PHYSICAL REVIEW D VOLUME 40, NUMBER 8 15 OCTOBER 1989

Time and the interpretation of canonical quantum gravity

William G. Unruh
Cosmology Program, Canadian Institute for Advanced Research, Department of Physics, University of British Columbia, Vancouver, British Columbia, Canada V6T 2A6 and Institute for Theoretical Physics, University of California, Santa Barbara, California 93106

Robert M. Wald
Enrico Fermi Institute and Department of Physics, 5640 S. Ellis Avenue, University of Chicago, Chicago, Illinois 60637
(Received 15 December 1988)

The unsatisfactory status of the interpretation of the wave function of the Universe in canonical quantum gravity is reviewed. The "naive interpretation" obtained by straightforwardly applying the standard interpretive rules to the canonical quantization of general relativity is manifestly unacceptable; the "WKB interpretation" has only a limited domain of applicability; and the "conditional probability interpretation" requires one to pick out a "preferred time variable" (or preferred class of such variables) from among the dynamical variables. Evidence against the possibility of using a dynamical variable to play the role of "time" in the conditional probability interpretation is provided by the fact (proven here) that in ordinary Schrödinger quantum mechanics for a system with a Hamiltonian bounded from below, no dynamical variable can correlate monotonically with the Schrödinger time parameter t, and thus the role of t in the interpretation of Schrödinger quantum mechanics cannot be replaced by that of a dynamical variable. We also argue that the interpretive problems of quantum gravity are not alleviated by the incorporation of observers into the theory. Faced with these difficulties, we seek a formulation of canonical quantum gravity in which an appropriate nondynamical time parameter is present. By analogy with a parametrized form of ordinary Schrödinger quantum mechanics, we make a proposal for such a formulation. A specific proposal considered in detail yields a theory which corresponds at the classical level to general relativity with an arbitrary, unspecified cosmological constant. In minisuperspace models, this proposal yields a quantum theory with satisfactory interpretive properties, although it is unlikely that this theory will admit sufficiently many observables for general spacetimes. Nevertheless, we feel that the approach suggested here is worthy of further investigation.

I. INTRODUCTION

Theories typically are formulated in terms of quantities taking values in abstract mathematical spaces. In order to relate these quantities to physical phenomena, one needs an interpretation of the theory. In this paper, by an "interpretation" we mean a description, in ordinary language, of what an observer would see or experience when the mathematical quantities used by the theory to describe the state of the system take on any of their allowed values. Thus, it should be noted that by our usage of the term, the Copenhagen and Everett "interpretations" of ordinary Schrödinger quantum mechanics are equivalent, at least for formulations of the Everett "interpretation" which are interpretations in our sense, since they give the same rules for what an observer "sees." An interpretation is an essential part of any theory and, indeed, interpretations of some classical theories, such as general relativity, are not entirely trivial to state. However, it is in the case of quantum theories that the issue of interpretations has attracted the most attention.

In fact, ordinary (nongravitational) quantum-mechanical theories such as standard Schrödinger quantum mechanics possess an interpretation which is entirely satisfactory in that their interpretive rules are well defined and yield physical predictions which agree with experiment. (The fact that these predictions are probabilistic in nature and, apparently, do not admit a consistent picture of "objective reality" are not valid objections to an "interpretation" according to our usage of the term.) However, as we shall elucidate further below, the notion of "time" plays a vital role in these interpretive rules;[1] in particular, measurements are made at "instants of time" and probabilities are assigned only to such measurements (not, in particular, to "histories"). Therefore, it should not be surprising that severe difficulties arise in the formulation of an interpretation of quantum theory corresponding to general relativity, since the nature of "time" and dynamical evolution in a generally covariant theory has important differences from the corresponding notion in theories which are not formulated in a diffeomorphism- (or reparametrization-) invariant manner. This appears to be the main reason why quantum gravity, even, say, in minisuperspace models where many technical difficulties such as renormalization can be avoided, does not, at present, possess a satisfactory interpretation.

