tilde A^\mu$ in (7.11), and reheating is expected to be efficient. We will therefore write the reheat temperature $T_{\rm reh}$ as the product of an efficiency factor ϵ and the maximum allowed reheat value:

$$\begin{aligned}T_{\rm reh}&=\epsilon\left[\frac{30}{\pi^2}g_{\rm eff}^{-1}U_\infty\right]^{1/4}\\&=\epsilon\left[\frac{15}{16\pi^2}g_{\rm eff}^{-1}N_I^{-2}\mathcal{P}_\zeta\right]^{1/4}m_P\\&=2.0\times10^{15}\epsilon\ {\rm GeV}\ , \end{aligned}\qquad (7.18)$$

where we have taken $\mathcal{P}_\zeta^{1/2}=7\times10^{-5}$, $N_I=60$, and the number of degrees of freedom, $g_{\rm eff}=160.75$. The efficiency factor must be determined numerically.

It is essential that $T_{\rm reh}$ should not exceed the symmetry-restoration temperature T_S otherwise in more realistic non-Abelian GUT models, monopoles will be produced in copious numbers as the Universe cools. In the simple U(1) model used in this section, the thermal correction to the potential is[75]

$$\begin{aligned}\Delta V_T&=\sum_i \tfrac{1}{24}m_i^2T^2\\&=\frac{T^2}{24}[(4\lambda|\phi|^2-2\lambda\sigma^2)+3g_G^2|\phi|^2]\ ;\end{aligned}\qquad (7.19)$$

the first term arises from the scalar field, $\phi=\phi_1+i\phi_2$, $\sum m_i^2=\partial^2V/\partial\phi_1^2+\partial^2V/\partial\phi_2^2$, whereas the second comes from the massive gauge boson A_μ which has three spin degrees of freedom. The U(1) symmetry is broken if $\partial^2(V+\Delta V_T)/\partial|\phi|^2<0$ at the origin, which is the case if the temperature is below

$$T_S=\frac{2\sigma}{(4/3+g_G^2/\lambda)^{1/2}}\ . \qquad (7.20)$$

[Actually, the above derivation of the symmetry-breaking temperature should replace ϕ and $V(\phi)$ by χ and $U(\chi)$, but at $\phi=0$, the results are very nearly the same if $\sigma\ll m$.]

One might also like to arrange for $T_{\rm reh}$ to exceed the low-energy superheavy boson mass, $m_A=g_G\sigma$, for then baryogenesis proceeds according to the standard picture: superheavy gauge bosons leave thermal equilibrium as the Universe cools and decay through baryon-nonconserving CP-violating channels to produce slightly more matter than antimatter. Since there are other ways of doing baryogenesis, this lower bound is not an essential requirement of the theory.

In order to solve the monopole problem and still retain the standard model of baryogenesis, we would then require

$$g_G<T_{\rm reh}/\sigma<\frac{2}{(4/3+g_G^2/\lambda)^{1/2}}\ , \qquad (7.21)$$

which always has a solution provided

$$\lambda>\frac{g_G^4}{4-4g_G^2/3}=1.7\times10^{-2}\quad\text{if } g_G=0.5\ . \qquad (7.22)$$

The value $\lambda=0.05$ adopted in Sec. VII A 2 satisfies this inequality. If we assume that perturbation theory is valid, $\lambda<1$, then σ is constrained to lie within the narrow limits

$$1.3\times10^{15}\epsilon\ {\rm GeV}<\sigma<4.0\times10^{15}\epsilon\ {\rm GeV}\ . \qquad (7.23)$$

The limits from a more realistic model are given in Sec. VII B.

However, if the Higgs particles rather than the gauge bosons generate the baryon asymmetry of the Universe,[62] we would only require that $m_\chi<T_{\rm reh}$. At the minimum of the potential, this always holds, since the mass of the χ field is

$$\begin{aligned}m_\chi&=\frac{\sqrt{2\lambda}\sigma}{(1+48\pi\xi^2\sigma^2/m_P^2)^{1/2}}\\&=1.5\times10^{13}\ {\rm GeV}\ .\end{aligned}$$

We have assumed that $\sigma\gg m_P/(48\pi\xi^2)^{1/2}$ and that the ratio λ/ξ^2 is given by the fluctuation formula, (7.5). Therefore only the lower bound in Eq. (7.23) must be satisfied, to avoid excessive monopole production.

In Fig. 13(a) we show the last five e-foldings of our numerical calculations of the variable Planck mass model, with $\lambda=0.05$, $\xi=-2.0\times10^4$, and $\sigma=10^{-4}m_P$. The radiation energy density reaches its maximum value shortly after the damping factor Γ and the mass of the χ field, $m(\chi)\equiv(|\partial^2U/\partial\chi^2|)^{1/2}$, exceed $3H$, for then the rate of rolldown and the particle creation rate are larger than the expansion rate. The value of the efficiency factor ϵ is determined primarily by the value of ϕ when

$2m(\chi)=3H$, provided Γ exceeds $3H$ and $2m$ at that time. For the case shown in Fig. 13(a), $\epsilon=0.48$. This depends somewhat on the choice of Γ. At the moment this is the greatest source of uncertainty in our models. We believe that it is natural for the coupling to bosons and fermions to be sufficiently strong as the scalar field drops into the well around $\phi\sim\sigma$ for the field to remain in the well, without crossing over the hump to negative ϕ. Otherwise, undesirable topological defects might arise.

For the calculations shown we have assumed

$$\Gamma(\chi)=f\sqrt{2\lambda}\sigma[m^2/m_P^2+8\pi|\xi|(1+6|\xi|)\phi^2(\chi)/m_P^2]^{-1/2} . \tag{7.24}$$

Our motivation for this choice is that when $\phi\approx\sigma$, Γ is just proportional to the mass of the χ field, while for $\phi > m_P/(8\pi|\xi|)^{1/2}$, it is negligible compared to the Hubble parameter, so that slow-rollover rate is determined by the $3H\dot{\chi}$ term of (4.6c). We used $f=5.0$ in Fig. 13(a); increasing or decreasing f by a factor of 4 does not alter ϵ or $\mathcal{P}_\zeta$ significantly. However, decreasing f to below about three leads to negative ϕ. Our handling of Γ is admittedly crude, but a more precise description from first principles is a difficult task. The perturbation calculations are started when $k/Ha=5$ instead of the usual value of 50 to minimize the effect of particle damping on the quantum fluctuations within the horizon where the damping expression is not valid.

Figure 13(b) shows the resulting fluctuation spectrum, $\mathcal{P}_\zeta(k)$. Apart from the way the variable Planck mass model alleviates the reheating problem found for the induced gravity model, the models are quite similar. The fiducial quantities displayed in Eq. (7.6) remain valid, except that σ is now an order of magnitude lower than σ_I and that m_A of (7.6) is now interpreted as the mass of the gauge boson in the $\phi\to\infty$ limit. In particular the estimates of $\mathcal{P}_\zeta$ made for case (1) still hold, and agree with the numerical results to within 10%.

