Wave function of the Universe

J. B. Hartle
Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637
and
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106

S. W. Hawking
Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England
and
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106

Originally published in Physical Review D, Volume 28, Number 12 on Dec. 15th 1983

(Received 29 July 1983)

(Refurbished by Stephen Crowley on May 6, 2024)

Table of contents

1 Introduction 1
II. QUANTUM GRAVITY 5
A. Wave functions 6
B. Wheeler-DeWitt equation 7
V. MINISUPERSPACE MODEL 11
VI. GROUND-STATE COSMOLOGICAL WAVE FUNCTION 15
A. The matter wave function 15

1 Introduction

In any attempt to apply quantum mechanics to the Universe as a whole the specification of the possible quantum-mechanical states which the Universe can occupy is of central importance. This specification determines the possible dynamical behavior of the Universe. Moreover, if the uniqueness of the present Universe is to find any explanation in quantum gravity it can only come from a restriction on the possible states available.

In quantum mechanics the state of a system is specified by giving its wave function on an appropriate configuration space. The possible wave functions can be constructed from the fundamental quantum-mechanical amplitude for a complete history of the system which may be regarded as the starting point for quantum theory. For example, in the case of a single particle a history is a path $x(t)$ and the amplitude for a particular path is proportional to

$$e^{iS[x(t)]} \tag{1}$$

where $S[x(t)]$ is the classical action. From this basic amplitude, the amplitude for more restricted observations can be constructed by superposition. In particular, the amplitude that the particle, having been prepared in a certain way, is located at position $x(t)$ and nowhere else at time t is

$$\psi(x(t)) = N(x(t)) \int e^{iS[x(t)]} \mathcal{D}\, x(t) \tag{2}$$

Here, $N(x(t))$ is a normalizing factor and the sum is over a class of paths which intersect x at time t and which are weighted in a way that reflects the preparation of the system. $\psi(x,t)$ is the wave function for the state determined by this preparation. As an example, if the particle were previously localized at x' at time t' one would sum over all paths which start at x' at t' and end at x at t thereby obtaining the propagator $\langle x,t|x',t'\rangle$. The oscillatory integral in Eq. (1.2) is not well defined but can be made so by rotating the time to imaginary values.

An alternative way of calculating quantum dynamics is to use the Schrödinger equation,

$$i\,\hbar\,\dot{\psi}(x(t)) = H\,\psi(x(t)) \tag{3}$$

This follows from Eq. (1.2) by varying the end conditions on the path integral. For a particular state specified by a weighting of paths C, the path integral (1.2) may be looked upon as providing the boundary conditions for the solution of Eq. (1.3).

A state of particular interest in any quantum-mechanical theory is the ground state, or state of minimum excitation. This is naturally defined by the path integral, made definite by a rotation to Euclidean time, over the class of paths which have vanishing action in the far past. Thus, for the ground state at $t = 0$ one would write

$$\psi_0(x(0)) = N \int e^{-I[x(\tau)]} \delta x(\tau) \tag{4}$$

where $I[x(\tau)]$ is the Euclidean action obtained from S by sending $t \rightarrow -i\,\tau$ and adjusting the sign so that it is positive.

In cases where there is a well-defined time and a corresponding time-independent Hamiltonian, this definition of the ground state coincides with the lowest eigenfunction of the Hamiltonian. To see this specialize the path-integral expression for the propagator $\langle x''$, $t''|x', t'\rangle$ to $t = 0$ and $x' = 0$ and insert a complete set of energy eigenstates between the initial and final state. One has

$$\begin{aligned}\langle x, 0|0, s\rangle &= \sum_n \psi_n(x)\, e^{-iE_n 0} \psi_n^*(0)\, e^{-iE_n s} = \int e^{iS[x(t)]} \mathcal{D}\, x(t) \\ &= \sum_n \psi_n(x)\, e^0\, \psi_n^*(0)\, e^{-iE_n s} = \int e^{iS[x(t)]} \mathcal{D}\, x(t) \\ &= \sum_n \psi_n(x) 1\, \psi_n^*(0)\, e^{-iE_n s} = \int e^{iS[x(t)]} \mathcal{D}\, x(t) \\ &= \sum_n \psi_n(x)\, \psi_n^*(0)\, e^{-iE_n s} = \int e^{iS[x(t)]} \mathcal{D}\, x(t)\end{aligned} \tag{5}$$

where $\psi_n(x)$ are the time-independent energy eigenfunctions. Rotate $t' \rightarrow -i\,\tau$ in (5) and take the limit as $\tau' \rightarrow -\infty$. *In the sum only the lowest eigenfunction (normalized to zero energy) survives.*[1] The path integral becomes the path integral on the right of (4) so that the equality is demonstrated.