In this paper we shall make a proposal for a version of quantum gravity in which a notion of "time" is present that allows a well-defined formulation of interpretive rules. However, our proposal fails to yield a quantum theory which corresponds classically to ordinary general relativity; rather, the specific proposal we shall investigate in Sec. III corresponds to Einstein's equation with an arbitrary, unspecified cosmological constant. Furthermore, although our proposal yields satisfactory interpretive rules and is fully satisfactory in minisuperspace models, it is far from clear that there are a sufficient number of allowed observables in general spacetimes. Nevertheless, we hope that this proposal, or, at the very least, the nature of the proposal, will be of some interest and that the related discussion will help elucidate the role which time plays in the interpretation of a quantum theory as well as some of the difficulties that arise from the notion of time in general relativity.

Classical general relativity admits a Hamiltonian formulation but it is a constrained Hamiltonian formulation. The configuration variable is taken to be a Riemannian metric h_{ab} on a three-dimensional manifold Σ and a spacetime is viewed as being comprised of the development of h_{ab} with time. However, h_{ab} and its canonically conjugate momentum π^{ab},

$$\pi^{ab}=h^{1/2}(K^{ab}-Kh^{ab}) \tag{1.1}$$

(where K_{ab} denotes the extrinsic curvature of Σ with the sign convention of Ref. 2), cannot be freely specified as initial data; rather, they are required to satisfy the so-called Hamiltonian and momentum constraints,

$$0=H_0(h_{ab},\pi^{cd})$$

$$=h^{1/2}[-{}^{(3)}R+h^{-1}(\pi_{ab}\pi^{ab}-\tfrac{1}{2}\pi^2)] , \tag{1.2}$$

$$0=H_a(h_{ab},\pi^{cd})=-2h^{1/2}D^b(h^{-1/2}\pi_{ab}) . \tag{1.3}$$

An appropriate Hamiltonian for classical general relativity then is

$$\mathcal{H}=\int(NH_0+N^aH_a) , \tag{1.4}$$

where N and N^a are, respectively, an arbitrary function and vector field on Σ which have the interpretation of lapse function and shift vector in the spacetime constructed from the time evolution. Variation of $\mathcal{H}$ with respect to N and N^a yields the constraints (1.2) and (1.3), whereas Hamilton's equations of motion for h_{ab} and π^{ab} yield the Einstein evolution equations; see, e.g., Ref. 2 for further discussion.

Note that the infinitesimal canonical transformations generated by the functions $\int_\Sigma \xi^a H_a$ on phase space (for arbitrary vector fields ξ^a) correspond simply to the changes of h_{ab} and π^{ab} resulting from infinitesimal diffeomorphisms of Σ. The situation is more subtle for the infinitesimal canonical transformations generated by functions on phase space of the form $\int_\Sigma \xi^0 H_0$ (for arbitrary functions ξ^0): the changes induced on h_{ab} and π^{ab} correspond to infinitesimal spacetime diffeomorphisms (of the spacetime metric constructed from h_{ab}, π^{ab}, N, and N^a) only when the field equations are satisfied. (Moreover, these spacetime diffeomorphisms are "field dependent.") Nevertheless, on the "constraint hypersurface" of phase space [i.e., the h_{ab} and π^{ab} satisfying Eqs. (1.2) and (1.3)], the canonical transformations generated by $\int_\Sigma \xi^a H_a$ and $\int_\Sigma \xi^0 H_0$ correspond to the spacetime diffeomorphisms of general relativity.

The standard canonical quantization rules for formulating a quantum theory in the "coordinate representation" based upon a classical Hamiltonian system suggest the following structure of the quantum theory in this case.[3] States of ths system are taken to be wave functions of the configuration variable, so $\Psi=\Psi(\tau;h_{ab})$ where τ denotes the time variable occurring in the Hamiltonian formulation. On the Hilbert space of such states, h_{ab} is represented by the multiplication operator and π^{ab} is to be represented by the functional differentiation $-i\delta/\delta h_{ab}$. The time evolution of Ψ is then given by the Schrödinger equation

$$i\frac{\partial\Psi}{\partial\tau}=\mathcal{H}\Psi , \tag{1.5}$$

where in $\mathcal{H}(h_{ab},\pi^{ab})$ we replace h_{ab} and π^{ab} by their operator expressions. If, as in our case, constraints are present, then following the Dirac prescription we impose these as conditions on the state vector. Thus, Ψ also is required to satisfy

