Probability of R^2 inflation

Don N. Page
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802
and Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125
(Received 3 June 1987)

The Gibbons-Hawking-Stewart canonical measure is applied to classical Friedmann-Robertson-Walker cosmologies with an $R+\epsilon R^2$ Lagrangian. Both the inflationary solutions and the noninflationary solutions have infinite measure, and the ratio is ambiguous. All but a finite measure of the $k=-1$ solutions have an arbitrarily long period of near spatial flatness, but for $k=+1$ there is also an infinite measure for solutions which have a small maximum radius, unlike the case with Einstein gravity coupled to a massive scalar field. For $k=-1$ there is a finite positive measure for complete nonsingular solutions with a nonzero minimum radius; all other $k=-1$ solutions expand from zero to infinite radius. For $k=+1$ all solutions but a set of measure zero expand from zero radius and eventually recollapse to zero radius. However, there is also an apparently fractal set of discrete (zero measure) $k=+1$ solutions which have no singularities but rather expand and recontract perpetually, with an arbitrary number of oscillations of the scalar curvature between each successive bounce.

I. INTRODUCTION

One particularly attractive form of the inflationary universe scenario[1] is Linde's chaotic inflation,[2,3] which does not require thermal equilibrium or a finely tuned effective potential in the early Universe. Instead, it relies on having some scalar field take values far from the minimum of its potential, say by a quantum fluctuation. This mechanism is effective for quite a wide class of possible potentials, but one might like to get inflation to work without assuming any scalar field at all. Indeed, this can be accomplished in higher-derivative gravitational theories, in which the scalar curvature acts rather like a massive scalar field.[4,5] An analysis of one of these theories, with an $R+\epsilon R^2$ Lagrangian, has shown that generic initial conditions within the Friedmann-Robertson-Walker (FRW) class can lead to inflation.[6]

In this paper the classical solutions of this model are examined in more detail, and the Gibbons-Hawking-Stewart (GHS) canonical measure[7] is applied to them. It is found that there are various classes of solutions, some countably discrete, some uncountably discrete but still of zero measure, and others of nonzero finite measure, and finally yet other solutions making up classes of infinite measure. Both inflationary and noninflationary solutions have infinite measure, so the ratio, which might be said to give the probability of inflation, is ambiguous, analogous to the case with a massive scalar field.[8]

II. FRW MODELS WITH AN $R+\epsilon R^2$ LAGRANGIAN

A homogeneous, isotropic (FRW) metric for the universe may be written as

$$ds_4{}^2=-N^2(t)dt^2+a^2(t)\tilde{g}_{ij}dx^i dx^j \ , \tag{2.1}$$

where $\tilde{g}_{ij}$ is a time-independent 3-metric with constant curvature

$${}^3\tilde{R}_{ijkl}=k(\tilde{g}_{ik}\tilde{g}_{jl}-\tilde{g}_{il}\tilde{g}_{jk}) \ , \tag{2.2}$$

normalized so that $k=-1$, 0, or $+1$. The dynamics will be assumed to be given purely by gravity, but with the Einstein-Hilbert action augmented by quadratic curvature terms so that the total action is, up to surface terms,

$$I=\frac{1}{16\pi G}\int d^4x\sqrt{-g}\,(R+\alpha R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}+\beta R_{\mu\nu}R^{\mu\nu}+\gamma R^2) \ . \tag{2.3}$$

One combination of the quadratic terms is the Euler density, a pure divergence, and another combination is the square of the Weyl tensor, which is zero in FRW metrics, so for classical solutions with the metric (2.1), the Lagrangian density is effectively $R+\epsilon R^2$. For microscopic stability, take $\epsilon>0$ (Ref. 9).

The second derivatives of the metric in the R^2 term cannot be canceled against a surface term by an integration by parts as they can for R, so extremizing the action leads to fourth-order equations of motion for the metric. In order to write the theory in canonical form with only first derivatives in the action (and hence leading to second-order equations of motion), it is convenient to regard $a(t)$ and

$$Q(t)\equiv a(1+2\epsilon R) \tag{2.4}$$

as independent dynamical degrees of freedom,[5,10] so that the action takes the form

$$I=\frac{3}{8\pi G}\int\sqrt{\tilde{g}}\,d^3x\int\frac{N}{a}dt\left[-\left(\frac{a}{N}\frac{da}{dt}\right)\left(\frac{a}{N}\frac{dQ}{dt}\right)+kaQ-\frac{1}{24\epsilon}a^2(Q-a)^2\right] . \tag{2.5}$$

Variation with respect to the lapse function $N(t)$ gives the Hamiltonian constraint equation

$$\left[\frac{a}{N}\frac{da}{dt}\right]\left[\frac{a}{N}\frac{dQ}{dt}\right]+kaQ-\frac{1}{24\epsilon}a^2(Q-a)^2=0\ . \qquad (2.6)$$

Once this is imposed, one may reparametrize the time coordinate t to make $N(t)=1$, which will be done henceforth. Then, for our purely classical analysis, we may also rescale the action to make the constant coefficient in front of the time integral in (2.5) equal to unity. These steps simplify the action to

$$I=\int dt\left[-a\dot{a}\dot{Q}+kQ-\frac{1}{24\epsilon}a(Q-a)^2\right]$$
$$=\int d\eta\left[-a'Q'+kaQ-\frac{1}{24\epsilon}a^2(Q-a)^2\right]\ , \qquad (2.7)$$

where the overdot represents d/dt with respect to the cosmological proper time t (since $N=1$ now) and the prime represents $d/d\eta$ with respect to the conformal time $\eta=\int dt/a$. The constraint (2.6) and the equations of motion for a and Q then become

$$a^2\dot{a}\dot{Q}\equiv a'Q'=\frac{1}{24\epsilon}a^2(Q-a)^2-kaQ\equiv f(a,Q)\ , \qquad (2.8)$$

$$\tfrac{1}{2}(a^2)^{\cdot\cdot}\equiv a\ddot{a}+\dot{a}^2=\frac{1}{12\epsilon}a(Q-a)-k\ , \qquad (2.9)$$

$$\ddot{Q}=\frac{1}{24\epsilon a}(Q-a)(Q-3a)\ . \qquad (2.10)$$

The definition (2.4) makes (2.9) become the standard FRW formula for the scalar curvature:

$$R=6\left[\frac{\ddot{a}}{a}+\frac{\dot{a}^2}{a^2}+\frac{k}{a^2}\right]=6\left[\dot{H}+2H^2+\frac{k}{a^2}\right]\ , \qquad (2.11)$$

where

$$H\equiv\frac{\dot{a}}{a}\equiv\frac{a'}{a^2} \qquad (2.12)$$

is the Hubble expansion rate. The constraint may then be rewritten as

$$0=2\epsilon H\dot{R}-\frac{1}{6}\epsilon R^2+\left[H^2+\frac{k}{a^2}\right](1+2\epsilon R)$$
$$=12\epsilon\left[H\ddot{H}-\tfrac{1}{2}\dot{H}^2+3H^2\dot{H}-\frac{k}{a^2}H^2+\frac{1}{2}\frac{k^2}{a^4}\right]$$
$$+H^2+\frac{k}{a^2}\ , \qquad (2.13)$$

which yields the single third-order equation for the scale factor a:

$$2a^2\dot{a}\dddot{a}-a^2\ddot{a}^2+2a\dot{a}^2\ddot{a}-3\dot{a}^4-2k\dot{a}^2+k^2$$
$$+\frac{1}{6\epsilon}a^2(\dot{a}^2+k)=0\ . \qquad (2.14)$$

This equation and its time derivative give the full fourth-order equations (except when $\dot{a}=0$, so $a^2=-6\epsilon k$ is a spurious solution). Equation (2.10) for $\ddot{Q}$ is not independent but may be obtained from (2.9) and the time derivative of the constraint. One particularly simple combination of (2.9) or (2.10) and the constraint is[6]

$$\ddot{R}+3H\dot{R}+\frac{1}{6\epsilon}R=0\ , \qquad (2.15)$$

which shows that R evolves according to the same equation as a scalar field of mass $(6\epsilon)^{-1/2}$ in the same metric, but the metric itself does not evolve quite as it would with a massive scalar field present instead of the R^2 term in the action.

Unlike the vacuum Einstein equations, which are scale free, the classical $R+\epsilon R^2$ equations contain a single dimensional coupling constant ϵ, of dimension length squared, which sets the scale. It is convenient to measure all lengths in units of $L\equiv(24\epsilon)^{1/2}$, in which case the equations all become dimensionless, with ϵ dropping out and no arbitrary dimensionless parameter appearing. (This would not be the case in the quantum theory, which would contain the dimensionless parameter $\epsilon/\hbar G$.) That is, one may define the dimensionless variables $t_*\equiv t/L$, $a_*\equiv a/L$, $R_*\equiv RL^2$, $Q_*\equiv Q/L$, $I_*\equiv I/L^2$, $\eta_*\equiv\eta$, $f_*\equiv f/L^2$, and $H_*\equiv HL$. Then the equations for the quantities with an asterisk are the same as for the corresponding quantities without an asterisk but with 24ϵ set equal to unity. For example, Eq. (2.7) becomes

$$I_*=\int d\eta_*[-a_*'Q_*'+ka_*Q_*-a_*^2(Q_*-a_*)^2]$$
$$\equiv\int d\eta[-a_*'Q_*'-f_*(a_*,Q_*)]\ , \qquad (2.16)$$

giving the Hamiltonian constraint and equation of motion as

$$a_*'Q_*'=f_*(a_*,Q_*)\equiv a_*^2(Q_*-a_*)^2-ka_*Q_*\ , \qquad (2.17)$$

$$a_*''=\partial f_*/\partial Q_*\ , \qquad (2.18)$$

$$Q_*''=\partial f_*/\partial a_*\ , \qquad (2.19)$$

where $F'\equiv dF/d\eta\equiv a\,dF/dt\equiv dF/d\eta_*\equiv a_*dF/dt_*$ for any time-dependent quantity F.

One method of analyzing these equations is to convert them into coupled first-order equations. Since Eq. (2.15) shows that R_* behaves as a scalar field of mass 2 in units of L, one may by analogy with the massive scalar field case[11] define

$$x\equiv\tfrac{1}{12}R_*\equiv2\epsilon R\ , \qquad (2.20)$$

$$y\equiv\frac{1}{24}\frac{dR_*}{dt_*}\equiv(24\epsilon^3)^{1/2}\dot{R}\ , \qquad (2.21)$$

$$z\equiv\tfrac{1}{2}H_*\equiv(6\epsilon)^{1/2}\frac{\dot{a}}{a}\ . \qquad (2.22)$$

Then one gets the coupled system

$$\frac{1}{2}\frac{dx}{dt_*}=y\ , \qquad (2.23)$$

$$\frac{1}{2}\frac{dy}{dt_*}=-x-3yz\ , \tag{2.24}$$

$$\frac{1}{2}\frac{dz}{dt_*}=\frac{2x+x^2+4yz}{4(1+x)}-z^2\ , \tag{2.25}$$

which has, as a first integral, the constraint

$$\frac{k}{4a_*^{\,2}}\equiv\frac{6\epsilon k}{a^2}=\frac{x^2-4yz}{4(1+x)}-z^2\ . \tag{2.26}$$

The solutions have three parameters (e.g., x, y, and z at $t_*=0$), but one combination of these three merely sets the origin of the time coordinate, so there is actually only a two-parameter set of physically distinct solutions, represented by the two-parameter congruence of trajectories in the (x,y,z) space which solve Eqs. (2.23)–(2.25).

As a consequence of Eq. (2.15), Eqs. (2.23) and (2.24) have precisely the same form as in the massive scalar-field case, but in the latter the right-hand sides of the analogues of Eqs. (2.25) and (2.26) are $x^2-2y^2-z^2$ and $x^2+y^2-z^2$, respectively, which are considerably simpler. For example, in the scalar-field case, the $k=0$ solutions lie on a cone, whereas here they lie on a more complicated surface.

Another method of analyzing the dynamical equations that is easier to grasp pictorially for a relativist such as the author, who finds it hard to visualize in more dimensions than has the surface of his retina, is to use the fact that Eqs. (2.17)–(2.19) give trajectories in the (a_*,Q_*) space which are timelike geodesics in the auxiliary (1 + 1)-dimensional metric

$$ds^2=-f_*(a_*,Q_*)da_*dQ_*$$
$$\equiv[ka_*Q_*-a_*^{\,2}(Q_*-a_*)^2]da_*dQ_*\ . \tag{2.27}$$

If τ is the proper time along the geodesic in this metric, the constraint Eq. (2.17) becomes

$$\frac{d\eta}{d\tau}\equiv\frac{1}{a}\frac{dt}{d\tau}=\pm f_*^{\,-1}\ , \tag{2.28}$$

which gives the variation in the conformal time η or in the cosmological time t along the trajectory.

The trajectories may also be interpreted as those of a particle of variable mass squared $f_*(a_*,Q_*)$ moving in the flat metric

$$d\bar{s}^2=-da_*dQ_*=-dT^2+dX^2\ , \tag{2.29}$$

where the new timelike and spacelike coordinates are

$$T\equiv\tfrac{1}{2}Q_*+\tfrac{1}{2}a_*,\quad X\equiv\tfrac{1}{2}Q_*-\tfrac{1}{2}a_*\ , \tag{2.30}$$

respectively (the same as $\frac{1}{2}y$ and $\frac{1}{2}x$ of Ref. 5), with $T>X$. In terms of these coordinates, the mass squared is

$$f_*=4X^2(T-X)^2-k(T^2-X^2)\ . \tag{2.31}$$

In either the curved metric (2.27) with a constant mass or in the flat metric (2.29) with a variable mass, eliminating the proper time parameter along the trajectory leaves one with the single second-order equation for $Q_*(a_*)$:

$$f_*\frac{d^2Q_*}{da_*^{\,2}}=\frac{\partial f_*}{\partial a_*}\frac{dQ_*}{da_*}-\frac{\partial f_*}{\partial Q_*}\left[\frac{dQ_*}{da_*}\right]^2\ . \tag{2.32}$$

The metric (2.27), the mass squared (2.31), and Eq. (2.32) all become degenerate on the curves $f_*=0$, but the trajectories generically pass through these curves in a nonsingular manner, with dQ_*/da_* crossing either 0 $(dX/dT=-1)$ or ∞ $(dX/dT=+1)$ and with no singularity in the physical metric (2.1), except at $a_*=0$, which is a physical singularity. A set of measure zero of trajectories have $\dot{a}_*=\dot{Q}_*=0$ at $f_*=0$ and so simply turn around and reverse their motion at that moment of time symmetry.

