PHYSICAL REVIEW D VOLUME 31, NUMBER 8 15 APRIL 1985

Quantum mechanics of inflation

Mark Evans and James G. McCarthy
Department of Physics, The Rockefeller University, New York, New York 10021
(Received 3 October 1984)

It is pointed out that the slow-roll equation of new-inflationary-universe models is not true in general for interacting field theories. We determine a set of conditions under which it is relevant, but argue that they are not likely to be satisfied by the early universe.

I. INTRODUCTION

It was first realized by Guth[1] that a period of exponential expansion in the early history of the universe (an "inflationary era") would go some way toward remedying certain naturalness problems associated with the initial conditions of the standard cosmological model.[2] If, in the standard model, initial conditions are defined when the universe is a few Planck lengths in size, its maximum radius (assuming a closed universe) will be of the same order of magnitude unless the initial expansion rate is carefully balanced against the initial energy density. The fact that the present universe is somewhat bigger than a Planck length is known as the "flatness problem." The homogeneity and isotropy of the cosmic microwave background radiation is puzzling in the standard model since the sources of this radiation from different parts of the sky have no event common to their pasts. This "horizon problem" is a consequence only of the initial spacelike singularity of the standard model. One can easily imagine that a quantum theory of gravity would eliminate this spacelike singularity and hence solve the horizon problem, but this would not obviously cure the flatness problem.

Inflation solves both problems. Blowing up spatial sections by an enormous factor ($>10^{28}$) reduces their curvature by the same factor, which solves the flatness problem, while points that are moderately separated after inflation must have been extremely close before (close enough to have points common to their pasts) resolving the horizon problem. Other advantages of inflationary models include a possible explanation for fluctuations that give rise to galaxies, a dilution of magnetic monopoles, and a mechanism for generating vast amounts of entropy.[3] Inflation places potentially useful constraints on models of particle physics,[4] and makes one unambiguous statement about the present universe—it is spatially flat.

The most successful method for generating an adequate period of inflation was proposed by Linde and by Albrecht and Steinhardt.[5] The prototypical model involves a first-order phase transition in which the Higgs field of some unified model of particle physics acquires a nonzero expectation value. If the Higgs potential is chosen to be of the Coleman-Weinberg type[6] (that is, the effective potential is very flat over a large region near the origin, see Fig. 1), the universe will supercool in the symmetric phase to essentially zero temperature, and then the expectation value of the Higgs field will roll slowly to the symmetry-breaking minimum. (Note that the *vacuum* expectation value does not roll. The universe is not in its vacuum state; if it were it would be a very boring place.) During the roll the roughly constant energy density (and pressure $p=-\rho$) act as an effective cosmological constant and the universe expands exponentially. It is thus crucial that the expectation value of the Higgs field takes a sufficient time to reach its symmetry-breaking minimum.

Linde and Albrecht and Steinhardt[5] proposed that the Higgs field be governed by the equation

$$\frac{\partial^2\phi}{\partial t^2}+3H\frac{\partial\phi}{\partial t}=-V'_{\rm eff}(\phi)\ , \tag{1}$$

where $\phi=\langle s\,|\,\hat{\phi}(x)\,|\,s\rangle$ is the expectation value of the Higgs field in the state $|\,s\rangle$ (assumed translation invariant) in which the universe finds itself, and H is the Hubble "constant" $\dot{R}/R$. We shall refer to (1) as the LAS equation, and to its right-hand side as the "driving force." If inflation depends on the smallness of the driving force, it is important to understand contributions to it from quantum effects. We shall therefore investigate the quantum-mechanical validity of Eq. (1). For simplicity we shall consider the flat-space case $H=0$.

In this paper we point out that the LAS equation is not true, in general, for interacting field theories (it is always true for free-field theory). We shall prove, however, that the LAS equation gives a lower bound for the rollover time (so that the usual picture goes through essentially unaltered) *if* the rolling starts in a particular state—that Gaussian which minimizes the expectation of the Hamiltonian subject to the constraint that the expectation of the spatial average of the field is zero.

We shall further argue that the universe is unlikely ever to be near this state if the usual picture of a hot primordial universe is correct.

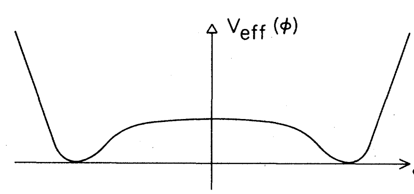

FIG. 1. Coleman-Weinberg effective potential.

It may be that theories with flat effective potentials do generate substantial inflation: our point is that the usual argument, based on Eq. (1), is inadequate.

We note that the question of the quantum mechanics of the slow rollover is currently attracting the attention of many physicists, some of whom are listed in Refs. 7 and 8.

We take the rather archaic point of view that a quantum field theory is defined by specifying equations of motion for a set of Hibert-space operators labeled by points in space (more carefully, a set of operator-valued distributions) and a set of equal-time commutation relations. For an interacting field theory the equations of motion will involve composite operators, the renormalization of which is generally considered arbitrary and ill defined. We therefore take the view that the theory should be regulated using a spatial lattice (i.e., continuous time).[9] This yields a quantum-mechanical system with no ultraviolet divergences, so that composite operators, and hence equations of motion, are well defined. Any divergences that reappear after the theory has been renormalized and the lattice spacing taken to zero are genuine pathologies of the theory.

Specifying the Hamilontian determines the equations of motion, and for definiteness we shall take it to be

$$\hat{H}=\int d^3x\hat{\mathscr{H}}=\int d^3x[\tfrac{1}{2}\hat{\pi}^2+\tfrac{1}{2}(\nabla\hat{\phi})^2+V(\hat{\phi})]\ ,$$
$$[\hat{\pi}(\mathbf{x},t),\hat{\phi}(\mathbf{y},t)]=-i\delta^{(3)}(\mathbf{x}-\mathbf{y})\ . \tag{2}$$

$\hat{\pi}(x)$ is the momentum density conjugate to $\hat{\phi}(x)$ (note that momentum here and throughout the paper does not have its usual field-theoretic meaning). Our results will depend only on the fact that $\hat{\mathscr{H}}$ is quadratic in $\hat{\pi}$, and so are more general than the specific choice of Eq. (2).