$$\left[\int_\Sigma \xi^0 H_0\right]\Psi=0 , \tag{1.6}$$

$$\left[\int_\Sigma \xi^a H_a\right]\Psi=0 , \tag{1.7}$$

where ξ^a and ξ^0 are, respectively, an arbitrary vector field and function on Σ, and in H and H^a, we again replace h_{ab} and π^{ab} by their operator representatives. Note that if we take $\xi^0=N$ and $\xi^a=N^a$, Eqs. (1.6) and (1.7) have the important consequence that

$$\mathcal{H}\Psi=0 \tag{1.8}$$

and hence by the Schrödinger equation (1.5), that Ψ is independent of τ, i.e., $\Psi=\Psi(h_{ab})$.

Roughly speaking, the quantum constraints (1.6) and (1.7) can be interpreted as requiring the invariance of Ψ under the infinitesimal canonical transformations generated by $\int_\Sigma \xi^0 H_0$ and $\int_\Sigma \xi^a H_a$, which, as discussed above, correspond to infinitesimal spacetime diffeomorphisms on the manifold of solutions. For the momentum constraint (1.7), this holds literally, since Eq. (1.7) implies that $\Psi(h_{ab})$ is unchanged when an infinitesimal diffeomorphism on Σ is applied to h_{ab}. This makes Ψ a functional on "superspace," the set of diffeomorphism equivalence classes of metrics on Σ. It does not appear possible to give as literal an interpretation of the quantum Hamiltonian constraint (1.6), known as the *Wheeler-DeWitt* equation, as corresponding to the invariance of Ψ under a variation of h_{ab} corresponding to an infinitesimal diffeomorphism in spacetime which moves points on Σ in the direction orthogonal to Σ. Nevertheless, this interpretation of the quantum constraints can be viewed as accounting for why $\partial\Psi/\partial\tau=0$ in the formalism.

In the following we shall consider canonical quantum gravity formulated by the above rules. To avoid some of the severe technical problems which arise, such as renormalization, and also for simplicity and definiteness, we will, for the most part, focus attention below on a simple minisuperspace model involving a Robertson-Walker cosmology and a (spatially homogeneous) Klein-Gordon scalar field ϕ, self-interacting via a potential $V(\phi)$. (However, our discussion will not depend upon the details of this model.) Thus, the configuration variables for this system are simply ϕ and the Robertson-Walker scale factor a. The "wave function of the Universe" is a function of these two variables: $\Psi=\Psi(a,\phi)$. For the case of a cosmologically flat ($k=0$) model,

$$ds^2=-N^2d\tau^2+a^2(\tau)(dx^2+dy^2+dz^2)\ , \tag{1.9}$$

a suitable classical Hamiltonian (in units where $8\pi G=1$) is

$$\mathcal{H}=N\left[-\frac{1}{12a}\pi_a^2+\frac{1}{2a^3}\pi_\phi^2+a^3V(\phi)\right]\ . \tag{1.10}$$

In quantum theory, the momentum constraints are automatically satisfied, so the only equation imposed upon Ψ is the Wheeler-DeWitt equation arising from the classical constraint

$$\mathcal{H}=0\ . \tag{1.11}$$

With the "Laplacian" factor ordering chosen for the "kinetic terms" in Eq. (1.10), this equation takes the explicit form

$$-\frac{1}{12a^2}\frac{\partial}{\partial a}\left[a\frac{\partial\Psi}{\partial a}\right]+\frac{1}{2a^3}\frac{\partial^2\Psi}{\partial\phi^2}-a^3V(\phi)\Psi=0\ . \tag{1.12}$$