Whitt[4] has given yet another way of analyzing the dynamics of $R+\epsilon R^2$ gravity, namely, by examining the conformally transformed metric

$$d\hat{s}_4^{\,2}=(1+2\epsilon R)ds_4^{\,2}\ . \tag{2.33}$$

In this metric one gets the equations of Einstein gravity minimally coupled to the scalar curvature interpreted as a scalar field with a somewhat unusual stress-energy tensor. One may put the kinetic-energy term of this tensor into canonical form by choosing the scalar field to be[12]

$$\phi\equiv\tfrac{1}{2}\ln|1+2\epsilon R|=\tfrac{1}{2}\ln|1+x|=\tfrac{1}{2}\ln|Q_*/a_*|\ . \tag{2.34}$$

Then if one writes the dimensionless scale factor of the conformal metric (2.33) as

$$\hat{a}_*\equiv e^{\beta}=e^{\phi}a_*\ , \tag{2.35}$$

so

$$a_*=e^{\beta-\phi},\quad Q_*=\pm e^{\beta+\phi}\ , \tag{2.36}$$

the action (2.16) becomes

$$I_*=\pm\int d\hat{t}_*e^{3\beta}\left[-\left[\frac{d\beta}{d\hat{t}_*}\right]^2+\left[\frac{d\phi}{d\hat{t}_*}\right]^2+ke^{-2\beta}-2V(\phi)\right]\ , \tag{2.37}$$

$$V(\phi)=\pm\tfrac{1}{2}(1\mp e^{-2\phi})^2\ , \tag{2.38}$$

where $d\hat{t}_*=e^{\phi}dt_*$ and the upper (lower) sign is for Q_* positive (negative). Thus, the constraint and dynamical equations become

$$\hat{H}_*^{\,2}\equiv\left[\frac{d\beta}{d\hat{t}_*}\right]^2=\left[\frac{d\phi}{d\hat{t}_*}\right]^2-ke^{-2\beta}+2V\ , \tag{2.39}$$

$$\frac{d^2\phi}{d\hat{t}_*^{\,2}}+3\frac{d\beta}{d\hat{t}_*}\frac{d\phi}{d\hat{t}_*}+\frac{dV}{d\phi}=0\ , \tag{2.40}$$

plus an equation for $d^2\beta/d\hat{t}_*^{\,2}$ that may be obtained from Eqs. (2.39) and (2.40) by differentiation and algebra. The solutions are geodesics of the auxiliary metric (2.27), which may be rewritten as

$$ds^2=(2e^{6\beta}V-ke^{4\beta})(-d\beta^2+d\phi^2)\ . \tag{2.41}$$

Alternatively, in the two separate regions $\pm Q_*>0$, the solutions are trajectories in the flat metric

$$d\tilde{s}^2=-da_*dQ_*=\pm e^{2\beta}(-d\beta^2+d\phi^2) \tag{2.42}$$

of a particle of variable mass squared

$$f_*=\pm e^{4\beta}(2V-ke^{-2\beta})=e^{4\beta}(1\mp e^{-2\phi})^2\mp ke^{2\beta} \tag{2.43}$$

or trajectories in the conformally related flat metric

$$-d(\ln a_*)d(\ln|Q_*|)=-d\beta^2+d\phi^2 \tag{2.44}$$

of a particle of mass squared

$$a_*Q_*f_*=e^{6\beta}(2V-ke^{-2\beta})$$
$$=\pm e^{6\beta}(1\mp e^{-2\phi})^2-ke^{4\beta}\ . \tag{2.45}$$

This approach gives the closest analogy with the massive scalar field equations, except that here the potential $V(\phi)$ has an unusual form corresponding to a cosmological constant when $\phi \gg 1$ and to an exponential potential when $\phi \ll -1$. One must bear in mind that singularities in the metric (2.33) arising from β and ϕ going to $-\infty$ with a finite difference between them merely correspond to $1+2\epsilon R$ passing through zero and, hence, no singularities in the physical metric (2.1).

III. THE CANONICAL MEASURE FOR THE FRW $R+\epsilon R^2$ MODELS

For a $2n$-dimensional phase space Γ_n with one Hamiltonian constraint (e.g., the models above with $n=2$), the GHS canonical measure[7] on an initial data surface Γ_{n-1} of dimension $2n-2$ is (up to sign) the $(n-1)$ power of the symplectic 2-form $\omega=dP_i\wedge dQ^i$ pulled back to Γ_{n-1}. Here, using the action (2.16) in which the canonical momenta are

$$P_{a_*}=-Q_*'\equiv -dQ_*/d\eta,\quad P_{Q_*}=-a_*'\equiv -da_*/d\eta\ , \tag{3.1}$$

the measure is simply the first power of

$$\begin{aligned}\omega&=dP_{a_*}\wedge da_*+dP_{Q_*}\wedge dQ_*=da_*\wedge dQ_*'+dQ_*\wedge da_*'\\&=da_*\wedge d[2a_*^2(1+x)z+2a_*^2y]+d[(1+x)a_*]\wedge d(2a_*^2z)\\&=2a_*^2da_*\wedge[-zdx+dy+2(1+x)dz]+2a_*^3dx\wedge dz\ .\end{aligned} \tag{3.2}$$

One must choose a two-dimensional initial data surface on the three-dimensional constraint surface [e.g., a two-dimensional surface in the x,y,z space for $k=\pm1$, with the constraint (2.26) determining a_*] and restrict ω to it to get the measure for the solutions crossing that surface.

For example, on a surface of fixed $a_*=\text{const}>0$, the measure (3.2) is simply proportional to $dx\wedge dz$ or $dR\wedge dH$ in terms of the initial data x and z or R and H, with the constraint (2.26) determining y or $\dot{R}$. Since the constraint may be solved for y at all values of x and z (except that $z=0$ gives $y=\infty$ unless $a_*^2x^2=k+kx$), their range is unrestricted and leads to an infinite total measure. Alternatively, if the data at $a_*=\text{const}$ are expressed in terms of Q_* and dQ_*/da_* of the trajectory there, using the constraint in the form (2.17) and inserting it into (3.2) gives the measure

$$\omega=f_*^{1/2}dQ_*\wedge d[(dQ_*/da_*)^{-1/2}]\ , \tag{3.3}$$

where at each value of Q_*, dQ_*/da_* is restricted to have the same sign as $f_*(a_*Q_*)$. Again, the measure is seen to diverge both as Q_* is taken to $\pm\infty$ and as dQ_*/da_* is taken to 0.

For $k=\pm1$ and a_* sufficiently small for nonzero Q_*, the metric (2.27) becomes approximately $ds^2=\frac{1}{4}kd(a_*^2)d(Q_*^2)$, so its geodesics become approximately linear in a_*^2 and Q_*^2 as a_* goes to zero. Thus, one may parametrize the solutions which to the singularity at $a_*=0$ with $Q_*\neq0$ by their values of Q_* and $d(Q_*^2)/d(a_*^2)$ there, and the measure (3.3) becomes

$$\omega=|Q_*|dQ_*\wedge d[-k(dQ_*^2/da_*^2)^{-1}]^{1/2} \tag{3.4}$$

at $a_*=0$. For each finite range of Q_*, the measure integrated over values of dQ_*^2/da_*^2 bounded away from zero is finite, but it diverges when integrated to $dQ_*^2/da_*^2=0$, which gives a limiting solution in which Q_* stays constant as a_* increases from 0 to ∞ at infinite speed with respect to t_* (so that at each nonzero but finite a_*, $x=2\epsilon R=Q_*a_*^{-1}-1$ is finite, but $y\propto\dot{R}$ and $z\propto H=\dot{a}/a$ are both infinite).

The FRW models with a massive scalar field almost all (i.e., all but a finite measure) start at zero size and expand monotonically from there to an arbitrarily large size.[8] That is, the solutions with $\dot{a}=0$ at finite a bounded above have only a finite measure for those models. However, in the present case, the measure at $\dot{a}=0$ (i.e., $a_*'=0$ or $z=0$) is, from (3.2),

$$\omega=2a_*^2da_*\wedge dy\ , \tag{3.5}$$

which diverges when integrated over the infinite range of y allowed even when a_* is restricted to a finite range.

One can also ask for the measure within a finite range of a_* at a fixed nonzero value of the expansion rate z. Solving Eq. (2.26) for y as a function of a_* and x at fixed z and inserting the result into the measure (3.2) gives

$$\omega=z^{-1}(a_*^2x-4a_*^2z^2-\tfrac{1}{2}k)da_*\wedge dx\ , \tag{3.6}$$

which also diverges when integrated to $x=\infty$ for each finite range of a_*. By contrast, the analogous integral converges for the massive scalar-field case, implying that almost all of those models are spatially flat (have k/a^2 smaller than any nonzero value) at any given value of the Hubble expansion rate H. In the $R+\epsilon R^2$ models, there is an infinite measure of solutions which are within any arbitrary range of spatial curvatures at any given value of the expansion rate. Thus, a uniform probability distribution relative to the GHS canonical measure would not unambiguously solve the flatness problem at a fixed Hubble constant for the $R+\epsilon R^2$ FRW models as it does for those with Einstein gravity containing a minimally coupled massive scalar field.[8]

On the other hand, if one restricts oneself to $1+x \equiv 1+2\epsilon R > 0$ and fixes the value of the "energy density" in Whitt's conformal metric (2.33), say

$$\rho_* \equiv \left[\frac{d\phi}{d\hat{t}_*}\right]^2 + 2V \equiv \frac{y^2}{(1+x)^3} + \frac{x^2}{(1+x)^2} , \tag{3.7}$$

then if $\rho_* < 1$, the measure integrated over a finite range of a_* is finite, and so almost all solutions at fixed $\rho_* < 1$ have a_* arbitrarily large and, hence, spatial curvature k/a^2 arbitrarily small. To see this, let

$$\frac{d\phi}{d\hat{t}_*} \equiv \frac{y}{(1+x)^{3/2}} = \rho_*^{\,1/2}\sin\psi , \tag{3.8}$$

$$(2V)^{1/2} \equiv 1-e^{-2\phi} \equiv \frac{x}{1+x} = \rho_*^{\,1/2}\cos\psi . \tag{3.9}$$

Then at fixed ρ_*, $d\phi/d\hat{t}_*$ is a (double-valued) function of ϕ, and the constraint (2.39) implies that $d\beta/d\hat{t}_*$ is a (double-valued) function of β, so the action (2.37) leads to the measure

$$\begin{aligned}\omega &= d(-2e^{3\beta}d\beta/d\hat{t}_*)\wedge d\beta + d(2e^{3\beta}d\phi/d\hat{t}_*)\wedge d\phi \\ &= d(e^{3\beta})\wedge(2d\phi/d\hat{t}_*)d\phi \\ &= d(e^{3\phi}a_*^{\,3})\wedge(2d\phi/d\hat{t}_*)d\phi \\ &= \frac{\rho_*\sin^2\psi\, d\psi\wedge d(a_*^{\,3})}{(1-\rho_*^{\,1/2}\cos\psi)^{5/2}} .\end{aligned} \tag{3.10}$$

When integrated over ψ from 0 to 2π with $\rho_* < 1$ and over a finite range of a_*, (3.10) gives a finite measure, though it diverges as $\psi \to \cos^{-1}\rho_*^{\,-1/2}$ if $\rho_* \geq 1$. Thus almost all solutions (i.e., all but a finite measure) at fixed $\rho_* < 1$ and with $x > -1$ have a_* arbitrarily large and, hence, spatial curvature k/a^2 arbitrarily small.

However, if one does not require $\rho_* < 1$, one does not have a solution of the flatness problem, because (3.10) indicates that there is an infinite measure of solutions which have $\rho_* > 1$, even within a finite range of a_*, and these solutions need never become flat. For example, if $k=+1$, the measure at an extremum of $e^{2\beta}=a_*Q_*$ (i.e., $\dot{\beta}=0$, which occurs at $\rho_*=e^{-2\beta}$) is

$$\omega = d(e^{\beta})\wedge\left[\frac{1-3\sin^2\psi}{1-e^{-\beta}\cos\psi}d\psi\right] , \tag{3.11}$$

which in any fixed range of $\beta<0$ diverges at $1-e^{-\beta}\cos\psi = e^{-2\phi}=0$, i.e., as the extremum at fixed β is moved to infinite $e^{2\phi}=1+x=Q_*/a_*$. If the extremum is at $e^{2\beta}<\frac{2}{3}$, it is an absolute maximum for $e^{2\beta}$ along the trajectory, and the entire trajectory also has a_* less than the resulting value of $(e^{2\beta}+e^{\beta})^{1/2}$ and, hence, always has significant spatial curvature.