4. *Relation to fluctuations in R^2 inflation*

The R^2 inflation introduced by Starobinskii[63] (see Sec. VI D) is actually mathematically equivalent to the induced gravity model. Using a conformal transformation, Whitt[76] has shown that Lagrangians of form $\mathcal{L}_G=(m_P^2/16\pi)[R+R^2/(6M^2)]$ can be written as regular Einstein gravity interacting with an additional scalar field (identified with R). If one defines a new field

$$\chi=\left[\frac{3m_P^2}{16\pi}\right]^{1/2}\ln\left[1+\frac{R}{3M^2}\right] , \tag{7.25}$$

the transformed Lagrangian takes the form

$$\tilde{\mathcal{L}}=\frac{m_P^2}{16\pi}\tilde{R}-\tfrac{1}{2}\tilde{g}^{\mu\nu}\partial_\mu\chi\partial_\nu\chi-U(\chi) , \tag{7.26}$$

where the potential is

$$U(\chi)=\frac{3m_P^2M^2}{32\pi}\left\{1-\exp\left[-\left[\frac{16\pi}{3m_P^2}\right]^{1/2}\chi\right]\right\}^2 . \tag{7.27}$$

Note that this is formally similar to Eq. (7.3), and in the large $|\xi|$ limit it is identical if we identify M/m_P with $[\lambda/(24\pi)]^{1/2}|\xi|^{-1}$. Applying the fluctuation formula (7.5), we therefore obtain

$$\mathcal{P}_\zeta^{1/2}=\left[\frac{3}{\pi}\right]^{1/2}\frac{M}{m_P}N_I(t_k) . \tag{7.28}$$

This gives the level of fluctuations corresponding to $b_\rho=1.44$ if $M\approx1.2\times10^{-6}m_P$.

With the large $|\xi|$ coupling, we expect that R^2 and other terms will be generated through quantum corrections. Since gravity is not renormalizable, this problem cannot

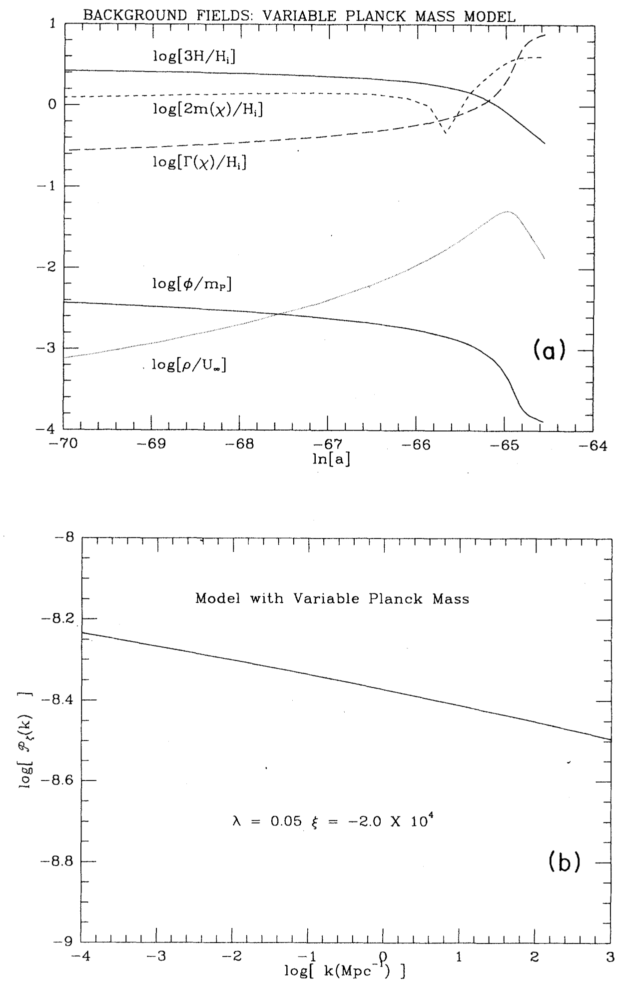

FIG. 13. The late-time behavior of the scalar field ϕ in the variable Planck mass model with the potential of Fig. 12(b) is shown in (a). When the rate of roll down, $2m(\chi)$, and the dissipation rate Γ grow larger than the Hubble expansion rate, $3H$, the radiation density, ρ grows to its maximum, yielding a reheating temperature efficiency $\epsilon=0.48$. The damping is chosen to be critical near the end of inflation to avoid ϕ reaching negative values. U_∞ and H_i are the asymptotic values of the potential and the Hubble parameter. Our calculated $\mathcal{P}_\zeta$ spectrum is shown in (b) for the model with $\lambda=0.05$ and $\xi=-2.0\times10^4$. Variations of λ and ξ yield similar scale-invariant spectra [$\mathcal{P}_\zeta\propto\lambda/\xi^2$, Eq. (7.5)].

really be addressed. Nonetheless, we can estimate the R^2 term to have a coefficient[49] $(\xi-1/6)^2/(384\pi^2)$. For the ξ value required to give the observed $\mathcal{P}_\zeta$, even larger perturbations might arise from the R^2 term, possibly giving double inflation naturally.

B. Chaotic inflation incorporated in a GUT

The methods of Sec. VII A may be generalized to any GUT model. For instance, in Appendix B, we consider in detail the 24-dimensional adjoint Higgs field of SU(5) and here we will only present the results. Because the non-Abelian character of SU(5) is irrelevant when one calculates the one-loop effective potential,[75] it is easy to show that radiative corrections do not destroy the flatness of the Higgs potential. If monopoles are not to be produced but baryogenesis is to occur through gauge-boson decays, then the symmetry-breaking parameter σ must now fall within the revised limits

$$3.0\times10^{15}\epsilon \text{ GeV} < \sigma < 5.9\times10^{15}\epsilon \text{ GeV} .$$

However, if baryogenesis occurs via the Higgs particle, then only the lower limit is necessary. These inequalities could prove useful for GUT builders. The range is tantalizingly close to the quoted value of 1.2×10^{15} GeV if we use the efficiency factor $\epsilon=0.48$ obtained in Sec. VII A. We have not tried to refine the model further to obtain exact agreement because this would require, among other things, a more careful treatment of the damping factor, Γ (Sec. VII A). Moreover, the SU(5) model is only illustrative of what would occur in more realistic GUT's, since it has been ruled out by the absence of proton decay.