The equations of motion from (2) are

$$\dot{\hat{\phi}}(\mathbf{x})=i[\hat{H},\hat{\phi}(\mathbf{x})]=\hat{\pi}(\mathbf{x})\ ,$$
$$\dot{\hat{\pi}}(\mathbf{x})=i[\hat{H},\hat{\pi}(\mathbf{x})]=\nabla^2\hat{\phi}(\mathbf{x})-V'(\hat{\phi}(\mathbf{x}))\ . \tag{3}$$

Combining these two equations and taking the expectation value in a translationally invariant state $|s\rangle$ yields

$$\frac{d^2}{dt^2}\langle s|\hat{\phi}|s\rangle=-\langle s|V'(\hat{\phi})|s\rangle\ . \tag{4}$$

Comparing Eq. (4), known as Ehrenfest's theorem, with the LAS equation (1) shows that the latter is true if and only if

$$\langle s|V'(\hat{\phi})|s\rangle=V'_{\text{eff}}(\langle s|\hat{\phi}|s\rangle)\ . \tag{5}$$

For free-field theories Eq. (5) is always true, since $V'(\hat{\phi})=m^2\hat{\phi}$ and $V'_{\text{eff}}(\langle s|\hat{\phi}|s\rangle)=m^2\langle s|\hat{\phi}|s\rangle$. However, for interacting field theories it is equally obvious that Eq. (5) is not true in general. Take, for example,

$$V(\hat{\phi})=\tfrac{1}{2}m^2\hat{\phi}^2+\tfrac{1}{4}\lambda\hat{\phi}^4\ .$$

Then condition (5) is

$$\langle s|m^2\hat{\phi}+\lambda\hat{\phi}^3|s\rangle=V'_{\text{eff}}(\langle s|\hat{\phi}|s\rangle)\ . \tag{5$'$}$$

The right-hand side is a function only of $\langle s|\hat{\phi}|s\rangle$, but states with equal values of $\langle\hat{\phi}\rangle$ may have different values of $\langle\hat{\phi}^3\rangle$, depending on the shape of the wave functional.

In Sec. II we shall characterize a set of states for which Eq. (5) holds. In Sec. III we shall show that the conditions of Eq. (5) are not preserved under the equation of motion, but that the LAS equation (1) nevertheless provides a lower bound on the rollover time if the universe starts rolling in one of the states described in Sec. II.

Section IV addresses the validity of the one-loop approximation to the effective potential with particular reference to the convexity problem. Finally, in Sec. V, we muse on whether the LAS equation is indeed relevant to cosmology.

II. ON THE EFFECTIVE POTENTIAL

We remind the reader of some standard results concerning the effective potential.[10] The effective action is the Legendre transform of $W[J]$, the generating functional of connected Green's functions. If we take J to be constant in space and time and take the Legendre transform, we obtain the effective potential multiplied by the volume of space-time. Taking J constant, we have, in Euclidean space,

$$e^{-W(J)}=\lim_{\substack{t_1\to\infty\\ t_0\to-\infty}}\int_{\phi_0}^{\phi_1}D\phi\exp\left[-\int_{t_0}^{t_1}dt\int d^3x[\mathscr{L}(x)-J\phi(x)]\right]\ . \tag{6}$$

The right-hand side of (6) is the transition matrix element $\langle\phi_1(x)|e^{-H_J(t_1-t_0)}|\phi_0(x)\rangle$ which leads, in the limit that $t_1-t_0\to\infty$, to

$$W(J)=\int dt\langle 0_J|\hat{H}_J|0_J\rangle\ ,$$
$$\hat{H}_J=\hat{H}+J\int d^3x\,\hat{\phi}(\mathbf{x})\ , \tag{7}$$

where $|0_J\rangle$ is the lowest eigenstate of $\hat{H}_J$. Interpreting J as a Lagrange multiplier, we see that $|0_J\rangle$ minimizes $\hat{H}$ subject to a constraint on the value of $\int d^3x\langle\hat{\phi}(x)\rangle$. Taking the Legendre transform yields

$$\int d^4x\,V_{\text{eff}}(\phi)=W(J)-J\frac{dW}{dJ}$$
$$=W(J)-J\int d^4x\langle 0_J|\hat{\phi}(x)|0_J\rangle\ , \tag{8}$$

i.e.,

$$V_{\text{eff}}(\phi)=\langle 0_\phi|\hat{\mathscr{H}}|0_\phi\rangle\ . \tag{9}$$

We have changed our notation for the state $|0_\phi\rangle=|0_J\rangle$ to indicate that it minimizes $\hat{H}$ for a given value of $\phi=\langle\hat{\phi}\rangle$, and have assumed that $|0_\phi\rangle$ is translationally invariant.

Thus we take as our definition of the effective potential

$$V_{\rm eff}(\phi)=\min_{\text{over }|s\rangle}\{\langle s|\hat{\mathscr{H}}|s\rangle \text{ with } \langle s|\hat{\phi}|s\rangle=\phi\}$$
$$=\langle 0_\phi|\hat{\mathscr{H}}|0_\phi\rangle\ . \tag{10}$$

Our first result is that Eq. (5) holds for the state $|0_\phi\rangle$.

