PHYSICAL REVIEW D

PARTICLES AND FIELDS

THIRD SERIES, VOLUME 31, NUMBER 8 — 15 APRIL 1985

Origin of structure in the Universe

J. J. Halliwell and S. W. Hawking

Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge CB3 9EW, United Kingdom and Max Planck Institut for Physics and Astrophysics, Foehringer Ring 6, Munich, Federal Republic of Germany

(Received 17 December 1984)

It is assumed that the Universe is in the quantum state defined by a path integral over compact four-metrics. This can be regarded as a boundary condition for the wave function of the Universe on superspace, the space of all three-metrics and matter field configurations on a three-surface. We extend previous work on finite-dimensional approximations to superspace to the full infinite-dimensional space. We treat the two homogeneous and isotropic degrees of freedom exactly and the others to second order. We justify this approximation by showing that the inhomogeneous or anisotropic modes start off in their ground state. We derive time-dependent Schrödinger equations for each mode. The modes remain in their ground state until their wavelength exceeds the horizon size in the period of exponential expansion. The ground-state fluctuations are then amplified by the subsequent expansion and the modes reenter the horizon in the matter- or radiation-dominated era in a highly excited state. We obtain a scale-free spectrum of density perturbations which could account for the origin of galaxies and all other structure in the Universe. The fluctuations would be compatible with observations of the microwave background if the mass of the scalar field that drives the inflation is 10^{14} GeV or less.

I. INTRODUCTION

Observations of the microwave background indicate that the Universe is very close to homogeneity and isotropy on a large scale. Yet we know that the early Universe cannot have been completely homogeneous and isotropic because in that case galaxies and stars would not have formed. In the standard hot big-bang model the density perturbations required to produce these structures have to be assumed as initial conditions. However, in the inflationary model of the Universe[1–4] it was possible to show that the ground-state fluctuations of the scalar field that causes the exponential expansion would lead to a spectrum of density perturbations that was almost scale free.[5–7] In the simplest grand-unified-theory (GUT) inflationary model the amplitude of the density perturbations was too large but an amplitude that was consistent with observation could be obtained in other models with a different potential for the scalar field.[8] Similarly, ground-state fluctuations of the gravitational-wave modes would lead to a spectrum of long-wavelength gravitational waves that would be consistent with observation provided that the Hubble constant H in the inflationary period was not more than about 10^{-4} of the Planck mass.[9]

One cannot regard these results as a completely satisfactory explanation of the origin of structure in the Universe because the inflationary model does not make any assumption about the initial or boundary conditions of the Universe. In particular, it does not guarantee that there should be a period of exponential expansion in which the scalar field and the gravitational-wave modes would be in the ground state. In the absence of some assumption about the boundary conditions of the Universe, any present state would be possible: one could pick an arbitrary state for the Universe at the present time and evolve it backward in time to see what initial conditions it arose from. It has recently been proposed[10–13] that the boundary conditions of the Universe are that it has no boundary. In other words, the quantum state of the Universe is defined by a path integral over compact four-metrics without boundary. The quantum state can be described by a wave function Ψ which is a function on the infinite-dimensional space W called superspace which consists of all three-metrics h_{ij} and matter field configurations Φ_0 on a three-surface S. Because the wave function does not depend on time explicitly, it obeys a system of zero-energy Schrödinger equations, one for each choice of the shift N_i and the lapse N on S. The Schrödinger equations can be decomposed into the momentum constraints, which imply that the wave function is the same at all points of W that are related by coordinate transformations, and the Wheeler-DeWitt equations, which can be regarded as a system of second-order differential equations for Ψ on W. The requirement that the wave function be given by a path integral over compact four-metrics then becomes a set of boundary conditions for the Wheeler-DeWitt equations which determines a unique solution for Ψ.

It is difficult to solve differential equations on an infinite-dimensional manifold. Attention has therefore been concentrated on finite-dimensional approximations to W, called "minisuperspaces." In other words, one restricts the number of gravitational and matter degrees of freedom to a finite number and then solves the Wheeler-DeWitt equations on a finite-dimensional manifold with boundary conditions that reflect the fact that the wave function is given by a path integral over compact four-metrics. In particular,[12–15] it has been shown that in the case of a homogeneous isotropic closed universe of radius a with a massive scalar field ϕ the wave function corresponds in the classical limit to a family of classical solutions which have a long period of exponential or "inflationary" expansion and then go over to a matter-dominated expansion, reach a maximum radius, and then collapse in a time-symmetric manner. This model would be in agreement with observation but, because it is so restricted, the only prediction it can make is that the observed value of the density parameter Ω should be exactly one.[15] The aim of this paper is to extend this minisuperspace model to the full number of degrees of freedom of the gravitational and scalar fields. We treat the 2 degrees of freedom of the minisuperspace model exactly and we expand the other inhomogeneous and anisotropic degrees of freedom to second order in the Hamiltonian. In the region of W in which Ψ oscillates rapidly, one can use the WKB approximation to relate the wave function to a family of classical solutions and so introduce a concept of time. As in the minisuperspace case, the family includes solutions with a long period of exponential expansion. We show that the gravitational-wave and density-perturbation modes obey decoupled time-dependent Schrödinger equations with respect to the time parameter of the classical solution. The boundary conditions imply that these modes start off in the ground state. While they remain within the horizon of the exponentially expanding phase, they can relax adiabatically and so they remain in the ground state. However, when they expand outside the horizon of the inflationary period, they become "frozen" until they reenter the horizon in the matter-dominated era. They then give rise to gravitational waves and a scale-free spectrum of density perturbations. These would be consistent with the observations of the microwave background and could be large enough to explain the origins of galaxies if the mass of the scalar field were about 10^{-5} of the Planck mass. Thus the proposal that the quantum state of the Universe is defined by a path integral over compact four-metrics seems to be able to account for the origin of structure in the Universe: it arises, not from arbitrary initial conditions, but from the ground-state fluctuations that have to be present by the Heisenberg uncertainty principle.

In Sec. II we review the Hamiltonian formalism of classical general relativity, and in Sec. III we show how this leads to the canonical treatment of the quantum theory. In Sec. IV we summarize earlier work[13] on a homogeneous isotropic minisuperspace model with a massive scalar field. We extend this to all the matter and gravitational degrees of freedom in Sec. V, treating the inhomogeneous modes to second order in the Hamiltonian. In Sec. VI we decompose the wave function into a background term which obeys an equation similar to that of the unperturbed minisuperspace model, and perturbation terms which obey time-dependent Schrödinger equations. We use the path-integral expression for the wave function in Sec. VII to show that the perturbation wave functions start out in their ground states. Their subsequent evolution is described in Sec. VIII. In Sec. IX we calculate the anisotropy that these perturbations would produce in the microwave background and compare with observation. In Sec. X we summarize the paper and conclude that the proposed quantum state could account not only for the large-scale homogeneity and isotropy but also for the structure on smaller scales.

II. CANONICAL FORMULATION OF GENERAL RELATIVITY

We consider a compact three-surface S which divides the four-manifold M into two parts. In a neighborhood of S one can introduce a coordinate t such that S is the surface $t=0$ and coordinates x^i ($i=1,2,3$). The metric takes the form

$$ds^2=-(N^2-N_iN^i)dt^2+2N_i dx^i dt+h_{ij}dx^i dx^j \,. \tag{2.1}$$

N is called the lapse function. It measure the proper-time separation of surfaces of constant t. N_i is called the shift vector. It measures the deviation of the lines of constant x^i from the normal to the surface S. The action is

$$I=\int (L_g+L_m)d^3x\,dt \,, \tag{2.2}$$

where

$$L_g=\frac{m_P{}^2}{16\pi}N(G^{ijkl}K_{ij}K_{kl}+h^{1/2}\,{}^3R)\,, \tag{2.3}$$

$$K_{ij}=\frac{1}{2N}\left[-\frac{\partial h_{ij}}{\partial t}+2N_{(i|j)}\right], \tag{2.4}$$

is the second fundamental form of S, and

$$G^{ijkl}=\tfrac{1}{2}h^{1/2}(h^{ik}h^{jl}+h^{il}h^{jk}-2h^{ij}h^{kl})\,. \tag{2.5}$$

In the case of a massive scalar field Φ

$$L_m=\tfrac{1}{2}Nh^{1/2}\left[N^{-2}\left[\frac{\partial\Phi}{\partial t}\right]^2-2\frac{N^i}{N^2}\frac{\partial\Phi}{\partial t}\frac{\partial\Phi}{\partial x^i}-\left[h^{ij}-\frac{N^iN^j}{N^2}\right]\frac{\partial\Phi\partial\Phi}{\partial x^i\partial x^j}-m^2\Phi^2\right]. \tag{2.6}$$

In the Hamiltonian treatment of general relativity one regards the components h_{ij} of the three-metric and the field Φ as the canonical coordinates. The canonically conjugate momenta are

$$\pi^{ij}=\frac{\partial L_g}{\partial \dot{h}_{ij}}=-\frac{h^{1/2}m_P{}^2}{16\pi}(K^{ij}-h^{ij}K)\ , \tag{2.7}$$

$$\pi_\Phi=\frac{\partial L_m}{\partial \dot{\Phi}}=N^{-1}h^{1/2}\left[\dot{\Phi}-N^i\frac{\partial \Phi}{\partial x^i}\right] . \tag{2.8}$$

The Hamiltonian is

$$H=\int (\pi^{ij}\dot{h}_{ij}+\pi_\Phi\dot{\Phi}-L_g-L_m)d^3x$$
$$=\int (NH_0+N_iH^i)d^3x\ , \tag{2.9}$$

where

$$H_0=16\pi m_P{}^{-2}G_{ijkl}\pi^{ij}\pi^{kl}-\frac{m_P{}^2}{16\pi}h^{1/2}\,{}^3R$$
$$+\tfrac{1}{2}h^{1/2}\left[\frac{\pi_\Phi{}^2}{h}+h^{ij}\frac{\partial\Phi\partial\Phi}{\partial x^i\partial x^j}+m^2\Phi^2\right] , \tag{2.10}$$

$$H^i=-2\pi^{ij}{}_{|j}+h^{ij}\frac{\partial \Phi}{\partial x^j}\pi_\Phi\ , \tag{2.11}$$

and

$$G_{ijkl}=\tfrac{1}{2}h^{-1/2}(h_{ik}h_{jl}+h_{il}h_{jk}-h_{ij}h_{kl})\ . \tag{2.12}$$

The quantities N and N_i are regarded as Lagrange multipliers. Thus the solution obeys the momentum constraint

$$H^i=0 \tag{2.13}$$

and the Hamiltonian constraint

$$H_0=0\ . \tag{2.14}$$

For given fields N and N^i on S the equations of motion are

$$\dot{h}_{ij}=\frac{\partial H}{\partial \pi^{ij}},\quad \dot{\pi}^{ij}=-\frac{\partial H}{\partial h_{ij}}\ ,$$
$$\dot{\Phi}=\frac{\partial H}{\partial \pi_\Phi},\quad \dot{\pi}_\Phi=-\frac{\partial H}{\partial \Phi}\ . \tag{2.15}$$