It is natural to consider the cosmological consequences of the Higgs boson H having so many components. 20 degrees of freedom may be gauged away, leaving four behind, one radial, $\phi=(\mathrm{Tr}H^2)^{1/2}$, and the other three angular components. ϕ corresponds to the field of Sec. VII A. The parameters may be chosen so that $SU(3)\times SU(2)\times U(1)$ is the absolute minimum of the angular potential (fixed ϕ), in which case $SU(4)\times U(1)$ is the absolute maximum. Although the potential structure is rich, containing, for example, many saddle points, it is improbable that spectra with large-scale power will be produced. The scalar field rolls down to the $SU(3)\times SU(2)\times U(1)$ trough along angular directions in potential space in a very short time (less than 1% of the Hubble time). Only then does it begin to inflate, slowly rolling down the trough in the radial direction. Since there is no structure in the trough, neither double-inflation plateaus nor negative mass-squared mountains will occur in this model. The general prediction of the model is therefore a scale-invariant spectrum. This conclusion is modified if additional multiplets besides the **24** Higgs representation are added. Double inflation results if the λ parameter of the new field is much larger than that of the adjoint Higgs representation, so it dominates the energy density at the start of the chaotic era. It then rolls down to its minimum and the adjoint Higgs representation leads to a second inflation phase. Other couplings might lead to mountain spectra, if moguls can be added in the trough.

Our mechanism for incorporating inflation in a GUT assumes the potential $V(\phi)$ of Eq. (3.1) is proportional to ϕ^4 for large values of ϕ, so that the effective potential $U(\phi)\propto V(\phi)/\phi^4$ is asymptotically flat. If $V(\phi)$ is less steep, say of form $m^2\phi^2/2$, then $U(\phi)\propto\phi^{-2}$; hence the scalar field would grow to very large values and, since the couplings to fermions and other bosons becomes weaker [Eq. (7.11) and sequel], reheating is very unlikely. We would also lose one of our original motivations for considering $\xi\neq0$, namely, to cancel infinities arising from $\lambda\phi^4$ in curved space. Potentials steeper than ϕ^4 would lead to a picture similar to the original $\xi=0$ chaotic inflation theory. For all non-ϕ^4 asymptotic cases, a fundamental mass scale would appear in the theory, which is unattractive. We therefore view the case we have treated as the most natural.

It may be argued that a large number of terms will arise from quantum corrections from the ϕ graviton and ϕ-ϕ interaction, which could steepen the asymptotic ϕ dependence above ϕ^4. However, we could calculate quantum corrections using the conformally transformed action (3.6) instead of the action (3.1) with the $-\xi R\phi^2/2$ explicitly appearing. The resulting residual graviton coupling to the scalar field is small. Further, the self-interactions are also small in the conformally transformed system, leading to negligible quantum corrections to $U(\chi)$, at least at the one-loop level, $\Delta U_{\rm cor}\approx m^4\ln(m^2)/(64\pi^2)$ where $m^2=\partial^2U/\partial^2\chi$. We do not know if higher loops will steepen the potential, but we suspect that the asymptotic decoupling of the scalar from other fermions and gauge bosons will survive. Nonetheless, confronted with the slippery ground of not having a renormalizable theory of gravity, we should be cautious about over-interpreting the results of our simplified treatment.

VIII. CONCLUSIONS

In this paper, we have presented the appropriate equations along with their initial conditions which must be solved to follow the evolution of density fluctuations when there is more than one scalar field present. We showed that special attention must be paid to the way the initially uncorrelated zero-point oscillations of the fields develop cross correlations as evolution proceeds through cross couplings in the interaction potential and through the metric perturbations. With N fields one must do altogether $2N$ runs for each wave number to properly construct $\mathcal{P}_\zeta$, two for each of the N independent annihilation operators, one being for the real part and the other for the imaginary part of the mode function. We solve $2N+1$ ordinary differential equations (ODE's) to get the background field evolution and must solve $4N(N+1)$ real ODE's for each wave vector to get the perturbations, since there are $2N$ replications of the $2(N+1)$ coupled mode equations. The replication runs differ only in their initial conditions. For numerical purposes we have found that working in either the longitudinal or synchronous gauge gives accurate results, although some care must be taken with the choice of variables to follow in the latter case.

We are in general agreement with the spirit of the Kofman and Linde[16] and Kofman and Pogosyan[19] papers: practically any form for the fluctuations can be generated in inflation provided we allow ourselves complete freedom in the structure of the potential surface and the initial conditions for the scalar fields. The issue is what forms are probable. Although it is not yet possible to assess the likelihood for required sets of initial conditions that gave rise to our unique patch of the Universe, we have a strong prejudice towards models that are not sensitive to them.

We have shown that initial fluctuation spectra with the ramp plus plateau shape can appear in inflation models with more than one dynamically important scalar field. However, the tuning of the initial condition $\phi_2(t_i)$ and the potential parameters to be just right so that the ramp in the power spectrum nearly coincides with cluster scale, does not rise too high to violate the microwave anisotropy constraints, yet rises high enough to give the cluster correlations, requires an unlikely set of circumstances. Gaussian mountains may also be produced in multiple scalar field models, but they share the same problems if we restrict ourselves to potentials consisting of quadratic and quartic interaction terms.

If the necessity for mountains of large-scale power above the CDM density spectrum is really thrust upon us by the observations, then we are confronted with a number of relatively unpalatable options. Large-scale power can be obtained with scale-invariant initial conditions by utilizing the natural scales defining the shape of the transfer function. However, since mountains of extra power cannot be obtained with flat initial conditions no matter how the transfer function is modified (only plateaus are possible), mountains could only arise by breaking scale invariance in the fluctuation generator, either to give CDM+mt initial conditions, or power-law initial conditions, as in the isocurvature baryon models. The power laws that are attainable in power-law inflation for isocurvature perturbations go in the wrong direction, with more power at large scales than the scale-invariant spectra which are already ruled out by excessive large-angle CMB anisotropies. Although for the adiabatic mode the power-law spectra (obtained with an exponential potential) do rise with increasing wavelengths compared with Zeldovich spectra, more of the power would reside at large scales, where the CMB constraints restrict the amplitude of the rise, than at intermediate scales, where the extra power could do some good.

Small mountains (moguls) on the potential surface generically lead to mountain fluctuations that are independent of initial field conditions. We saw that these could be Gaussian provided the difference between the background field and the ridge line of the mountain exceeds the level of the quantum fluctuations of the field. An explicit construction of such a case was given in Sec. VI C 2 by adding a cubic interaction to our generic quartic potential. On the other hand, if the quantum fluctuations dominate, the density field will be non-Gaussian, as we showed in Sec. VI D. This subject certainly warrants further investigation since it has such profound implications for the large-scale texture of the Universe.

Adiabatic perturbations will always arise from the inflaton. If the potential is sufficiently flat in the isocon direction, isocurvature perturbations could also rise, perhaps being dominant. Placing a single mogul at just the right location to solve our large-scale structure problems seems contrived: potential parameters defining the position, the height, and the width of the mogul to place extra power with the "observed" amplitude just beyond the scale of clusters of galaxies adds much more fine-tuning to the already unnatural tiny λ requirement of single field inflation.