Proof:

$$\langle 0_\phi|V'(\hat{\phi})|0_\phi\rangle=-i\langle 0_\phi|[\hat{H},\hat{\pi}(x)]|0_\phi\rangle \tag{11}$$

but $|0_\phi\rangle=|0_J\rangle$ is an eigenstate of $\hat{H}_J$ [Eq. (7)], that is,

$$\hat{H}|0_\phi\rangle=-J\int d^3y\,\hat{\phi}(y)|0_\phi\rangle+W(J)|0_\phi\rangle\ . \tag{12}$$

Inserting Eq. (12) in Eq. (11) yields

$$\langle 0_\phi|V'(\hat{\phi})|0_\phi\rangle=-J\int d^3y\langle 0_\phi|[\hat{\phi}(y),\hat{\pi}(x)]|0_\phi\rangle$$
$$=V'_{\rm eff}(\phi)\ . \tag{13}$$

The last equality comes from differentiating Eq. (8).

We have therefore identified one state, $|0_\phi\rangle$, for which the LAS equation (1) holds at one point in time. It is, however, insufficient for our purposes, since any state which satisfies Eq. (12) at some time t must yield $\dot{\phi}(t)=0$, and so cannot be rolling, i.e.,

$$\frac{d}{dt}\langle 0_\phi|\hat{\phi}|0_\phi\rangle=\langle 0_\phi|\hat{\pi}|0_\phi\rangle=0\ . \tag{14}$$

Proof:

$$\frac{d}{dt}\langle 0_\phi|\hat{\phi}(\mathbf{x})|0_\phi\rangle=i\langle 0_\phi|[\hat{H},\hat{\phi}(\mathbf{x})]|0_\phi\rangle$$
$$=-iJ\int d^3y\langle 0_\phi|[\hat{\phi}(\mathbf{y}),\hat{\phi}(\mathbf{x})]|0_\phi\rangle$$
$$=0\ ,$$

where use was made of Eq. (12).

We may generalize the above argument and characterize a set of states for which Eq. (5) holds, but for which Eq. (14) does not.

Define the quantity

$$\Gamma(\phi,\pi)=\min_{|s\rangle}\{\langle s|\hat{\mathscr{H}}|s\rangle \text{ with } \langle s|\hat{\phi}|s\rangle=\phi \text{ and } \langle s|\hat{\pi}|s\rangle=\pi\} \tag{15}$$

and denote the state $|s\rangle$ that achieves that minimization by $|0_{\phi,\pi}\rangle$. Again, we drop the labels on the field operators because we assume that $|0_{\phi,\pi}\rangle$ is translationally invariant.

As before, $\int d^4x\,\Gamma(\phi,\pi)$ is the Legendre transform of

$$W(J,K)=\int dt\langle 0_{\phi,\pi}|\hat{H}+J\int d^3x\,\hat{\phi}(\mathbf{x})+K\int d^3x\,\hat{\pi}(\mathbf{x})|0_{\phi,\pi}\rangle\ , \tag{16}$$

$$\int d^4x\,\Gamma(\phi,\pi)=W(J,K)-J\frac{\partial W}{\partial J}-K\frac{\partial W}{\partial K}\ . \tag{17}$$

$|0_{\phi,\pi}\rangle$ is the lowest eigenstate of $\hat{H}_{J,K}$,

$$\hat{H}|0_{\phi,\pi}\rangle+J\int d^3x\,\hat{\phi}(\mathbf{x})|0_{\phi,\pi}\rangle+K\int d^3x\,\hat{\pi}(\mathbf{x})|0_{\phi,\pi}\rangle=W|0_{\phi,\pi}\rangle\ . \tag{18}$$

It is also easy to prove from Eqs. (15)–(18) that

$$\frac{\partial\Gamma}{\partial\phi}=-J,\quad \frac{\partial\Gamma}{\partial\pi}=-K\ . \tag{19}$$

For the Hamiltonian of Eq. (2) it is also true that

$$\Gamma(\phi,\pi)=\tfrac{1}{2}\pi^2+V_{\rm eff}(\phi)\ . \tag{20}$$

Proof:

$$\Gamma(\phi,\pi)=\langle 0_{\phi,\pi}|\hat{\mathscr{H}}|0_{\phi,\pi}\rangle=\langle 0_{\phi,\pi}|\tfrac{1}{2}\hat{\pi}^2+\tfrac{1}{2}(\nabla\hat{\phi})^2+V(\hat{\phi})|0_{\phi,\pi}\rangle$$
$$=\tfrac{1}{2}\pi^2+\langle 0_{\phi,\pi}|\tfrac{1}{2}(\hat{\pi}-\pi)^2+\tfrac{1}{2}(\nabla\hat{\phi})^2+V(\hat{\phi})|0_{\phi,\pi}\rangle\ . \tag{21}$$

Define

$$\hat{\pi}'=\hat{\pi}-\pi,\quad [\hat{\pi}'(\mathbf{x}),\hat{\phi}(\mathbf{y})]=-i\delta^{(3)}(\mathbf{x}-\mathbf{y})\ , \tag{22}$$

then, from the definition of Γ [Eq. (15)],

$$\langle 0_{\phi,\pi}|\tfrac{1}{2}(\hat{\pi}-\pi)^2+\tfrac{1}{2}(\nabla\hat{\phi})^2+V(\hat{\phi})|0_{\phi,\pi}\rangle=\min_{|s\rangle}\{\langle s|\tfrac{1}{2}\hat{\pi}'^2+(\nabla\hat{\phi})^2+V(\hat{\phi})|s\rangle\}\ ,$$
$$\langle s|\hat{\phi}|s\rangle=\phi,\quad \langle s|\hat{\pi}'|s\rangle=0\ . \tag{23}$$

Now the right-hand side of Eq. (23) is determined by the commutation relation, Eq. (22), which is the same as that for $\hat{\pi}$

and $\hat{\phi}$. Thus, if we were to relax the constraint on $\langle s\,|\,\hat{\pi}'\,|\,s\rangle$, the right-hand side would be $V_{\rm eff}(\phi)$, from the definition, Eq. (10). However, we have shown [Eq. (14)] that the constraint on $\langle s\,|\,\hat{\pi}'\,|\,s\rangle$ is redundant, so that

$$\langle 0_{\phi,\pi}\,|\,\tfrac{1}{2}(\hat{\pi}-\pi)^2+\tfrac{1}{2}(\nabla\hat{\phi})^2+V(\hat{\phi})\,|\,0_{\phi,\pi}\rangle=V_{\rm eff}(\phi)$$

and Eq. (20) is proved. Note that the only property of $\hat{H}$ that we used was that it is quadratic in $\hat{\pi}$.