III. QUANTIZATION

The quantum state of the Universe can be described by a wave function Ψ which is a function on the infinite-dimensional manifold W of all three-metrics h_{ij} and matter fields Φ on S. A tangent vector to W is a pair of fields (γ_{ij},μ) on S where γ_{ij} can be regarded as a small change of the metric h_{ij} and μ can be regarded as a small change of Φ. For each choice of $N>0$ on S there is a natural metric $\Gamma(N)$ on W:[15]

$$ds^2=\int N^{-1}\left[\frac{m_P{}^2}{32\pi}G^{ijkl}\gamma_{ij}\gamma_{kl}+\tfrac{1}{2}h^{1/2}\mu^2\right]d^3x\ . \tag{3.1}$$

The wave function Ψ does not depend explicitly on the time t because t is just a coordinate which can be given arbitrary values by different choices of the undetermined multipliers N and N_i. This means that Ψ obeys the zero-energy Schrödinger equation:

$$H\Psi=0\ . \tag{3.2}$$

The Hamiltonian operator H is the classical Hamiltonian with the usual substitutions:

$$\pi^{ij}(x)\to -i\frac{\delta}{\delta h_{ij}(x)},\quad \pi_\phi(x)\to -i\frac{\delta}{\delta\phi(x)}\ . \tag{3.3}$$

Because N and N_i are regarded as independent Lagrange multipliers, the Schrödinger equation can be decomposed into two parts. There is the momentum constraint

$$H_-\Psi\equiv\int N_iH^id^3x\ \Psi$$
$$=\int h^{1/2}N_i\left[2\left[\frac{\delta}{\delta h_{ij}(x)}\right]_{|j}-h^{ij}\frac{\partial\Phi}{\partial x^j}\frac{\delta}{\delta\Phi(x)}\right]d^3x\ \Psi$$
$$=0\ . \tag{3.4}$$

This implies that Ψ is the same on three-metrics and matter field configurations that are related by coordinate transformations in S. The other part of the Schrödinger equation, corresponding to $H_|\Psi=0$, where $H_|=\int NH_0d^3x$ is called the Wheeler-DeWitt equation. There is one Wheeler-DeWitt equation for each choice of N on S. One can regard them as a system of second-order partial differential equations for Ψ on W. There is some ambiguity in the choice of operator ordering in these equations but this will not affect the results of this paper. We shall assume that $H_|$ has the form[15]

$$(-\tfrac{1}{2}\nabla^2+\xi\mathbb{R}+V)\Psi=0\ , \tag{3.5}$$

where ∇^2 is the Laplacian in the metric $\Gamma(N)$. R is the curvature scalar of this metric and the potential V is

$$V=\int h^{1/2}N\left[-\frac{m_P{}^2}{16\pi}{}^3R+\epsilon+U\right]d^3x\ , \tag{3.6}$$

where $U=T^{00}-\frac{1}{2}\pi_\Phi{}^2$. The constant ϵ can be regarded as a renormalization of the cosmological constant Λ. We shall assume that the renormalized Λ is zero. We shall also assume that the coefficient ξ of the scalar curvature $\mathbb{R}$ of W is zero.

Any wave function Ψ which satisfies the momentum constraint and the Wheeler-DeWitt equation for each choice of N and N_i on S describes a possible quantum state of the Universe. We shall be concerned with the particular solution which represents the quantum state defined by a path integral over compact four-metrics without boundary. In this case[11-13]

$$\Psi=\int d[g_{\mu\nu}]d[\Phi]\exp[-\hat{I}(g_{\mu\nu},\Phi)]\ , \tag{3.7}$$

where $\hat{I}$ is the Euclidean action obtained by setting N negative imaginary and the path integral is taken over all compact four-metrics $g_{\mu\nu}$ and matter fields Φ which are bounded by S on which the three-metric is h_{ij} and the matter field is Φ. One can regard (3.7) as a boundary condition on the Wheeler-DeWitt equations. It implies that Ψ tends to a constant, which can be normalized to one, as h_{ij} goes to zero.

IV. UNPERTURBED FRIEDMANN MODEL

References 12–14 considered the minisuperspace model which consisted of a Friedmann model with metric

$$ds^2=\sigma^2(-N^2dt^2+a^2d\Omega_3{}^2)\ , \tag{4.1}$$

where $d\Omega_3{}^2$ is the metric of the unit three-sphere. The normalization factor $\sigma^2=2/3\pi m_P{}^2$ has been included for convenience. The model contains a scalar field $(2^{1/2}\pi\sigma)^{-1}\phi$ with mass $\sigma^{-1}m$ which is constant on surfaces of constant t. One can easily generalize this to the case of a scalar field with a potential $V(\phi)$. Such generalizations include models with higher-derivative quantum corrections.[16] The action is

$$I=-\tfrac{1}{2}\int dt\, Na^3\left[\frac{1}{N^2a^2}\left[\frac{da}{dt}\right]^2-\frac{1}{a^2}-\frac{1}{N^2}\left[\frac{d\phi}{dt}\right]^2+m^2\phi^2\right]\ . \tag{4.2}$$

The classical Hamiltonian is

$$H=\tfrac{1}{2}N(-a^{-1}\pi_a{}^2+a^{-3}\pi_\phi{}^2-a+a^3m^2\phi^2)\ , \tag{4.3}$$

where

$$\pi_a=-\frac{a\,da}{N\,dt},\quad \pi_\phi=\frac{a^3d\phi}{N\,dt}\ . \tag{4.4}$$

The classical Hamiltonian constraint is $H=0$. The classical field equations are

$$N\frac{d}{dt}\left[\frac{1}{N}\frac{d\phi}{dt}\right]+\frac{3}{a}\frac{da}{dt}\frac{d\phi}{dt}+N^2m^2\phi=0\ , \tag{4.5}$$

$$N\frac{d}{dt}\left[\frac{1}{N}\frac{da}{dt}\right]=N^2am^2\phi^2-2a\left[\frac{d\phi}{dt}\right]^2\ . \tag{4.6}$$

The Wheeler-DeWitt equation is

$$\tfrac{1}{2}N\,e^{-3\alpha}\left[\frac{\partial^2}{\partial\alpha^2}-\frac{\partial^2}{\partial\phi^2}+2V\right]\Psi(\alpha,\phi)=0\ , \tag{4.7}$$

where

$$V=\tfrac{1}{2}(e^{6\alpha}m^2\phi^2-e^{4\alpha}) \tag{4.8}$$

and $\alpha=\ln a$. One can regard Eq. (4.7) as a hyperbolic equation for Ψ in the flat space with coordinates (α,ϕ) with α as the time coordinate. The boundary condition that gives the quantum state defined by a path integral over compact four-metrics is $\Psi\to 1$ as $\alpha\to-\infty$. If one integrates Eq. (4.7) with this boundary condition, one finds that the wave function starts oscillating in the region $V>0$, $|\phi|>1$ (this has been confirmed numerically[14]). One can interpret the oscillatory component of the wave function by the WKB approximation:

$$\Psi=\mathrm{Re}(C\,e^{iS})\ , \tag{4.9}$$

where C is a slowly varying amplitude and S is a rapidly varying phase. One chooses S to satisfy the classical Hamilton-Jacobi equation:

$$H(\pi_\alpha,\pi_\phi,\alpha,\phi)=0\ , \tag{4.10}$$

where

$$\pi_\alpha=\frac{\partial S}{\partial\alpha},\quad \pi_\phi=\frac{\partial S}{\partial\phi}\ . \tag{4.11}$$

One can write (4.10) in the form

$$\tfrac{1}{2}f^{ab}\frac{\partial S\partial S}{\partial q^a\partial q^b}+e^{-3\alpha}V=0\ , \tag{4.12}$$

where f^{ab} is the inverse to the metric $\Gamma(1)$:

$$f^{ab}=e^{-3\alpha}\mathrm{diag}(-1,1)\ . \tag{4.13}$$

The wave function (4.9) will then satisfy the Wheeler-DeWitt equation if

$$\nabla^2C+2if^{ab}\frac{\partial C\partial S}{\partial q^a\partial q^b}+iC\nabla^2S=0\ , \tag{4.14}$$

where ∇^2 is the Laplacian in the metric f_{ab}. One can ignore the first term in Eq. (4.14) and can integrate the equation along the trajectories of the vector field $X^a=dq^a/dt=f^{ab}\partial S/\partial q^b$ and so determine the amplitude C. These trajectories correspond to classical solutions of the field equations. They are parametrized by the coordinate time t of the classical solutions.

The solutions that correspond to the oscillating part of the wave function of the minisuperspace model start out at $V=0$, $|\phi|>1$ with $d\alpha/dt=d\phi/dt=0$. They expand exponentially with

$$S=-\tfrac{1}{3}e^{3\alpha}m\,|\phi|(1-m^{-2}e^{-2\alpha}\phi^{-2})$$
$$\approx-\tfrac{1}{3}e^{3\alpha}m\,|\phi|\ , \tag{4.15}$$

$$\frac{d\alpha}{dt}=m\,|\phi|,\quad \frac{d\,|\phi|}{dt}=-\tfrac{1}{3}m\ . \tag{4.16}$$

After a time of order $3m^{-1}(|\phi_1|-1)$, where ϕ_1 is the initial value of ϕ, the field ϕ starts to oscillate with frequency m. The solution then becomes matter dominated and expands with e^α proportional to $t^{2/3}$. If there were other fields present, the massive scalar particles would decay into light particles and then the solution would expand with e^α proportional to $t^{1/2}$. Eventually the solution would reach a maximum radius of order $\exp(9\phi_1{}^2/2)$ or $\exp(9\phi_1{}^2)$ depending on whether it is radiation or matter dominated for most of the expansion. The solution would then recollapse in a similar manner.

V. THE PERTURBED FRIEDMANN MODEL

We assume that the metric is of the form (2.1) except the right hand side has been multiplied by a normalization factor σ^2. The three-metric h_{ij} has the form

$$h_{ij}=a^2(\Omega_{ij}+\epsilon_{ij})\ , \tag{5.1}$$

where Ω_{ij} is the metric on the unit three-sphere and ϵ_{ij} is a perturbation on this metric and may be expanded in harmonics:

$$\epsilon_{ij}=\sum_{n,l,m}[6^{1/2}a_{nlm}\tfrac{1}{3}\Omega_{ij}Q^n_{lm}+6^{1/2}b_{nlm}(P_{ij})^n_{lm}+2^{1/2}c^0_{nlm}(S^0_{ij})^n_{lm}+2^{1/2}c^e_{nlm}(S^e_{ij})^n_{lm}+2\,d^0_{nlm}(G^0_{ij})^n_{lm}+2\,d^e_{nlm}(G^e_{ij})^n_{lm}]\ . \tag{5.2}$$

The coefficients $a_{nlm}, b_{nlm}, c^0_{nlm}, c^e_{nlm}, d^0_{nlm}, d^e_{nlm}$ are functions of the time coordinate t but not the three spatial coordinates x^i.

The $Q(x^i)$ are the standard scalar harmonics on the three-sphere. The $P_{ij}(x^i)$ are given by (suppressing all but the i,j indices)

$$P_{ij}=\frac{1}{(n^2-1)}Q_{|ij}+\tfrac{1}{3}\Omega_{ij}Q\ . \tag{5.3}$$

They are traceless, $P_i{}^i=0$. The S_{ij} are defined by

$$S_{ij}=S_{i|j}+S_{j|i}\ , \tag{5.4}$$

where S_i are the transverse vector harmonics, $S_i{}^{|i}=0$. The G_{ij} are the transverse traceless tensor harmonics $G_i{}^i=G_{ij}{}^{|j}=0$. Further details about the harmonics and their normalization can be found in Appendix A.