However, as we emphasized in Fig. 4, current observations of structure in the Universe probe a surprisingly small region of the potential surface. In standard chaotic inflation, the range $3.7m_P$ to $4.4m_P$ model encompasses scales from about 1 kpc, corresponding to the Jean's mass at recombination, to about 5000 h^{-1} Mpc, our current horizon scale. The extra large-scale power suggested by the current state of the data would correspond to a mogul covering the region from about 4.1 to $4.3m_P$. Rather than specially placing one mogul, a more natural situation would be to consider many moguls, possibly of varying scales, littering the potential surface. How strong is the case for scale-invariant spectra in our narrow observable window of potential space? The RELICT CMB bounds are only upper bounds, and cover about 4.3 to $4.4m_P$. The strongest evidence for a flat regime would be the success of the CDM model is giving the galaxy-galaxy correlation function, covering $4.0m_P$ to $4.1m_P$. It is also possible to argue that even this is a fortuitous coincidence, and that the true fluctuations are steeper than scale invariant.[66] We certainty could have a spectrum radically different than scale invariant from $3.7m_P$ to $4.1m_P$ with our current level of knowledge of subgalactic masses. On very short distances, the constraints are quite weak; for example, the degree of hilliness must not lead to a prolonged cessation of inflation. Many moguls could also lead to spectra with plateaus or even power laws over the narrow window of observation. In all cases, non-Gaussian fluctuations could arise. Thus the difficulty does not lie in breaking scale invariance to generate any prescribed fluctuation spectrum provided the potential surface is treated phenomenologically, but in limiting the possibilities of mogul structures that may arise *naturally* in particle theories of the early Universe.

Within the standard chaotic inflation scenario, there are three problems: a finely-tuned λ, a constrained curvature coupling constant $\xi \lesssim 0.002$, and inefficient reheating. By allowing ξ to be negative and large, and recognizing that grand unified models are more likely to undergo chaotic rather than new inflation, we showed how these three problems could be solved, although $\sqrt{\lambda}/|\xi|$ still has to be tuned to get the "observed" fluctuation level. If we expand the Higgs sector of the GUT to include more than one Higgs multiplet, then all of the methods we have used to generate broken scale spectra can be applied to these theories.

ACKNOWLEDGMENTS

We would like to thank B. Holdom, A. Linde, L. Kofman, and A. Starobinskii for informative discussions on various aspects of this work. J.M.B. was supported by DOE Grant No. DE-AC0G-81ER40048 at the University of Washington, J.R.B. by the Canadian Institute for Advanced Research and the Sloan Foundation. J.R.B. and D.S.S. were also supported by the Natural Sciences and Engineering Research Council of Canada.

APPENDIX A: COMMUTATION RELATIONS FOR NONMINIMALLY COUPLED SCALAR FIELDS

In the presence of a $f(\phi_k)R$ term in the Lagrangian (3.1), the usual commutation relations, $[\phi(x,\tau),(\partial\phi/\partial\tau)(x',\tau)]=ia^{-2}\delta^3(x-x')$ are modified (τ is conformal time). Our development here will be nonrigorous. The rigorous and elegant method uses the transformed action (3.6) and is given implicitly in (3.7c). For simplicity then, choose $N=1$, $T^{ij}(\phi_k)=1$, and write the metric as $ds^2=a^2(-d\tau^2+dx^2+dy^2+dz^2)$. Making use of the conformal transformation (3.5) and then integrating by parts, the action becomes (note that all indices are raised and lowered by $\eta_{\mu\nu}$)

$$S=\int d^4x\left[6fa_{,\mu}a^{,\mu}+6a\frac{\partial f}{\partial\phi}a_{,\mu}\phi^{,\mu}-\tfrac{1}{2}a^2\phi_{,\mu}\phi^{,\mu}-a^4V(\phi)\right], \tag{A1a}$$

with corresponding momenta

$$P_\phi\equiv\frac{\delta S}{\delta(\partial_0\phi)}=-a^2\partial^0\phi+6a\frac{\partial f}{\partial\phi}\partial^0 a\ ,$$
$$P_a\equiv\frac{\delta S}{\delta(\partial_0 a)}=6a\frac{\partial f}{\partial\phi}\partial^0\phi+12f\partial^0 a\ , \tag{A1b}$$

which satisfy the classical Poisson-brackets relations[77]

$$\{\phi(x,\tau),P_\phi(x',\tau)\}=\{a(x,\tau),P_a(x',\tau)\}=\delta^3(x-x')\ , \tag{A1c}$$

with all other Poisson brackets vanishing. Note that the integration by parts eliminates second derivatives, produces the nondiagonal term $6a(\partial f/\partial\phi)a_{,\mu}\phi^{,\mu}$ and, consequently, the momenta now contain nonstandard terms $\propto\partial f/\partial\phi$. The Lagrangian (A1a) describes a minisuperspace model since only the scale factor of the metric is treated. Inverting (A1b) in favor of $\partial^0\phi$ and $\partial^0 a$, we find that

$$\left\{\phi(x,\tau),\frac{\partial\phi}{\partial\tau}(x',\tau)\right\}=\frac{fa^{-2}}{f+3\left[\dfrac{\partial f}{\partial\phi}\right]^2}\delta^3(x-x')\ . \tag{A2}$$

Only if $\partial f/\partial\phi=0$ is the standard Poisson brackets obtained. In particular, if $f=m_P^2/16\pi-\xi\phi^2/2$ then

$$\left\{\phi(x,\tau),\frac{\partial\phi}{\partial\tau}(x',\tau)\right\}=a^{-2}\frac{1-8\pi\xi\dfrac{\phi^2}{m_P^2}}{1+8\pi\xi(6\xi-1)\dfrac{\phi^2}{m_P^2}}\delta^3(x-x')\ . \tag{A3}$$

In the quantum treatment, we would replace $\delta^3(x-x')$ by $i\delta^3(x-x')$, and the resulting effect on the normalization of the mode functions is crucial if $\phi>m_P/(8\pi|\xi|)^{1/2}$, a case we treat in Sec. VII using a different method. The right-hand side of (A3) is time dependent and it is therefore possible to arrange that it be small at the onset of inflation. This forces the quantum fluctuations leaving the horizon to be small and produces a suitable level of metric fluctuations even if $\lambda=5\times10^{-2}$. If at the end of inflation, $\phi<|8\pi\xi(6\xi-1)|^{-1/2}m_P$, (A3) reduces to the regular equation. The final result is a standard grand unified theory (see Appendix B).