The states $|\,0_{\phi,\pi}\rangle$ satisfy Eq. (5).

Proof:

$$\begin{aligned}\langle 0_{\phi,\pi}\,|\,V'(\hat{\phi})\,|\,0_{\phi,\pi}\rangle&=-i\langle 0_{\phi,\pi}\,|\,[\hat{H},\hat{\pi}(x)]\,|\,0_{\phi,\pi}\rangle\\&=iJ\int d^3y\,\langle 0_{\phi,\pi}\,|\,[\hat{\phi}(\mathbf{y}),\hat{\pi}(\mathbf{x})]\,|\,0_{\phi,\pi}\rangle\ \text{[from Eq. (18)]}\\&=-J=\frac{\partial\Gamma}{\partial\phi}\ \text{[from Eq. (19)]}\\&=V'_{\rm eff}(\phi)\ \text{[from Eq. (20)]}\,,\end{aligned}$$

i.e.,

$$\langle 0_{\phi,\pi}\,|\,V'(\hat{\phi})\,|\,0_{\phi,\pi}\rangle=V'_{\rm eff}(\langle 0_{\phi,\pi}\,|\,\hat{\phi}\,|\,0_{\phi,\pi}\rangle)\ .$$

Thus we have characterized a set of states which satisfy Eq. (5), and hence the LAS equation (1), at some particular moment, and also permit the field to roll. They are those states which minimize the expectation value of the Hamiltonian subject to constraints on the spatially averaged expectation value of the field and its conjugate momentum.

We shall refer to these conditions as the LAS conditions, and the states $|\,0_{\phi,\pi}\rangle$ as LAS states. That $|\,s\,\rangle$ is an LAS state is, of course, sufficient for Eq. (5) to hold but not obviously necessary.

III. EVOLUTION OF THE LAS CONDITIONS

We are using the Heisenberg picture, so that although the states do not evolve in time, $\hat{\phi}$ and $\hat{\pi}$ do. Thus a state that satisfies the LAS condition at one time need not satisfy it at a later time. We shall show that in general it does not.

Suppose that the LAS condition is satisfied at time t, then, from Eq. (18)

$$[\hat{\mathscr{H}}+J(t)\hat{\phi}(t)+K(t)\hat{\pi}(t)]\,|\,\Psi\rangle=E\,|\,\Psi\rangle\ . \tag{24}$$

Denoting operators at time $t+\delta t$ with a prime and operators at time t without, we have

$$\hat{\phi}=\hat{\phi}'-\hat{\pi}'\delta t,\quad \hat{\pi}=\hat{\pi}'+V'(\hat{\phi}')\delta t\ . \tag{25}$$

Substituting Eqs. (25) into Eq. (24) yields

$$\{\hat{\mathscr{H}}+[J\hat{\phi}'+KV'(\hat{\phi}')dt]+(K-Jdt)\hat{\pi}'\}\,|\,\Psi\rangle=E\,|\,\Psi\rangle\ . \tag{26}$$

Equation (26) is compatible with the LAS condition at time $t+\delta t$ only if there exist A and B such that

$$[J\hat{\phi}'+KV'(\hat{\phi}')]\,|\,\Psi\rangle=(A+B\hat{\phi}')\,|\,\Psi\rangle \tag{27}$$

[since $K(t)=-\langle\hat{\pi}(t)\rangle$ and $J(t)=\langle -V'(\hat{\phi})\rangle$, $K(t)-J(t)dt=K(t+\delta t)$].

We note that in fact (27) is satisfied, and the LAS condition preserved, in free-field theory. This is even less interesting than statements about free-field theory usually are, since, as we have remarked, the LAS condition is redundant; the LAS equation is always obeyed. For an interacting field theory, however, Eq. (27) holds only if $|\,\Psi\rangle$ is an eigenstate of $\hat{\phi}'$, which is not possible for an LAS state.

Thus the LAS condition is not preserved, and even though one may prepare a system such that Eq. (5) is true, there is no reason to believe that it remains true. In the Appendix we show that indeed Eq. (5) is not preserved, through an explicit one-loop calculation in the simplest model we know—the quantum-mechanical anharmonic oscillator. [Note that the fact that the LAS condition is not preserved is a necessary, but not sufficient, condition for Eq. (5) to be violated.]