IV. SOLUTIONS WITH $k=+1$

Now the dynamical equations and canonical measure will be used to analyze the FRW models with the $R+\epsilon R^2$ Lagrangian. First, consider the case with positive spatial curvature (2.2), $k=+1$. The trajectories in the flat metric (2.29), (2.42), or (2.44) are null at the vanishing of the mass squared $f_*(a_*,Q_*)$, given by Eq. (2.17) with $a_* \equiv (24\epsilon)^{-1/2}a$ and $Q_* \equiv (24\epsilon)^{-1/2}Q \equiv (24\epsilon)^{-1/2}a(1+2\epsilon R)$, which occurs at the singularity $a_*=0$ and on the two other curves $f_*=0$ which may be alternatively represented as the two branches of

$$\begin{aligned} a_*(x) &= \frac{(1+x)^{1/2}}{|x|} , \\ Q_*(x) &= \frac{(1+x)^{3/2}}{|x|} = \frac{[(1+4a_*^{\,2})^{1/2}\mp 1]^2}{4a_*} , \end{aligned} \tag{4.1}$$

as

$$T(X) = X\left[\frac{4X^2+1}{4X^2-1}\right] , \tag{4.2}$$

or, using the coordinates β and ϕ in the region $Q_*>0$ where the curves occur, as

$$\beta(\phi) = -\ln|1-e^{-2\phi}| . \tag{4.3}$$

Along these curves $a_*'Q_*'=f_*=0$, so that either a_* (if $dX/dT=d\phi/d\beta=+1$) or Q_* (if $dX/dT=d\phi/d\beta=-1$) has an extremum there. These two curves divide the $(a_*,\,Q_*)$ parameter space into three regions (cf. Fig. 2 of Ref. 5):

$$\text{I:}\quad Q_* < \frac{[(1+4a_*^{\,2})^{1/2}-1]^2}{4a_*} , \tag{4.4}$$

$$\text{II:}\quad \frac{[(1+4a_*^{\,2})^{1/2}-1]^2}{4a_*} < Q_* < \frac{[(1+4a_*^{\,2})^{1/2}+1]^2}{4a_*} , \tag{4.5}$$

$$\text{III:}\quad \frac{[(1+4a_*^{\,2})^{1/2}+1]^2}{4a_*} < Q_* . \tag{4.6}$$

In regions I and III, the mass squared f_* is positive, so the trajectories are timelike in the flat metric (2.29) or (2.42). That is, $dQ_*/da_*>0$, $|dX/dT|<1$, or $|d\phi/d\beta|<1$ for $Q_*>0$ and $|d\phi/d\beta|>1$ for $Q_*<0$ (which occurs only in I). In region II, f_* is negative, so the trajectories are spacelike: $dQ_*/da_*<0$, $|dX/dT|>1$, or $|d\phi/d\beta|>1$ (since $Q_*>0$).

The boundary between regions I and II, given by Eq. (4.1) for $-1<x<0$ (so the $\mp$ sign is $-$) or (4.2) for $-\frac{1}{2}<X<0$ or (4.3) for $\phi<0$, is timelike in the flat metric (2.29) or (2.42). Thus, a trajectory which leaves I to enter II, and hence becomes spacelike, cannot return to I unless it first traverses II into III and then returns across II with the opposite sign of $\dot{a}_*$, $\dot{Q}_*$, $\dot{X}$, and $\dot{\phi}$. In the other direction, a trajectory which enters I from II may return to II (with the opposite sign of $dX/dT=\pm 1$ or $d\phi/d\beta=\pm 1$ at the boundary). Equation (2.32) implies that $d^2a_*/dQ_*{}^2<0$ in region I, so a trajectory which enters I will inevitably return to II unless it instead hits $a_*=0$ with $Q_*<0$. In particular, unless a trajectory is null within I and, hence, has $\dot{a}_*$ or $\dot{Q}_*$ infinite, it cannot have Q_* going to $\pm\infty$ in region I, but instead must reach $a_*=0$ at finite $Q_*\leq 0$ or else cross into II at finite $Q_*>0$.

The boundary between regions II and III, given by Eq. (4.1) for $x>0$ (so the $\mp$ sign is $+$) or (4.2) for $X>\frac{1}{2}$ or (4.3) for $\phi>0$, is timelike for $x<2$, $a_*>\frac{1}{2}\sqrt{3}$, $X<\frac{1}{2}\sqrt{3}$, $\beta>\ln\frac{3}{2}$, or $\phi<\frac{1}{2}\ln 3$ and is spacelike if these inequalities are reversed. A trajectory crossing the timelike part of the boundary, going from III into II, becomes spacelike in II and must cross II to enter I, whereupon it will inevitably return to II or else go to $a_*=0$ at $Q_*<0$ as discussed above. A trajectory going from III to II across the spacelike part of the boundary with $a'_*=0$ $(dX/dT=d\phi/d\beta=+1)$ has a_* increasing and Q_* decreasing in II and hence must either return to the spacelike part of the boundary with $Q'_*=0$ $(dX/dT=d\phi/d\beta=-1)$ there and reenter III or else cross II and enter I. A trajectory crossing the spacelike part of the boundary into II with $Q'_*=0$ there, either must return to the boundary, with $a'_*=0$, at a smaller value of a_* and a larger value of Q_* (e.g., the time reverse of the first possibility in the previous sentence), or else it must hit $a_*=0$ at finite $Q_*(a_*=0)>0$, except for a single intermediate case on the border between these two classes of possibilities.

This borderline curve is the limit of the previous curves as the end-point value $Q_*(a_*=0)$ is taken to infinity. For large Q_*, where the spacelike part of the boundary has the asymptotic behavior $a_*\sim Q_*{}^{-1}$ or $\beta\sim 0$, the limiting curve has $a_*\sim\frac{2}{3}Q_*{}^{-1}\sim\pm 3^{-1/2}(t_{*0}-t_*)$ or $\beta\sim-\frac{1}{2}\ln\frac{3}{2}$ and, thus, is precisely tuned to avoid going either back to the boundary or to $a_*=0$ at any finite Q_*, though it does reach $a_*=0$ at finite time, unlike the analogous solution in the massive scalar-field case.[13] There is one borderline curve for each point on the boundary at which it enters region II, so there is a one-parameter family of such curves, a set of measure zero compared with the two-parameter set of all possible trajectories.

A trajectory going in the opposite direction, from II into III, has as a consequence of Eqs. (2.39) and (2.40),

$$\frac{d}{d\hat{t}_*}\left[e^{3\beta}\frac{d\phi}{d\hat{t}_*}\right]=-e^{3\beta}\frac{dV}{d\phi}\ , \tag{4.7}$$

which is negative in III, where ϕ is positive and $V(\phi)$ is monotonically increasing, so, unless $\dot{\phi}=\infty$, the trajectory eventually has $d\phi/d\hat{t}_*<0$, and then ϕ decreases to its value $-\frac{1}{2}\ln(1-e^{-\beta})$ on the boundary with II (unless the trajectory returns to the boundary even before ϕ starts to decrease). The same conclusion may be derived from Eq. (2.15) written in the form

$$(a^3\dot{R})^{\cdot}=-\frac{1}{6\epsilon}a^3R\ . \tag{4.8}$$

Thus, except for degenerate null trajectories with infinite time derivatives, trajectories cannot remain in region III indefinitely, just as in region I, but must return to region II.

Trajectories in II are spacelike in the metric (2.29) or (2.42), so they generically traverse II in going between I and III (which must happen for all trajectories in II with $Q_*>\frac{3}{2}\sqrt{3}$, the part of II in the domain of dependence of both boundaries in the metric $-d\bar{s}^2$) or else go between I or III and the singularity at $a_*=0$. A one-parameter family of borderline curves in II was discussed above, which goes between a point on the boundary with III and the limit $Q_*=\infty$ with $a_*\sim\frac{2}{3}Q_*{}^{-1}$ asymptotically. This family may be extended to include curves which each enter II at some point on the boundary with I (this point now giving the single free parameter) and staying in II to become $a_*\sim\frac{2}{3}Q_*{}^{-1}$ asymptotically as Q_* increases indefinitely.

Except for certain one-parameter or discrete solutions to be discussed below, the generic (i.e., two-parameter) solutions in the $k=+1$ case correspond to trajectories which have both ends at $a_*=0$, i.e., both a big bang and a big crunch. The simplest such solutions go between $a_*=0$ at $Q_*<0$ and $a_*=0$ at $Q_*>0$ with Q_* changing monotonically and without ever entering region III. The discussion at the end of Sec. III shows that there is an infinite measure for such solutions with β having a maximum value $<-\frac{1}{2}\ln\frac{3}{2}$ (i.e., a_*Q_* having a maximum $<\frac{2}{3}$) in region II. None of these solutions have any inflationary phase or even ever have a_* get larger than some upper limit of order unity, so it is not the case that almost all solutions with the $R+\epsilon R^2$ Lagrangian exhibit inflation.

All other generic solutions enter region III as well as I and II and have one or more traverses of II, where a *traverse* may be defined as a crossing of region II between its boundaries with I and III (or vice versa) without reaching the singularity $a_*=0$. Trajectories which enter III (or merely turn around at its boundary) must have their maximum value of β (which will occur during a traverse or at its end point) be >1, so by Eq. (3.11) there is only a finite measure for such solutions with β bounded above. In other words, almost all solutions which reach the boundary of III have β eventually become arbitrarily large.

The constraint (2.39) implies that for $k=+1$, the "energy density" (3.7) obeys $\rho_*\geq e^{-2\beta}$, with equality at the maximum for β. Equation (2.40) leads to

$$\frac{d\rho_*}{d\beta}=-6\left[\frac{d\phi}{d\hat{t}_*}\right]^2\leq 0\ , \tag{4.9}$$

so ρ_* will decrease as β increases, eventually becoming very small if β becomes very large. Then $x\equiv 2\epsilon R\equiv e^{2\phi}-1$ will also be very small, so the potential (2.38) will have the approximate form $V(\phi)\approx 2\phi^2$, that of a scalar field of dimensionless mass $m_*=2$. Hence, Eq. (2.40) will lead to oscillations of very small magnitude in ϕ. Averaged over each oscillation, the kinetic and potential pieces on the right-hand side of Eq. (3.7) will be approximately equal, so for $\rho_* \ll 1$, Eq. (4.9) will be

$$\frac{d}{d\beta}\langle\rho_*\rangle \approx -3\langle\rho_*\rangle \ , \tag{4.10}$$

leading to dustlike FRW behavior with an approximately constant "energy"

$$E_* \equiv \rho_* \hat{a}_*{}^3 \equiv \rho_* e^{3\beta} \equiv \frac{y^2+x^2+x^3}{(1+x)^{3/2}} a_*{}^3 \approx \text{const} \tag{4.11}$$

in the constraint equation

$$\left[\frac{d\hat{a}_*}{d\hat{t}_*}\right]^2 = \frac{E_*}{\hat{a}_*} - k \ . \tag{4.12}$$

For $k=+1$, the evolution of $\hat{a}_* \equiv e^{\beta}$ has approximately the standard cycloid solution, which may be written in parametric form in terms of an angle $0<\hat{\theta}<\pi$ as

$$\hat{a}_* \approx E_* \sin^2\hat{\theta} \ , \tag{4.13}$$

$$\hat{t}_* \approx E_*(\hat{\theta} - \sin\hat{\theta}\cos\hat{\theta}) \ . \tag{4.14}$$

Meanwhile, the oscillations of ϕ will go roughly as

$$\phi \approx \tfrac{1}{2}(E_*/\hat{a}_*{}^3)^{1/2}\cos(2\hat{t}_* + \theta_0) \ , \tag{4.15}$$

where θ_0 is an arbitrary phase angle. E_* and θ_0 may be regarded as the two arbitrary parameters of the solution in the dustlike regime. One sees that for $\rho_* \ll 1$, $a_* = \hat{a}_* e^{-\phi}$ will be very near $\hat{a}_*$, except for some very small oscillations, and $t_* = \int e^{-\phi} d\hat{t}_*$ will be nearly equal to $\hat{t}_*$.

Solutions which have $\beta \equiv \ln\hat{a}_*$ become arbitrarily large will, thus, also have arbitrarily large $E_* = \hat{a}_{*\max}$ and will have an arbitrarily large number of oscillations of ϕ, of order E_* periods, or $2E_*$ traverses of region II. The energy density $\rho_* = E_*/\hat{a}_*{}^3$ will decrease to the extremely low minimum of $1/E_*{}^2$, but long before it reaches this, there will be a very long period in which the dimensionless spatial curvature $k/a_*{}^2 \approx k/\hat{a}_*{}^2$ will be negligible compared with ρ_*. For example, one is in both the dustlike and spatially flat regime for

$$1 \ll E_*{}^{1/3} \ll \hat{a}_* \approx a_* \ll E_* \ . \tag{4.16}$$

Because almost all of the $k=+1$ solutions which reach $\beta>0$ have an arbitrarily long period of approximately spatial flatness, the character of these solutions and the measure for inflationary and noninflationary classes (both of which turn out to be infinite) will be discussed below, in Sec. VI on $k=0$ solutions. In the remainder of the present section, the class of solutions in which β has bounded positive extrema will be considered. Generically, these trajectories have $a_*=0$ at both ends and various traverses in between, as the solution oscillates between regions I and III by crossing back and forth across II.

Similar to the case of Einstein gravity with a massive scalar field,[13,14] a subclass of these $R+\epsilon R^2$ solutions will have one or more bounces, a *bounce* being defined as a trajectory segment which enters region II from the spacelike part of the boundary with III and then returns to that boundary with III without going to I (as a traverse does) or to $a_*=0$. The limiting case of a trajectory which turns around with $\dot{a}_* = \dot{Q}_* = 0$ at a single point on the spatial part of the boundary of region III without actually entering II will also be called a bounce. Bounces are always confined to the region $\phi > \frac{1}{2}\ln 3$, $-\frac{1}{2}\ln\frac{3}{2} < \beta < \ln\frac{3}{2}$. Out of the finite measure of solutions which have a maximum for β which is greater than zero but less than some given upper bound, only a fraction less than unity (but greater than zero) will have one bounce, and only a fraction of these two bounces, and so on. Bounces basically convert decreasing β to increasing β so that the conformal metric (2.33) can begin reexpanding, leading to a new sequence of traverses in which β reaches a maximum and begins recontracting again (perhaps with a nearby minimum and another maximum, since bounces do not give the only local minima for β, though they apparently provide the only minima effective for converting contractions that persist over more than part of a single traverse into expansions that also persist over more than part of a single traverse). The nonzero measure of solutions which have only a finite number of bounces interspersing its traverses will eventually have a_* collapse to 0 (in both directions along the trajectories).