The case of quantum fields is a straightforward generalization of quantum particle mechanics. The wave function is a functional of the field configuration on a space-like surface of constant time, that is $\Psi[\phi(x(t))]$. The functional Ψ gives the amplitude that a particular field distribution $\phi(x(t))$ occurs on this space-like surface. The rest of the formalism is similarly generalized. For example, for the ground-state wave function one has

$$\Psi_0[\phi(x(t)), 0] = N \int e^{-I[\varphi(x)]} \mathcal{D}\, \phi(x(t)) \tag{6}$$

where the integral is over all Euclidean field configurations for $\tau < 0$ which match $\phi(x(0))$ on the surface $\tau = 0$ and leave the action finite at Euclidean infinity.

In the case of quantum gravity new features enter. For definiteness and simplicity we shall restrict our attention throughout this paper to spatially closed universes. For these there is no well-defined intrinsic measure of the location of a spacelike surface in the spacetime beyond that contained in the intrinsic or extrinsic geometry of the surface itself. One therefore labels the wave function by the three-metric h_{ij} writing $\Psi(x(t)) = \Psi[h_{ij}(x(t))]$. Quantum dynamics is supplied by the functional integral

$$\Psi[h_{ij}(x(t))] = N \int e^{iS_E[g_{ij}(x(t))]} \mathcal{D}\, g_{ij}(x(t)) \tag{7}$$

1. In the long-time limit, all terms with $e^{-iE_n s}$ for $E_n > 0$ vanish, leaving only the ground state term $e^0 = 1$, which does not decay.

S_E is the classical action for gravity including a cosmological constant A and *the functional integral is over all four-geometries with a spacelike boundary on which the induced metric is* h_{ij} *and which to the past of that surface satisfy some appropriate condition to define the state*. In particular for the amplitude to go from a three-geometry h'_{ij} on an initial spacelike surface to a three-geometry h''_{ij} on a final spacelike surface is

$$\langle h''_{ij} | h'_{ij} \rangle = \int e^{iS_E[g_{ij}(x(t))]} \mathcal{D}\, g_{ij}(x(t)) \tag{8}$$

where the sum is over all four-geometries which match h'_{ij} on the initial surface and h''_{ij} on the final surface. Here one clearly sees that one cannot specify time in these states. The proper time between the surfaces depends on the four-geometries in the sum.

As in the mechanics of a particle the functional integral (7) implies a differential equation on the wave function. This is the Wheeler-DeWitt equation[2] which we shall derive from this point of view in Sec. II. With a simple choice of factor ordering it is

$$\left[-G_{ijkl} \frac{\delta^2}{\delta\, h_{ij}\,(x(t)) \delta\, h_{kl}(x(t))} - \sqrt{h(x(t))}\, ({}^3R(h(x(t))) + 2\,\Lambda) \right] \Psi[h_{ij}(x(t))] = 0 \tag{9}$$

where G_{ijkl} is the metric on superspace,

$$G_{ijkl}(x(t)) = \frac{h_{ik}\,(x(t))h_{jl}(x(t)) + h_{il}(x(t))\, h_{jk}(x(t)) - h_{ij}\,(x(t))h_{kl}(x(t))}{2\sqrt{h(x(t))}} \tag{10}$$

and ${}^{(3)}R$ is the scalar curvature of the intrinsic geometry of the three-surface. *The problem of specifying cosmological states is the same as specifying boundary conditions for the solution of the Wheeler-DeWitt equation*. A natural first question to ask is what boundary conditions specify the ground state?

In the quantum mechanics of closed universes we do not expect to find a notion of ground state as a state of lowest energy. There is no natural definition of energy for a closed universe just as there is no independent standard of time. Indeed in a certain sense the total energy for a closed universe is always zero—the gravitational energy cancelling the matter energy. It is still reasonable, however, to expect to be able to define a state of minimum excitation corresponding to the classical notion of a geometry of high symmetry. This paper contains a proposal for the definition of such a ground-state wave function for closed universes. The proposal is to extend to gravity the Euclidean-functional-integral construction of nonrelativistic quantum mechanics and field theory [Eqs. (4) and (6)]. Thus, we write for the ground-state wave function

$$\Psi_0[h_{ij}(x(0)), 0] = N \int e^{-I_E[g(x(0))]} \delta\, g(x(0)) \tag{11}$$

where I_E is the Euclidean action for gravity including a cosmological constant Λ. The Euclidean four-geometries summed over must have a boundary on which the induced metric is h_{ij}. The remaining specification of the class of geometries which are summed over determines the ground state. Our proposal is that the sum should be over compact geometries. This means that the Universe does not have any boundaries in space or time (at least in the Euclidean regime) (cf. Ref. 3). There is thus no problem of boundary conditions. One can interpret the functional integral over all compact four-geometries bounded by a given three-geometry as giving the amplitude for that three-geometry to arise from a zero three-geometry, i.e., a single point. In other words, the ground state is the amplitude for the Universe to appear from nothing.[4] In the following we shall elaborate on this construction and show in simple models that it indeed supplies reasonable wave functions for a state of minimum excitation.The specification of the ground-state wave function is a constraint on the other states allowed in the theory. They must be such, for example, as to make the Wheeler-DeWitt equation Hermitian in an appropriate norm. In analogy with ordinary quantum mechanics one would expect to be able to use these constraints to extrapolate the boundary conditions which determine the excited states of

the theory from those fixed for the ground state by Eq. (7). Thus, one can in principle determine all the allowed cosmological states.