The lapse, shift, and the scalar field $\Phi(x^i,t)$ can be expanded in terms of harmonics:

$$N=N_0\left[1+6^{-1/2}\sum_{n,l,m}g_{nlm}Q^n_{lm}\right]\ , \tag{5.5}$$

$$N_i=e^{\alpha}\sum_{n,l,m}[6^{-1/2}k_{nlm}(P_i)^n_{lm}+2^{1/2}j_{nlm}(S_i)^n_{lm}]\ , \tag{5.6}$$

$$\Phi=\sigma^{-1}\left[\frac{1}{2^{1/2}\pi}\phi(t)+\sum_{n,l,m}f_{nlm}Q^n_{lm}\right]\ , \tag{5.7}$$

where $P_i=[1/(n^2-1)]Q_{|i}$. Hereafter, the labels n, l, m, o, and e will be denoted simply by n. One can then expand the action to all orders in terms of the "background" quantities a,ϕ,N_0 but only to second order in the "perturbations" $a_n,b_n,c_n,d_n,f_n,g_n,k_n,j_n$:

$$I=I_0(a,\phi,N_0)+\sum_n I_n\ , \tag{5.8}$$

where I_0 is the action of the unperturbed model (4.2) and I_n is quadratic in the perturbations and is given in Appendix B.

One can define conjugate momenta in the usual manner. They are

$$\pi_\alpha=-N_0{}^{-1}e^{3\alpha}\dot\alpha+\text{quadratic terms}\ , \tag{5.9}$$

$$\pi_\phi=N_0{}^{-1}e^{3\alpha}\dot\phi+\text{quadratic terms}\ , \tag{5.10}$$

$$\pi_{a_n}=-N_0{}^{-1}e^{3\alpha}[\dot a_n+\dot\alpha(a_n-g_n)+\tfrac{1}{3}e^{-\alpha}k_n]\ , \tag{5.11}$$

$$\pi_{b_n}=N_0{}^{-1}e^{3\alpha}\frac{(n^2-4)}{(n^2-1)}(\dot b_n+4\dot\alpha b_n-\tfrac{1}{3}e^{-\alpha}k_n)\ , \tag{5.12}$$

$$\pi_{c_n}=N_0{}^{-1}e^{3\alpha}(n^2-4)(\dot c_n+4\dot\alpha c_n-e^{-\alpha}j_n)\ , \tag{5.13}$$

$$\pi_{d_n}=N_0{}^{-1}e^{3\alpha}(\dot d_n+4\dot\alpha d_n)\ , \tag{5.14}$$

$$\pi_{f_n}=N_0{}^{-1}e^{3\alpha}[\dot f_n+\dot\phi(3a_n-g_n)]\ . \tag{5.15}$$

The quadratic terms in Eqs. (5.9) and (5.10) are given in Appendix B. The Hamiltonian can then be expressed in terms of these momenta and the other quantities:

$$H=N_0\left[H_{|0}+\sum_n H^n_{|2}+\sum_n g_nH^n_{|1}\right]+\sum_n(k_n{}^SH^n_{\_1}+j_n{}^VH^n_{\_1})\ . \tag{5.16}$$

The subscripts 0,1,2 on the $H_|$ and $H_\_$ denote the orders of the quantities in the perturbations and S and V denote the scalar and vector parts of the shift part of the Hamiltonian. $H_{|0}$ is the Hamiltonian of the unperturbed model with $N=1$:

$$H_{|0}=\tfrac{1}{2}e^{-3\alpha}(-\pi_\alpha{}^2+\pi_\phi{}^2+e^{6\alpha}m^2\phi^2-e^{4\alpha})\ . \tag{5.17}$$

The second-order Hamiltonian is given by

$$H_{|2}=\sum_n H^n_{|2}=\sum_n({}^SH^n_{|2}+{}^VH^n_{|2}+{}^TH^n_{|2})\ ,$$

where

$$\begin{aligned}{}^SH^n_{|2}=\tfrac{1}{2}e^{-3\alpha}\Bigg[&\left[\tfrac{1}{2}a_n{}^2+\frac{10(n^2-4)}{(n^2-1)}b_n{}^2\right]\pi_\alpha{}^2+\left[\tfrac{15}{2}a_n{}^2+\frac{6(n^2-4)}{(n^2-1)}b_n{}^2\right]\pi_\phi{}^2\\&-\pi_{a_n}{}^2+\frac{(n^2-1)}{(n^2-4)}\pi_{b_n}{}^2+\pi_{f_n}{}^2+2a_n\pi_{a_n}\pi_\alpha+8b_n\pi_{b_n}\pi_\alpha-6a_n\pi_{f_n}\pi_\phi\\&-e^{4\alpha}\left[\tfrac{1}{3}(n^2-\tfrac{5}{2})a_n{}^2+\frac{(n^2-7)}{3}\frac{(n^2-4)}{(n^2-1)}b_n{}^2+\tfrac{2}{3}(n^2-4)a_nb_n-(n^2-1)f_n{}^2\right]\\&+e^{6\alpha}m^2(f_n{}^2+6a_nf_n\phi)+e^{6\alpha}m^2\phi^2\left[\tfrac{3}{2}a_n{}^2-\frac{6(n^2-4)}{(n^2-1)}b_n{}^2\right]\Bigg]\ ,\end{aligned} \tag{5.18}$$

$${}^VH^n_{|2}=\tfrac{1}{2}e^{-3\alpha}\left[(n^2-4)c_n{}^2(10\pi_\alpha{}^2+6\pi_\phi{}^2)+\frac{1}{(n^2-4)}\pi_{c_n}{}^2+8c_n\pi_{c_n}\pi_\alpha+(n^2-4)c_n{}^2(2\,e^{4\alpha}-6\,e^{6\alpha}m^2\phi^2)\right]\ , \tag{5.19}$$

$${}^TH^n_{|2}=\tfrac{1}{2}e^{-3\alpha}\{d_n{}^2(10\pi_\alpha{}^2+6\pi_\phi{}^2)+\pi_{d_n}{}^2+8d_n\pi_{d_n}\pi_\alpha+d_n{}^2[(n^2+1)e^{4\alpha}-6\,e^{6\alpha}m^2\phi^2]\}\ . \tag{5.20}$$

The first-order Hamiltonians are

$$H^n_{|1}=\tfrac{1}{2}e^{-3\alpha}\{-a_n(\pi_\alpha{}^2+3\pi_\phi{}^2)+2(\pi_\phi\pi_{f_n}-\pi_\alpha\pi_{a_n})+m^2e^{6\alpha}(2f_n\phi+3a_n\phi^2)-\tfrac{2}{3}e^{4\alpha}[(n^2-4)b_n+(n^2+\tfrac{1}{2})a_n]\}\ . \tag{5.21}$$

The shift parts of the Hamiltonian are

$$^{S}H^n_{-1}=\tfrac{1}{3}e^{-3\alpha}\left[-\pi_{a_n}+\pi_{b_n}+\left[a_n+\frac{4(n^2-4)}{(n^2-1)}b_n\right]\pi_\alpha+3f_n\pi_\phi\right]\ , \tag{5.22}$$

$$^{V}H^n_{-1}=e^{-\alpha}[\pi_{c_n}+4(n^2-4)c_n\pi_\alpha]\ . \tag{5.23}$$

The classical field equations are given in Appendix B.

Because the Lagrange multipliers N_0, g_n, k_n, j_n are independent, the zero energy Schrödinger equation

$$H\Psi=0 \tag{5.24}$$

can be decomposed as before into momentum constraints and Wheeler-DeWitt equations. As the momentum constraints are linear in the momenta, there is no ambiguity in the operator ordering. One therefore has

$$^{S}H^n_{-1}\Psi=-\tfrac{1}{3}e^{-3\alpha}\left[\frac{\partial}{\partial a_n}-\left[a_n+\frac{4(n^2-4)}{(n^2-1)}b_n\right]\frac{\partial}{\partial\alpha}-\frac{\partial}{\partial b_n}-3f_n\frac{\partial}{\partial\phi}\right]\Psi=0\ , \tag{5.25}$$

$$^{V}H^n_{-1}\Psi=e^{-\alpha}\left[\frac{\partial}{\partial c_n}+4(n^2-4)c_n\frac{\partial}{\partial\alpha}\right]\Psi=0\ . \tag{5.26}$$

The first-order Hamiltonians $H^n_{|1}$ give a series of finite dimensional second-order differential equations, one for each n. In the order of approximation that we are using, the ambiguity in the operator ordering will consist of the possible addition of terms linear in $\partial/\partial\alpha$. The effect of such terms can be compensated for by multiplying the wave function by powers of e^{α}. This will not affect the relative probabilities of different observations at a given value of α. We shall therefore ignore such ambiguities and terms:

$$\tfrac{1}{2}e^{-3\alpha}\left[a_n\left[\frac{\partial^2}{\partial\alpha^2}+3\frac{\partial^2}{\partial\phi^2}\right]-2\left[\frac{\partial^2}{\partial f_n\partial\phi}-\frac{\partial^2}{\partial a_n\partial\alpha}\right]+m^2e^{6\alpha}[2\phi f_n+3a_n\phi^2]-\tfrac{2}{3}e^{4\alpha}[(n^2-4)b_n+(n^2+\tfrac{1}{2})a_n]\right]\Psi=0\ . \tag{5.27}$$