Even though the LAS equation (1) is not true, it does provide a rigorous lower bound on the rollover time, provided the rolling commences in an LAS state. The argument is based on nothing more than conservation of energy. From the LAS equation (1) one may prove that

$$\dot{\phi}_{\rm LAS}(t)=\sqrt{2}[V_{\rm eff}(\phi_0)-V_{\rm eff}(\phi)]^{1/2} \tag{28}$$

(the initial conditions are $\phi=\phi_0,\pi=0$). The true version of Eq. (28) is [Eq. (21)]

$$\dot{\phi}(t)=\sqrt{2}[M(0)-M(t)]^{1/2}\ , \tag{29}$$

$$M(t)=\langle s\,|\,\tfrac{1}{2}[\hat{\pi}(t)-\pi(t)]^2+\tfrac{1}{2}[\nabla\hat{\phi}(t)]^2+V(\hat{\phi}(t))\,|\,s\rangle\ . \tag{30}$$

By assumption, $|\,s\,\rangle$ is an LAS state at time $t=0$ so that [Eq. (20)]

$$M(0)=V_{\rm eff}(\phi_0)\ . \tag{31}$$

Furthermore, from the definition of the effective potential [Eq. (10)],

$$M(t)\geq V_{\rm eff}(\phi(t))\ . \tag{32}$$

Combining Eqs. (28), (29), (31), and (32) yields

$$\dot{\phi}(\phi)\leq\dot{\phi}_{\rm LAS}(\phi) \tag{33}$$

and we see that the LAS equation does indeed give a lower bound on the rollover time. This is all that is required for most cosmological purposes. The physics of this result is not hard to understand. "Kinetic energy" of the rolling spatially averaged field is transferred elsewhere. In part it excites other Fourier modes of the field, a phenomenon more commonly known as particle production. This is usually taken into account by adding a "friction term" to the LAS equation.[11] However, even when there are no other Fourier modes coupled in, as in the case of the anharmonic oscillator in quantum mechanics, kinetic energy is still converted into "heat" or "internal energy" associated with the spatial average itself, because the state ceases to satisfy the minimizing LAS condition.

IV. ON THE ONE-LOOP APPROXIMATION, AND CONVEXITY

The results that we have derived so far all relate to the exact effective potential, and, as such, are somewhat academic. Nobody can calculate an effective potential exactly for an interacting field theory and so a one-loop approximation is universally substituted. Unfortunately, the one-loop approximation is not convex, in violation of a by now well-known theorem.[12] This might not seem so disturbing until one realizes that convexity, the symmetry $\hat{\phi}(x)\to-\hat{\phi}(x)$ (present in many models of inflation) and the LAS equation are together inconsistent with inflation. It is indeed fortunate that the LAS equation is not generally true. We are thus led to modify the question addressed by this paper. Under what circumstances is the one-loop LAS equation approximately true, and the exact LAS equation (i.e., with the exact effective potential on the right-hand side) completely wrong?

For our purposes, the statement that the effective potential is convex is simply that its second derivative is never negative. The one-loop approximation displayed in Fig. 1 clearly violates this condition in the neighborhood of $\phi=0$, and we assert that the exact result looks something like Fig. 2, with the potential exactly flat near the origin. It is easy to understand why this should be from our definition of the effective potential, Eq. (10). Since the symmetry $\hat{\phi}\to-\hat{\phi}$ is spontaneously broken we can choose either of two vacuum states. $|0_+\rangle$ has its wave functional peaked around $\phi=\phi_0$ while $|0_-\rangle$ peaks around $\phi=-\phi_0$. The fact that the symmetry is broken implies that

$$\langle 0_+|\hat{\theta}|0_-\rangle=0 \tag{34}$$

for any observable $\hat{\theta}$. Consider now the state

$$|s\rangle=a|0_+\rangle+b|0_-\rangle,\quad |a|^2+|b|^2=1\ . \tag{35}$$

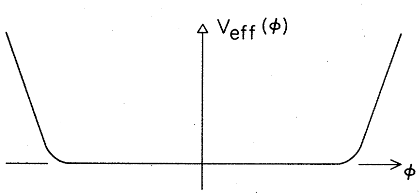

FIG. 2. Exact (convex) effective potential.

Then, from Eq. (34)

$$\langle s|\hat{\phi}|s\rangle=|a|^2\phi_0-|b|^2\phi_0\ ,$$

$$\langle s|\hat{H}|s\rangle=\langle 0_+|\hat{H}|0_+\rangle=\langle 0_-|\hat{H}|0_-\rangle\ .$$

Thus there are states in which $\langle\hat{\phi}\rangle$ can take any value between $-\phi_0$ and ϕ_0 and in which $\langle\hat{H}\rangle$ takes its vacuum value. From its definition, then, the effective potential takes its minimum value everywhere from $-\phi_0$ to ϕ_0.

Note that although the vacuum expectation value of $\hat{\phi}$ is, strictly speaking, somewhat arbitrary, this has no effect on the outcome of experiments. The Hilbert space is the sum of two separate spaces, one built on the vacuum $|0_+\rangle$ and the other on $|0_-\rangle$. The outcome of any experiment is the same in either space, and there are no transitions from one space to the other. If we build our Hilbert space on the vacuum of Eq. (35), each state is a linear combination of corresponding states in the spaces built on $|0_+\rangle$ and $|0_-\rangle$. There are no interference effects arising from this superposition because of Eq. (34), so that the outcome of an experiment is the same as in the spaces built on $|0_+\rangle$ and $|0_-\rangle$.

Where does the one-loop approximation go wrong? In calculating $V_{\rm eff}(\phi)$ to one loop, the bare potential is expanded about ϕ to second order, and the constrained ground state found in that quadratic potential. This state is inevitably a Gaussian centered on ϕ, and cannot be anything like the bimodal state, peaked about ϕ_0 and $-\phi_0$, described above.

Thus the one-loop effective potential $V_{\rm eff}(\phi)$ is (approximately) the minimum expectation value of the Hamiltonian density among all *Gaussian* states centered on ϕ.

Furthermore, our foregoing exact discussion of the LAS equation holds at the one-loop level. In particular, the one-loop LAS equation provides a lower bound on the rollover time provided rolling commences in a Gaussian state that minimizes $\langle\hat{H}\rangle$.

V. THE LAS EQUATION AND COSMOLOGY

We wish to emphasize that we believe that the LAS equation, as an equation for the expectation value $\langle\phi\rangle$ is valid *only* if rolling commences in a Gaussian LAS state, and, correspondingly, is only approximately valid if this condition is approximately met.