However, just as in the massive scalar-field case,[13,14] there will be a set of solutions of measure zero which have an infinite number of bounces. A discrete but apparently uncountable and even fractal family of one-parameter solutions will have $a_*=0$ at some value of Q_* (the one continuous parameter) at one end but will have no end in the other direction, rather an infinite sequence of traverses interspersed with bounces. The number of traverses between each bounce gives an infinite sequence of integers, which are the discrete parameters. If the trajectory starts at $a_*=0$ with $Q_*<0$, the first integer (the number of traverses before the first bounce) will be odd, and all of the other integers will be even (since the trajectory must always return to region III with its spacelike boundary before another bounce can occur). If the trajectory starts at $a_*=0$ with $Q_*>0$, but crosses II to I (not a traverse, since not from III) before traversing to III, again the first integer will be odd and all the rest even. If the trajectory goes from $a_*=0$ and $Q_*>0$ to III first, all of the integers will be even. One might conjecture that for each value of Q_* at $a_*=0$ and for each allowed infinite sequence of positive integers (by the rules above, i.e., the first integer is odd if $Q_*<0$ and arbitrary of $Q_*>0$, and the remaining integers are all even), there exists a corresponding solution, and that it is unique, but the proof or refutation of this conjecture will be left as an exercise for the reader.

There will also be another apparently uncountable and fractal discrete set of solutions with no continuous parameters which never reach $a_*=0$ and so give complete nonsingular spacetime metrics (2.1), now with a doubly sequence of traverses and bounces, just as apparently occurs for the massive scalar-field case.[13] Again, the discrete parameters may be taken to be the numbers of traverses between successive bounces, which this time forms a doubly infinite sequence of even integers. One may conjecture that there is a one-to-one correspondence between the doubly infinite sequences of positive, even integers and these discrete solutions. The evidence for this would be similar to that for the scalar-field case.[13] The existence of these discrete nonsingular solutions answers a question posed in Ref. 15.

One countably infinite subset of the perpetually bouncing $R+\epsilon R^2$ universes consists of periodic solutions, characterized by a finite sequence $(n_1,n_2,\ldots,n_l)$ of positive, even integers of arbitrary positive-integer length l which repeats indefinitely in both directions, so $n_i=n_{i(\mathrm{mod}\,l)}$ for all i. If $n_i=n_{j-1}$ for some j and all i, the trajectory will have two moments of time symmetry per period where it momentarily halts on one of the curves $f_*=0$ with $\dot a_*=\dot Q_*=0$ and then reverses its path, but for $l>2$ there need be no moments of time symmetry.

V. SOLUTIONS WITH $k=-1$

Now consider Friedmann-Robertson-Walker models with negative spatial curvature (2.2), $k=-1$. In this case, the trajectories of the flat metric (2.39), (2.42), or (2.44) are null on the single curve $f_*=0$ (for $a_*>0$) given by

$$a_*(x)=\left[\frac{-1-x}{x^2}\right]^{1/2}<\tfrac{1}{2}\,,$$
$$Q_*(x)=(1+x)\left[\frac{-1-x}{x^2}\right]^{1/2}=-\frac{[1\pm(1-4a_*{}^2)^{1/2}]^2}{4a_*}<0 \qquad (5.1)$$

for $x\equiv 2\epsilon R<-1$, or by

$$T(X)\equiv\tfrac{1}{2}(a_*+Q_*)\equiv\tfrac{1}{2}a_*(2+x)=X\left[\frac{4X^2-1}{4X^2+1}\right] \qquad (5.2)$$

for $X\equiv\frac{1}{2}(Q_*-a_*)\equiv\frac{1}{2}a_*x<0$, or, using $a_*\equiv(24\epsilon)^{-1/2}a=e^{\beta-\phi}$ and $Q_*\equiv(24\epsilon)^{-1/2}a(1+2\epsilon R)=-e^{\beta+\phi}$, by

$$\beta=-\ln(1+e^{-2\phi})<0 \qquad (5.3)$$

for $\phi=\frac{1}{2}\ln(-1-x)$ arbitrary. This curve divides the (a_*,Q_*) parameter space into two regions:

$$\text{IV:}\quad f_*=a_*{}^2(Q_*-a_*)^2+a_*Q_*<0$$
$$\text{or } T<X\left[\frac{4X^2-1}{4X^2+1}\right], \qquad (5.4)$$

$$\text{V:}\quad f_*>0 \text{ or } T>X\left[\frac{4X^2-1}{4X^2+1}\right]. \qquad (5.5)$$

The boundary curve $f_*=0$ is timelike for $x<-2$, $Q_*<-\frac{1}{2}$, $T<0$, $X<-\frac{1}{2}$, $\beta>-\ln 2$, or $\phi>0$, and it is spacelike if these inequalities are reversed. A trajectory entering region IV from V by crossing the timelike part of the boundary, or by crossing the spacelike part with $Q'_*=0$ $(dX/dT=d\phi/d\beta=-1)$ there, becomes spacelike in IV with a_* decreasing and Q_* increasing and, hence, must either go to $a_*=0$ or else emerge back into V across the spacelike part of the boundary. On the other hand, a trajectory crossing the spacelike part of the boundary into IV with $a'_*=0$ $(dX/dT=d\phi/d\beta=+1)$ at the boundary has a_* increasing and Q_* decreasing in IV and, hence, must eventually emerge from IV back into V.

In the opposite direction, a trajectory emerging from IV into V across the spacelike part of the boundary becomes timelike in V with $a'_*>0$ and $Q'_*>0$ and so cannot return to IV. Instead, a_* and Q_* both simply continue to increase indefinitely, and the Universe expands without limit.

However, a trajectory crossing the timelike part of the boundary into V can recross the timelike part back into VI. Indeed, if it crosses into V with $a'_*=0$ at the boundary and, hence, has Q_* decreasing in V, Eq. (2.32) implies that $d^2a_*/dQ_*{}^2<0$ in the part of V with $x<-2$, $Q_*<-a_*$, or $T<0$, so such a trajectory must bend back to reenter IV, this time crossing the boundary with $Q'_*=0$, and then Q_* increases thereafter in IV and also later in V if the trajectory returns there (across the spacelike part of the boundary), rather than hitting $a_*=0$.

On the other hand, a trajectory leaving region IV by crossing the timelike part of the boundary with $Q'_*=0$ may either expand indefinitely in V (and inevitably must if it ever reaches $a_*>\frac{1}{2}$, so that IV is no longer within its future null lines) or may bend around to reenter IV with $a'_*=0$. In the latter case, one may use Eqs. (2.39) and (2.40) combined into the form

$$\frac{d^2\phi}{d\beta^2}=\left[\frac{1-(d\phi/d\beta)^2}{2V-ke^{-2\beta}}\right]\left[-\frac{dV}{d\phi}-(6V-2ke^{-2\beta})\frac{d\phi}{d\beta}\right]$$
$$=\left[\frac{1-(d\phi/d\beta)^2}{e^{-2\beta}-(1+e^{-2\phi})^2}\right]\left[-2e^{-2\phi}-2e^{-4\phi}-(2e^{-2\beta}-3-6e^{-2\phi}-3e^{-4\phi})\frac{d\phi}{d\beta}\right] \qquad (5.6)$$

to show that after the trajectory enters IV with $d\phi/d\beta=+1$, $d\phi/d\beta$ decreases somewhat but stays positive as β and ϕ decrease, but then, at least for $\beta<-\ln 2$ (if not before), one must have $d^2\phi/d\beta^2<0$, so the trajectory must thereafter have $d\phi/d\beta$ increase again in IV and cross the spacelike part of the boundary rather than ever reaching $a_*=0$. Thereafter, it has both a_* and Q_* increasing monotonically, as discussed above.

The fact that the preceding trajectory could not go to $a_*=0$ in the future means that in the opposite direction of time, a trajectory cannot start at $a_*=0$ and cross the timelike part of the boundary of region IV with $a_*'=0$. Instead, if it reaches the timelike part of the boundary it must have $Q_*'=0$ there. Once it thus gets into region V, the preceding argument implies that Q_* will increase monotonically thereafter, even though a_* may suffer a temporary decrease while the trajectory cuts through part of region IV again, but, in any case, a_* cannot decrease to zero again. Of course, the same is true (except for now no possibility of a decrease of a_*) if the trajectory goes directly in region IV from $a_*=0$ at $Q_*<0$ to the spacelike part of the boundary or if it starts from $a_*=0$ at $Q_*>0$ and, hence, has its entire expansion in region V. In particular, there are no trajectories in the $k=-1$ case which are singular (have $a_*=0$) at both ends; all solutions expand forever in at least one or the other direction of time.

However, unlike the case of Einstein gravity with a massive scalar field, the $R+\epsilon R^2$ Lagrangian allows $k=-1$ solutions which have no singularities at all. In one direction of time, these nonsingular solutions contract from infinite a_* and Q_* in region V, cross either the timelike or spacelike part of the boundary into IV at a minimum for Q_*, have Q_* increase but a_* continue to decrease in IV, emerge back into V across the spacelike part of the boundary where a_* has its minimum, and then expand forever again to infinite a_* and Q_*. Of course, there are also the time reversals of these trajectories. Like the singular trajectories, they are generic in the sense of having two continuous parameters, which, for example, could be chosen to be coordinates for the two points at which they each cross the boundary between regions IV and V. If these two points coincide, one gets a one-parameter subset of solutions which do not really enter region IV but instead have $\dot{a}_*=0$ and $\dot{Q}_*=0$ there at a point on the spacelike part of the boundary as a moment of time symmetry. The expanding and contracting parts of the trajectory coincide in the (a_*,Q_*) plane, and in this case the slope is not null at the boundary $f_*=0$ but rather is orthogonal to it in the flat Lorentz metric (2.29), (2.42), or (2.44).

The fact that the nonsingular solutions are generic means that they have a nonzero measure, which may be evaluated on the initial data surface $d\phi/d\hat{t}_*=0$, which occurs in region IV. There the constraint (2.39) with $k=-1$ gives $d\beta/d\hat{t}_*=-(e^{-2\beta}+2V)^{1/2}$, where the direction of time has been arbitrarily chosen to agree with the original description of these trajectories in the previous paragraph. Equation (2.38) with the lower sign for this $Q_*<0$ region implies that $V\le -\frac{1}{2}$, and the constraint gives $-\frac{1}{2}e^{-2\beta}\le V$. Then the action (2.37) leads to the measure

$$\begin{aligned}\omega_*&=d(2e^{3\beta}d\beta/d\hat{t}_*)\wedge d\beta+d(-2e^{3\beta}d\phi/d\hat{t}_*)\wedge d\phi\\&=d\beta\wedge d[2e^{3\beta}(e^{-2\beta}+2V)^{1/2}]\\&=2e^{3\beta}(e^{-2\beta}+2V)^{-1/2}d\beta\wedge dV\ .\end{aligned}\tag{5.7}$$

Integrating this over region IV gives

$$\int_{-\infty}^{0}d\beta\int_{-e^{-2\beta}/2}^{-1/2}2e^{3\beta}(e^{-2\beta}+2V)^{-1/2}dV=\int_{-\infty}^{0}2e^{3\beta}(e^{-2\beta}-1)^{1/2}d\beta=\tfrac{2}{3}\ ,\tag{5.8}$$

which is finite.

On the other hand, the measure of the singular solutions, which may be obtained from Eq. (3.4) in terms of the parameters Q_* and $d(Q_*{}^2)/d(a_*{}^2)$ at the singularity, is infinite. Thus, almost all (i.e., all but a finite measure) of the $k=-1$ solutions are singular.

Because all $k=-1$ solutions expand to (or contract from) arbitrarily large sizes and arbitrarily low values of the "energy density" ρ_* given by Eq. (3.7), unlike the case for $k=+1$ solutions, we can evaluate the measure (3.10) at any fixed $\rho_*<1$ and find that then almost all solutions have an arbitrarily large size a_* and, thus, are arbitrarily near spatial flatness. Thus, we can wait until Sec. VI on $k=0$ solutions to calculate the measure for the inflationary and noninflationary solutions, both of which will be infinite. Of course, for any given solution, even though it has arbitrarily large "energy" E_* given by Eq. (4.11) in the dustlike regime, one can integrate the constraint equation (4.12) with $k=-1$ forward in time until $a_*\approx\hat{a}_*>E_*$ and find that over all of the remaining infinite expansion time the spatial curvature will dominate the dynamics. However, this will occur only when ρ_* has dropped to the extremely low value $<1/E_*{}^2$, and so, at fixed ρ_*, only the finite measure of solutions with $\hat{a}_*<\rho_*{}^{-1/2}$ will have entered that asymptotic stage.