Finally, one has an infinite-dimensional second-order differential equation

$$\left[H_{|0}+\sum_n({}^{S}H^n_{|2}+{}^{V}H^n_{|2}+{}^{T}H^n_{|2})\right]\Psi=0\ , \tag{5.28}$$

where $H_{|0}$ is the operator in the Wheeler-DeWitt equation of the unperturbed Friedmann minisuperspace model:

$$H_{|0}=\tfrac{1}{2}e^{-3\alpha}\left[\frac{\partial^2}{\partial\alpha^2}-\frac{\partial^2}{\partial\phi^2}+e^{6\alpha}m^2\phi^2-e^{4\alpha}\right] \tag{5.29}$$

and

$$\begin{aligned}^{S}H^n_{|2}=\tfrac{1}{2}e^{-3\alpha}\Bigg[&-\left[\tfrac{1}{2}a_n{}^2+\frac{10(n^2-4)}{(n^2-1)}b_n{}^2\right]\frac{\partial^2}{\partial\alpha^2}-\left[\tfrac{15}{2}a_n{}^2+\frac{6(n^2-4)}{(n^2-1)}b_n{}^2\right]\frac{\partial^2}{\partial\phi^2}\\ &+\frac{\partial^2}{\partial a_n{}^2}-\frac{(n^2-1)}{(n^2-4)}\frac{\partial^2}{\partial b_n{}^2}-\frac{\partial^2}{\partial f_n{}^2}-2a_n\frac{\partial^2}{\partial a_n\partial\alpha}-8b_n\frac{\partial^2}{\partial b_n\partial\alpha}+6a_n\frac{\partial^2}{\partial f_n\partial\phi}\\ &-e^{4\alpha}\left[\tfrac{1}{3}(n^2-\tfrac{5}{2})a_n{}^2+\frac{(n^2-7)}{3}\frac{(n^2-4)}{(n^2-1)}b_n{}^2+\tfrac{2}{3}(n^2-4)a_nb_n-(n^2-1)f_n{}^2\right]\\ &+e^{6\alpha}m^2(f_n{}^2+6a_nf_n\phi)+e^{6\alpha}m^2\phi^2\left[\tfrac{3}{2}a_n{}^2-\frac{6(n^2-4)}{(n^2-1)}b_n{}^2\right]\Bigg]\ ,\end{aligned} \tag{5.30}$$

$$^{V}H^n_{|2}=\tfrac{1}{2}e^{-3\alpha}\left[-(n^2-4)c_n{}^2\left[10\frac{\partial^2}{\partial\alpha^2}+6\frac{\partial^2}{\partial\phi^2}\right]-\frac{1}{(n^2-4)}\frac{\partial^2}{\partial c_n{}^2}-8c_n\frac{\partial^2}{\partial c_n\partial\alpha}+(n^2-4)c_n{}^2(2e^{4\alpha}-6e^{6\alpha}m^2\phi^2)\right]\ , \tag{5.31}$$

$$^{T}H^n_{|2}=\tfrac{1}{2}e^{-3\alpha}\left[-d_n{}^2\left[10\frac{\partial^2}{\partial\alpha^2}+6\frac{\partial^2}{\partial\phi^2}\right]-\frac{\partial^2}{\partial d_n{}^2}-8d_n\frac{\partial^2}{\partial d_n\partial\alpha}+d_n{}^2[(n^2+1)e^{4\alpha}-6e^{6\alpha}m^2\phi^2]\right]\ . \tag{5.32}$$

We shall call Eq. (5.28) the master equation. It is not hyperbolic because, as well as the positive second derivatives $\partial^2/\partial\alpha^2$ in $H_{|0}$, there are the positive second derivatives $\partial^2/\partial a_n{}^2$ in each ${}^S H^n_{|2}$. However, one can use the momentum constraint (5.25) to substitute for the partial derivatives with respect to a_n and then solve the resultant differential equation on $a_n=0$. Similarly, one can use the momentum constraint (5.26) to substitute for the partial derivatives with respect to c_n and then solve on $c_n=0$. One thus obtains a modified equation which is hyperbolic for small f_n. If one knows the wave function on $a_n=0=c_n$, one can use the momentum constraints to calculate the wave function at other values of a_n and c_n.

VI. THE WAVE FUNCTION

Because the perturbation modes are not coupled to each other, the wave function can be expressed as a sum of terms of the form

$$\Psi=\mathrm{Re}\left[\Psi_0(\alpha,\phi)\prod_n \Psi^{(n)}(\alpha,\phi,a_n,b_n,c_n,d_n,f_n)\right]$$
$$=\mathrm{Re}(C\,e^{iS}) , \tag{6.1}$$

where S is a rapidly varying function of α and ϕ and C is a slowly varying function of all the variables. If one substitutes (6.1) into the master equation and divides by Ψ, one obtains

$$-\frac{\nabla_2{}^2\Psi_0}{2\Psi_0}-\sum_n \frac{\nabla_2{}^2\Psi^{(n)}}{2\Psi^{(n)}}-\sum_{n\lessgtr m}\frac{(\nabla_2\Psi^{(n)})\cdot(\nabla_2\Psi^{(m)})}{2\Psi^{(n)}\Psi^{(m)}}$$
$$-\frac{(\nabla_2\Psi_0)}{\Psi_0}\cdot\left[\sum_n \frac{\nabla_2\Psi^{(n)}}{\Psi^{(n)}}\right]$$
$$+\sum_n \frac{H^n_{|2}\Psi}{\Psi}+e^{-3\alpha}V(\alpha,\phi)=0 , \tag{6.2}$$

where $\nabla_2{}^2$ is the Laplacian in the minisuperspace metric $f_{ab}=e^{3\alpha}\mathrm{diag}(-1,1)$ and the dot product is with respect to this metric.

An individual perturbation mode does not contribute a significant fraction of the sums in the third and fourth terms in Eq. (6.2). Thus these terms can be replaced by

$$-\frac{(\nabla_2\Psi)}{\Psi}\cdot\sum_n \frac{(\nabla_2\Psi^{(n)})}{\Psi^{(n)}}+\frac{1}{2}\left[\sum_n \frac{\nabla_2\Psi^{(n)}}{\Psi^{(n)}}\right]^2$$
$$\approx -i(\nabla_2 S)\cdot\sum_n \frac{(\nabla_2\Psi^{(n)})}{\Psi^{(n)}}+\frac{1}{2}\left[\sum_n \frac{\nabla_2\Psi^{(n)}}{\Psi^{(n)}}\right]^2 . \tag{6.3}$$

In order that the ansatz (6.1) be valid, the terms in (6.2) that depend on a_n,b_n,c_n,d_n,f_n have to cancel out. This implies

$$\frac{(\nabla_2\Psi)}{\Psi}\cdot(\nabla_2\Psi^{(n)})+\tfrac{1}{2}\nabla_2{}^2\Psi^{(n)}=\frac{H^n_{|2}\Psi}{\Psi}\Psi^{(n)} , \tag{6.4}$$

$$(-\tfrac{1}{2}\nabla_2{}^2+e^{-3\alpha}V+\tfrac{1}{2}J\cdot J)\Psi_0=0 , \tag{6.5}$$

where

$$J=\sum_n \frac{\nabla_2\Psi^{(n)}}{\Psi^{(n)}} .$$

In regions in which the phase S is a rapidly varying function of α and ϕ, one can neglect the second term in (6.4) in comparison with the first term. One can also replace the π_α and π_ϕ which appear in $H^n_{|2}$ by $\partial S/\partial\alpha$ and $\partial S/\partial\phi$, respectively. The vector $X^a=f^{ab}\partial S/\partial q^b$ obtained by raising the covector $\nabla_2 S$ by the inverse minisuperspace metric f^{ab} can be regarded as $\partial/\partial t$ where t is the time parameter of the classical Friedmann metric that corresponds to Ψ by the WKB approximation. One then obtains a time dependent Schrödinger equation for each mode along a trajectory of the vector field X^a:

$$i\frac{\partial\Psi^{(n)}}{\partial t}=H^n_{|2}\Psi^{(n)} . \tag{6.6}$$

Equation (6.5) can be interpreted as the Wheeler-DeWitt equation for a two-dimensional minisuperspace model with an extra term $\frac{1}{2}J\cdot J$ arising from the perturbations. In order to make J finite, one will have to make subtractions. Subtracting out the ground-state energies of the $H^n_{|2}$ corresponds to a renormalization of the cosmological constant Λ. There is a second subtraction which corresponds to a renormalization of the Planck mass m_P and a third one which corresponds to a curvature-squared counterterm. The effect of such higher-derivative terms in the action has been considered elsewhere.[16]

One can write $\Psi^{(n)}$ as

$$\Psi^{(n)}={}^S\Psi^{(n)}(\alpha,\phi,a_n,b_n,f_n)\,{}^V\Psi^{(n)}(\alpha,\phi,c_n)\,{}^T\Psi^{(n)}(\alpha,\phi,d_n) , \tag{6.7}$$

where ${}^S\Psi^{(n)}$, ${}^V\Psi^{(n)}$, and ${}^T\Psi^{(n)}$ obey independent Schrödinger equations with ${}^SH^n_{|2}$, ${}^VH^n_{|2}$, and ${}^TH^n_{|2}$, respectively.

VII. THE BOUNDARY CONDITIONS

We want to find the solution of the master equation that corresponds to

$$\Psi[h_{ij},\Phi]=\int d[g_{\mu\nu}]d[\Phi]\exp(-\hat{I}) , \tag{7.1}$$

where the integral is taken over all compact four-metrics and matter fields which are bounded by the three-surface S. If one takes the scale parameter α to be very negative but keeps the other parameters fixed, the Euclidean action $\hat{I}$ tends to zero like $e^{2\alpha}$. Thus one would expect Ψ to tend to one as α tends to minus infinity.

One can estimate the form of the scalar, vector, and tensor parts ${}^S\Psi^{(n)}$, ${}^V\Psi^{(n)}$, ${}^T\Psi^{(n)}$ of the perturbation $\Psi^{(n)}$ from the path integral (7.1) One takes the four-metric $g_{\mu\nu}$ and the scalar field Φ to be of the background form

$$ds^2=\sigma^2(-N^2dt^2+e^{2\alpha(t)}d\Omega_3{}^2) \tag{7.2}$$

and $\phi(t)$, respectively, plus a small perturbation described by the variables (a_n,b_n,f_n), c_n, and d_n as functions of t. In order for the background four-metric to be compact, it has to be Euclidean when $\alpha=-\infty$, i.e., N has to be purely negative imaginary at $\alpha=-\infty$, which we shall take to be $t=0$. In regions in which the metric is Lorentzian, N

will be real and positive. In order to allow a smooth transition from Euclidean to Lorentzian, we shall take N to be of the form $-i\,e^{i\mu}$ where $\mu=0$ at $t=0$. In order that the four-metric and the scalar field be regular at $t=0, a_n, b_n, c_n, d_n, f_n$ have to vanish there.

The tensor perturbations d_n have the Euclidean action

$$^T\hat{I}_n=\tfrac{1}{2}\int dt\, d_n\, {}^TD d_n + \text{boundary term}\ , \tag{7.3}$$

where

$$^TD=\left[-\frac{d}{dt}\left[\frac{e^{3\alpha}d}{iN_0 dt}\right]+iN_0e^{\alpha}(n^2-1)\right]+4iN_0e^{3\alpha}\left[+\tfrac{1}{2}e^{-2\alpha}-\tfrac{3}{2}m^2\phi^2-\frac{3\dot{\phi}^2}{2(iN_0)^2}-\frac{3\dot{\alpha}^2}{2(iN_0)^2}-\frac{1d}{iN_0 dt}\left[\frac{\dot{\alpha}}{iN_0}\right]\right] . \tag{7.4}$$

The last term in (7.4) vanishes if the background metric satisfies the background field equations. The action is extremized when d_n satisfies the equation

$$^TDd_n=0\ . \tag{7.5}$$

For a d_n that satisfies (7.5), the action is just the boundary term

$$^T\hat{I}^{\,\text{cl}}_n=\frac{1}{2iN_0}e^{3\alpha}(d_n\dot{d}_n+4\dot{\alpha}d_n^2)\ . \tag{7.6}$$

The path integral over d_n will be

$$\int d[d_n]\exp(-{}^T\hat{I}_n)=(\det{}^TD)^{-1/2}\exp(-{}^T\hat{I}^{\,\text{cl}}_n)\ . \tag{7.7}$$