For example, by shrinking the width of the LAS Gaussian to zero, the potential whose derivative is the driving force interpolates smoothly between the effective and bare potentials. The latter, with its parameters taking their singular bare values, is, in general, very different from the former. (We do not suggest that a wave functional of zero width is a physically plausible state—it has infinite energy density—we are merely pointing out that a different initial state leads to a different driving force, and consequently some physical principle has to be invoked to determine it.) Furthermore, in contradistinction to suggestions by other authors, slowness of roll (i.e., $\pi=0$) does not, in our analysis, seem to be related to the validity of the LAS equation; what matters is the shape of the wave functional. As we have seen, there are LAS states with arbitrary values of momentum, and one may imagine (one may certainly do this in quantum mechanics) constructing states with $\pi=0$, but with a wave functional very different from that for the corresponding LAS state, which will consequently strongly violate Eq. (5).

Other authors have not, to the best of our knowledge, found the particular state of the system to play such a crucial role, and appear to ascribe to the LAS equation an approximate validity more general than that found here; it is therefore important to understand how the various approaches differ. A detailed examination of this question is beyond the scope of this paper, but we shall nevertheless make a few tentative comments, which, in all probability, will fail to do justice to the work of others.

Previous analyses have generally fallen into two classes: (i) those based on Euclidean path integrals and (ii) those based on the classical equation of motion derived from variation of the full effective action. We have used neither approach and this makes comparison difficult, although all approaches are presumably equivalent (we do not believe our results are incompatible with these other calculations, for the reasons given below).

The canonical formalism that we have used is, of course, not directly equivalent to the calculus of Euclidean functional integrals. The initial state should, rather, by evolved using the *Minkowski*-space functional integral which has, as its time translation operator, e^{iHt}. The Euclidean integral, on the other hand, is equivalent to the use of e^{-Ht} as the evolution operator. While the Euclidean path integral is an excellent tool for the calculation of *vacuum* expectation values (since, for large times, it projects out the ground state), we feel that it should be used cautiously in the determination of the properties of other states. As we have remarked, the Universe is *not* in its vacuum state.

Arguments based on the full effective action go roughly as follows. The exact equation of motion for ϕ is,

$$\frac{\delta\Gamma}{\delta\phi}=0\,, \tag{36}$$

where

$$\Gamma=V_{\rm eff}(\phi)+\text{derivative terms} \tag{37}$$

is the full effective action. It is then argued that if ϕ changes only slowly (and field renormalization is chosen appropriately), the leading derivative term will be the classical $(\partial_\mu\phi)^2/2$. Equation (36) is then the LAS equation.

However, the derivative terms in Eq. (37) are not just higher powers of $(\partial_\mu\phi)^2$, but include all higher derivatives of ϕ as well. Thus to argue that the classical term is dominant one needs to *assume* that *all* derivatives of ϕ are small, which is equivalent to saying that ϕ moves slowly not just initially, but for all time. That, of course, is what we are trying to show. To summarize, if ϕ moves slowly for all time, then it satisfies the LAS equation, but it need not move slowly for all time even if it starts off doing so.

Let us make two further remarks concerning the effective-action approach. Since Γ contains all higher derivatives of ϕ, we must specify as initial conditions for Eq. (36) all derivatives of ϕ. As this is equivalent to specifying the complete trajectory for ϕ, Eq. (36) contains (essentially) no information. We conjecture that specifying these derivatives is equivalent to specifying the quantum state.

Even if the higher-derivative terms have small coefficients, they may profoundly alter the trajectory, since higher-derivative theories are generally afflicted with exponentially growing solutions.[13] A conventional theory, such as one defined by the Lagrangian

$$L=(\partial_t x)^2/2-V(x) \tag{38}$$

is protected from such diseases because the conserved energy

$$H=(\partial_t x)^2/2+V(x) \tag{39}$$

has a positive-definite kinetic piece. Thus a suitable choice of the potential confines x to a finite interval. For a higher-derivative theory this argument fails, because the kinetic piece is not in general positive definite. Consider the system defined by

$$L=\mu(\partial_t{}^2x)^2+(\partial_t x)^2/2-V(x)\,. \tag{40}$$

The Noether procedure yields a conserved energy

$$H=\mu(\partial_t{}^2x)^2+(\partial_t x)^2/2-2\mu\partial_t x\partial_t{}^3x+V(x)\,. \tag{41}$$

No matter what form the potential takes, all values of x are accessible by making the kinetic energy arbitrarily negative. Indeed, for quadratic $V(x)$ and small positive μ, there exist solutions that grow like $e^{t/\sqrt{2\mu}}$.

Having argued that the LAS equation is valid only if the universe starts rolling in an LAS state, we now turn to the question of whether this state, or something close to it, is likely to arise from a hot primordial universe. Our thoughts on this matter have benefited from a remark by Brandenberger.[14]

The standard model starts with matter in thermodynamic equilibrium at a temperature $T>T_{\rm GUT}$, and as long as it remains in equilibrium its equation of state is tolerably approximated by that for a perfect gas: $p=\rho/3$. Under these circumstances the radius of the universe grows only as $\sqrt{t}$, and there is no inflation.

For that (which requires an equation of state $p=-\rho$) and for the irreversible processes which generate entropy (the need for inflation can be formulated as a need for entropy generation[1]), we need the matter to be out of thermodynamic equilibrium. This clearly happens in all models of inflation, where a substantial energy density is associated with the spatial average of the Higgs field, but all other modes are in their ground states—a distribution of energy that is certainly not that of equilibrium thermodynamics.

Departures from equilibrium occur when the time for transitions between states is not small compared to the time it takes for the equilibrium temperature to change significantly. For a perfect gas the equilibrium temperature scales with the radius of the Universe such that $RT=$const, so that the time for a significant change is $\tau=R/\dot{R}$. It is the fact that the spatial average takes much longer than this time to roll to its equilibrium value that leads to departures from equilibrium and to inflation. Thus the rolling does not take place in thermal equilibrium, and it is incorrect to use the formalism of finite-temperature field theory to describe it.