VI. SOLUTIONS WITH $k=0$

We have seen that almost all FRW solutions with the $R+\epsilon R^2$ Lagrangian (i.e., all but a finite measure) that expand to $aQ\equiv 24\epsilon a_*Q_*>16\epsilon$ continue to expand to an arbitrarily large a and Q and include an arbitrarily long phase in which the spatial curvature is negligible, thus acting effectively as $k=0$ solutions even if actually $k=\pm 1$. (There is also an infinite measure for the class of $k=+1$ solutions with a_*Q_* bounded above by $\frac{2}{3}$ which always have significant spatial curvature, but these and the various classes of finite positive or zero measure will be considered no longer in this section.) When $k=0$, $f_*=a_*{}^2(Q_*-a_*)^2$ is homogeneous in a_* and Q_*, so the auxiliary metric (2.27), $ds^2=-f_*da_*dQ_*$, has a conformal Killing vector $a_*\partial/\partial a_*+Q_*\partial/\partial Q_*$. Thus, its timelike geodesics, which represent solutions of the dynamical equations (2.17)–(2.19) for $a_*\equiv(24\epsilon)^{-1/2}a$ and $Q_*\equiv(24\epsilon)^{-1/2}Q$, are invariant under the scale transformation $a_*\to\lambda a_*$, $Q_*\to\lambda Q_*$, $\tau\to\lambda^3\tau$, which gives $f_*\to\lambda^4 f_*$ and $\eta_*\equiv\eta\to\lambda^{-1}\eta_*$ but leaves $t_*\equiv(24\epsilon)^{-1/2}t$, $R_*\equiv 24\epsilon R$, $H_*\equiv(24\epsilon)^{1/2}\dot{a}/a$, $x\equiv 2\epsilon R$, $y\equiv(24\epsilon^3)^{1/2}\dot{R}$, and $z\equiv(6\epsilon)^{1/2}\dot{a}/a$ unchanged.

After taking out the scale freedom, there is only a one-parameter family of physically distinct $k=0$ solutions. For example, the solutions can be given by curves in the (x,y,z) space obeying Eqs. (2.23)–(2.25) and lying on the constraint surface (2.26), which for $k=0$ is

$$x^2-4yz-4(1+x)z^2=0\ . \tag{6.1}$$

One may use Eq. (6.1) to eliminate one of the three variables (x,y,z) from the dynamical equations and then take the ratio of the remaining two time derivatives to eliminate t_* and get a single autonomous first-order differential equation for the remaining two variables. The solutions are then a one-parameter family of curves in the space of these two variables. Once such a curve is found, one can integrate the reciprocal of one of the time derivatives to get t_* along the curve, and the expansion rate may be integrated over time to get

$$\ln a_* = \int 2z\, dt_* + \text{const}\ , \tag{6.2}$$

where the constant of integration represents the scale parameter, which is the second arbitrary parameter of the $k=0$ solutions. (The constant of integration for t_* may be considered as the third arbitrary parameter for the solutions of the full fourth-order system with one constraint, but since it is trivial for all k, it was already dropped in Sec. II.)

It is simplest to eliminate y by means of the constraint (6.1), since it is linear in that variable and, hence, leads to a unique solution (when $z\neq 0$). Inserting $y=(x^2/4z)-(1+x)z$ into Eqs. (2.23) and (2.25) and taking the ratio gives

$$\frac{dx}{dz}=\frac{x}{2z}-\frac{2z}{x-4z^2}\ . \tag{6.3}$$

For each of the one-parameter sets of solution curves in the (x,z) plane, one may integrate the reciprocal of Eq. (2.25) to get

$$t_* = \int \frac{dz}{x-4z^2} = \int \frac{2z\, dx}{x-4(1+x)z^2} \tag{6.4}$$

and then use Eq. (6.2) to evaluate

$$\ln a_* = \int \frac{2z\, dz}{x-4z^2} = \int \frac{4z^2 dx}{x-4(1+x)z^2}\ . \tag{6.5}$$

Although Eq. (6.3) is fairly simple algebraically, its singularity at $z=0$ (i.e., $H=0$ or $\dot a=0$) poses some hindrance to an immediate qualitative understanding of the solutions, all of which go to $z=0$ at $x=0$ (i.e., $R=0$) an infinite number of times. As a solution approaches $z=0$, Eq. (6.1) implies that it has $z\sim x^2/4y$, so it is the value of y which distinguishes the solution there, but that variable has been eliminated from the differential equation (6.3). It would be better to find solution curves in terms of two variables such that different solutions remain distinct at all times. After some trial and error, it was found that the following transformation $(x,y,z)\to(r,u,v)$ is useful:

$$x\equiv(r-u)^2+v,\quad y\equiv\tfrac{1}{2}(r-u)v+u,\quad z\equiv\tfrac{1}{2}(r-u)\ , \tag{6.6}$$

whose inverse transformation gives

$$\begin{aligned} r &\equiv y-xz+2z+4z^3=\frac{1}{2}\frac{d}{dt_*}(x-z^2+2\ln a_*)=(24\epsilon)^{1/2}\left[4\epsilon\frac{(a^{3/2})^{\cdot\cdot}}{a^{3/2}}+\ln a\right]=(6\epsilon)^{1/2}\left[6\epsilon\frac{\dot H^2}{H}+H\right]\ , \\ u &\equiv y-xz+4z^3=\frac{1}{2}\frac{d}{dt_*}(x-z^2)=(24\epsilon)^{1/2}\left[4\epsilon\frac{(a^{3/2})^{\cdot\cdot}}{a^{3/2}}\right]=(6\epsilon)^{1/2}\left[6\epsilon\frac{\dot H^2}{H}-H\right]\ , \\ v &\equiv x-4z^2=\frac{dz}{dt_*}\equiv\frac{1}{2}\frac{d}{dt_*}\left[\frac{1}{a_*}\frac{da_*}{dt_*}\right]\equiv 12\epsilon\left[\frac{\dot a}{a}\right]^{\cdot}\equiv 12\epsilon\dot H\ , \end{aligned} \tag{6.7}$$

where the $k=0$ constraint (6.1), which takes the simple form

$$r^2=u^2+v^2\ , \tag{6.8}$$

was assumed in all equalities except the $\equiv$ ones in Eq. (6.7). Then the dynamical equations (2.23)–(2.25) become

$$\frac{dr}{dt_*}=-3v^2=-3(r^2-u^2)\ , \tag{6.9}$$

$$\frac{du}{dt_*}=-2v-3v^2=-v(2+3v)\ , \tag{6.10}$$

$$\frac{dv}{dt_*}=2u+3uv-3rv=u(2+3v)-3v\sqrt{u^2+v^2}\ , \tag{6.11}$$

which, as always, preserve the constraint.

These equations give curves in the unrestricted (u,v) plane which do not intersect at any finite t_*, though they all spiral in to the origin at $t_*=\infty$ (if the direction of time is chosen, as shall be assumed, so that the Universe is expanding, which implies that both z and r, neither of which can change sign during the evolution, are positive). If one takes

$$u\equiv r\cos\theta,\quad v\equiv r\sin\theta\ , \tag{6.12}$$

then Eqs. (6.9)–(6.11) imply that

$$\frac{dr}{dt_*}=-3r^2\sin^2\theta\ , \tag{6.13}$$

$$\frac{d\theta}{dt_*}=2+3r\sin\theta(1-\cos\theta)\ , \tag{6.14}$$

so r always decreases monotonically, and θ increases monotonically for $r<3^{-5/2}8$. One can also divide Eq. (6.11) by (6.10) and use (6.8) to get

$$\frac{dv}{du}=-\frac{u}{v}+\frac{3\sqrt{u^2+v^2}}{2+3v}\ . \tag{6.15}$$

This gives curves with $dv/du=0$ at

$$u=\tfrac{3}{2}v^2(1+3v)^{-1/2}\operatorname{sgn}v \tag{6.16}$$

and with $dv/du=\infty$ at $v=0$ and at $v=-\frac{2}{3}$.

In order to calculate the approximate behavior of the solution curves in the (u,v) plane in parametric form, it is convenient also to introduce two new variables

$$\begin{aligned}s&\equiv\pm(2r)^{1/2}\sin\frac{\theta}{2}\equiv\pm(r-u)^{1/2}\equiv\pm(2z)^{1/2}\equiv\pm H_*^{\ 1/2}\equiv\pm(24\epsilon)^{1/4}H^{1/2}\equiv\pm(24\epsilon)^{1/4}(\dot a/a)^{1/2}\ ,\\ w&\equiv\pm\left[(2r)^{3/2}\sin^3\frac{\theta}{2}+(2r)^{1/2}\cos\frac{\theta}{2}\right]\equiv s^3+\frac{v}{s}\equiv\frac{x}{s}\equiv\frac{\pm R_*}{12H_*^{\ 1/2}}\equiv\pm(\tfrac{2}{3}\epsilon^3)^{1/4}\frac{R}{H^{1/2}}\ ,\end{aligned} \tag{6.17}$$

so that in terms of these, the previous variables may be written as

$$x=sw,\quad y=\tfrac{1}{2}(-s^2-s^3w+w^2),\quad z=\tfrac{1}{2}s^2\ , \tag{6.18}$$

$$\begin{aligned}r&=\tfrac{1}{2}[(s^3-w)^2+s^2],\quad u=\tfrac{1}{2}[(s^3-w)^2-s^2]\ ,\\ v&=-s(s^3-w)\ .\end{aligned} \tag{6.19}$$

The joint sign of s and w is irrelevant, so there is a 2-1 map from (s,w) to (u,v). The dynamical equations now take the simple form

$$\frac{ds}{dt_*}=w-s^3\ , \tag{6.20}$$

$$\frac{dw}{dt_*}=-s\ . \tag{6.21}$$

These may be combined into a single second-order equation for the square root of the Hubble expansion rate,

$$\frac{d^2s}{dt_*^{\ 2}}+3s^2\frac{ds}{dt_*}+s=0\ , \tag{6.22}$$

which is equivalent to Eq. (2.8) of Mijić *et al.*,[6] or into a single equation for w,

$$\frac{d^2w}{dt_*^{\ 2}}+\left[\frac{dw}{dt_*}\right]^3+w=0\ , \tag{6.23}$$

each of which is a nonlinearly damped harmonic-oscillator equation. One may then note that Eq. (6.19) becomes

$$\begin{aligned}r&=\frac{1}{2}\left[\left[\frac{ds}{dt_*}\right]^2+s^2\right]=\frac{1}{2}\left[\left[\frac{ds}{dt_*}\right]^2+\left[\frac{dw}{dt_*}\right]^2\right]\ ,\\ u&=\frac{1}{2}\left[\left[\frac{ds}{dt_*}\right]^2-s^2\right]=\frac{1}{2}\left[\left[\frac{ds}{dt_*}\right]^2-\left[\frac{dw}{dt_*}\right]^2\right]\ ,\\ v&=s\frac{ds}{dt_*}=-\frac{ds}{dt_*}\frac{dw}{dt_*}\ .\end{aligned} \tag{6.24}$$

Alternatively, one may combine Eqs. (6.20) and (6.21) into the single first-order equation

$$\frac{ds}{dw}=s^2-s^{-1}w\Longleftrightarrow\frac{dw}{ds}=\frac{s}{s^3-w}\ , \tag{6.25}$$

which by the change of variables

$$w=\pm2^{3/4}X^{1/2},\quad s=2^{1/4}Y^{-1}\ , \tag{6.26}$$

becomes an Abel equation of the first kind in standard form:[16]

$$\frac{dY}{dX}=Y^3\mp X^{-1/2}\ . \tag{6.27}$$

Once $s(w)$ or $w(s)$ is found, one can use Eq. (6.20) or (6.21) to integrate t_*, and then

$$\begin{aligned}\ln a_*&=\int s^2dt_*=-\int s\,dw=\int\frac{-s^2ds}{s^3-w}\\ &=\int\frac{s^3ds}{v}=-\frac{1}{3}\int\frac{dr}{r+u}\ .\end{aligned} \tag{6.28}$$

After taking out the scale freedom, which comes from the constant of integration in Eq. (6.28), the single remaining physical parameter of the $k=0$ solution may be taken to be w_0, the value of w at $a_*=0$ (and, hence, necessarily at $s=\pm\infty$). Without loss of generality, one may fix the arbitrary joint sign of s and w so that s starts at $+\infty$ at $a_*=0$. For later convenience, we define

$$s_0\equiv w_0^{1/3}\equiv w^{1/3}\quad(a_*=0)\ , \tag{6.29}$$

which can have either sign, as w_0 can, and is not to be confused with the positive infinite initial value of s.

If $w_0\equiv s_0^{\ 3}=\pm\infty$, one gets a one-parameter (the scale parameter) set of $k=0$ solutions, of zero measure, which evolve along the separatrix of Eq. (6.25). This solution has the form

$$w(s)=w_\infty(s)=s^3-\tfrac{1}{3}s^{-1}+\tfrac{1}{27}s^{-5}-\tfrac{2}{81}s^{-9}+O(s^{-13}) \tag{6.30}$$

or

$$v(s)=v_\infty(s)=-\tfrac{1}{3}+\tfrac{1}{27}s^{-4}-\tfrac{2}{81}s^{-8}+O(s^{-12}) \tag{6.31}$$

for large s and takes an infinite amount of time to expand from $a_* = 0$ with a rough time dependence:

$$s \approx (-\tfrac{2}{3}t_*)^{1/2}, \quad w \approx (-\tfrac{2}{3}t_*)^{3/2} , \tag{6.32}$$

$$x \approx \tfrac{4}{9}t_*^{\,2}, \quad y \approx \tfrac{4}{9}t_*, \quad z \approx r \approx -u \approx -\tfrac{1}{3}t_* , \quad v \approx -\frac{1}{3} + \frac{1}{12t_*^{\,2}} , \tag{6.33}$$

$$a_* \approx a_{*\infty} s^{-1/3} e^{-3s^4/4} \approx a_{*\infty}(-\tfrac{2}{3}t_*)^{-1/6} e^{-t_*^{\,2}/3} , \quad R_* \approx \tfrac{16}{3}t_*^{\,2} , \tag{6.34}$$

$$a \approx (24\epsilon)^{5/12} a_{*\infty}(-\tfrac{2}{3}t)^{-1/6} e^{-t^2/72\epsilon}, \quad R \approx \frac{t^2}{108\epsilon^2} , \tag{6.35}$$

for large negative t_* if the origin of the coordinate t_* is appropriately chosen (e.g., t_* is roughly zero when s first crosses zero). Although the resulting spacetimes are nonsingular throughout an infinite range of the proper time t of a comoving observer, any geodesic with nonzero spatial momentum will reach $a = 0$ and $R = \infty$ at a finite value (in the past) of its affine parameter, so the separatrix solutions are geodesically incomplete and represent singular spacetimes, as do all other $k = 0$ solutions (which reach $a = 0$ at finite t).