APPENDIX B: INDUCED GRAVITY IN A GUT FRAMEWORK

In Sec. VII we claimed that if the curvature coupling constant was chosen to be $\xi=-2\times10^4$, inflation could be incorporated within a grand unified theory. However, the simplified analysis of that section utilized a single complex scalar field coupled to a U(1) gauge field, and we now wish to consider the physically more interesting case of non-Abelian gauge fields. We concentrate on the well-studied minimal SU(5) model, although any gauge group could be incorporated. In SU(5), the Higgs field, H, and the gauge field, A_μ, both in the adjoint representation, are described by the Lagrangian density

$$\mathcal{L}=-\tfrac{1}{2}\mathrm{tr}(\partial_\mu H+ig_G[A_\mu,H])(\partial^\mu H+ig_G[A^\mu,H])-V(H)-\tfrac{1}{2}\mathrm{tr}G_{\mu\nu}G^{\mu\nu}\ , \tag{B1a}$$

with the potential chosen to be

$$V(H)=-m_1^2\mathrm{tr}H^2+\lambda_1(\mathrm{tr}H^2)^2+\lambda_2\mathrm{tr}H^4\ . \tag{B1b}$$

The conventions are those of Itzykson and Zuber,[78] Cheng and Li,[74] and Brandenberger:[75]

$$H=\sqrt{2}\phi_i\tau^i,\quad A_\mu=A_{\mu i}\tau^i,\quad \mathrm{tr}\tau^i\tau^j=\tfrac{1}{2}\delta^{ij}\ . \tag{B1c}$$

We work with the variable Planck mass model of Sec. VII, where $f=m^2/(16\pi)-\xi\,\mathrm{tr}H^2/2$, and $m^2\approx m_P^2$, so that when H reaches the bottom of its potential, $\mathrm{tr}H^2\ll m_P^2/(8\pi|\xi|)$. The conformal transformation (3.7a) leads to a new Lagrangian density:

$$\tilde{\mathcal{L}}=-\frac{1}{2}\frac{m_P^2}{16\pi f^2}\{3f_{,\mu}f^{,\mu}+f\,\mathrm{tr}(\partial_\mu H+ig_G[A_\mu,H])\times(\partial^\mu H+ig_G[A^\mu,H])\}-\left[\frac{m_P^2}{16\pi f}\right]^2V(H)-\tfrac{1}{2}\mathrm{tr}G_{\mu\nu}G^{\mu\nu}\ . \tag{B2}$$

The gauge boson masses are generated through the term

$$\frac{1}{2}\left[\frac{m_P^2}{16\pi f}\right]g_G^2\mathrm{tr}[A_\mu,H][A^\mu,H]\ . \tag{B3}$$

If H is diagonalized, $H=\mathrm{diag}[h_1,\ldots,h_5]$, then the massive gauge bosons may be labeled by two indices, i,j, $1\le i,j\le5, i\ne j$, whose masses are given by

$$m_{ij}^2 = \frac{1}{2}\left[\frac{m_P^2}{16\pi f}\right] g_G^2 (h_i - h_j)^2 \ . \tag{B4}$$

The formula, (7.16), for the gauge-boson contribution to the effective potential generalizes to a sum over all gauge bosons, and if H is in the SU(3)×SU(2)×U(1) trough, $H = \phi_1 \mathrm{diag}[2,2,2,-3,-3]/\sqrt{30}$, the radiative potential is

$$V_{\mathrm{rad}} = \frac{25 g_G^4}{256\pi^2}\left\{\frac{\sigma^4}{2} + \left[\frac{m_P^2 \phi_1^2}{16\pi f}\right]^2 \left[\ln\left[\frac{m_P^2}{16\pi f}\frac{\phi_1^2}{\sigma^2}\right] - \frac{1}{2}\right]\right\} . \tag{B5}$$

The normalization scale t of (7.16) and the cosmological constant were chosen so that at $\phi_1 = \sigma$, V_{rad} reaches its minimum of zero. Radiative corrections to the potential are negligible if the coefficient of ϕ_1^4 in Eq. (B2) exceeds that in V_{rad}:

$$\lambda_1 + \tfrac{7}{30}\lambda_2 > \lim_{\phi_1 \to \infty} \frac{25 g_G^4}{256\pi^2}\left[\ln\left[\frac{m_P^2}{16\pi}\frac{\phi_1^2}{\sigma^2}\right] - \tfrac{1}{2}\right] . \tag{B6}$$

In particular, if $\xi = -2\times 10^4$, $\sigma = 1.2\times 10^{15}$ GeV, and $g_G = 0.5$, we find that $\lambda_1 + \frac{7}{30}\lambda_2 > 3.0\times 10^{-3}$. This constraint is quite plausible and is easily achieved (see below), whereas in the minimally coupled case it was devastating.

To ensure the suppression of monopoles and yet have baryogenesis proceed through the decay of gauge bosons, we impose the condition

$$M_X < T_{\mathrm{reh}} < T_S \ . \tag{B7}$$

Additional constraints beyond $M_X < T_{\mathrm{reh}}$ are required if the theoretically calculated baryon asymmetry is to agree with observations. In fact, minimal SU(5) does not give the correct value,[62] so we use the SU(5) model to illustrate what would have to been done in other models. T_{reh} depends primarily on the level of primordial metric fluctuations and is given in Eq. (7.18): $T_{\mathrm{reh}} = 2.0\times 10^{15}\epsilon$ GeV; according to (B4), $M_X = \sqrt{5/12}\, g_G \sigma$, where $\sigma = m_1/[2(\lambda_1 + 7\lambda_2/30)]^{1/2}$ is where the minimum value of the ϕ_1 potential is achieved; T_S has been derived by Guth and Tye,[79] and has the form

$$T_S = 4\sigma\left[\frac{\lambda_1 + \frac{7}{30}\lambda_2}{5 g_G^2 + \frac{104}{3}\lambda_1 + \frac{188}{15}\lambda_2}\right]^{1/2} . \tag{B8}$$

The inequality (B7) then becomes

$$\frac{1}{4}\left[\frac{5 g_G^2 + \frac{104}{3}\lambda_1 + \frac{188}{15}\lambda_2}{\lambda_1 + \frac{7}{30}\lambda_2}\right]^{1/2} < \sigma / T_{\mathrm{reh}} < \sqrt{\tfrac{12}{5}}\,\frac{1}{g_G} , \tag{B9}$$

which always has a solution if the following consistency constraint is satisfied:

$$\left[\frac{192}{5 g_G^2} - \frac{104}{3}\right]\lambda_1 + \left[\frac{224}{25 g_G^2} - \frac{188}{15}\right]\lambda_2 > 5 g_G^2 \ . \tag{B10}$$

Additional constraints on the Higgs parameters λ_1, λ_2 come from the requirement that the SU(3)×SU(2)×U(1)-breaking pattern be the minimum of the potential:[74]

$$\lambda_1 + \tfrac{7}{30}\lambda_2 > 0 \ , \tag{B11a}$$

$$\lambda_2 > 0 \ . \tag{B11b}$$

We will also apply the constraint that $\lambda < 1$ in $\lambda\phi^4/4$ which ensures that perturbation theory is valid:

$$\lambda_1 + \tfrac{7}{30}\lambda_2 < 0.25 \ . \tag{B12}$$

These inequalities imply that λ_1 and λ_2 must fall within a triangular region bounded by (B10)–(B12) (see Fig. 14); (B11a) is actually redundant and the constraint (B6) turns out to be very weak.