The wave functions which result from this specification will not vanish on the singular, zero-volume three-geometries which correspond to the big-bang singularity. This is analogous to the behavior of the wave function of the electron in the hydrogen atom. In a classical treatment, the situation in which the electron is at the proton is singular. However, in a quantum-mechanical treatment the wave function in a state of zero angular momentum is finite and nonzero at the proton. This does not cause any problems in the case of the hydrogen atom. In the case of the Universe we would interpret the fact that the wave function can be finite and nonzero at the zero three-geometry as allowing the possibility of topological fluctuations of the three-geometry. This will be discussed further in Sec. VIII.

After a general discussion of this proposal for the ground-state wave function we shall implement it in a minisuperspace model. The geometrical degrees of freedom in the model are restricted to spatially homogeneous, isotropic, closed universes with S^3 topology, the matter degrees of freedom to a single, homogeneous, conformally invariant scalar field and the cosmological constant is assumed to be positive. A semiclassical evaluation of the functional integral for the ground-state wave function shows that it indeed does possess characteristics appropriate to a "state of minimum excitation."

II. QUANTUM GRAVITY

In this section we shall review the basic principles and machinery of quantum gravity with which we shall explore the wave functions for closed universes. For simplicity we shall represent the matter degrees of freedom by a single scalar field ϕ, more realistic cases being straightforward generalizations. We shall approach this review from the functional-integral point of view although we shall arrive at many canonical results. None of these are new and for different approaches to the same ends the reader is referred to the standard literature.

A. Wave functions

Our starting point is the quantum-mechanical amplitude for the occurrence of a given spacetime and a given field history. This is

$$e^{iS[\boldsymbol{g}(x(t)),\boldsymbol{\phi}(x(t))]} \tag{12}$$

where $S[\boldsymbol{g}(x(t)), \boldsymbol{\phi}(x(t))]$ is the total classical action for gravity coupled to a scalar field. We are envisaging here a fixed manifold although there is no real reason that amplitudes for different manifolds may not be considered provided a rule is given for their relative phases. Just as the interesting observations of a particle are not typically its entire history but rather observations of position at different times, so also the interesting quantum-mechanical questions for gravity correspond to observations of spacetime and field on different spacelike surfaces. Following the general rules of quantum mechanics the amplitudes for these more restricted sets of observations are obtained from (12) by summing over the unobserved quantities.

It is easy to understand what is meant by fixing the field on a given spacelike surface. What is meant by fixing the four-geometry is less obvious. Consider all four-geometries in which a given spacelike surface occurs but whose form is free to vary off the surface. By an appropriate choice of gauge near the surface (e.g., Gaussian normal coordinates) all these four-geometries can be expressed so that the only freedom in the four-metric is the specification of the three-metric $h_{ij}(x(t))$ in the surface. Specifying the three-metric is therefore what we mean by fixing the four-geometry on a spacelike surface. The situation is not unlike gauge theories. There a history is specified by a vector potential $A_\mu(x(t))$ but by an appropriate gauge transformation $A_0(x(t))$ can be made to vanish so that the field on a surface can be completely specified by the $A_i(x(t))$.

As an example of the quantum-mechanical superposition principle the amplitude for the three-geometry and field to be fixed on two spacelike surfaces is

$$\langle h''_{ij}, \boldsymbol{\phi}''|h'_{ij}, \boldsymbol{\phi}'\rangle = \int e^{iS[\boldsymbol{g}(x(t)),\boldsymbol{\phi}(x(t))]}\mathcal{D}\,g(x(t))\,\mathcal{D}\,\boldsymbol{\phi}(x(t)) \tag{13}$$

where the integral is over all four-geometries and field configurations which match the given values on the two spacelike surfaces. This is the natural analog of the propagator $K(x'', t''|x', t')$ in the quantum mechanics of a single particle. We note again that the proper time between the two surfaces is not specified. Rather it is summed over in the sense that the separation between the surfaces depends on the four-geometry being summed over. It is not that one could not ask for the amplitude to have the three-geometry and field fixed on two surfaces and the proper time between them. One could. Such an amplitude, however, would not correspond to fixing observations on just two surfaces but rather would involve a set of intermediate observations to determine the time. It would therefore not be the natural analog of the propagator.