(This form holds even if the lapse function N is allowed to depend upon a and ϕ; i.e., the ordinary wave operator is conformally invariant in two dimensions; for an n-dimensional minisuperspace, the appropriate factor-ordering choice would appear to be the conformally invariant wave operator.) Although this is a highly simplified model of quantum gravity, the features leading to interpretative difficulties remain fully present. It should be noted that some technical issues also remain in the model; in particular, there are some "factor ordering" ambiguities in the definition of the Wheeler-DeWitt equation. (As already mentioned, the factor ordering we have chosen has the advantage of being conformally invariant as well as invariant under redefinition of the "coordinates" a,ϕ on superspace.) In addition, for quantum cosmology based upon the model, the issue arises as to what boundary conditions should be imposed on solutions of the Wheeler-DeWitt equation in order to obtain the solution corresponding to our Universe; several such proposals have recently been given[4,5] and widely discussed. However, we shall not address any of these issues here, since our entire concern in this paper is the interpretation of Ψ. If the wave function of the Universe is given by the mathematical expression $\Psi(a,\phi)$, what does an observer of the Universe "see" or experience?

A possible approach toward formulating an interpretation of Ψ is to treat Ψ in essentially the same way as we treat wave functions in ordinary quantum mechanics. We will refer to this as the "naive interpretation" of canonical quantum gravity. Thus, we could give a Hilbert-space structure to the wave functions Ψ by demanding that they be square integrable with respect to a and ϕ. Indeed, such a Hilbert-space structure is implicit in the above canonical quantization rules. We then could interpret Ψ, as in ordinary quantum mechanics, as giving the amplitude that an observer, making a measurement at parameter time τ, will find the values a and ϕ. However, the fact that Ψ is forced to be independent of τ by the Wheeler-DeWitt equation makes this interpretation not viable. Indeed, the Wheeler-DeWitt equation must hold exactly at all times, including after measurements are made. Thus, the situation is analogous to demanding in ordinary quantum mechanics, not only that the wave function be in an eigenstate of zero energy (so that the wave function is time independent), but also that no measurements can be made which disturb this condition. This restricts the allowed observables in the theory to be those which commute with the Hamiltonian, i.e., to those which are time independent. Hence, in this interpretation of Ψ, dynamical variables such as a and ϕ would not be measurable (since they fail to commute with $\mathcal{H}$), and the Universe would appear to be strictly time independent with respect to those very few observables which are measurable.

In this connection, it should be noted that for some dynamical systems, the requirement that an observable commute with the Hamiltonian H can be extremely restrictive. In particular, for a classical system which is ergodic in the sense that the dynamical trajectories are "mixing" on each surface of constant energy, any classical observable (i.e, measurable function on phase space) which has vanishing Poisson brackets with the Hamiltonian must be constant on each dynamical trajectory and hence constant on each surface of constant energy. Thus, for such systems, the only classical observables which have vanishing Poisson brackets with H are functions of H. Hence, in the quantum theory, apart from functions of H, there presumably are no quantum observables corresponding to classical observables which commute with H. Thus, the suggestion by Page and Wooters[6] that dynamical evolution can be described in terms of stationary observables as a dependence upon internal clock readings is manifestly not viable for sufficiently complicated (i.e., ergodic) systems: There are no nontrivial stationary observables and "internal clocks" would correlate with other observables in a random manner.

The difficulty with the naive interpretation can be traced directly to the conflict between the role of time in quantum theory and the nature of time in general relativity. In quantum mechanics, all measurements are made at "instants of time"; *only quantities referring to the instantaneous state of a system have physical meaning.* In particular, "histories" are unmeasurable in quantum theory. On the other hand, in general relativity "time" is merely an arbitrary label assigned to a spacelike hypersurface. The physically meaningful quantities must be independent of such labels; they must be diffeomorphism invariant. In other words, only the spacetime geometry is measurable; i.e., *only histories have physical meaning.* Thus, it should not be surprising that when one naively combines quantum theory and general relativity, the only meaningful quantities which survive are those which are both instantaneously measurable (i.e., refer to quantities defined on a spacelike hypersurface) and yet depend only on the spacetime geometry (i.e., are independent of choice of spacelike hypersurface). These are precisely the conserved quantities, i.e., the observables which commute with the Hamiltonian. Note that the "naive quantization" of any classical theory which has a time reparametrization invariance will share this feature of quantum general relativity. Thus, further insight into what has gone wrong can be obtained from comparison with the parametrized form of ordinary Schrödinger quantum mechanics, which will be discussed below.