Within the allowed region, the minimum value of the left-hand side of (B9) is 1.575, which occurs at $\lambda_1 = -1.247$, and $\lambda_2 = 6.415$, and the resulting constraint on σ is

$$3.0\times 10^{15}\epsilon \ \mathrm{GeV} < \sigma < 5.9\times 10^{15}\epsilon \ \mathrm{GeV} \ , \tag{B13}$$

where we use the efficiency factor determined in Sec. VII: $\epsilon = 0.48$. On the other hand, if the Higgs field produces the baryon asymmetry, then only the left-hand side of (B13) is necessary; we have already shown in Sec. VII A that $m_H < T_{\mathrm{reh}}$ occurs quite naturally. The value normally quoted for σ is 1.2×10^{15} GeV, which is based on the extrapolation of running coupling constants over 13 orders of magnitude. Although this number is outside the

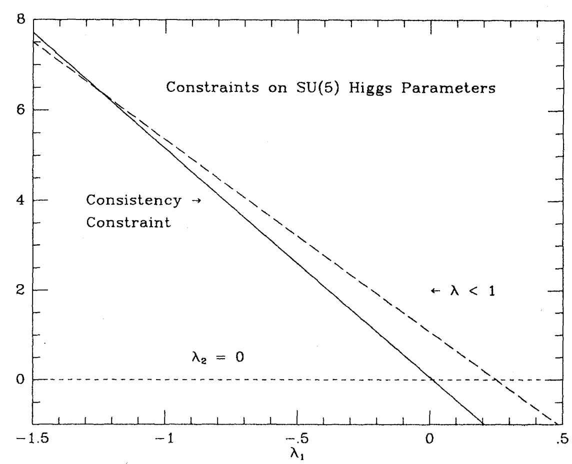

FIG. 14. The sort of constraints that arise in large $-\xi$ variable Planck mass models are illustrated in this figure for an SU(5) GUT model. The SU(5) Higgs parameters λ_1 and λ_2 [defined by Eq. (B1b)] must fall within the triangle shown to (1) give no monopole production and yet have baryogenesis occur through gauge-boson decays [solid line, from Eq. (B10)], (2) ensure that symmetry breaking gives SU(3)×SU(2)×U(1) as the absolute minimum ($\lambda_2 < 0$), and (3) allow perturbation theory to be used [Eq. (B12)]. If Higgs bosons generate the baryon asymmetry, the solid line is replaced by a line parallel to the long-dashed line which passes through the origin, as given by Eq. (B11a).

allowed region (B13), we do not believe that this failure is serious as further refinements and/or change of GUT model could possibly give agreement.

We now consider the cosmological consequences of the four fields that parametrize the diagonal Higgs field, $H=\sqrt{2}\phi_i\tau^i$. It proves convenient to choose the diagonal basis

$$\begin{aligned}\tau^1&=\text{diag}[2,2,2,-3,-3]/\sqrt{60}\ ,\\ \tau^2&=\text{diag}[0,0,0,1,-1]/2\ ,\\ \tau^3&=\text{diag}[1,-1,0,0,0]/2\ ,\\ \tau^4&=\text{diag}[1,1,-2,0,0]/(2\sqrt{3})\ ,\end{aligned} \tag{B14}$$

and then express the ϕ_i in hyperspherical coordinates, $\phi,\theta_1,\theta_2,\theta_3$:

$$\begin{aligned}\phi_1&=\phi\cos\theta_1\ ,\\ \phi_2&=\phi\sin\theta_1\cos\theta_2\ ,\\ \phi_3&=\phi\sin\theta_1\sin\theta_2\cos\theta_3\ ,\\ \phi_4&=\phi\sin\theta_1\sin\theta_2\sin\theta_3\ .\end{aligned} \tag{B15}$$

We find that the scalar field part of the Lagrangian density, (B2), is

$$\begin{aligned}\tilde{\mathcal{L}}_\phi=-\tfrac{1}{2}\chi_{,\mu}\chi^{,\mu}-\frac{1}{2}\frac{m_P^2\phi^2}{16\pi f}(\theta_{1,\mu}\theta_1^{,\mu}+\sin^2\theta_1\theta_{2,\mu}\theta_2^{,\mu}\\ +\sin^2\theta_1\sin^2\theta_2\theta_{3,\mu}\theta_3^{,\mu})\\ -\left[\frac{m_P^2}{16\pi f}\right]^2 V(H)\ ,\end{aligned} \tag{B16a}$$

where

$$V(H)=-m_1^2\phi^2+\phi^4[\lambda_1+\lambda_2 g(\theta_i)]\ , \tag{B16b}$$

$$\begin{aligned}g(\theta_i)=\tfrac{7}{30}+\tfrac{4}{3}\sin^2\theta_1-\tfrac{16}{15}\sin^4\theta_1+\sin^2\theta_1\sin^2\theta_2\\ \times\Bigg[-1+\sin^2\theta_1\sin^2\theta_2\\ +\frac{4\sqrt{5}}{15}\cos\theta_1\sin\theta_1\sin\theta_2\sin(3\theta_3)\Bigg]\ ,\end{aligned} \tag{B16c}$$

and $\chi(\phi)$ is defined in (7.2a). The angular potential, $g(\theta_i)$, possesses an absolute minima at $\theta_1=0$ corresponding to SU(3)×SU(2)×U(1), $H\propto\text{diag}[2,2,2,-3,-3]$, and an absolute maximum at $\theta_2=0$, $\tan(\theta_1)=(\frac{5}{3})^{1/2}$ where $H\propto\text{diag}[1,1,1,1,-4]$. Permutations of the diagonal elements of the Higgs field would also lead to absolute extrema. All other critical points are saddle points, so if the conditions (B11a) and (B11b) are satisfied, it is inevitable that the Higgs field will roll into the SU(3)×SU(2)×U(1) trough. In fact, if the Universe begins chaotically, the time to do so is extremely short, measured in fractions of a Hubble time. To see this, we need to compare the effective mass with the Hubble parameter. Since an order-of-magnitude estimate is sufficient, we will restrict ourselves to $\theta_2=0$, in which case $g(\theta_i)$ is a function only of θ_1. If $\phi\gg m_P/(8\pi|\xi|)^{1/2}$, the coefficient of $-\frac{1}{2}\theta_{1,\mu}\theta_1^{,\mu}$ in (B16a) is essentially a constant, $m_P^2/(8\pi|\xi|)$, and a rescaling of the field, $y=[m_P/(8\pi|\xi|)^{1/2}]\theta_1$, yields the standard kinetic term. The mass of this degree of freedom is of order $\lambda_2 m_P^2/(8\pi|\xi|)$, which is approximately $10|\xi|H_i^2\approx 10^5 H_i^2$.