Wave functions Ψ are defined by

$$\Psi[h_{ij}(x(t)), \boldsymbol{\phi}(x(t))] = \int_{\mathcal{C}} e^{iS[\boldsymbol{g},\boldsymbol{\phi}]}\mathcal{D}\,g(x(t))\,\mathcal{D}\,\boldsymbol{\phi}(x(t)) \tag{14}$$

The sum is over a class $\mathcal{C}$ of spacetimes with a compact boundary on which the induced metric is h_{ij} and field configurations which match ϕ on the boundary. The remaining specification of the class C is the specification of the state.

If the Universe is in a quantum state specified by a wave function Ψ then that wave function describes the correlations between observables to be expected in that state. For example, in the semiclassical wave function describing a universe like our own, one would expect Ψ to be large when ϕ is small and the spatial volume is big, and small when ϕ denotes quantities are oppositely correlated[2]. This is the only interpretative structure we shall propose or need.

B. Wheeler-DeWitt equation

A differential equation for Ψ can be derived by varying the end conditions on the path integral (14) which defines it. To carry out this derivation first recall that the gravitational action appropriate to keeping the three-geometry fixed on a boundary is

$$l^2 S_E = 2\int_{\partial M} \sqrt{h(x(t))}\, K\,(x(t)) d^3 x(t) + \int_M \sqrt{-g(x(t))}\,(R - 2\,\Lambda) d^4 x(t) \tag{15}$$

The second term is integrated over spacetime and the first over its boundary. K is the trace of the extrinsic curvature K_{ij} of the boundary three-surface. If its unit normal is n^i, $K_{ij} = -\nabla_i n_j$ in the usual Lorentzian convention. l is the Planck length $\sqrt{16\,\pi\,G}$ in the units with $\hbar = c = 1$ we use throughout. Introduce coordinates so that the boundary is a constant t surface and write the metric in the standard 3 + 1 decomposition:

$$d\,s^2 = -(N^2 - N_i\,N^i)\,d\,t^2 + 2\,N_i\,d\,x^i\,d\,t + h_{ij}(x(t))\,d\,x^i\,d\,x^j \tag{16}$$

The action (15) becomes

$$l^2 S_E = \int \sqrt{h(x(t))}\,N\,[K_{ij}\,K^{ij} - K^2 +^3 R(h(x(t))) - 2\,\Lambda] d^4 x(t) \tag{17}$$

where explicitly

$$K_{ij} = \frac{1}{N}\left(-\frac{1}{2}\frac{\partial\,h_{ij}(x(t))}{\partial\,t} + N_{(i|j)}\right) \tag{18}$$

and a stroke and ${}^3R$ denote the covariant derivative and scalar curvature constructed from the three-metric h_{ij}. The matter action S_M can similarly be expressed as a function of N, N_i, h_{ij}, and the matter field.

The functional integral defining the wave function contains an integral over N. By varying N at the surface we push it forward or backward in time. Since the wave function does not depend on time we must have

$$0 = \int\left(\frac{\delta\,S}{\delta\,N}\right) e^{i S[g(x(t)),\,\phi(x(t))]} \delta\,S\,x(t) \delta\,\phi(x(t)) \tag{19}$$

2. This is similar to the relationship of a covariance function and the spectral density of a Gaussian process.

More precisely, the value of the integral (14) should be left unchanged by an infinitesimal translation of the integration variable N. If the measure is invariant under translation this leads to (19). If it is not, there will be in addition a divergent contribution to the relation which must be suitably regulated to zero or cancel divergences arising from the calculation of the right-hand side of (19).

Classically the field equation $H = \frac{\delta S}{\delta N} = 0$ is the Hamiltonian constraint for general relativity. It is

$$H(x(t)) = \sqrt{h(x(t))}\,(K^2(x(t)) - K_{ij}(x(t))K^{ij}(x(t)) + R^{(3)}(h(x(t))) - 2\,\Lambda) - \sqrt{h}\,T^n_m \qquad (20)$$

where T^n_m is the stress-energy tensor of the matter field projected in the direction normal to the surface. Equation (19) shows how $H = 0$ is enforced as an operator identity for the wave function. More explicitly one can note that the K_{ij} involve only first-time derivatives of the h_{ij} and therefore may be completely expressed in terms of the momentum π^{ij} conjugate to the h_{ij} which follow from the Lagrangian in (?):

$$\pi^{ij} = -\sqrt{h}\,(K^{ij} - h^{ij}\,K) \qquad (21)$$

In a similar manner the energy of the matter field can be expressed in terms of the momentum conjugate to the field ϕ and the field itself. Equation (19) thus implies the operator identity $H(\Psi|h_{ij}, \pi^{ij}, \phi, \pi_\phi) = 0$ with the replacements

$$\pi^{ij} = -i\,\frac{\delta}{\delta\,h_{ij}}, \quad \pi_\phi = -i\,\frac{\delta}{\delta\,\phi} \qquad (22)$$