Now turn to the generic $k = 0$ solutions, which have finite w_0. Set $t_* = 0$ at $a_* = 0$, where $s = +\infty$ and $w = w_0$. Then the initial evolution, for $t_* \ll w_0^{-2/3} = s_0^{-2}$, is roughly

$$w \approx w_0 - (2t_*)^{1/2}, \quad s \approx (2t_*)^{-1/2} , \tag{6.36}$$

$$r \approx u \approx \frac{1}{16t_*^{\,3}}, \quad v \approx -\frac{1}{4t_*^{\,2}} , \tag{6.37}$$

$$x \approx w_0(2t_*)^{-1/2}, \quad y \approx -\tfrac{1}{2}w_0(2t_*)^{-3/2} , \quad z \approx (4t_*)^{-1} , \tag{6.38}$$

$$a_* \approx a_{*0}(2t_*)^{1/2}, \quad Q_* \approx a_{*0}w_0[1 + \tfrac{1}{3}w_0(2t_*)^{3/2}] , \tag{6.39}$$

$$a \approx (96\epsilon)^{1/4} a_{*0} t^{1/2}, \quad R \approx 3(6\epsilon)^{-3/4} w_0 t^{-1/2} . \tag{6.40}$$

This approximate power-law expansion, the same as for a radiation-dominated universe, gives $\dot{H} \approx -2H^2 \approx -\frac{1}{2}t^{-2}$, which is noninflationary. Here the positive dimensionless quantity a_{*0} represents the arbitrary scale parameter from the constant of integration in Eq. (6.2) or (6.28). It transforms as $a_{*0} \to \lambda a_{*0}$ under the action of the conformal Killing vector of the auxiliary metric (2.27) discussed above, which does not alter any of the variables in Eqs. (6.36)–(6.38). One sees that the sign of w_0 or s_0 is the same as the sign of Q_* or of Ra at $a_* = 0$. In fact, the two parameters of the geodesic in the metric (2.27), which near $a_* = 0$ has the form

$$ds^2 \sim -d(\tfrac{1}{3}a_*^{\,3}) d(\tfrac{1}{3}Q_*^{\,3}) , \tag{6.41}$$

and, hence, leads to $Q_*^{\,3}$ varying linearly with $a_*^{\,3}$ there, may be taken to be

$$\lim_{a_* \to 0} Q_* = \lim_{a \to 0} (\tfrac{1}{6}\epsilon)^{1/2} aR = a_{*0} w_0 , \tag{6.42}$$

$$\lim_{a_* \to 0} \frac{d(Q_*^{\,3})}{d(a_*^{\,3})} = \lim_{a \to 0} \tfrac{2}{3}\epsilon^3 \frac{R^4}{H^2} = w_0^{\,4} . \tag{6.43}$$

These limits are 0 and ∞, respectively, for the separatrix solution (6.30)–(6.35), which has $a_{*0} = 0$ and $w_0 = \pm\infty$ in such a way that, for a sufficiently small a_*,

$$Q_* \approx \tfrac{4}{3} a_* \ln \left[\frac{a_{*\infty}}{a_*} \left[\tfrac{4}{3} \ln \frac{a_{*\infty}}{a_*} \right]^{1/12} \right] , \tag{6.44}$$

which goes to zero with an infinite derivative at $a_* = 0$.

Now consider the evolution for values of t_* larger than roughly s_0^{-2}, where it leaves the "radiation" regime. If $w_0 \equiv s_0^{\,3}$ has a magnitude of order unity or smaller, the solution will only leave the $a \propto t^{1/2}$ regime when u and v, s and w, or H and R begin oscillating and the solution begins spiraling around the origin in the (u,v) or (s,w) plane. This spiraling behavior occurs asymptotically with time for all $k = 0$ solutions and represents a "dustlike" regime with asymptotically constant "energy" E_* defined by Eq. (4.11), which is of order $a_{*0}^{\,3}$ for w_0 fairly near zero. The transition between the "radiation" and "dustlike" regimes is hard to describe accurately by an explicit closed-form analytic solution and so will not be described here, but the main point is that there is not a long inflationary period ($|\dot{H}| \ll H^2$ for many e folds of expansion of a) if $|w_0| \lesssim 1$.

Therefore, turn to the consideration of the solutions with $|w_0| \gg 1$, which do turn out to have a long period of inflation between the "radiation" and "dustlike" regimes. By Eq. (6.25), w will change only negligibly compared with s (which decreases from $+\infty$), until s drops near $w^{1/3}$. Hence, one may replace w by $w_0 \equiv s_0^{\,3}$ on the right-hand side and integrate to get

$$w = w_0 + \int_\infty^s \frac{s'ds'}{s'^3 - w(s')} \approx s_0^{\,3} + \int_\infty^s \frac{s'ds'}{s'^3 - s_0^{\,3}} = s_0^{\,3} + \frac{1}{3s_0} \left[\tfrac{1}{2} \ln \frac{(s - s_0)^2}{s^2 + s_0 s + s_0^{\,2}} - \sqrt{3} \arctan \frac{\sqrt{3}s_0}{2s + s_0} \right] , \tag{6.45}$$

which is valid for

$$|w_0 - w| \ll |s^3 - w| \tag{6.46}$$

or

$$s - s_0 \gg |s_0|^{-3} \ln |s_0| . \tag{6.47}$$

Similarly, in this same region one can integrate Eqs. (6.20) and (6.28) and fit the results to Eqs. (6.36) and (6.39) in the overlap region $s \gg |s_0|$ to get

$$t_* \approx \frac{1}{3s_0^{\,2}} \left[-\frac{1}{2} \ln \frac{(s - s_0)^2}{s^2 + s_0 s + s_0^{\,2}} - \sqrt{3} \arctan \frac{\sqrt{3}s_0}{2s + s_0} \right] , \tag{6.48}$$

$$a_* \approx a_{*0}(s^3-w)^{-1/3} \approx a_{*0}(s^3-s_0^3)^{-1/3} . \quad (6.49)$$

The solution is highly inflationary when

$$|\dot{H}| \equiv \left| \frac{-2s(s^3-w)}{24\epsilon} \right| \ll H^2 \equiv \frac{s^4}{24\epsilon} , \quad (6.50)$$

which occurs for $|s^3-w| \ll |s^3|$ or $|s-s_0| \ll |s_0|$. In this part of the region (6.47), Eq. (6.48) reduces to

$$t_* \approx t_{*0} + \frac{1}{3s_0^2}\ln\left|\frac{s_0}{s-s_0}\right| , \quad (6.51)$$

where

$$t_{*0} \approx \frac{1}{6s_0^2}[\ln 3 - 3^{-1/2}\pi + 2\pi 3^{1/2}\theta(-s_0)] , \quad (6.52)$$

with the Heaviside step function $\theta(-s_0)$ needed because as s decreases from $+\infty$ to near s_0, the arctangent in Eq. (6.48) increases from 0 to near $\pi/6$ if $s_0>0$ but decreases from 0 to near $-5\pi/6$ if $s_0<0$. Similarly, Eq. (6.45) reduces to

$$w \approx s_0^3 + s_0 t_{*0} - \frac{1}{3s_0}\ln\left|\frac{3s_0}{s-s_0}\right|$$

$$\approx s_0^3 + 2s_0 t_{*0} - \tfrac{1}{3}s_0^{-1}\ln 3 - s_0 t_* . \quad (6.53)$$

To continue the evolution as s gets very near s_0, change the dependent variable from w to $v=-s(s^3-w)$, so that

$$\frac{dv}{ds} = \frac{v}{s} - s^3\left[\frac{1+3v}{v}\right] . \quad (6.54)$$

For $s^{-4} \ll |1+3v| \ll s^4$, which begins to be true when $s-s_0$ drops well below $|s_0|$ and ceases to be true only when the curve gets very near the separatrix (6.31), the first term on the right of Eq. (6.54) is negligible compared to the second term. Dropping it and integrating the separable equation that remains gives

$$s^4 \approx s_1^4 - \tfrac{4}{3}v + \tfrac{4}{9}\ln|1+3v| , \quad (6.55)$$

which matches to Eq. (6.53) in the overlap region $1 \ll |v| \ll s^4$ if $\mathrm{sgn}(1+3v) = -\mathrm{sgn} s_0$ and if

$$s_1^4 \approx s_0^4 - \tfrac{16}{9}\ln|s_0| - \tfrac{10}{9}\ln 3 - \frac{2\pi}{9\sqrt{3}}[1-6\theta(-s_0)] . \quad (6.56)$$

Integrating t_* in this region and matching to Eq. (6.51) gives

$$t_* \approx t_{*0} + \frac{1}{3s_1^2}\ln\left|\frac{9s_1^4}{1+3v}\right| ,$$
$$v \approx -\tfrac{1}{3} - 3(\mathrm{sgn} s_1)s_1^4 e^{-3s_1^2(t_*-t_{*0})} . \quad (6.57)$$

These approximations are valid for

$$s_1^4 \gg |1+3v| \gg s_1^{-4} \Longleftrightarrow 1 \ll 3s_1^2(t_*-t_{*0}) \ll 8\ln|s_1| \quad (6.58)$$

which includes all of the highly inflationary part of region (6.47) but goes beyond it to include also small values of $|1+3v|$, which are not in (6.47).

In this region, one may solve for the various quantities in terms of v or t_* to get

$$s \approx s_1 + \frac{1}{9s_1^3}(\ln|1+3v| - 3v)$$
$$\approx s_0 + |s_1| e^{-3s_1^2(t_*-t_{*0})} - \frac{1}{3s_1}(t_*-t_{*0}) , \quad (6.59)$$

$$w \approx s_1^3 + \frac{1}{3s_1}\ln|1+3v| \approx s_0^3 - s_1(t_*-t_{*0}) , \quad (6.60)$$

$$x \approx s_1^4 + \tfrac{4}{9}\ln|1+3v| - \tfrac{1}{3}v$$
$$\approx s_0^4 - \tfrac{4}{3}s_1^2(t_*-t_{*0})$$
$$+ (\mathrm{sgn} s_1)s_1^4 e^{-3s_1^2(t_*-t_{*0})} , \quad (6.61)$$

$$y \approx \tfrac{1}{2}s_1^2(v-1)$$
$$\approx -\tfrac{2}{3}s_1^2 - \tfrac{3}{2}(\mathrm{sgn} s_1)s_1^6 e^{-3s_1^2(t_*-t_{*0})} , \quad (6.62)$$

$$z \equiv \tfrac{1}{2}s^2 \approx \tfrac{1}{2}s_1^2 + \frac{1}{9s_1^2}(\ln|1+3v| - 3v)$$
$$\approx \tfrac{1}{2}s_0^2 + (\mathrm{sgn} s_1)s_1^2 e^{-3s_1(t_*-t_{*0})}$$
$$- \tfrac{1}{3}(t_*-t_{*0}) , \quad (6.63)$$

$$r = \tfrac{1}{2}\left[s^2 + \frac{v^2}{s^2}\right]$$
$$\approx \tfrac{1}{2}s_1^2 + \frac{1}{9s_1^2}(\ln|1+3v| - 3v + \tfrac{9}{2}v^2)$$
$$\approx \tfrac{1}{2}s_0^2 + 2(\mathrm{sgn} s_1)s_1^2 e^{-3s_1^2(t_*-t_{*0})}$$
$$+ \tfrac{9}{2}s_1^6 e^{-6s_1^2(t_*-t_{*0})} - \tfrac{1}{3}(t_*-t_{*0}) , \quad (6.64)$$

$$u = \tfrac{1}{2}\left[-s^2 + \frac{v^2}{s^2}\right]$$
$$\approx -\tfrac{1}{2}s_1^2 + \frac{1}{9s_1^2}(-\ln|1+3v| + 3v + \tfrac{9}{2}v^2)$$
$$\approx -\tfrac{1}{2}s_0^2 + \tfrac{9}{2}s_1^6 e^{-6s_1^2(t_*-t_{*0})} + \tfrac{1}{3}(t_*-t_{*0}) , \quad (6.65)$$

$$a_* \approx a_{*0}\left[\frac{-3s_1}{1+3v}\right]^{1/3}$$
$$\approx 3^{-1/3}a_{*0}|s_1|^{-1}e^{s_1^2(t_*-t_{*0})} , \quad (6.66)$$

$$Q_* = a_*(1+x)$$
$$\approx s_1^4 a_* \approx 3^{-1/3} a_{*0} |s_1|^3 e^{s_1^2(t_* - t_{*0})} , \quad (6.67)$$

$$a \equiv (24\epsilon)^{1/2} a_*, \quad H \equiv (24\epsilon)^{-1/2} s^2, \quad R \equiv \frac{1}{2\epsilon} x . \quad (6.68)$$

In the (s,w) plane, the curves are nearly horizontal, $w(s) \approx w_0$, for $|1+3v| \gg 1$ or $3s_1^2(t_* - t_{*0}) \ll \ln(s_1^4/\ln|s_1|)$, but then they approach closely to the separatrix $w_\infty(s) \approx s^3$ when these inequalities are reversed. In the (u,v) plane they are roughly parabolas, focused nearly on the origin, for $|1+3v| \gg 1$. At

$$t_* \approx t_{*1} \approx t_{*0} + \frac{1}{3s_1^2} \ln(9es_1^4)$$
$$\approx \frac{1}{6s_0^2}[8 \ln|s_0| + 5 \ln 3 - 3^{-1/2}\pi$$
$$+ 1 + 2\pi 3^{1/2} \theta(-s_0)] , \quad (6.69)$$

$|1+3v|$ drops below unity and u reaches its minimum value $\approx -\frac{1}{2}s_1^2$. Thereafter, u increases as the curve approaches the separatrix, $v_\infty(s) \approx -\frac{1}{3}$.