Therefore, the time scale for motion along the θ_1 direction is extremely rapid, hence the Higgs field will roll rapidly in the angular direction into the proper trough; only then will the slow roll-down phase begin. This holds even if the angular variables start very near the maxima of their potential. Thus, although the potential (B16b) contains a peak as well as saddle structures, double inflation spectra do not arise in this model. Further, mountains also do not occur, since the fields do not pass through a region of negative mass squared during slow rollover. Double inflation might arise with additional multiplets. One would only require that this new chaotic field have a value of the λ parameter much larger than that of the **24**, Higgs field so that it may dominate the energy density initially.

APPENDIX C: NUMERICAL SOLUTION OF INFLATION EQUATIONS

The numerical computations of the background equations (4.6) and the perturbation equations, (4.7) for the synchronous gauge and (4.8) for the longitudinal gauge, form the backbone of this paper. These ordinary differential equations are integrated using a fourth-order Runge-Kutta scheme. We express all quantities in units of m_P in the code and use $x=\ln a$ for the integration variable. For the case of two scalar fields, the integration variables are ρ, ϕ_1, $\dot\phi_1$, ϕ_2, and $\dot\phi_2$. The equations we solve for ϕ_1, for example, are therefore

$$\frac{d\phi_1}{dx}=\dot\phi_1/H\ ,$$

$$\frac{d\dot\phi_1}{dx}=-\left[(3H+\Gamma_1)\dot\phi_1+\frac{\partial V}{\partial\phi_1}\right]\Big/H\ .$$

The initial conditions for the background fields $\phi_j(t_i)$ are always chosen to give at least 60 e-foldings of inflation, and $\dot\phi_j(t_i)$ and $\rho(t_i)$ are determined by the slow-rollover condition (Sec. V). The timestep, Δx, is chosen to be

$$\Delta x=\varepsilon X\ ,$$

$$X=\min\left[2\pi H\Big/\left|\frac{\partial^2 V}{\partial\phi_i^2}\right|^{1/2},3H/(3H+\Gamma_i)\right]\ .$$

The minimum of the kinetic and damping time scales is taken over all fields. We typically choose the conservative value $\varepsilon\approx\frac{1}{40}$ to ensure our steps are small compared with the characteristic times of the problem. For some problems larger values ε are quite sufficient. The background run is stopped when the Universe reheats: $\rho_\phi<10^{-3}\rho_{\text{tot}}$.

Other model-dependent criteria must be used to supplement the above choice of time step. For example, the

Coleman-Weinberg potential has an inflection point; hence, $2\pi H/|\partial^2 V/\partial\phi^2|^{1/2}$ is a poor time scale choice near there; for $\phi/\sigma > 0.4$, we replace it by $2\pi H/|(\partial^2 V/\partial\phi^2)(\sigma)|^{1/2}$. For multiple scalar fields, one often encounters the problem of two disparate time scales: say m_1 and m_2. As ϕ_1 rolls down first and damps away in its trough, we are justified in killing this degree of freedom once $\rho_{\phi_1} < 10^{-5}\rho_{\phi_2}$. For simple $\nu=0$ double inflation models, we then set the background values ϕ_1 and $\dot{\phi}_1$ to zero as well as the two perturbation values ψ_{11} and ψ_{12}. For models with $\nu<0$, we set $\phi_1=\phi_{1tr}$, $\dot{\phi}_1=\dot{\phi}_{1tr}$, $\psi_{11}=0=\psi_{12}$, but, in addition, we insist that this occurs once $\phi_2 < 0.9(m_1^2/\nu)^{1/2}$ to avoid turning off the field before it has reached a region of negative mass squared.

We use the χ variable, Eq. (7.2a), to evolve the variable Planck mass model. To calculate the potential $U(\phi)$ in terms of χ, we invert (7.2a) by a four-point Lagrange polynomial interpolation on a table of ordered pairs, (ϕ,χ), with ϕ values ranging from $10^{-3}\sigma$ to $0.1m_P$.

To integrate the perturbation equations, we require the values of the background fields, interpolated from our computed values using once again a four-point Lagrange polynomial technique. The values of the perturbation variables are given as in Sec. V when $k > fHa$, with f usually 50, although for some problems it can be smaller. Once $k/Ha < f$, the integration begins, with the time step choice

$$\Delta x = \min(\varepsilon 2\pi Ha/k, \Delta x_{bac}) ,$$

where Δx_{bac} is the time step of the background calculation at that same time.

[1]A. D. Linde, Phys. Lett. **129B**, 177 (1983).

[2]M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, Astrophys. J. **292**, 371 (1985).

[3]J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).

[4]N. Bahcall and R. Soneira, Astrophys. J. **270**, 70 (1983).

[5]J. M. Bardeen, J. R. Bond, and G. Efstathiou, Astrophys. J. **321**, 28 (1987).

[6]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).

[7]N. Kaiser and U. Lahav, in *Large Scale Motions in the Universe,* A Vatican Study Week, edited by V. C. Rubin and G. V. Coyne (Princeton University Press, Princeton, New Jersey, 1988).

[8]J. R. Bond, and H. Couchman, in *General Relativity and Relativistic Astrophysics,* proceedings of the 2nd Canadian Conference, Toronto, Canada, edited by A. Coley, C. Dyer, and B. Tupper (World Scientific, Singapore, 1988).

[9]S. Maddox and G. Efstathiou, in *Evolution and Large Scale Structure in the Universe,* edited by J. Audouze and A. Szalay (Reidel, Dordrecht, 1988).

[10]J. R. Bond and H. Couchman (in preparation).

[11]G. Efstathiou and P. Lilje, Mon. Not. R. Astron. Soc. **231**, 635 (1988).

[12]J. R. Bond, in *Large Scale Motions in the Universe* (Ref. 7).

[13]J. R. Bond, D. S. Salopek, and J. M. Bardeen, in *Large Scale Motions in the Universe* (Ref. 7).

[14]T. Futamase and K. Maeda, Phys. Rev. D **39**, 399 (1989).

[15]A. A. Starobinskii, Pis'ma Zh. Eksp. Teor. Fiz. **42**, 124 (1985) [JETP Lett. **42**, 152 (1985)].

[16]L. A. Kofman and A. D. Linde, Nucl. Phys. **B282**, 555 (1987).

[17]J. I. Silk and M. S. Turner, Phys. Rev. D **35**, 419 (1987).

[18]J. M. Bardeen, J. R. Bond, and D. Salopek, in *Proceedings of the Second Canadian Conference on General Relativity and Relativistic Astrophysics* (Ref. 8).

[19]L. A. Kofman and D. Y. Pogosyan, Tallinn Report No. A-8, 1988 (unpublished).