These replacements may be viewed as arising directly from the functional integral, e.g., from the observation that when the time derivatives in the exponent are written in differenced form

$$-i\int \delta\,h_{ij}(x(t))\,\delta\,K^{ij}(x(t)) = -i\int \delta\,\phi(x(t))\,\delta\,\pi_\phi(x(t)) \qquad (23)$$

Alternatively, *they are the standard representation of the canonical commutation relations of* ϕ *and* π_ϕ.

In translating a classical equation like $\frac{\delta S}{\delta N} = 0$ into an operator identity there is always the question of factor ordering. This will not be important for us so making a convenient choice we obtain

$$\frac{\delta^2}{\delta\,h_{ij}\,\delta\,h_{kl}}\left(\sqrt{h}\left(2\,\Lambda + l^2 T_{nn}\left[-i\frac{\delta}{\delta\,\phi}, \phi\right] - {}^3R(h)\right)\Psi[h_{ij}, \phi] - G_{ijkl}\right) \times \Psi(h_{ij}, \phi) = 0 \qquad (24)$$

This is the Wheeler-DeWitt equation which wave functions for closed universes must satisfy. There are also the other constraints of the classical theory, but the operator versions of these express the gauge invariance of the wave function rather than any dynamical information. We should emphasize that the ground-state wave function constructed by a Euclidean functional-integral prescription [Eq. (11)] will satisfy the Wheeler-DeWitt equation in the form (24). Indeed, this can be demonstrated explicitly by repeating the steps in the above demonstration starting with the Euclidean functional integral.

C. Boundary conditions

The quantity G_{ijkl} can be viewed as a metric on superspace—the space of all three-geometries (no connection with supersymmetry). It has signature (...,+,+,...,+,+) and the Wheeler-DeWitt equation is therefore a "hyperbolic" equation on superspace. It would be natural, therefore, to expect to impose boundary conditions on two "spacelike surfaces" in superspace. A convenient choice for the timelike direction is $\sqrt{h}$ and we therefore expect to impose boundary conditions at the upper and lower limits of the range of $\sqrt{h}$. The upper limit is infinity. The lower limit is zero because if h is positive definite or degenerate, $\sqrt{h} \geq 0$. Positive-definite metrics are everywhere spacelike surfaces; degenerate metrics may signal topology change. Summarizing the remaining functions of h_{ij} by the conformal metric $\tilde{h}_{ij} = \frac{h_{ij}}{h^{1/3}}$ we may write an important boundary condition on Ψ as

$$\Psi[\tilde{h}_{ij}(x), \sqrt{h(x)}, \phi(x)] = 0, \quad \sqrt{h(x)} < 0 \tag{25}$$

Because $\sqrt{h}$ has a semidefinite range it is for many purposes convenient to introduce a representation in which $\sqrt{h}$ is replaced by its canonically conjugate variable $-\frac{4\pi}{K^2}$ which has an infinite range. The advantages of this representation have been extensively discussed. In the case of pure gravity since $-4\pi K^{-2}$ and $h^{1/2}$ are conjugate, we can write for the transformation to the representation where $\tilde{h}_{ij}$ and K are definite

$$\Phi[\tilde{h}_{ij}(x), K(x)] = \int_0^\infty e^{\frac{i}{3}4\pi^{-2}\int\sqrt{h(x)}K(x)d^3x}\, \Psi[\tilde{h}_{ij}(x)]\delta\sqrt{h(x)} \tag{26}$$

and inversely,

$$\Psi[\tilde{h}_{ij}(x)] = \int_{-\infty}^\infty e^{-\frac{i}{3}4\pi^{-2}\int\sqrt{h(x)}K(x)d^3x}\, \Phi[\tilde{h}_{ij}(x), K(x)]\delta K(x) \tag{27}$$

In each case the functional integrals are over the values of $h^{1/2}$ or K at each point of the spacelike hypersurface and we have indicated limits of integration. The condition (25) implies through (26) that $\Phi[\tilde{h}_{ij}, K]$ is analytic in the lower-half K plane. The contour in (27) can thus be distorted into the lower-half K plane. Conversely, if we are given $\Phi[\tilde{h}_{ij}, K]$ we can reconstruct the wave function Ψ which satisfies the boundary condition (25) by carrying out the integration in (27) over a contour which lies below any singularities of $\Phi[\tilde{h}_{ij}, K]$ in K. In the presence of matter K and $\tilde{h}_{ij}$ remain convenient labels for the wave functional provided the labels for the matter-field amplitudes ϕ are chosen so that a multiple of K is canonically conjugate to $h^{1/2}$. In cases where the matter-field action itself involves the scalar curvature this means that the label ϕ will be the field amplitude rescaled by some power of $h^{1/2}$. For example, in the case of a conformally invariant scalar field the appropriate label is $\phi = \varphi\, h^{1/6}$. With this understanding we can write for the functionals

$$\Psi(x) = \Psi[\tilde{h}_{ij}(x), \phi(x)], \quad \Phi(x) = \Phi[\tilde{h}_{ij}(x), K(x), \phi(x)] \tag{28}$$

and the transformation formulas (26) and (27) remain unchanged.