One now has to integrate (7.7) over different background metrics to obtain the wave function $^T\Psi^{(n)}$. One expects the dominant contribution to come from background metrics that are near a solution of the classical background field equations. For such metrics one can employ the adiabatic approximation in which one regards α to be a slowly varying function of t. Then the solution of (7.5) which obeys the boundary condition $d_n=0$ at $t=0$ is

$$d_n=A(e^{\nu\tau}-e^{-\nu\tau})\ , \tag{7.8}$$

where $\nu=e^{-\alpha}(n^2-1)^{1/2}$ and $\tau=\int iN_0 dt$. This approximation will be valid for background fields which are near a solution of the background field equations and for which

$$\left|\frac{\dot{\alpha}}{N_0}\right| \ll n\,e^{-\alpha}\ . \tag{7.9}$$

For a regular Euclidean metric, $|\dot{\alpha}/N_0|=e^{-\alpha}$ near $t=0$. If the metric is a Euclidean solution of the background field equations, then $|\dot{\alpha}/N_0|<e^{-\alpha}$. Thus the adiabatic approximation should hold for large values of n into the region in which the solution of the background field equations becomes Lorentzian and the WKB approximation can be used. The wave function $^T\Psi^{(n)}$ will then be

$$^T\Psi^{(n)}=B\exp\left[-\left[\tfrac{1}{2}n\,e^{2\alpha}\coth(\nu\tau)+\frac{2}{iN_0}\dot{\alpha}\,e^{3\alpha}\right]d_n^2\right] . \tag{7.10}$$

In the Euclidean region, τ will be real and positive. For large values of n, $\coth(\nu\tau)\approx 1$. In the Lorentzian region where the WKB approximation applies, τ will be complex but it will still have a positive real part and $\coth(\nu\tau)$ will still be approximately 1 for large n. Thus

$$^T\Psi^{(n)}=B\exp\left[-2i\frac{\partial S}{\partial\alpha}d_n^2-\tfrac{1}{2}n\,e^{2\alpha}d_n^2\right] . \tag{7.11}$$

The normalization constant B can be chosen to be 1. Thus, apart from a phase factor, the gravitational-wave modes enter the WKB region in their ground state.

We now consider the vector part $^V\Psi^{(n)}$ of the wave function. This is pure gauge as the quantities c_n can be given any value by gauge transformations parametrized by the j_n. The freedom to make gauge transformations is reflected quantum mechanically in the constraint

$$e^{-\alpha}\left[\frac{\partial}{\partial c_n}+4(n^2-4)c_n\frac{\partial}{\partial\alpha}\right]\Psi=0\ . \tag{7.12}$$

One can integrate (7.12) to give

$$\Psi(\alpha,\{c_n\})=\Psi\left[\alpha-2\sum_n(n^2-4)c_n^2,0\right] , \tag{7.13}$$

where the dependence on the other variables has been suppressed. One can also replace $\partial\Psi/\partial\alpha$ by $i(\partial S/\partial\alpha)\Psi$. One can then solve for $^V\Psi^{(n)}$:

$$^V\Psi^{(n)}=\exp\left[2i(n^2-4)c_n^2\frac{\partial S}{\partial\alpha}\right] . \tag{7.14}$$

The scalar perturbation modes a_n, b_n, and f_n involve a combination of the behavior of the tensor and vector perturbations. The scalar part of the action is given in Appendix B. The action is extremized by solutions of the classical equations

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{a}_n}{N_0}\right]+\tfrac{1}{3}(n^2-4)N_0^2e^{\alpha}(a_n+b_n)+3\,e^{3\alpha}(\dot{\phi}\dot{f}_n-N_0^2m^2\phi f_n)$$
$$=N_0^2[3\,e^{3\alpha}m^2\phi^2-\tfrac{1}{3}(n^2+2)e^{\alpha}]g_n+e^{3\alpha}\dot{\alpha}\dot{g}_n-\tfrac{1}{3}N_0\frac{d}{dt}\left[e^{2\alpha}\frac{k_n}{N_0}\right] \tag{7.15}$$

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{b}_n}{N_0}\right]-\tfrac{1}{3}(n^2-1)N_0{}^2e^{\alpha}(a_n+b_n)=\tfrac{1}{3}(n^2-1)N_0{}^2e^{\alpha}g_n+\tfrac{1}{3}N_0\frac{d}{dt}\left[e^{2\alpha}\frac{k_n}{N_0}\right], \tag{7.16}$$

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{f}_n}{N_0}\right]+3e^{3\alpha}\dot{\phi}\,\dot{a}_n+N_0{}^2[m^2e^{3\alpha}+(n^2-1)e^{\alpha}]f_n=e^{3\alpha}(-2N_0{}^2m^2\phi g_n+\dot{\phi}\,\dot{g}_n-e^{-\alpha}\dot{\phi}k_n)\ . \tag{7.17}$$

There is a three-parameter family of solutions to (7.15)—(7.17) which obey the boundary condition $a_n=b_n=f_n=0$ at $t=0$. There are however, two constraint equations:

$$\dot{a}_n+\frac{(n^2-4)}{(n^2-1)}\dot{b}_n+3f_n\dot{\phi}=\dot{\alpha}g_n-\frac{e^{-\alpha}}{(n^2-1)}k_n\ , \tag{7.18}$$

$$3a_n(-\dot{\alpha}^2+\dot{\phi}^2)+2(\dot{\phi}\dot{f}_n-\dot{\alpha}\dot{a}_n)+N_0{}^2m^2(2f_n\phi+3a_n\phi^2)-\tfrac{2}{3}N_0{}^2e^{-2\alpha}[(n^2-4)b_n+(n^2+\tfrac{1}{2})a_n]$$
$$=\tfrac{2}{3}\dot{\alpha}e^{-\alpha}k_n+2g_n(-\dot{\alpha}^2+\dot{\phi}^2)\ . \tag{7.19}$$

These correspond to the two gauge degrees of freedom parametrized by k_n and g_n, respectively. The Euclidean action for a solution to Eqs. (7.15)—(7.19) is

$$^{S}\hat{I}^{\,\mathrm{cl}}_n=\frac{1}{2iN_0}e^{3\alpha}\left[-a_n\dot{a}_n+\frac{(n^2-4)}{(n^2-1)}b_n\dot{b}_n+f_n\dot{f}_n+\dot{\alpha}\left[-a_n{}^2+\frac{4(n^2-4)}{(n^2-1)}b_n{}^2\right]+3\dot{\phi}a_nf_n+g_n(\dot{\alpha}a_n-\dot{\phi}f_n)\right.$$
$$\left.-\tfrac{1}{3}e^{-\alpha}k_n\left[a_n+\frac{(n^2-4)}{(n^2-1)}b_n\right]\right], \tag{7.20}$$

where the background field equations have been used.

In many ways the simplest gauge to work in is that with $g_n=k_n=0$. However, this gauge does not allow one to find a compact four-metric which is bounded by a three-surface with arbitrary values of a_n, b_n, and f_n and which is a solution of the Eqs. (7.15)—(7.17) and the constraint equations. Instead, we shall use the gauge $a_n=b_n=0$ and shall solve the constraint Eqs. (7.18) and (7.19) to find g_n and k_n:

$$g_n=3\frac{(n^2-1)\dot{\alpha}\dot{\phi}f_n+\dot{\phi}\dot{f}_n+N_0{}^2m^2\phi f_n}{(n^2-4)\dot{\alpha}^2+3\dot{\phi}^2}\ , \tag{7.21}$$

$$k_n=3(n^2-1)e^{\alpha}\frac{\dot{\alpha}\dot{\phi}\dot{f}_n+N_0{}^2m^2\phi f_n\dot{\alpha}-3f_n\dot{\phi}(-\dot{\alpha}^2+\dot{\phi}^2)}{(n^2-4)\dot{\alpha}^2+3\dot{\phi}^2}\ . \tag{7.22}$$

With these substituted, (7.17) becomes a second-order equation for f_n,

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{f}_n}{N_0}\right]+N_0{}^2[m^2e^{3\alpha}+(n^2-1)e^{\alpha}]f_n=e^{3\alpha}(-2N_0{}^2m^2\phi g_n+\dot{\phi}\,\dot{g}_n-e^{-\alpha}\dot{\phi}k_n)\ . \tag{7.23}$$

For large n we can again use the adiabatic approximation to estimate the solution of (7.23) when $|\phi|>1$:

$$f_n=A\sinh(\nu\tau)\ , \tag{7.24}$$

where $\nu^2=e^{-2\alpha}(n^2-1)$. Thus for these modes

$$^{S}\Psi^{(n)}(\alpha,\phi,0,0,f_n)\approx\exp\left[-\tfrac{1}{2}n\,e^{2\alpha}f_n{}^2-\tfrac{1}{2}i\frac{\partial S}{\partial\phi}g_nf_n\right]\ . \tag{7.25}$$

This is of the ground-state form apart from a small phase factor. The value of $^{S}\Psi^{(n)}$ at nonzero values of a_n and b_n can be found by integrating the constraint equations (5.25) and (5.27).

The tensor and scalar modes start off in their ground states, apart possibly from the modes at low n. The vector modes are pure gauge and can be neglected. Thus the total energy

$$E=\sum_n\frac{H^{(n)}_{|2}\Psi^{(n)}}{\Psi^{(n)}}$$

of the perturbations will be small when the ground-state energies are subtracted. But $E=i(\nabla_2S)\cdot J$ where $J=\sum_n\nabla_2\Psi^{(n)}/\Psi^{(n)}$. Thus J is small. This means that the wave function Ψ_0 will obey the Wheeler-DeWitt equation of the unperturbed minisuperspace model and the phase factor S will be approximately $-i\ln\Psi_0$. However the homogeneous scalar field mode ϕ will not start out in its ground state. There are two reasons for this: first, regularity at $t=0$ requires $a_n=b_n=c_n=d_n=f_n=0$, but does not require $\phi=0$. Second, the classical field equation for ϕ is of the form of a damped harmonic oscillator with a constant frequency m rather than a decreasing frequency $e^{-\alpha}n$. This means that the adiabatic approximation is not valid at small t and that the solution of the classical field equation is ϕ approximately constant. The action of such solutions is small, so large values of $|\phi|$ are not damped as they are for the other variables. Thus the WKB trajectories which start out from large values of $|\phi|$ have high probability. They will correspond to classical solutions which have a long inflationary period and then go over to a matter-dominated expansion. In a realistic model which included other fields of low rest mass, the matter energy in the oscillations of the massive scalar field would decay into light particles with a thermal spectrum. The model would then expand as a radiation-dominated universe.