[20]A. Zee, Phys. Rev. Lett. **42**, 417 (1979).

[21]S. Adler, Rev. Mod. Phys. **54**, 729 (1982).

[22]F. S. Accetta, D. J. Zoller, and M. S. Turner, Phys. Rev. D **31**, 3046 (1985).

[23]B. L. Spokoiny, Phys. Lett. **129B**, 39 (1984); R. Fakir and W. G. Unruh, University of British Columbia report, 1988 (unpublished).

[24]J. R. Bond, in *Large Scale Structure in the Universe,* proceedings of the IAU Symposium No. 130, Lake Balaton, Hungary, edited by J. Audouze and A. Szalay (Reidel, Dordrecht, 1988).

[25]I. A. Strukov, D. P. Skulachev, and A. A. Klypin, in *Large Scale Structures of the Universe,* edited by J. Audouze *et al.* (Kluwar, Dordrecht, 1988).

[26]P. J. E. Peebles, *The Large Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980).

[27]J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).

[28]J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. D **28**, 679 (1983).

[29]J. M. Bardeen, in *Second Guo Shou-jing Summer School on Particle Physics and Cosmology,* edited by A. Zee (Gordon and Breach, London, 1988).

[30]J. R. Bond and G. Efstathiou, Astrophys. J. Lett. **285**, L45 (1984).

[31]J. R. Bond, in *Proceedings of the NATO Advanced Study Institute on the Epoch of Galaxy Formation, Durham, England,* edited by C. Frenk *et al.* (Kluwer, Dordrecht, 1988).

[32]J. R. Bond and G. Efstathiou (in preparation).

[33]R. D. Davies *et al.*, Nature (London) **326**, 462 (1987).

[34]J. Yang, M. S. Turner, G. Steigman, D. N. Schramm, and K. A. Olive, Astrophys. J. **281**, 493 (1984).

[35]P. Steinhardt and M. S. Turner, Phys. Lett. **129B**, 51 (1983).

[36]G. Efstathiou and J. R. Bond, Mon. Not. R. Astron. Soc. **218**, 103 (1986).

[37]G. Efstathiou and J. R. Bond, Mon. Not. R. Astron. Soc. **227**, 33P (1987).

[38]P. J. E. Peebles, in *Early Universe,* proceedings of the NATO Advanced Study Institute, Victoria, Canada, 1986, edited by W. G. Unruh and G. W. Semenoff (NATO ASI Series C, Vol. 219) (Reidel, Boston, 1988).

[39]P. J. E. Peebles, Astrophys. J. Lett. **277**, L1 (1987).

[40]P. J. E. Peebles, Nature (London) **327**, 210 (1987).

[41]J. R. Gott and M. Rees, Astron. Astrophys. **45**, 365 (1975).

[42]N. Vittorio, F. Lucchin, and S. Matarrese, Astrophys. J. **328**, 69 (1988).

69 (1988).
[43]M. S. Turner and L. Widrow, Fermilab report, 1987 (unpublished).
[44]L. A. Kofman, A. D. Linde, and A. A. Starobinskii, Phys. Lett. **157B**, 361 (1985).
[45]C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[46]P. Jordan, Z. Phys. **157**, 112 (1959).
[47]M. Green, J. Schwarz, and E. Witten, *Superstrings* (Cambridge University Press, Cambridge, England, 1987).
[48]S. L. Adler, Rev. Mod. Phys. **54**, 729 (1982).
[49]N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
[50]F. Lucchin, S. Mataresse, and M. D. Pollock, Phys. Lett. **167B**, 163 (1986).
[51]C. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[52]J. J. Halliwell and S. Hawking, Phys. Rev. D **31**, 1777 (1985).
[53]A. V. Veryaskin, V. A. Rubakov, and M. V. Sazhin, Astron. Zh. **60**, 26 (1983) [Sov. Astron. **27**, 16 (1983)].
[54]A. A. Starobinskii, Pis'ma Astron. Zh. **11**, 323 (1985) [Sov. Astron. Lett. **11**, 133 (1985)].
[55]V. N. Lukash, Zh. Eksp. Teor. Fiz. **79**, 1601 (1980) [Sov. Phys. JETP **52**, 807 (1980)].
[56]G. V. Chibisov and V. F. Mukhanov, Mon. Not. R. Astron. Soc. **200**, 535 (1982).
[57]W. Fischler, B. Ratra, and L. Susskind, Nucl. Phys. **B259**, 730 (1985).
[58]J. B. Hartle and S. Hawking, Phys. Rev. D **28**, 2960 (1983).
[59]E. M. Lifshitz, JETP Lett. **16**, 587 (1946).
[60]M. Den and K. Tomita, Prog. Theor. Phys. **72**, 989 (1984).
[61]M. Sasaki, Prog. Theor. Phys. **70**, 394 (1983).
[62]M. S. Turner, in *Early Universe* (Ref. 38).
[63]A. A. Starobinskii, in *Quantum Gravity,* edited by M. A. Markov and P. West (Plenum, New York, 1983), p. 103.
[64]L. F. Abbott and M. B. Wise, Nucl. Phys. **B244**, 541 (1984).
[65]F. Lucchin and S. Matarrese, Phys. Rev. D **32**, 1316 (1985).
[66]P. J. E. Peebles, in *Large Scale Motions in the Universe* (Ref. 7).
[67]G. Masenko, W. Unruh, and R. Wald, Phys. Rev. D **31**, 273 (1984); R. H. Brandenberger and H. A. Feldman, Brown University Report No. HET-686, 1988 (unpublished).
[68]J. M. Bardeen, in *Proceedings of the CCAST Symposium on Particle Physics and Cosmology,* edited by A. Zee (Gordon and Breach, New York, 1988).
[69]T. J. Allen, B. Grinstein, and M. Wise, Phys. Lett. B **197**, 66 (1987).
[70]L. A. Kofman, A. D. Linde, and J. Einasto, Nature (London) **326**, 48 (1987).
[71]J. M. Bardeen and G. J. Bublik, Class. Quantum Gravit. **4**, 573 (1987); J. R. Bond and D. S. Salopek (in preparation); A. Ortolan, F. Lucchin, and S. Matarrese, Phys. Rev. D **38**, 465 (1988).
[72]J. J. Halliwell, Phys. Lett. B **185**, 341 (1987).
[73]S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).
[74]T. Cheng and L. Li, *Gauge Theory of Elementary Particle Physics* (Oxford University Press, New York, 1984).
[75]R. H. Brandenberger, Rev. Mod. Phys. **57**, 1 (1985).
[76]B. Whitt, Phys. Lett. **145B**, 176 (1984).
[77]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1981).
[78]C. Itzykson and J. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).
[79]A. H. Guth and S. H. Tye, Phys. Rev. Lett. **44**, 631 (1980); **44**, 963 (1980).