D. Hermiticity

The introduction of wave functions as functional integrals [Eq. (14)] allows the definition of a scalar product with a simple geometric interpretation in terms of sums over spacetime histories. Consider a wave function Ψ defined by the integral

$$\Psi[h_{ij}(x), \phi(x)] = N \int_C e^{iS[g(x),\phi(x)]} \delta\, g(x(t))\, \delta\, \phi(x(t)) \tag{29}$$

over a class of four-geometries and fields C, and a second wave function Ψ' defined by a similar sum over a class C'. The scalar product

$$(\Psi', \Psi) = \int \Psi'[h_{ij}(x(t)), \phi(x(t))]^* \, \Psi[h_{ij}(x(t)), \phi(x(t))] \delta\, h\, (x(t)) \delta\, \phi(x(t)) \tag{30}$$

has the geometric interpretation of a sum over all histories

$$(\Psi', \Psi) = N'\, N \int e^{iS[g(x(t)),\phi(x(t))]} \delta\, g\, (x(t)) \delta\, \phi(x(t)) \tag{31}$$

where the sum is over histories which lie in class C to the past of the surface and in the time reversed of class C' to its future.

The scalar product (30) is not the product that would be required by canonical theory to define the Hilbert space of physical states. That would presumably involve integration over a hypersurface in the space of all three-geometries rather than over the whole space as in (30). Rather, Eq. (30) is a mathematical construction made natural by the functional-integral formulation of quantum gravity.

In gravity we expect the field equations to be satisfied as identities. An extension of the argument leading to Eq. (19) will give

$$\int H(x)\, e^{iS[g(x),\phi(x)]} \delta\, g(x)\, \delta\, \phi(x) = 0 \tag{32}$$

for any class of geometries summed over and for any intermediate spacelike surface on which $H(x)$ is evaluated. Equation (32) can be evaluated for the particular sum which enters Eq. (31). $H(x)$ can be interpreted in the scalar product as an operator acting on either Ψ' or Ψ. Thus,

$$(H\, \Psi', \Psi) = (\Psi', H\, \Psi) = 0. \tag{33}$$

The Wheeler-DeWitt operator must therefore be Hermitian in the scalar product (2.19).

Since the Wheeler-DeWitt operator is a second-order functional-differential operator, the requirement of Hermiticity will essentially be a requirement that certain surface terms on the boundary of the space of three-metrics vanish and, in particular, at $h^{1/2}=0$ and $h^{1/2}=\infty$. As in ordinary quantum mechanics these conditions will prove useful in providing boundary conditions for the solution of the equation.

III. GROUND-STATE WAVE FUNCTION

In this section, we shall put forward in detail our proposal for the ground-state wave function for closed cosmologies. The wave function depends on the topology and the three-metric of the spacelike surface and on the values of the matter field on the surface. For simplicity we shall begin by considering only S^3 topology. If the topology of the four-geometry is not that of a product of a three-surface with the real line or the circle, the volume will vanish. As a result, the topology of the three-geometry will change. One cannot calculate the amplitude for such topology change from the Wheeler-DeWitt equation but one can do so using the Euclidean functional integral. We shall estimate the amplitude in some simple cases in Sec. VIII.

A qualitative discussion of the expected behavior of the wave function at large three-volumes can be given on the basis of the semiclassical approximation when $\Lambda>0$ as follows. The four-sphere has the largest volume of any real solution to (4.2). As the volume of the three-geometry becomes large one will reach three-geometries with no longer fit anywhere in the four-sphere. We then expect that the stationary-phase geometries become complex. The ground-state wave function will be a real combination of two expansions like (4.1) evaluated at the complex-conjugate stationary-phase four-geometries. We thus expect the wave function to oscillate as the volume of the three-geometry becomes large. If it oscillates without being strongly damped this corresponds to a universe which expands without limit.

The above considerations are only qualitative but do suggest how the behavior of the ground-state wave function determines the boundary conditions for the Wheeler-DeWitt equation. In the following we shall make these considerations concrete in a minisuperspace model.