VIII. GROWTH OF PERTURBATIONS

The tensor modes will obey the Schrödinger equation

$$i\frac{\partial\,{}^T\Psi^{(n)}}{\partial t}={}^TH^n_{|2}\,{}^T\Psi^{(n)} \tag{8.1}$$

$$=\tfrac{1}{2}e^{-3\alpha}\left\{+d_n{}^2\left[10\left[\frac{\partial S}{\partial\alpha}\right]^2+6\left[\frac{\partial S}{\partial\phi}\right]^2\right]-\frac{\partial^2}{\partial d_n{}^2}-8d_n i\frac{\partial S}{\partial\alpha}\frac{\partial}{\partial d_n}+d_n{}^2[(n^2+1)e^{4\alpha}-6e^{6\alpha}m^2\phi^2]\right\}. \tag{8.2}$$

One can write

$${}^T\Psi^{(n)}=\exp(-2\alpha)\exp\left[-2i\frac{\partial S}{\partial\alpha}d_n{}^2\right]{}^T\Psi_0^{(n)}\,, \tag{8.3}$$

then

$$i\frac{\partial\,{}^T\Psi_0^{(n)}}{\partial t}=\tfrac{1}{2}e^{-3\alpha}\left[-\frac{\partial^2}{\partial d_n{}^2}+d_n{}^2(n^2-1)e^{4\alpha}\right]{}^T\Psi_0^{(n)}\,. \tag{8.4}$$

The WKB approximation to the background Wheeler-DeWitt equation has been used in deriving (8.4). Then (8.4) has the form of the Schrödinger equation for an oscillator with a time-dependent frequency $\nu=(n^2-1)^{1/2}e^{-\alpha}$. Initially the wave function ${}^T\Psi_0^{(n)}$ will be in the ground state (apart from a normalization factor) and the frequency ν will be large compared to $\dot\alpha$. In this case one can use the adiabatic approximation to show that ${}^T\Psi_0^{(n)}$ remains in the ground state

$${}^T\Psi_0^{(n)}\approx\exp(-\tfrac{1}{2}n\,e^{2\alpha}d_n{}^2)\,. \tag{8.5}$$

The adiabatic approximation will break down when $\nu\approx\dot\alpha$, i.e., the wave length of the gravitational mode becomes equal to the horizon scale in the inflationary period. The wave function ${}^T\Psi_0^{(n)}$ will then freeze

$${}^T\Psi_0^{(n)}\approx\exp(-\tfrac{1}{2}n\,e^{2\alpha_*}d_n{}^2)\,, \tag{8.6}$$

where α_* is the value of α at which the mode goes outside the horizon. The wave function ${}^T\Psi_0^{(n)}$ will remain of the form (8.6) until the mode reenters the horizon in the matter- or radiation-dominated era at the much greater value α_e of α. One can then apply the adiabatic approximation again to (8.4) but ${}^T\Psi_0^{(n)}$ will no longer be in the ground state; it will be a superposition of a number of highly excited states. This is the phenomenon of the amplification of the ground-state fluctuations in the gravitational-wave modes that was discussed in Refs. 9, 17, and 18.

The behavior of the scalar modes is rather similar but their description is more complicated because of the gauge degrees of freedom. In the previous section we evaluated the wave function ${}^S\Psi^{(n)}$ on $a_n=b_n=0$ by the path-integral prescription. The ground-state form (in f_n) that we found will be valid until the adiabatic approximation breaks down, i.e., until the wavelength of the mode exceeds the horizon distance during the inflationary period. In order to discuss the subsequent behavior of the wave function. It is convenient to use the first-order Hamiltonian constraint (5.27) to evaluate ${}^S\Psi^{(n)}$ on $a_n\neq0, b_n=f_n=0$. One finds that

$${}^S\Psi^{(n)}(\alpha,\phi,a_n,0,0)=B\exp[iCa_n{}^2]\,{}^S\Psi_0^{(n)}(\alpha,\phi,a_n)\,. \tag{8.7}$$

The normalization and phase factors B and C depend on α and ϕ but not a_n:

$$C=\frac{1}{2}\left[\frac{\partial S}{\partial\alpha}\right]^{-1}\left[\left[\frac{\partial S}{\partial\alpha}\right]^2-\tfrac{1}{3}(n^2-4)e^{4\alpha}\right]. \tag{8.8}$$

At the time the wavelength of the mode equals the horizon distance during the inflationary period, the wave function ${}^S\Psi_0^{(n)}$ has the form

$${}^S\Psi_0^{(n)}=\exp(-\tfrac{1}{2}ny_*{}^{-2}e^{2\alpha_*}a_n{}^2)\,, \tag{8.9}$$

where y_* is the value of $y=(\partial S/\partial\alpha)[\partial S/\partial\phi]^{-1}$ when the mode leaves the horizon, $y_*=3\phi_*$. More generally, in the case of a scalar field with a potential $V(\phi)$, $y=6V(\partial V/\partial\phi)^{-1}$.

One can obtain a Schrödinger equation for ${}^S\Psi_0^{(n)}$ by putting $b_n=f_n=0$ in the scalar Hamiltonian ${}^SH^n_{|2}$ and substituting for $\partial/\partial b_n$ and $\partial/\partial f_n$ from the momentum constraint (5.25) and the first-order Hamiltonian constraint (5.27), respectively. This gives

$$i\frac{\partial\,{}^S\Psi_0^{(n)}}{\partial t}=\tfrac{1}{2}e^{-3\alpha}\left\{-y^2\frac{\partial^2}{\partial a_n{}^2}+e^{4\alpha}(n^2-4)\times\left[\frac{1}{y^2}-\tfrac{1}{3}e^{4\alpha}\left[\frac{\partial S}{\partial\alpha}\right]^{-2}\right]a_n{}^2\right\}{}^S\Psi_0^{(n)}\,, \tag{8.10}$$

where terms of order $1/n^2$ have been neglected. The term $e^{4\alpha}[\partial S/\partial\alpha]^{-2}$ will be small compared to $1/y^2$ except near the time of maximum radius of the background solution. The Schrödinger equation for ${}^S\Psi_0^{(n)}(a_n)$ is very similar to the equation for ${}^T\Psi_0^{(n)}(d_n)$, (8.4), except that the kinetic term is multiplied by a factor y^2 and the potential term is divided by a factor y^2. One would therefore expect that for wavelengths within the horizon. ${}^S\Psi_0^{(n)}$ would have the ground-state form $\exp(-\frac{1}{2}ny^{-2}e^{2\alpha}{a_n}^2)$ and this is borne out by (8.9). On the other hand, when the wavelength becomes larger than the horizon, the Schrödinger equation (8.10) indicates that ${}^T\Psi_0^{(n)}$ will freeze in the form (8.9) until the mode reenters the horizon in the matter-dominated era. Even if the equation of state of the Universe changes to radiation dominated during the period that the wavelength of the mode is greater than the horizon size, it will still be true that ${}^S\Psi_0^{(n)}$ is frozen in the form (8.9). The ground-state fluctuations in the scalar modes will therefore be amplified in a similar manner to the tensor modes. At the time of reentry of the horizon the rms fluctuation in the scalar modes, in the gauge in which $b_n=f_n=0$, will be greater by the factor y_* than the rms fluctuation in the tensor modes of the same wavelength.

IX. COMPARISON WITH OBSERVATION

From a knowledge of ${}^T\Psi_0^{(n)}$ and ${}^S\Psi_0^{(n)}$ one can calculate the relative probabilities of observing different values of d_n and a_n at a given point on a trajectory of the vector field X^i, i.e., at a given value of α and ϕ in a background metric which is a solution of the classical field equations. In fact, the dependence on ϕ will be unimportant and we shall neglect it. One can then calculate the probabilities of observing different amounts of anisotropy in the microwave background and can compare these predictions with the upper limits set by observation.

The tensor and scalar perturbation modes will be in highly excited states at large values of α. This means that we can treat their development as an ensemble evolving according to the classical equations of motion with initial distributions in d_n and a_n proportional to $|{}^T\Psi_0^{(n)}|^2$ and $|{}^S\Psi_0^{(n)}|^2$, respectively. The initial distributions in $\dot{d}_n$ and $\dot{a}_n$ will be proportional to $|{}^T\Psi_0^{(n)}\pi_{d_n}{}^T\Psi_0^{(n)}|$ and $|{}^S\Psi_0^{(n)}\pi_{a_n}{}^S\Psi_0^{(n)}|$, respectively. In fact, at the time that the modes reenter the horizon, the distributions will be concentrated at $\dot{d}_n=\dot{a}_n=0$.

The surfaces with $b_n=f_n=0$ will be surfaces of constant energy density in the classical solution during the inflationary period. By local conservation of energy, they will remain surfaces of constant energy density in the era after the inflationary period when the energy is dominated by the coherent oscillations of the homogeneous background scalar field ϕ. If the scalar particles decay into light particles and heat up the Universe, the surfaces with $b_n=f_n=0$ will be surfaces of constant temperature. The surface of last scattering of the microwave background will be such a surface with temperature T_s. The microwave radiation can be considered to have propagated freely to us from this surface. Thus the observed temperature will be

$$T_0=\frac{T_s}{1+z}\,, \tag{9.1}$$

where z is the red-shift of the surface of last scattering. Variations in the observed temperature will arise from variations in z in different directions of observation. These are given by

$$1+z=l^\mu n_\mu \tag{9.2}$$

evaluated at the surface of last scattering where n_μ is the unit normal to the surfaces of constant t in the gauge $g_n=k_n=j_n=0$ and $b_n=f_n=0$ on the surface of last scattering and l^μ is the parallel propagated tangent vector to the null geodesic from the observer normalized by $l^\mu n_\mu=1$ at the present time. One can calculate the evolution of $l^\mu n_\mu$ down the past light cone of the observer:

$$\frac{d}{d\lambda}[l^\mu n_\mu]=n_{\mu;\nu}l^\mu l^\nu\,, \tag{9.3}$$

where λ is the affine parameter on the null geodesic. The only nonzero components of $n_{\mu;\nu}$ are

$$n_{i;j}=e^{2\alpha}\left[\dot{\alpha}\Omega_{ij}+\sum_n(\dot{a}_n+\dot{\alpha}a_n)\tfrac{1}{3}\Omega_{ij}Q+\sum_n(\dot{b}_n+\dot{\alpha}b_n)P_{ij}+\sum_n(\dot{d}_n+\dot{\alpha}d_n)G_{ij}\right]. \tag{9.4}$$

In the gauge that we are using, the dominant anisotropic terms in (9.4) on the scale of the horizon, will be those involving $\dot{\alpha}a_n$ and $\dot{\alpha}d_n$. These will give temperature anisotropies of the form

$$\langle(\Delta T/T)^2\rangle\approx\langle {a_n}^2\rangle \quad\text{or}\approx\langle {d_n}^2\rangle\,. \tag{9.5}$$

The number of modes that contribute to anisotropies on the scale of the horizon is of the order of n^3. From the results of the last section

$$\langle {a_n}^2\rangle={y_*}^2n^{-1}e^{-2\alpha_*}\,, \tag{9.6}$$

$$\langle {d_n}^2\rangle=n^{-1}e^{-2\alpha_*}\,. \tag{9.7}$$

The dominant contribution comes from the scalar modes which give

$$\langle(\Delta T/T)^2\rangle\approx{y_*}^2n^2e^{-2\alpha_*}\,. \tag{9.8}$$

But $n\,e^{-\alpha_*}\approx\dot{\alpha}_*$, the value of the Hubble constant at the time that the present horizon size left the horizon during the inflationary period. The observational upper limit of about 10^{-8} on $\langle(\Delta T/T)^2\rangle$ restricts this Hubble constant to be less than about $5\times10^{-5}m_P$ (Ref. 8) which in turn restricts the mass of the scalar field to be less than 10^{14} GeV.

X. CONCLUSION AND SUMMARY

We started from the proposal that the quantum state of the Universe is defined by a path integral over compact four-metrics. This can be regarded as a boundary condition for the Wheeler-DeWitt equation for the wave function of the Universe on the infinite-dimensional manifold, superspace, the space of all three-metrics and matter field configurations on a three-surface S. Previous papers had considered finite-dimensional approximations to superspace and had shown that the boundary condition led to a wave function which could be interpreted as corresponding to a family of classical solutions which were homogeneous and isotropic and which had a period of exponential or inflationary expansion. In the present paper we extended this work to the full superspace without restrictions. We treated the two basic homogeneous and isotropic degrees of freedom exactly and the other degrees of freedom to second order. We justified this approximation by showing that the inhomogeneous or anisotropic modes started out in their ground states.