A possible modification of the above rules would be to postulate that quantities, such as a and ϕ, which fail to commute with the Hamiltonian constraint (i.e., Wheeler-DeWitt) operator are, nevertheless, measurable. However, one then would be faced with the problem of how to incorporate the information obtained from the results of measurements of such quantities into future predictions. If one "reduces" the wave function in the standard way, the Hamiltonian constraint will be violated. If one projects the resulting reduced wave function onto the constraint subspace (or does not reduce it at all), an immediate repetition of the same measurement could yield a totally different result. Thus, there does not appear to be any way of making the naive interpretation viable.

A possible reaction to the failure of the naive interpretation would be to object that our simplified model of the universe does not include the presence of observers within the model, and that it is not fair (and, perhaps, even not necessary) to require an interpretation of the theory until observers are explicitly incorporated into the model. However, this same comment could be made with equal force in ordinary (nongravitational) quantum theory, where human observers also have not been properly incorporated into any model system. It is true that because of the universal nature of gravitation, in quantum gravity the influence of the observer on the observed system (i.e., gravitational field) cannot be strictly zero. However, it seems clear that the influence of a human observer on the global state of the Universe should, in fact, be far more negligible than the influence of such an observer on a typical laboratory experiment in ordinary quantum theory. Thus, in the following discussion, we shall proceed on the assumption that, as in ordinary quantum theory, it should not be necessary to explicitly incorporate observers into the model in order to obtain an interpretation. We will return to this issue at the end of this section and argue that, in any case, such an incorporation is not likely to alleviate any of the interpretive problems which occur without their explicit presence and, indeed, it creates additional difficulties.

In the literature on quantum cosmology, statements about physical phenomena usually are extracted from the wave function of the Universe, $\Psi(a,\phi)$, by one of two means, which we shall refer to, respectively, as the "WKB interpretation" and the "conditional probability interpretation." We now discuss these two interpretations in turn.

The WKB interpretation is applicable only if Ψ takes the WKB form

$$\Psi=\mathcal{A}\exp(i\mathcal{S}) , \tag{1.13}$$

where $\mathcal{S}$ is a real function which is rapidly varying compared with $\mathcal{A}$. In that case, the Wheeler-DeWitt equation implies, to a good approximation, that $\mathcal{S}$ satisfies the classical Hamilton-Jacobi equation. As in ordinary particle mechanics, associated with any solution of the Hamilton-Jacobi equation is a family of classical trajectories on configuration space. These trajectories are tangent to the "current vector" $j^A=|\mathcal{A}|^2\nabla^A\mathcal{S}$ on superspace, which is conserved (in the WKB approximation) as a consequence of the Wheeler-DeWitt equation. (In this formula, the superspace index A is raised using the natural metric on superspace arising from the "kinetic terms" in the Hamiltonian.) The WKB interpretation consists of saying that, if the wave function of the Universe takes the form (1.13), then an observer in the Universe who makes measurements of a and ϕ, which are not so precise as to significantly disturb the Universe, would obtain results consistent with one of these classical trajectories. It would be natural to supplement this statement by using the conserved current j^A to assign a probability density for observing a given classical trajectory. This can be done if superspace can be foliated by hypersurfaces such that each classical trajectory crosses each hypersurface once and only once.[7] If the metric on superspace were positive definite, or if the "potential term" ${}^{(3)}R$ in the Wheeler-DeWitt equation had a definite sign, the surfaces of constant $\mathcal{S}$ would automatically satisfy this property (in the "classically allowed" region) and the conserved probability density (which is proportional to $\nabla^A\mathcal{S}\nabla_A\mathcal{S}$) would be non-negative. However, since the metric on superspace is not positive definite and ${}^{(3)}R$ also is, in general, of indefinite sign, there is no reason why $\nabla^A\mathcal{S}\nabla_S\mathcal{S}$ need be non-negative. Thus, there appear to be serious difficulties in obtaining an everywhere non-negative probability density for the classical trajectories valid in all circumstances.