V. MINISUPERSPACE MODEL

It is particularly straightforward to construct minisuperspace models using the functional-integral approach to quantum gravity. One simply restricts the functional integral to the restricted degrees of freedom to be quantized. In this and the following sections, we shall illustrate the general discussion of those proceeding with a particularly simple minisuperspace model. In it we restrict the cosmological constant to be positive and the four-geometries to be spatially homogeneous, isotropic, and closed so that they are characterized by a single scale factor. An explicit metric in a useful coordinate system is

$$d s^2=\sigma^2 \left[-N(t)^2 \, d t^2+a(t)^2 \, d \Omega_3^2\right], \tag{34}$$

where $N(t)$ is the lapse function and $\sigma^2 = 1/2\,\Lambda$. For the matter degrees of freedom, we take a single conformally invariant scalar field which, consistent with the geometry, is always spatially homogeneous, $\phi(t)$. The wave function is then a function of only two variables:

$$\begin{aligned} \Psi &= \Psi(a, \phi) \\ \Phi &= \Phi(K, \phi) \end{aligned} \tag{35}$$

Models of this general structure have been considered previously by DeWitt, Isham and Nelson, and Blyth and Isham.

To simplify the subsequent discussion we introduce the following definitions and rescalings of variables:

$$\begin{aligned} \phi &= \sqrt{a\,(2\,\pi\,\sigma^2)} \\ \Lambda &= 3\,\frac{A}{\sigma^2} \\ H^2 &= |\Lambda| \end{aligned} \tag{36}$$

The Lorentzian action keeping X and a fixed on the boundaries is

$$S = \frac{1}{2}\int \frac{a}{N}\,(\dot{X}^2 - X^2)\,d\,t - \frac{1}{2}\int \frac{a\,\dot{a}^2}{N} + \frac{N}{a} - \frac{\Lambda\,a^3}{3}\,d\,t \tag{37}$$

From this action the momenta $\pi_a = -\frac{a\,\dot{a}}{N}$ and $\pi_X = \frac{a\,\dot{X}}{N}$ conjugate to a and X can be constructed in the usual way. The Hamiltonian constraint then follows by varying the action with respect to the lapse function and expressing the result in terms of a, X, and their conjugate momenta. One finds

$$\left(1 - \tau^2\,a^2 + \frac{a^4}{4}\,X^2 + \frac{X^2}{2}\right) = 0 \tag{38}$$

The Wheeler-DeWitt equation is the operator expression of this classical constraint. There is the usual operator-ordering problem in passing from classical to quantum relations but its particular resolution will not be central to our subsequent semiclassical considerations. A class wide enough to remind oneself that the issue exists can be encompassed by writing

$$\tau^2 = -\frac{1}{\sqrt{p}}\,\frac{\partial^2}{\partial\,a^2}\,\sqrt{p} \tag{39}$$

although this is certainly not the most general form possible. In passing from the classical constraint to its quantum operator one must also consider the possibility of a matter-energy correction term. This will lead to an addition to the quantum constant in the equation. Thus we write for the arbitrary version of S^2,

$$\left(\frac{p}{a}\,\frac{\partial}{\partial\,a}\,\frac{p}{a}\,\frac{\partial}{\partial\,a} - a^2 + a^4 - \frac{X^2}{2}\right)\Psi(a, X) = 0 \tag{40}$$

A useful property stemming from the conformal invariance of the scalar field is that this equation separates. If we assume reasonable behavior for the function Ψ in the amplitude of the scale field we can expand in harmonic-oscillator eigenstates

$$\Psi(a,X) = \sum_n c_n(a)\, u_n(X) \tag{41}$$

where

$$u_n(X) = \left(\frac{1}{2} d^2\, X^2 + X^2\right) u_X = (n + \frac{1}{2})\, \hbar\, u_X \tag{42}$$

The consequent equation for the $c_n(a)$ is

$$\frac{1}{2}\left(-\frac{1}{a} p \frac{d}{d\,a} p \frac{d\,c_n}{d\,a} + (a^2 - \Lambda\, a^4)\, c_n\right) = (n + \frac{1}{2})\, \epsilon\, c_n \tag{43}$$

For small a this equation has solutions of the form

$$c_n = a^{p-1} \tag{44}$$

If p is an integer then there may be a $\log(a)$ factor. For large α the possible behaviors are

$$c_n \sim a^{-p/2+1} \exp\left(\pm i \frac{1}{3} H\, a^3\right) \tag{45}$$

To construct the solution of Eq. (5.11) which corresponds to the ground state of the minisuperspace model we turn to our Euclidean functional-integral prescription as applied to this minisuperspace model, the prescription of Sec. III for $\Psi_0(x_0)$ which sums over all those Euclidean geometries and field configurations which are represented in the minisuperspace and which satisfy the ground-state boundary conditions. The geometrical sum would be over compact geometries of the form