We derived time-dependent Schrödinger equations for each mode. We showed that they remained in the ground state until their wavelength exceeded the horizon size during the inflationary period. In the subsequent expansion the ground-state fluctuations got frozen until the wavelength reentered the horizon during the radiation- or matter-dominated era. This part of the calculation is similar to earlier work on the development of gravitational waves[9] and density perturbations[5,6] in the inflationary Universe but it has the advantage that the assumptions of a period of exponential expansion and of an initial ground state for the perturbations are justified. The perturbations would be compatible with the upper limits set by observations of the microwave background if the scalar field that drives the inflation has a mass of 10^{14} GeV or less.

In Sec. VIII we calculated the scalar perturbations in a gauge in which the surfaces of constant time are surfaces of constant density. There are thus no density fluctuations in this gauge. However, one can make a transformation to a gauge in which $a_n = b_n = 0$. In this gauge the density fluctuation at the time that the wavelength comes within the horizon is

$$\langle (\Delta\rho/\rho)^2 \rangle \approx y^2 \frac{\dot{\rho}_e^{\,2}}{\dot{\alpha}_e^{\,2}\rho_e^{\,2}} \dot{\alpha}_*^{\,2} . \tag{10.1}$$

Because y and $\dot{\alpha}_*$ depend only logarithmically on the wavelength of the perturbations, this gives an almost scale-free spectrum of density fluctuations. These fluctuations can evolve according to the classical field equations to give rise to the formation of galaxies and all the other structure that we observe in the Universe. Thus all the complexities of the present state of the Universe have their origin in the ground-state fluctuations in the inhomogeneous modes and so arise from the Heisenberg uncertainty principle.

APPENDIX A: HARMONICS ON THE THREE-SPHERE

In this appendix we describe the properties of the scalar, vector, and tensor harmonics on the three-sphere S^3. The metric on S^3 is Ω_{ij} and so the line element is

$$\begin{aligned} dl^2 &= \Omega_{ij} dx^i dx^j \\ &= d\chi^2 + \sin^2\chi (d\theta^2 + \sin^2\theta \, d\phi^2) . \end{aligned} \tag{A1}$$

A vertical bar will denote covariant differentiation with respect to the metric Ω_{ij}. Indices i,j,k are raised and lowered using Ω_{ij}.

Scalar harmonics

The scalar spherical harmonics $Q^n_{lm}(\chi,\theta,\phi)$ are scalar eigenfunctions of the Laplacian operator on S^3. Thus, they satisfy the eigenvalue equation

$$Q^{(n)}{}_{|k}{}^{|k} = -(n^2-1)Q^{(n)}, \quad n=1,2,3,\ldots . \tag{A2}$$

The most general solution to (A2), for given n, is a sum of solutions

$$Q^{(n)}(\chi,\theta,\phi) = \sum_{l=0}^{n-1} \sum_{m=-l}^{l} A^n_{lm} Q^n_{lm}(\chi,\theta,\phi) , \tag{A3}$$

where A^n_{lm} are a set of arbitrary constants. The Q^n_{lm} are given explicitly by

$$Q^n_{lm}(\chi,\theta,\phi) = \Pi^n_l(\chi) Y_{lm}(\theta,\phi) , \tag{A4}$$

where $Y_{lm}(\theta,\phi)$ are the usual harmonics on the two-sphere, S^2, and $\Pi^n_l(\chi)$ are the Fock harmonics.[19,20] The spherical harmonics Q^n_{lm} constitute a complete orthogonal set for the expansion of any scalar field on S^3.

Vector harmonics

The transverse vector harmonics $(S_i)^n_{lm}(\chi,\theta,\phi)$ are vector eigenfunctions of the Laplacian operator on S^3 which are transverse. That is, they satisfy the eigenvalue equation

$$S_i^{(n)}{}_{|k}{}^{|k} = -(n^2-2)S_i^{(n)}, \quad n=2,3,4,\ldots \tag{A5}$$

and the transverse condition

$$S_i^{(n)|i} = 0 . \tag{A6}$$

The most general solution to (A5) and (A6) is a sum of solutions

$$S_i^{(n)}(\chi,\theta,\phi) = \sum_{l=1}^{n-1} \sum_{m=-l}^{l} B^n_{lm} (S_i)^n_{lm}(\chi,\theta,\phi) , \tag{A7}$$

where B^n_{lm} are a set of arbitrary constants. Explicit expressions for the $(S_i)^n_{lm}$ are given in Ref. 20 where it is also explained how they are classified as odd (o) or even (e) using a parity transformation. We thus have two linearly independent transverse vector harmonics S_i^o and S_i^e (n,l,m suppressed).

Using the scalar harmonics Q^n_{lm} we may construct a third vector harmonics $(P_i)^n_{lm}$. defined by (n,l,m suppressed)

$$P_i = \frac{1}{(n^2-1)} Q_{|i}, \quad n=2,3,4,\ldots . \tag{A8}$$

It may be shown to satisfy

$$P_{i|k}{}^{|k} = -(n^2-3)P_i \quad \text{and} \quad P_i{}^{|i} = -Q . \tag{A9}$$

The three vector harmonics S_i^o, S_i^e, and P_i constitute a complete orthogonal set for the expansion of any vector field on S^3.

Tensor harmonics

The transverse traceless tensor harmonics $(G_{ij})^n_{lm}(\chi,\theta,\phi)$ are tensor eigenfunctions of the Laplacian operator on S^3 which are transverse and traceless. That is, they satisfy the eigenvalue equation

$$G_{ij}^{(n)}{}_{|k}{}^{|k} = -(n^2-3)G_{ij}^{(n)}, \quad n=3,4,5,\ldots \tag{A10}$$

and the transverse and traceless conditions

$$G_{ij}^{(n)|i}=0, \quad G_i^{(n)i}=0 . \tag{A11}$$

The most general solution to (A11) and (A12) is a sum of solutions

$$G_{ij}^{(n)}(\chi,\theta,\phi)=\sum_{l=2}^{n-1}\sum_{m=-l}^{l} C_{lm}^n (G_{ij})_{lm}^n(\chi,\theta,\phi) , \tag{A12}$$

where C_{lm}^n are a set of arbitrary constants. As in the vector case they may be classified as odd or even. Explicit expressions for $(G_{ij}^o)_{lm}^n$ and $(G_{ij}^e)_{lm}^n$ are given in Ref. 20.

Using the transverse vector harmonics $(S_i^o)_{lm}^n$ and $(S_i^e)_{lm}^n$, we may construct traceless tensor harmonics $(S_{ij}^o)_{lm}^n$ and $(S_{ij}^e)_{lm}^n$ defined, both for odd and even, by $(n,l,m$ suppressed)

$$S_{ij}=S_{i\,|\,j}+S_{j\,|\,i} \tag{A13}$$

and thus $S_i{}^i=0$ since S_i is transverse. In addition, the S_{ij} may be shown to satisfy

$$S_{ij}{}^{|j}=-(n^2-4)S_i , \tag{A14}$$

$$S_{ij}{}^{|ij}=0 , \tag{A15}$$

$$S_{ij\,|\,k}{}^{|k}=-(n^2-6)S_{ij} . \tag{A16}$$

Using the scalar harmonics Q_{lm}^n, we may construct two tensors $(Q_{ij})_{lm}^n$ and $(P_{ij})_{lm}^n$ defined by $(n,l,m$ suppressed)

$$Q_{ij}=\tfrac{1}{3}\Omega_{ij}Q, \quad n=1,2,3 \tag{A17}$$

and

$$P_{ij}=\frac{1}{(n^2-1)}Q_{\,|\,ij}+\tfrac{1}{3}\Omega_{ij}Q, \quad n=2,3,4 . \tag{A18}$$

The P_{ij} are traceless, $P_i{}^i=0$, and in addition, may be shown to satisfy

$$P_{ij}{}^{|j}=-\tfrac{2}{3}(n^2-4)P_i , \tag{A19}$$

$$P_{ij\,|\,k}{}^{|k}=-(n^2-7)P_{ij} , \tag{A20}$$

$$P_{ij}{}^{|ij}=\tfrac{2}{3}(n^2-4)Q . \tag{A21}$$

The six tensor harmonics Q_{ij}, P_{ij}, S_{ij}^o, S_{ij}^e, G_{ij}^o, and G_{ij}^e constitute a complete orthogonal set for the expansion of any symmetric second-rank tensor field on S^3.

Orthogonality and normalization

The normalization of the scalar, vector, and tensor harmonics is fixed by the orthogonality relations. We denote the integration measure on S^3 by $d\mu$. Thus

$$d\mu=d^3x(\det\Omega_{ij})^{1/2}=\sin^2\chi\,\sin\theta\,d\chi\,d\theta\,d\phi . \tag{A22}$$

The Q_{lm}^n are normalized so that

$$\int d\mu Q_{lm}^n Q_{l'm'}^{n'}=\delta^{nn'}\delta_{ll'}\delta_{mm'} . \tag{A23}$$

This implies

$$\int d\mu (P_i)_{lm}^n (P^i)_{l'm'}^{n'}=\frac{1}{(n^2-1)}\delta^{nn'}\delta_{ll'}\delta_{mm'} \tag{A24}$$

and

$$\int d\mu (P_{ij})_{lm}^n (P^{ij})_{l'm'}^{n'}=\frac{2(n^2-4)}{3(n^2-1)}\delta^{nn'}\delta_{ll'}\delta_{mm'} . \tag{A25}$$

The $(S_i)_{lm}^n$, both odd and even, are normalized so that

$$\int d\mu (S_i)_{lm}^n (S^i)_{l'm'}^{n'}=\delta^{nn'}\delta_{ll'}\delta_{mm'} . \tag{A26}$$

This implies

$$\int d\mu (S_{ij})_{lm}^n (S^{ij})_{l'm'}^{n'}=2(n^2-4)\delta^{nn'}\delta_{ll'}\delta_{mm'} . \tag{A27}$$

Finally, the $(G_{ij})_{lm}^n$, both odd and even, are normalized so that

$$\int d\mu (G_{ij})_{lm}^n (G^{ij})_{l'm'}^{n'}=\delta^{nn'}\delta_{ll'}\delta_{mm'} . \tag{A28}$$

The information given in this appendix about the spherical harmonics is all that is needed to perform the derivations presented in the main text. Further details may be found in Refs. 19 and 20.