When $|1+3v| \ll 1$, so that the solution is near the separatrix (6.30), one can linearize the equation

$$\frac{d}{ds}\delta w = \frac{s\delta w}{[s^3 - w_\infty(s)][s^3 - w_\infty(s) - \delta w]} \quad (6.70)$$

for the deviation $\delta w \equiv w(s) - w_\infty(s)$, integrate with the asymptotic expansion (6.30), and match to Eq. (6.60) in the overlap region $s^{-4} \ll |1+3v| \ll 1$ to obtain

$$\delta w \approx \exp \int \frac{s\,ds}{(s^3 - w_\infty)^2} \approx \frac{-s^2}{3s_1^3} \exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.71)$$

This is valid for all $s_1^4 - s^4 \gg 1$ as s^4 decreases, until s^4 itself becomes of order unity. For $s^4 \lesssim s_1^4 - \frac{16}{3}\ln|s_1|$, $|\delta w| \le O(s^{-13})$, so that for smaller values of s^4 the correction to the separatrix is smaller than the accuracy to which it is given in Eq. (6.30). In this region one also obtains, using Eq. (6.69),

$$t_* \approx t_{*1} + \tfrac{3}{2}(s_1^2 - s^2) + O(s^{-2}) , \quad (6.72)$$

$$s \approx (\mathrm{sgn} s_1)[s_1^2 - \tfrac{2}{3}(t_* - t_{*1})]^{1/2} , \quad (6.73)$$

$$\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1 \approx (t_* - t_{*1})^2 - 3s_1^2(t_* - t_{*1}) - 1 , \quad (6.74)$$

$$w \approx s^3 - \tfrac{1}{3}s^{-1} + O(s^{-5}) - \frac{s^2}{3s_1^3}\exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.75)$$

$$x \approx s^4 - \tfrac{1}{3} + O(s^{-4}) - \frac{s^3}{3s_1^3}\exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.76)$$

$$y \approx -\tfrac{2}{3}s^2 + \tfrac{2}{27}s^{-2} + O(s^{-6}) - \frac{s^5}{6s_1^3}\exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.77)$$

$$z \equiv \tfrac{1}{2}s^2 \approx \tfrac{1}{2}s_1^2 - \tfrac{1}{3}(t_* - t_{*1}) , \quad (6.78)$$

$$r \approx \tfrac{1}{2}s^2 + \tfrac{1}{18}s^{-2} + O(s^{-6}) + \frac{s}{9s_1^3}\exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.79)$$

$$u \approx -\tfrac{1}{2}s^2 + \tfrac{1}{18}s^{-2} + O(s^{-6})$$
$$+ \frac{s}{9s_1^3}\exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.80)$$

$$v \approx -\tfrac{1}{3} + \tfrac{1}{27}s^{-4} + O(s^{-8}) - \frac{s^3}{3s_1^3}\exp(\tfrac{9}{4}s^4 - \tfrac{9}{4}s_1^4 - 1) , \quad (6.81)$$

$$a_* \approx a_{*0}(3s_1^2)^{1/3}|s|^{-1/3}\exp(-\tfrac{3}{4}s^4 + \tfrac{3}{4}s_1^4 + \tfrac{1}{3})$$
$$\approx a_{*0}(3es_1^2)^{1/3}[s_1^2 - \tfrac{2}{3}(t_* - t_{*0})]^{-1/6}$$
$$\times \exp[s_1^2(t_* - t_{*1}) - \tfrac{1}{3}(t_* - t_{*1})^2] , \quad (6.82)$$

$$Q_* \approx a_{*0}(3es_1^2)^{1/3}|s|^{11/3}\exp(-\tfrac{3}{4}s^4 + \tfrac{3}{4}s_1^4) . \quad (6.83)$$

This represents inflationary expansion with an approximately linearly decreasing Hubble expansion rate:[6]

$$H \equiv (24\epsilon)^{-1/2}s^2 \approx (24\epsilon)^{-1/2}[s_1^2 - \tfrac{2}{3}(t_* - t_{*1})]$$
$$= H_1 - \frac{1}{36\epsilon}(t - t_1) . \quad (6.84)$$

For $t_* \approx \frac{3}{2}s_1^2 \approx \frac{3}{2}s_0^2 = \frac{3}{2}w_0^{2/3}$, the inflationary expansion will come to an end, and the solution will begin spiraling around the origin in the (u,v) or (s,w) plane. In the transition region where r is of order unity, there is no useful small parameter available (such as $1/s_0$ is for $r \gg 1$) to expand in to give good approximate solutions. However, for $r \ll 1$, the curves spiral counterclockwise, inward toward the origin of the (u,v) plane with

$$r \approx \tfrac{4}{3}(\theta - \theta_0 - \tfrac{1}{2}\sin 2\theta)^{-1} , \quad (6.85)$$

where now the free parameter is the phase constant θ_0, which in principle could be determined numerically as a function of s_0. For large s_0 this phase can be shown from Eqs. (6.71) and (6.56) to have the asymptotic functional dependence

$$\theta_0(s_0) \approx \theta_\infty + \theta_1 s_1^{-3} e^{-(9/4)s_1^4} \approx \theta_\infty + 3^{5/2}\exp\left[\frac{\pi}{2\sqrt{3}}[1 - 6\theta(-s_0)]\right]\theta_1 s_0 e^{-(9/4)s_0^4} , \quad (6.86)$$

where θ_∞ is the phase constant for the separatrix, and θ_1 is another constant that could, in principle, be determined numerically.

In the region $r \ll 1$ or $t_* \gg s_0{}^2$, where $r(\theta)$ is given by Eq. (6.85), one obtains, with

$$\tau \equiv t_* - t_{*2}, \quad t_{*2} = \tfrac{3}{2} s_1{}^2 + O(1) = \tfrac{3}{2} s_0{}^2 + O(1) , \tag{6.87}$$

the following asymptotic behavior of the various quantities:

$$\theta = 2\tau + \tau^{-1}(-\cos 2\tau + \tfrac{1}{4}\cos 4\tau) + O(\tau^{-2}) , \tag{6.88}$$

$$s = (\tfrac{3}{4}\tau)^{-1/2}[\sin\tau + \tau^{-1}(\tfrac{1}{4}\theta_0 \sin\tau - \tfrac{1}{4}\cos\tau - \tfrac{1}{8}\cos 3\tau) + O(\tau^{-2})] , \tag{6.89}$$

$$w = (\tfrac{3}{4}\tau)^{-1/2}[\cos\tau + \tau^{-1}(\tfrac{1}{4}\theta_0 \cos\tau + \tfrac{3}{4}\sin\tau + \tfrac{1}{24}\sin 3\tau) + O(\tau^{-2})] , \tag{6.90}$$

$$x = \tfrac{2}{3}\tau^{-1}[\sin 2\tau + \tau^{-1}(\tfrac{1}{2}\theta_0 \sin 2\tau + \tfrac{1}{2} - \tfrac{13}{12}\cos 2\tau - \tfrac{1}{6}\cos 4\tau) + O(\tau^{-2})] , \tag{6.91}$$

$$y = \tfrac{2}{3}\tau^{-1}[\cos 2\tau + \tau^{-1}(\tfrac{1}{2}\theta_0 \cos 2\tau + \tfrac{7}{12}\sin 2\tau + \tfrac{1}{3}\sin 4\tau) + O(\tau^{-2})] , \tag{6.92}$$

$$z = \tfrac{1}{3}\tau^{-1}[1 - \cos 2\tau + \tau^{-1}(\tfrac{1}{2}\theta_0 - \tfrac{1}{2}\theta_0 \cos 2\tau - \tfrac{1}{4}\sin 2\tau - \tfrac{1}{4}\sin 4\tau) + O(\tau^{-2})] , \tag{6.93}$$

$$r = \tfrac{2}{3}\tau^{-1}[1 + \tau^{-1}(\tfrac{1}{2}\theta_0 + \tfrac{1}{4}\sin 4\tau) + O(\tau^{-2})] , \tag{6.94}$$

$$u = \tfrac{2}{3}\tau^{-1}[\cos 2\tau + \tau^{-1}(\tfrac{1}{2}\theta_0 \cos 2\tau + \tfrac{1}{4}\sin 2\tau + \tfrac{1}{2}\sin 4\tau) + O(\tau^{-2})] , \tag{6.95}$$

$$v = \tfrac{2}{3}\tau^{-1}[\sin 2\tau + \tau^{-1}(\tfrac{1}{2}\theta_0 \sin 2\tau - \tfrac{1}{2} + \tfrac{1}{4}\cos 2\tau - \tfrac{1}{2}\cos 4\tau) + O(\tau^{-2})] , \tag{6.96}$$

$$a_* = a_{*0}\alpha(s_1) s_1{}^{2/3} e^{3s_1{}^4/4} \tau^{2/3}[1 + \tau^{-1}(-\tfrac{1}{3}\theta_0 - \tfrac{1}{3}\sin 2\tau) + \tau^{-2}(\tfrac{1}{18}\theta_0{}^2 - \tfrac{1}{18}\theta_0 \sin 2\tau + \tfrac{1}{36} + \tfrac{1}{4}\cos 2\tau + \tfrac{1}{72}\cos 4\tau) + O(\tau^{-3})] , \tag{6.97}$$

$$Q_* = a_{*0}\alpha(s_1) s_1{}^{2/3} e^{3s_1{}^4/4} \tau^{2/3}[1 + \tau^{-1}(-\tfrac{1}{3}\theta_0 + \tfrac{1}{3}\sin 2\tau) + \tau^{-2}(\tfrac{1}{18}\theta_0{}^2 + \tfrac{1}{18}\theta_0 \sin 2\tau + \tfrac{1}{4} - \tfrac{17}{36}\cos 2\tau + \tfrac{1}{72}\cos 4\tau) + O(\tau^{-3})] , \tag{6.98}$$

where $\alpha(s_1)$ is a function of s_1 (or of s_0) which appears to be asymptotically constant and of order unity for large $|s_1|$ (or at least varying more slowly than a power law); it could, in principle, be determined by numerical integration. One may use Eq. (6.56) to write

$$s_1{}^{2/3} e^{3s_1{}^4/4} \approx 3^{-5/6} \exp\left[-\frac{\pi}{6\sqrt{3}}[1 - 6\theta(-s_0)] \right] \times s_0{}^{-2/3} e^{3s_0{}^4/4} \tag{6.99}$$

in terms of s_0 for large $|s_0|$. Here, as in Eqs. (6.52), (6.56), and (6.86), $\theta(-s_0)$ is the Heaviside step function, which should not be confused with the phase constant θ_0 or the angle θ of Eq. (6.88).

The behavior of Eqs. (6.88)–(6.98) is that of small oscillation about "dustlike" or "matter-dominated" evolution[17] with the asymptotically constant "energy" from Eq. (4.11) being

$$E_* = \tfrac{4}{9} a_{*0}{}^3 \alpha^3 s_1{}^2 e^{9s_1{}^4/4} + O(\tau^{-1}) . \tag{6.100}$$

Using the definitions (2.20)–(2.22), (6.7), and (6.17) of the various quantities above, one sees that the leading-order behavior agrees with that of Eqs. (2.23)–(2.25) of Mijić *et al.*[6] if $\tau = \omega(t - t_{os}) + \pi/2$, but the $O(\tau^{-1})$ corrections differ.

Thus, we see that the $R + \epsilon R^2$ $k=0$ solutions with large positive or negative initial values $w_0 \equiv s_0{}^3$ of $w = (2\epsilon^3/3)^{1/4} R / H^{1/2}$ expand initially [$t_* \equiv (24\epsilon)^{-1/2} t \lesssim s_0{}^{-2}$] with $a \propto t^{1/2}$, like radiation-dominated Einstein solutions, and then enter a period of exponential or inflationary expansion for $s_0{}^{-2} \lesssim t_* \lesssim \tfrac{3}{2} s_0{}^2$. For $\tfrac{3}{2} s_0{}^2 \lesssim t_* < \infty$, the solutions behave as dust-dominated Einstein solutions with $a \propto t^{2/3}$ when averaged over each period in the oscillation of $x \equiv 2\epsilon R$. However, for solutions with $|s_0| \lesssim 1$, there is little or no inflation between the "radiation" era ($a \propto t^{1/2}$) and the "dustlike" era ($a \propto t^{2/3}$). In the (u,v) plane, these noninflationary solutions have $v \sim -(2u)^{2/3}$ as u decreases from $+\infty$ all the way down to some number of order unity, where these solutions begin spiralling around the origin without first approaching the separatrix.

Now we may consider how the measure associated with these solutions depends upon the amount of inflation. First, we calculate how much inflation occurs. For concreteness, define inflation to be the first continuous epoch during which

$$|H^2/\dot{H}| \equiv |\tfrac{1}{2}s^4/v| > p \tag{6.101}$$

for some positive p which is not too small. For example, $a_* \propto t^p$ gives $H^2/\dot{H} = -p$, as does a $k=0$ Einstein solution with the ratio of pressure to energy density being $(2-3p)/(3p)$ (and, hence, near -1 for highly inflationary behavior with p large). By counting only the first continuous epoch in which (6.101) holds, we are excluding the short period of its validity during each oscillation of $\dot{H} = (12\epsilon)^{-1} v$ throughout the "dustlike" regime.