$$d\,s^2 = \sigma^2\, (\tau^2\, d\,\tau^2 + a^2(\tau)\, d\,\Omega_3^2) \tag{46}$$

for which $a(\tau)$ matches the prescribed value of a_0 on the hypersurface of interest. The prescription for the matter field would be to sum over homogeneous fields $X(\tau)$ which match the prescribed value X_0 on the surface and which are regular on the compact geometry. Explicitly we could write

$$\Psi_0(a_0, X_0) = \int \mathcal{D}\, a \exp\left(-I[a, X]\right), \tag{47}$$

where, defining $d\eta = d\tau/a$, the action is

$$I = \frac{1}{2}\int d\eta\left[\left(\frac{da}{d\eta}\right)^2 - a^2 + a^4\left(\frac{dX}{d\eta}\right)^2 + X^2\right] \tag{48}$$

A conformal rotation [in this case of $a(\eta)$] is necessary to make the functional integral in (5.15) converge.

An alternative way of constructing the ground-state wave function for the minisuperspace model is to work in the K representation. Here, introducing

$$k = -\frac{\partial F}{\partial a} \tag{49}$$

as a simplifying measure of K, one would have

$$\Psi_0(a_0, X_0) = \int \mathcal{D}a \exp\left(-I'[a, X]\right) \tag{50}$$

The sum is over the same class of geometries and fields as in (5.15) except they must now assume the given value of k on the bounding three-surface. That is, on the boundary they must satisfy

$$k_0 = \frac{1}{3}\frac{da}{d\tau} \tag{51}$$

The action I' appropriate for holding k fixed on the boundary is

$$I' = k_0\, a_0^3 + H \tag{52}$$

[cf. Eq. (3.6)]. Once $\Psi_0(a_0, X_0)$ has been computed, the ground-state wave function $\Psi_0(k_0, X_0)$ may be recovered by carrying out the contour integral

$$\Psi_0(a_0, X_0) = \frac{1}{2\pi i}\int dk_0\, e^{k_0 a_0^3}\,\Psi_0(k_0, X_0) \tag{53}$$

where the contour runs from $-\infty$ to $+\infty$ to the right of any singularities of $\Psi_0(k_0, X_0)$.

From the general point of view there is no difference between computing $\Psi_0(k_0, X_0)$ directly from (5.15) or via the K representation from (5.21). In Sec. VI we shall calculate the semiclassical approximation to $\Psi_0(a_0, X_0)$ both ways with the aim of advancing arguments that the rules of Sec. III define a wave function which may reasonably be considered as the state of minimal excitation and of displaying the boundary conditions under which Eq. (5.11) is to be solved.

VI. GROUND-STATE COSMOLOGICAL WAVE FUNCTION

In this section, we shall evaluate the ground-state wave function for our minisuperspace model and show that it possesses properties appropriate to a state of minimum excitation. We shall first evaluate the steepest-descents semiclassical approximation from the functional integral as approximation to IV. We shall then solve the Wheeler-DeWitt equation with the boundary conditions implied by the semiclassical approximation to obtain the precise wave function.

It is the exponent of the semiclassical approximation which will be most important in its interpretation. We shall calculate only this exponent of the prefactor of action and leave the determination of the detailed equation (4.14) to the solution of the differential equation itself. Thus, for example, if there were a single real Euclidean extremum of least action we would write for the semiclassical-approximation to the functional integral in Eq. (5.15)

$$\Psi_0(a_0, X_0) = N e^{-I} \tag{54}$$

Here, $I(a_0, X_0)$ is the action (5.16) which is evaluated at the extremum of the functional integral and under the boundary conditions specified out in (5.11) and which match the arguments of the wave function on a fixed-η hypersurface.

A. The matter wave function

A considerable simplification in evaluating the ground-state wave function arises from the fact that the energy-momentum tensor of an extremizing conformally invariant field vanishes in the compact geometries considered over as a consequence of the ground-state boundary conditions. One can see this because the conformal transformations of the class we are considering are space. A conformal invariant wave function is therefore flat space which is a constant scalar field relies only on solution of the conformally invariant wave function in flat space which is a constant on the boundary three-sphere. This energy-momentum tensor of this field is zero. The implies that it is zero in any geometry of the class (5.14) because the energy-momentum tensor of a conformally invariant field scales by a power of the conformal factor under a conformal transformation.

More explicitly in the minisuperspace model we can show that the matter and gravitational functional integrals in (5.15) may be evaluated separately. The ground-state boundary conditions imply that geometries in the sum are conformal to half of a Euclidean Einstein-static universe, i.e., that the range of η is $(-\infty, 0)$. The boundary conditions at infinity are that $X(\eta)$ and $a(\eta)$ vanish. The boundary conditions at $\eta = 0$ are that $a(0)$ and $X(0)$ match.