In its range of applicability, the WKB interpretation is a genuine interpretation in our sense, and essentially all "predictions" given in the recent literature in quantum cosmology have been made by means of it. However, it suffers from the obvious shortcoming of having only a very limited range of applicability; i.e., it applies only to cases where the wave function of the Universe is "very nearly classical." Furthermore, even in cases where the WKB approximation holds, it typically will be valid only in a limited region of a-ϕ space. What is the probability that an observer will measure values of a and ϕ in this non-WKB region? [Here, by "non-WKB region" we include any region where the "Euclidean WKB form" $\Psi=\mathcal{A}e^{\mathcal{S}}$ holds, since no satisfactory interpretation exists in that case either, as there are no classical trajectories, and the conserved current $j^A=\frac{1}{2}(\Psi^*\nabla^A\Psi-\Psi\nabla^A\Psi^*)$ vanishes there.] What would the Universe look like to an observer who finds himself in a non-WKB regime?

The WKB interpretation plausibly can be extended to the case where the wave function of the Universe takes the WKB form only with respect to some dynamical variables and all the non-WKB variables can be treated as "small perturbations."[8,9] However, there is no reason, *a priori,* why Ψ should take this form on all of superspace, nor is there any evidence from the solutions of the Wheeler-DeWitt equation in minisuperspace models that it does. The WKB interpretation also has been applied to cases where Ψ takes the form of a sum of WKB solutions. For example, in the solutions of Ref. 4, Ψ is real and hence cannot take the form (1.13); at best, it can be a sum of two WKB solutions. The interpretation which has been proposed in that case is that the Universe corresponds to a classical trajectory of either of the two WKB solutions. This is analogous to interpreting the wave function resulting from a two-slit experiment in ordinary quantum mechanics as corresponding to classical trajectories in which the particle goes through one (and only one) of the two slits. In fact, in ordinary quantum mechanics the predictions made from this interpretation of the wave function for the two-slit experiment will be correct for the case in which accurate enough position measurements are made near the slits so that it can be determined through which slit the particle has passed—through not so accurate that the particle motion is significantly disturbed. (Interactions of the particle with the environment near the slits would have the same effect as a measurement.) Furthermore, even if no measurements or interactions occur, highly precise position or momentum measurements may be needed to observe the interference effects. However, in principle, highly nonclassical behavior can be observed. The inability of the WKB interpretation to account for this and thus to properly extend even to the case of a sum of two WKB solutions highlights its very limited range of applicability.

One might attempt to rescue the WKB interpretation from its shortcoming of limited applicability by postulating that observers can exist only when the Universe is very nearly classical, so an interpretation is required only in that case. However, even if one were to accept this rather radical postulate, one would still be in the highly unsatisfactory situation of having only an approximate interpretation. How close to a WKB solution must the Universe be before observers can exist? How large can the departures be from the predictions of the WKB interpretation due to the fact that the WKB form does not hold exactly? These equations must be answered if the theory is to have any predictive power, but it does not appear that such answers are possible unless there exists an "exact" interpretation of the theory which does not rely upon the WKB approximation for its validity.

We turn, now to a discussion of the conditional probability interpretation. The conditional probability interpretation asserts that if an observer measures a particular value of one of the dynamical variables, then the wave function of the Universe (evaluated at that value of that variable) yields the amplitude for measuring the various possible values of the remaining dynamical variables. Thus, for example, in our minisuperspace model, if an observer measures the radius of the Universe to be a_0, then the probability density for measuring the value of the scalar field to be ϕ would be given by

$$P(\phi)=|\Psi(a_0,\phi)|^2 / \int |\psi(a_0,\phi)|^2 d\mu_\phi \ , \qquad (1.14)$$

where $d\mu_\phi$ is a measure on ϕ space (and the "probability density" is specified with respect to this measure). Alternatively, if the observer measures the value ϕ_0 of the scalar field, then the probability for measuring the value of a would be given by

$$P(a)=|\Psi(a,\phi_0)|^2 / \int |\Psi(a,\phi_0)|^2 d\mu_a \ , \qquad (1.15)$$

where $d\mu_a$ is a measure on a space.