APPENDIX B: ACTION AND FIELD EQUATIONS

The action (5.8) is

$$I=I_0(\alpha,\phi,N_0)+\sum_n I_n , \tag{B1}$$

where I_0 is the action of the unperturbed model (4.2):

$$I_0=-\tfrac{1}{2}\int dt\, N_0 e^{3\alpha}\left[\frac{\dot\alpha^2}{N_0{}^2}-e^{-2\alpha}-\frac{\dot\phi^2}{N_0{}^2}+m^2\phi^2\right] . \tag{B2}$$

I_n is quadratic in the perturbations and may be written

$$I_n=\int dt(L_g^n+L_m^n) , \tag{B3}$$

where

$$L_g^n=\tfrac{1}{2}e^{\alpha}N_0\left[\tfrac{1}{3}(n^2-\tfrac{5}{2})a_n{}^2+\frac{(n^2-7)}{3}\frac{(n^2-4)}{(n^2-1)}b_n{}^2-2(n^2-4)c_n{}^2-(n^2+1)d_n{}^2+\tfrac{2}{3}(n^2-4)a_nb_n\right.$$
$$\left.+g_n[\tfrac{2}{3}(n^2-4)b_n+\tfrac{2}{3}(n^2+\tfrac{1}{2})a_n]+\frac{1}{N_0{}^2}\left[-\frac{1}{3(n^2-1)}k_n{}^2+(n^2-4)j_n{}^2\right]\right]$$
$$+\frac{1}{2}\frac{e^{3\alpha}}{N_0}\left\{-\dot a_n{}^2+\frac{(n^2-4)}{(n^2-1)}\dot b_n{}^2+(n^2-4)\dot c_n{}^2+\dot d_n{}^2\right.$$

$$+\dot{\alpha}\left[-2a_n\dot{a}_n+8\frac{(n^2-4)}{(n^2-1)}b_n\dot{b}_n+8(n^2-4)c_n\dot{c}_n+8d_n\dot{d}_n\right]$$
$$+\dot{\alpha}^2\left[-\tfrac{3}{2}a_n{}^2+6\frac{(n^2-4)}{(n^2-1)}b_n{}^2+6(n^2-4)c_n{}^2+6d_n{}^2\right]+g_n[2\dot{\alpha}\,\dot{a}_n+\dot{\alpha}^2(3a_n-g_n)]$$
$$+e^{-\alpha}\left[k_n\left[-\tfrac{2}{3}\dot{a}_n-\tfrac{2}{3}\frac{(n^2-4)}{(n^2-1)}\dot{b}_n+\tfrac{2}{3}\dot{\alpha}g_n\right]-2(n^2-4)\dot{c}_nj_n\right]\Bigg\} \qquad \text{(B4)}$$

and

$$L_m^n=\tfrac{1}{2}N_0e^{3\alpha}\Bigg[\frac{1}{N_0{}^2}(\dot{f}_n{}^2+6a_n\dot{f}_n\dot{\phi})-m^2(f_n{}^2+6a_nf_n\phi)-e^{-2\alpha}(n^2-1)f_n{}^2$$
$$+\frac{3}{2}\left[\frac{\dot{\phi}^2}{N_0{}^2}-m^2\phi^2\right]\left[a_n{}^2-\frac{4(n^2-4)}{(n^2-1)}b_n{}^2-4(n^2-4)c_n{}^2-4d_n{}^2\right]+\frac{\dot{\phi}^2}{N_0{}^2}g_n{}^2$$
$$-g_n\left[2m^2f_n\phi+3m^2a_n\phi^2+2\frac{\dot{f}_n\dot{\phi}}{N_0{}^2}+3\frac{a_n\dot{\phi}^2}{N_0{}^2}\right]-2\frac{e^{-\alpha}}{N_0{}^2}k_nf_n\dot{\phi}\Bigg] . \qquad \text{(B5)}$$

The full expressions for π_α and π_ϕ are

$$\pi_\alpha=\frac{e^{3\alpha}}{N_0}\Bigg[-\dot{\alpha}+\sum_n\left[-a_n\dot{a}_n+\frac{4(n^2-4)}{(n^2-1)}b_n\dot{b}_n+4(n^2-4)c_n\dot{c}_n+4d_n\dot{d}_n\right]$$
$$+\dot{\alpha}\sum_n\left[-\tfrac{3}{2}a_n{}^2+\frac{6(n^2-4)}{(n^2-1)}b_n{}^2+6(n^2-4)c_n{}^2+6d_n{}^2\right]+\sum_n g_n[\dot{a}_n+\dot{\alpha}(3a_n-g_n)+\tfrac{1}{3}e^{-\alpha}k_n]\Bigg] , \qquad \text{(B6)}$$

$$\pi_\phi=\frac{e^{3\alpha}}{N_0}\Bigg\{\dot{\phi}+\sum_n\left[3a_n\dot{f}_n+\dot{\phi}\left[\tfrac{3}{2}a_n{}^2-\frac{4(n^2-4)}{(n^2-1)}b_n{}^2-4(n^2-4)c_n{}^2-4d_n{}^2\right]\right]$$
$$+\sum_n[\dot{\phi}g_n{}^2-g_n(\dot{f}_n+3a_n\dot{\phi})-e^{-\alpha}k_nf_n]\Bigg\} . \qquad \text{(B7)}$$

The classical field equations may be obtained from the action (B1) by varying with respect to each of the fields in turn. Variation with respect to α and ϕ gives two field equations, similar to those obtained in Sec. IV, but modified by terms quadratic in the perturbations:

$$N_0\frac{d}{dt}\left[\frac{1}{N_0}\frac{d\phi}{dt}\right]+3\frac{d\alpha}{dt}\frac{d\phi}{dt}+N_0{}^2m^2\phi=\text{quadratic terms} , \qquad \text{(B8)}$$

$$N_0\frac{d}{dt}\left[\frac{\dot{\alpha}}{N_0}\right]+3\dot{\phi}^2-N_0{}^2e^{-2\alpha}-\tfrac{3}{2}(-\dot{\alpha}^2+\dot{\phi}^2-N_0{}^2e^{-2\alpha}+N_0{}^2m^2\phi^2)=\text{quadratic terms} . \qquad \text{(B9)}$$

Variation with respect to the perturbations a_n, b_n, c_n, d_n, and f_n leads to five field equations:

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{a}_n}{N_0}\right]+\tfrac{1}{3}(n^2-4)N_0{}^2e^{\alpha}(a_n+b_n)+3e^{3\alpha}(\dot{\phi}\dot{f}_n-N_0{}^2m^2\phi f_n)=N_0{}^2[3e^{3\alpha}m^2\phi^2-\tfrac{1}{3}(n^2+2)e^{\alpha}]g_n$$
$$+e^{3\alpha}\dot{\alpha}\dot{g}_n-\tfrac{1}{3}N_0\frac{d}{dt}\left[e^{2\alpha}\frac{k_n}{N_0}\right] , \qquad \text{(B10)}$$

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{b}_n}{N_0}\right]-\tfrac{1}{3}(n^2-1)N_0{}^2e^{\alpha}(a_n+b_n)=\tfrac{1}{3}(n^2-1)N_0{}^2e^{\alpha}g_n+\tfrac{1}{3}N_0\frac{d}{dt}\left[e^{2\alpha}\frac{k_n}{N_0}\right] , \qquad \text{(B11)}$$

$$\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{c}_n}{N_0}\right]=\frac{d}{dt}\left[e^{2\alpha}\frac{j_n}{N_0}\right] , \qquad \text{(B12)}$$

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{d}_n}{N_0}\right]+(n^2-1)N_0{}^2e^{\alpha}d_n=0\ , \tag{B13}$$

$$N_0\frac{d}{dt}\left[e^{3\alpha}\frac{\dot{f}_n}{N_0}\right]+3\,e^{3\alpha}\dot{\phi}\,\dot{a}_n+N_0{}^2[m^2e^{3\alpha}+(n^2-1)e^{\alpha}]f_n=e^{3\alpha}(-2N_0{}^2m^2\phi g_n+\dot{\phi}\,\dot{g}_n-e^{-\alpha}\phi k_n)\ . \tag{B14}$$

In obtaining (B10)—(B14), the field equations (B8) and (B9) have been used and terms cubic in the perturbations have been dropped.

Variation with respect to the Lagrange multipliers k_n, j_n, g_n, and N_0 leads to a set of constraints. Variation with respect to k_n and j_n leads to the momentum constraints:

$$\dot{a}_n+\frac{(n^2-4)}{(n^2-1)}\dot{b}_n+3f_n\dot{\phi}=\dot{\alpha}g_n-\frac{e^{-\alpha}}{(n^2-1)}k_n\ , \tag{B15}$$

$$\dot{c}_n=e^{-\alpha}j_n\ . \tag{B16}$$

Variation with respect to g_n gives the linear Hamiltonian constraint:

$$3a_n(-\dot{\alpha}^2+\dot{\phi}^2)+2(\dot{\phi}\dot{f}_n-\dot{\alpha}\dot{a}_n)+N_0{}^2m^2(2f_n\phi+3a_n\phi^2)-\tfrac{2}{3}N_0{}^2e^{-2\alpha}[(n^2-4)b_n+(n^2+\tfrac{1}{2})a_n]$$
$$=\tfrac{2}{3}\dot{\alpha}\,e^{-\alpha}k_n+2g_n(-\dot{\alpha}^2+\dot{\phi}^2)\ . \tag{B17}$$

Finally, variation with respect to N_0 yields the Hamiltonian constraint, which we write as

$$\tfrac{1}{2}e^{3\alpha}\left[-\frac{\dot{\alpha}^2}{N_0{}^2}+\frac{\dot{\phi}^2}{N_0{}^2}-e^{-2\alpha}+m^2\phi^2\right]=\text{quadratic terms}\ . \tag{B18}$$

[1]A. H. Guth, Phys. Rev. D **23**, 347 (1981).
[2]A. D. Linde, Phys. Lett. **108B**, 389 (1982).
[3]S. W. Hawking and I. G. Moss, Phys. Lett. **110B**, 35 (1982).
[4]A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 120 (1982).
[5]S. W. Hawking, Phys. Lett. **115B**, 295 (1982).
[6]A. H. Guth and S. Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982).
[7]J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. D **28**, 679 (1983).
[8]S. W. Hawking, Phys. Lett. B **150B**, 339 (1985).
[9]V. A. Rubakov, M. V. Sazhin, and A. V. Veryaskin, Phys. Lett. **115B**, 189 (1982).
[10]S. W. Hawking, Pontif. Accad. Sci. Varia **48**, 563 (1982).
[11]J. B. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983).
[12]S. W. Hawking, in *Relativity, Groups and Topology II,* Les Houches 1983, Session XL, edited by B. S. DeWitt and R. Stora (North-Holland, Amsterdam, 1984).
[13]S. W. Hawking, Nucl. Phys. **B239**, 257 (1984).
[14]S. W. Hawking and Z. C. Wu, Phys. Lett. **151B**, 15 (1985).
[15]S. W. Hawking and D. N. Page, DAMTP report, 1984 (unpublished).
[16]S. W. Hawking and J. C. Luttrell, Nucl. Phys. **B247**, 250 (1984).
[17]L. P. Grishchuk, Zh. Eksp. Teor. Fiz. **67**, 825 (1974) [Sov. Phys. JETP **40**, 409 (1975)]; Ann. N.Y. Acad. Sci. **302**, 439 (1977).
[18]A. A. Starobinsky, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 719 (1979) [JETP Lett. **30**, 682 (1979)].
[19]E. M. Lifschitz and I. M. Khalatnikov, Adv. Phys. **12**, 185 (1963).
[20]U. H. Gerlach and U. K. Sengupta, Phys. Rev. D **18**, 1773 (1978).