For $1 \ll p \ll s_0{}^4$, which will be assumed henceforth, inflation will begin at a time when Eq. (6.55) is valid, at

$$v \approx \frac{9s_1{}^4 + 4\ln(\tfrac{3}{2}s_1{}^4/p)}{-18p\,\mathrm{sgn}s_1 + 12} \approx -\frac{s_1{}^4}{2p}\mathrm{sgn}s_1 , \tag{6.102}$$

and end at a time Eq. (6.81) is valid, at

$$s^4 \approx \tfrac{2}{3}p - \tfrac{1}{9} \approx \tfrac{2}{3}p \ . \quad (6.103)$$

Then using Eqs. (6.66) and (6.82) to evaluate the linear size of the Universe at these two times and taking the ratio gives an expansion factor during inflation of

$$1+Z \equiv \frac{a_* \text{ (end)}}{a_* \text{ (beginning)}} \approx \left[\frac{3es_1{}^4}{2p}\right]^{5/12} \exp\left[\tfrac{3}{4}s_1{}^4 - \tfrac{1}{2}p\right] \ . \quad (6.104)$$

One can now insert Eq. (6.56) to reexpress this in terms of $s_0 \equiv w_0{}^{1/3}$:

$$1+Z \approx (2p)^{-5/12} e^{-p/2} \exp\left[\frac{5}{12} - \frac{\pi}{6\sqrt{3}}[1-6\theta(-s_0)]\right] |s_0{}^{1/3}| e^{3s_0{}^4/4} \ . \quad (6.105)$$

For $\ln(1+X) \gg \frac{1}{2}p \gg 1$, this can be inverted approximately to give

$$s_0 \approx \pm[\tfrac{4}{3}\ln(1+Z) - \tfrac{1}{9}\ln\ln(1+Z) + \tfrac{2}{3}p + \tfrac{5}{9}\ln p]^{1/4} \approx \pm[\tfrac{4}{3}\ln(1+Z)]^{1/4} \ . \quad (6.106)$$

We may now define an inflationary solution as one in which the expansion factor $1+Z$ during the regime (6.101) of inflation is at least some minimum amount, say $1+Z_m$. Then if this is much larger than $e^{p/2}$, we find that the solutions are inflationary if and only if

$$|w_0| = |s_0{}^3| \geq w_m \approx [\tfrac{4}{3}\ln(1+Z_m)]^{3/4} \ . \quad (6.107)$$

Next, we calculate the measure for these various solutions. In terms of the parameters of Eqs. (6.42) and (6.43), the measure two-form (3.2) at $a_* = 0$ is

$$\omega = Q_*{}^2 dQ_* \wedge d\{[d(Q_*{}^3)/d(a_*{}^3)]^{-1/2}\}$$
$$= \tfrac{2}{3} dw_0 \wedge d(a_{*0}{}^3) = 6s_0{}^2 ds_0 \wedge a_{*0}{}^2 da_{*0} \ . \quad (6.108)$$

This of course gives an infinite measure when integrated over all real w_0 and all positive a_{*0}. It also diverges for any nonzero range of w_0 when integrated over all a_{*0}, so the total measure for either the inflationary solutions, satisfying the inequality (6.107), or the noninflationary solutions, not satisfying (6.107), is each infinite. The ratio of these two infinite quantities is then arbitrary, depending on how it is taken, so, just as for the Einstein solutions with a massive scalar field,[8] the canonical measure gives an ambiguous probability for inflation in FRW models with the $R+\epsilon R^2$ Lagrangian.

As an example of how to get different values for the ratio of the measure of inflationary and noninflationary solutions, consider the procedure of taking the initial solutions to be at fixed r, so that the initial data are a_* and $\theta \bmod 2\pi$. Before fixing r, one can use the definitions (6.7), (6.12), and (6.17) and insert (for $k=0$)

$$a'_* \equiv \frac{da_*}{d\eta} \equiv a_* \frac{da_*}{dt_*} \equiv 2a_*{}^2 z \equiv a_*{}^2 s^2 \ , \quad (6.109)$$

$$Q_* \equiv a_*(1+x) \equiv a_*(1+sw) \ , \quad (6.110)$$

$$Q'_* \equiv \frac{dQ_*}{d\eta} \equiv a_* \frac{dQ_*}{dt_*} = \frac{a_*{}^2 x^2}{2z} \equiv a_*{}^2 w^2 \ , \quad (6.111)$$

into the measure two-form (3.2) to get

$$\omega = -2a_*{}^2[a_* s^2 ds \wedge dw + (w-s^3) dw \wedge da_* - s\, da_* \wedge ds]$$

$$= -2a_*{}^2 \left[\frac{da_*}{dt_*} ds \wedge dw + \frac{ds}{dt_*} dw \wedge da_* + \frac{dw}{dt_*} da_* \wedge ds\right]$$

$$= \frac{a_*{}^2}{r}\left[\frac{da_*}{dt_*} du \wedge dv + \frac{du}{dt_*} dv \wedge da_* + \frac{dv}{dt_*} da_* \wedge du\right]$$

$$= a_*{}^2 \left[\frac{da_*}{dt_*} dr \wedge d\theta + \frac{dr}{dt_*} d\theta \wedge da_* + \frac{d\theta}{dt_*} da_* \wedge dr\right] \ . \quad (6.112)$$

On a surface of fixed r, Eq. (6.13) reduces this to

$$\omega = r^2 d(a_*{}^3) \wedge \sin^2\theta\, d\theta \ . \quad (6.113)$$

The integral of (6.113) diverges for each range of θ, giving an infinite measure for both inflationary and noninflationary solutions. However, one gets a finite measure for each finite range of a_*, so one might calculate the probability of inflation as the fraction of solutions within each range of a_* which have inflation. Since the measure (6.113) factorizes, the fraction of inflationary solutions for each range of a_* is

$$f_I(r,p,1+Z_m) = \frac{1}{\pi} \int \sin^2\theta\, d\theta \ , \quad (6.114)$$

where θ is integrated over the range for which the Universe has an expansion by at least a factor of $1+Z_m$ while the inflationary criterion (6.101) holds. For $2\ln(1+Z_m) \gg p \gtrsim 1$, this means to integrate (6.114) over the range of θ for which the parameter w_0 obeys (6.107). Now, the point is that even for fixed criteria p and $1+Z_m$ defining which solutions are inflationary, the range of θ and the resulting integral (6.114) depend on the value of r, defined by Eq. (6.7), at which it is evaluated.

For example, at $r \gg w_m{}^{2/3} \approx [\frac{4}{3}\ln(1+Z_m)]^{1/2}$, almost all values of θ give $w_0{}^2 > w_m{}^2$, and hence sufficient inflation, except for a narrow range with θ just below 0.

There the solution has $w \approx w_0$ and so Eq. (6.19) implies that for $r \gg 1+w_0^{2/3}$ or $u \gg 1+|v|$ the curve in the (u,v) plane has the approximate form

$$u \approx r-[(2r)^{1/2}+w_0]^{2/3} ,$$
$$v \approx -(2r)^{1/2}[(2r)^{1/2}+w_0]^{1/3} , \tag{6.115}$$

thus giving

$$\theta \approx \sin^{-1}(v/r) \approx -2(2r)^{-1/2}[(2r)^{1/2}+w_0]^{1/3} . \tag{6.116}$$

Then integrating (6.114) over all θ in $-\pi<\theta<\pi$ except for the values (6.116) which give $|w_0|<w_m$ yields

$$f_I(r \gg 1, 1 \lesssim p \ll r^2, e^{p/2} \ll 1+Z_m \ll e^{r^2/2}) \approx 1-\frac{16}{3\pi}w_m(2r)^{-3/2} \approx 1-\frac{16}{3\pi}[\tfrac{1}{3}\ln(1+Z_m)]^{3/4}r^{-3/2} . \tag{6.117}$$

The answer depends upon r, but for $r \gg w_m{}^{2/3}$, as assumed, it is very near unity, so one might conclude that nearly all solutions have the required amount of inflation.

On the other hand, at $r \ll 1$ only a narrow range of θ near the separatrix corresponds to solutions with significant inflation (which of course occurs at larger values of r, i.e., earlier in time). From Eqs. (6.85), (6.86), and (6.104), one obtains

$$\theta-\tfrac{1}{2}\sin 2\theta-\frac{4}{3r} \approx \theta_0 \approx \theta_\infty \pm \theta_1 \left[\frac{3e}{2p}\right]^{5/4} e^{-3p/2}[\tfrac{4}{3}\ln(1+Z)]^{1/2}(1+Z)^{-3} , \tag{6.118}$$

where θ_1 is the undetermined constant in Eq. (6.86). Differentiating this at fixed $r \ll 1$ and inserting the result into Eq. (6.114) with the restriction $1+Z>1+Z_m$ gives

$$f_I(r \ll 1, p \gtrsim 1, 1+Z_m \gg e^{p/2}) \approx \frac{1}{2\pi}\int d\theta_0 \approx 2\theta_1 \left[\frac{3e}{2p}\right]^{5/4} e^{-3p/2}[\tfrac{4}{3}\ln(1+Z_m)]^{1/2}(1+Z_m)^{-3} . \tag{6.119}$$

Therefore, when the criterion for inflation requires a large minimum-expansion factor $1+Z_m$ in a, the fraction of inflationary solutions within any fixed range of the scale factor a is small when evaluated at small r given by Eq. (6.7).

A similar conclusion would result if one evaluated the ratio of inflationary and noninflationary solutions at fixed values of t_*, ρ_*, s^2+w^2, or virtually any other monotonically varying function of time. The analysis above was done at fixed values of $r=(6\epsilon)^{1/2}(H+6\epsilon\dot{H}^2/H)$ because this is a fairly simple monotonically decreasing variable which is locally defined (unlike t_*, which has a locally arbitrary constant of integration) and which is positive and finite at all $0<a_*<\infty$ (unlike ρ_*, which diverges at $1+x \equiv 1+2\epsilon R=0$), though

$$s^2+w^2=(24\epsilon)^{1/2}(H+\tfrac{1}{6}\epsilon R^2/H)$$
$$=(24\epsilon)^{1/2}[H+6\epsilon H^{-1}(\dot{H}+2H^2)^2] \tag{6.120}$$

also has the same desirable properties.

The ambiguity of the fraction of inflationary solutions within a fixed range of a under a change of r is not in contradiction with the fact that the GHS canonical measure is preserved under the Hamiltonian evolution of the data from one initial hypersurface to another in the constrained phase space.[7] Rather, what is happening is that the solutions at one r within a fixed range of $a_*{}^3$ (that is, a range that does not depend on θ) do not remain within a fixed range of $a_*{}^3$ as r decreases with time. Those in the inflationary range of θ, which appear to dominate at large r, undergo more expansion in going from large r to small r, so they spread out into a much bigger range of $a_*{}^3$ than their noninflationary counterparts. In the large range of $a_*{}^3$ that these inflationary solutions exist at small r, they are dominated by a different set of noninflationary solutions which evolved to that range of sizes from a much larger range of $a_*{}^3$ at large r than the inflationary solutions had then . Therefore, the set that will dominate within a given range of sizes depends upon the value of r at which the measure is evaluated.

ACKNOWLEDGMENTS

This work was stimulated by conversations with Stephen Hawking, Milan Mijić, and Michael Morris. Financial support was provided in part by National Science Foundation Grants Nos. PHY-8316811, PHY-8701860, and AST-8414911 and by the John Simon Guggenheim Memorial Foundation.

[1] A. H. Guth, Phys. Rev. D **23**, 347 (1981).
[2] A. D. Linde, Phys. Lett. **129B**, 177 (1983).
[3] A. D. Linde, Rep. Prog. Phys. **47**, 925 (1984).
[4] B. Whitt, Phys. Lett. **145B**, 176 (1984).
[5] S. W. Hawking and J. C. Luttrell, Nucl. Phys. **B247**, 250 (1984).
[6] L. A. Kofman, A. D. Linde, and A. A. Starobinsky, Phys. Lett. **157B**, 361 (1985); M. B. Mijić, M. S. Morris, and W.-M. Suen, Phys. Rev. D **34**, 2934 (1986).
[7] G. W. Gibbons, S. W. Hawking, and J. M. Stewart, Nucl. Phys. **B281**, 736 (1987).
[8] S. W. Hawking and D. N. Page, Penn State report, 1987 (un-

published).
[9]K. S. Stelle, Gen. Relativ. Gravit. **9**, 353 (1978); P. Teyssandier and Ph. Tourrenc, J. Math. Phys. **24**, 2793 (1983).
[10]D. Boulware, in *Quantum Theory of Gravity: Essays in Honor of the 60th Birthday of Bryce S. DeWitt,* edited by S. M. Christensen (Hilger, Bristol, 1984).
[11]V. A. Belinsky, L. P. Grishchuk, I. M. Khalatnikov, and Ya. B. Zeldovich, Phys. Lett. **155B**, 232 (1985); Zh. Eksp. Teor. Fiz. **89**, 346 (1985) [Sov. Phys. JETP **62**, 195 (1985)].
[12]M. B. Mijić, M. S. Morris, and W.-M. Suen (in preparation); A. A. Starobinsky (private communication).
[13]D. N. Page, Class. Quantum Gravit. **1**, 417 (1984).
[14]S. W. Hawking, Nucl. Phys. **B239**, 257 (1984).
[15]A. Frenkel and K. Brecher, Phys. Rev. D **26**, 368 (1982).
[16]G. M. Murphy, *Ordinary Differential Equations and Their Solutions* (Van Nostrand, Princeton, 1960), p. 23.
[17]V. Ts. Gurovich and A. A. Starobinsky, Zh. Eksp. Teor. Fiz. **77**, 1683 (1979) [Sov. Phys. JETP **50**, 844 (1979)].