Having made these preliminary remarks let us attempt a qualitative, simple-minded description of what happens as the universe cools. Any discussion of the details of departure from equilibrium during a phase transition in curved space at such early times is a hazardous enterprise, and should be regarded with corresponding skepticism.[8] We assume that the universe starts in equilibrium at a temperature $T > T_{\rm GUT}$. At around some critical temperature T_c (on the order of $T_{\rm GUT}$), the matter goes out of equilibrium and inflation begins. By the assumption of inflation, and so departure from equilibrium, the long-wavelength modes are changing very slowly (in particular we are assuming that the spatial average is rolling slowly). Short-wavelength degrees of freedom are rapidly red-shifted to their ground state by the inflation, but longer-wavelength modes ($>M_{\rm GUT}{}^{-1}$) are not so affected (in particular, the spatial average is not red-shifted at all). These long-wavelength modes thus have an energy excess, over and above the minimum for a given value of ϕ, on the order of T_c throughout the inflationary era. This, of course, is just the statement that the universe enters the inflationary era in something other than an LAS state.

We may arrive at the same conclusion by another route. The minimal Gaussian that is the required LAS state can be expressed as a superposition of energy eigenstates. To obtain the desired state, the coefficients of the individual eigenstates must have not only the correct magnitude but also the correct relative phases; change the phases and the state looks completely different (for example, temporal evolution is simply a change in the relative phases). A thermal state, by contrast, specifies the magnitudes of the weights for the different eigenstates, but gives no information on the relative phases, which are completely random. It is therefore very hard to imagine how a hot primordial universe can yield the necessary phase information for the construction, even approximately, of an LAS state (Figs. 3 and 4). If N eigenstates contribute significantly, the probability of getting the phases roughly right is on the order of $(2\pi)^{-N}$.

The foregoing arguments lead us to believe that it is unlikely that the LAS equation correctly describes the evolution of the expectation value $\langle\phi\rangle$, even approximately. However, this does not mean that said equation is not relevant for a physical description of the roll. For example, it may be possible to argue that while the universe is never in an LAS state, it may commence rolling in a superposition of such states (a superposition of LAS states is not, in general, an LAS state). Since the time evolution operator acts linearly on the Hilbert space, each of these LAS states would evolve independent of the others, essentially according to the LAS equation. Thus although $\langle\hat{\phi}\rangle$ would not obey the LAS equation, the roll time would be derived from an average over an ensemble of LAS trajectories. (This argument clearly owes much to the work of Guth and Pi.[7]) However, the LAS states do not form a complete set, and so this possibility requires a physical argument to justify it.

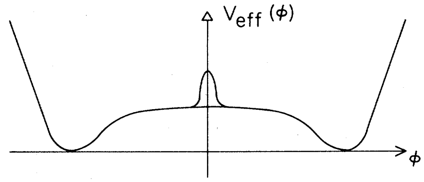

FIG. 3. Initial state required for the one-loop LAS equation to be valid.

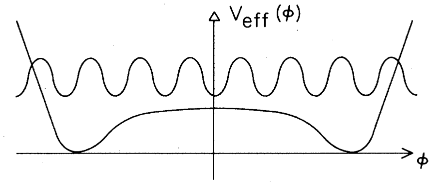

FIG. 4. Typical initial state from a hot early universe.

In this paper we have drawn attention to the importance, as we see it, of the quantum state of the universe at the beginning of the inflationary era.[15] We hope that future work will illuminate its properties and clarify the role of quantum effects in the early universe.

ACKNOWLEDGMENTS

The authors wish to acknowledge useful conversations with R. Brandenberger, B. A. Ovrut, H. Pagels, and S.-Y. Pi. Work supported in part under the Department of Energy Contract Grant No. AC02-81ER40033B.000.

APPENDIX

To demonstrate that Eq. (5) is not preserved under evolution we consider the anharmonic oscillator in ordinary quantum mechanics ($m=1$)

$$\hat{H}=\tfrac{1}{2}\hat{P}^2+V(\hat{X})=\tfrac{1}{2}\hat{P}^2+\tfrac{1}{2}\omega^2\hat{X}^2+\lambda\hat{X}^4 \,. \tag{A1}$$

We first compute $|0_{XP}^{(1)}\rangle$, the one-loop approximation to the state $|0_{XP}\rangle$ which minimizes $\langle\psi|\hat{H}|\psi\rangle$ subject to the constraints

$$\langle\psi|\psi\rangle=1,\;\; \langle\psi|\hat{X}|\psi\rangle=X\,,$$

$$\langle\psi|\hat{P}|\psi\rangle=P\,.$$

Clearly the state $|0_{XP}\rangle$ satisfies

$$\hat{H}_{JK}|0_{KP}\rangle\equiv(\hat{H}+J\hat{X}+K\hat{P})|0_{XP}\rangle=E|0_{XP}\rangle\,, \tag{A2}$$

where J and K are to be determined by applying the constraints. From the discussion in Sec. III we know that $\hat{H}^{(1)}$, the appropriate Hamiltonian for the one-loop variational problem, is $\hat{H}_{JK}$ expanded to second order around X and P with J and K chosen to eliminate the linear terms. Up to irrelevant c-number constants, we have

$$\hat{H}^{(1)}=\tfrac{1}{2}(\hat{P}-P)^2+\tfrac{1}{2}\varpi^2(\hat{X}-X)^2\,,$$

where

$$\varpi^2=\omega^2+12\lambda X^2\,. \tag{A3}$$

Of course the ground state of this Hamiltonian is

$$\langle x \mid 0_{XP}^{(1)} \rangle = A_0 e^{-(\bar\omega/2\hbar)(x-X)^2 + iP_x} , \quad \text{(A4)}$$

where

$$A_0 = \left[\frac{\bar\omega}{\pi\hbar} \right]^{1/4} .$$

The reader may very simply verify that, to first order in $\hbar$,

$$V_{\text{eff}}^{(1)}(X) = \langle 0_{XP}^{(1)} \mid \hat H \mid 0_{XP}^{(1)} \rangle ,$$

$$V_{\text{eff}}^{(1)\prime}(X) = \langle 0_{XP}^{(1)} \mid V'(\hat X) \mid 0_{XP}^{(1)} \rangle .$$

Here the one-loop effective potential is easily calculated, as, for example, in Coleman's notes,[10] to be

$$V_{\text{eff}}^{(1)}(X) = \tfrac{1}{2}\omega^2 X^2 + \lambda X^4 + \frac{\hbar}{2}(\omega^2 + 12\lambda X^2)^{1/2} . \quad \text{(A5)}$$

Suppose now that the system has been prepared in the state $|0_{XP}\rangle$ at $t=0$, and denote (in the Heisenberg picture)

$$N(t) \equiv \langle 0_{XP} \mid V'(\hat X(t)) \mid 0_{XP} \rangle . \quad \text{(A6)}$$

Then after an infinitesimal time δt, we have, to first order in δt,

$$\delta N = \frac{-i}{\hbar} \delta t \langle 0_{XP} \mid [V'(\hat X(0)), \hat H] \mid 0_{XP} \rangle .$$

Making use of Eqs. (A1) and (A2) we write this as

$$\delta N = \frac{-i}{\hbar} \delta t K [\omega^2 + 12\lambda \langle 0_{XP} \mid \hat X(0)^2 \mid 0_{XP} \rangle] . \quad \text{(A7)}$$

Here K is obtained from the equation of motion

$$\hat P(0) = \left. \frac{d\hat X}{dt} \right|_{t=0} = \frac{-i}{\hbar} [\hat X(0), \hat H]$$

by taking the matrix element between $|0_{XP}\rangle$ and using (A2) again to get

$$K = P . \quad \text{(A8)}$$

Substituting the one-loop approximation $|0_{XP}^{(1)}\rangle$ the matrix element in (A7) is a simple integration, and with (A8) we obtain

$$\delta N = P\delta t \left[\omega^2 + 12\lambda X^2 + \hbar \frac{6\lambda}{(\omega^2 + 12\lambda X^2)^{1/2}} \right] .$$

If the LAS is preserved under evolution to $O(\hbar)$ and $O(\delta t)$, we must have

$$\delta N = \delta V'_{\text{eff}} = V''_{\text{eff}}(X) P \delta t .$$

But

$$\delta V'_{\text{eff}} = P\delta t \left[\omega^2 + 12\lambda X^2 + \hbar \frac{6\lambda}{(\omega^2 + 12\lambda X^2)^{1/2}} - \hbar \frac{72\lambda^2 X^2}{(\omega^2 + 12\lambda X^2)^{3/2}} \right]$$

differing in the last term.

[1]A. Guth, Phys. Rev. D **23**, 347 (1981).

[2]For the standard model see S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972); J. Peebles, *Physical Cosmology* (Princeton University Press, Princeton, New Jersey, 1971); *The Large Scale Structure of the Universe* (Princeton University Press, Princeton, New Jersey, 1980).

[3]For a review see R. Brandenberger, Rev. Mod. Phys. **57**, 1 (1985); P. Steinhardt and M. Turner, Fermilab Report No. 84/19-A, 1984 (unpublished).

[4]B. Ovrut and P. Steinhardt, University of Pennsylvania Report No. UPR-0266T, 1984 (unpublished); Q. Shafi and A. Vilenkin, Phys. Rev. Lett. **52**, 691 (1984); S.-Y. Pi, *ibid.* **52**, 1725 (1984); J. Ellis, D. Nanopoulos, K. Olive, and K. Tamvakis, Nucl. Phys. **B221**, 524 (1983).

[5]A. Albrecht and P. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982); A. Linde, Phys. Lett. **108B**, 389 (1982).

[6]S. Coleman and E. Weinberg, Phys. Rev. D **7**, 1888 (1973); G. Denardo and E. Spallucci, Nuovo Cimento **83A**, 35 (1984).

[7]A. Guth and S.-Y. Pi, MIT Report No. CTP 1181, 1984 (unpublished) and references therein to unpublished work of J. Bardeen; W. Bardeen and C. Hill; R. Brandenberger, C. Bender, F. Cooper, B. Freedman, and R. Haymaker, Los Alamos report, 1984 (unpublished).

[8]G. Mazenko, W. Unruh and R. Wald, Phys. Rev. D **31**, 273 (1985).

[9]J. Kogut and L. Susskind, Phys. Rev. D **11**, 395 (1975).

[10]E. Abers and B. W. Lee, Phys. Rep. **9C**, 1 (1973); S. Coleman, in *Secret Symmetries,* edited by A. Zichichi (Academic, New York, 1975).

[11]L. Abbott, E. Farhi, and M. Wise, Phys. Lett. **117B**, 29 (1982).

[12]J. Iliopoulos, C. Itzykson, and A. Martin, Rev. Mod. Phys. **47**, 165 (1975); K. Symanzik, Commun. Math. Phys. **16**, 48 (1970); G. Borner and E. Seiler, Max Planck Institute report, 1983 (unpublished).

[13]A. Pais and G. Uhlenbeck, Phys. Rev. **79**, 145 (1950).

[14]R. Brandenberger, private communication.

[15]S. Hawking and J. Luttrell, Phys. Lett. **143B**, 83 (1984); A. Vilenkin, Phys. Rev. D **27**, 2848 (1983).