It is far from clear how $d\mu_\phi$ and $d\mu_a$ are to be chosen. (They, of course, contain as much information about the probability distribution as does Ψ.) However, an even more serious difficulty with the interpretation is that it is clear that not all quantities are suitable for playing the role of the variable which "sets the conditions" for the other variables. For example, even in the context of ordinary Schrödinger quantum mechanics of a particle, although at fixed $t=t_0$, the wave function $\psi(t_0,X,Y,Z)$ gives the amplitude for finding the particle at position (X,Y,Z), at fixed $X=X_0$, $\psi(t,X_0,Y,Z)$ is not in any sense proportional to the amplitude that the other coordinates of the particle are (Y,Z) and the time is t. (Note that for the purposes of making this point here, we have blurred the distinction between the time *parameter* t and the *dynamical* variables X,Y,Z. We will return to this point below in the context of a parametrized theory, where t is replaced by a dynamical variable T.) In minisuperspace models, this difficulty often manifests itself by the fact that in typical solutions $\Psi(a,\phi)$ chosen for study, $\Psi(a_0,\phi)$ fails to be square integrable in ϕ, and/or $\Psi(a,\phi_0)$ fails to be square integrable in a, thus making formulas (1.14) and (1.15) meaningless.

Since not all variables are suitable for "setting the conditions" for the other variables, the conditional probability interpretation must specify which variables are suitable to "fix," so that Ψ yields the conditional probabilities for the remaining variables. This problem has not, as yet, been solved, so the conditional probability interpretation remains seriously deficient.

An interesting variant of the above conditional probability interpretation which makes effective use of the available geometrical structure of superspace is implicit in the work of Misner.[10] Consider the case of a spatially homogeneous class of models, with an n-dimensional minisuperspace. The "kinetic terms" in the Hamiltonian (1.4) define a metric on minisuperspace (or, really , a conformal metric, since the lapse function N could be chosen to depend arbitrarily on the spatial geometry). This metric has Lorentz signature because the square of the momentum π conjugate to the "conformal degree of freedom" enters the Hamiltonian with sign opposite that of all the other squared momentum terms [see Eq. (1.2)]. Consequently, the Wheeler-DeWitt equation (1.6) for the wave function of the Universe Ψ on minisuperspace is formally identical to the equation for a linear scalar field Φ on an n-dimensional curved spacetime [with an additional "external potential" corresponding to the term involving the scalar curvature of space in Eq. (1.2)]. Of course, there is a crucial physical difference between the two cases: the scalar field Φ on spacetime is physically measurable and should be represented as an operator on the Hilbert space of states, whereas Ψ is not, in any sense, measurable on minisuperspace; rather the measurable quantities now would be mathematical analogs of the spacetime coordinates. Thus, it would not appear to make any sense to define a quantum theory for Ψ analogous to a "second quantized" theory for Φ. (Such a theory for Ψ often is referred to as a "third quantized" theory.) However, it would make sense to define for Ψ the analog of a "first quantized" theory for Φ, i.e., to make quantum cosmology the mathematical analog of the theory of a relativistic particle (as opposed to a relativistic field) on curved spacetime. The basic idea would be to take the Hilbert space of states in quantum cosmology to be the analog of the Hilbert space of single-particle states for a scalar field on spacetime. The conditional probability interpretation would then be used in the following sense: A configuration variable q, on minisuperspace would be taken as appropriate for "setting the conditions" if its "level surfaces" (i.e., surfaces of constant value) are Cauchy surfaces in minisuperspace. One then would hope to define operators (analogous to the Newton-Wigner position and momentum operators for a scalar field in flat spacetime) to represent the remaining configuration variables and their conjugate momenta on minisuperspace at "time" q.

From the theory of quantum fields on curved spacetime, it is known that if minisuperspace is globally hyperbolic, then mathematically consistent prescriptions can be given for constructing a Hilbert space of states from solutions to the Wheeler-DeWitt equation. However, in order to single out a prescription, additional mathematical structure must be specified. This additional structure is most conveniently expressed in terms of a bilinear map μ on the space of solutions (see proposition 3.1 of Ref. 11). Unfortunately for this program, there does not appear to be any natural choice for such a μ for minisuperspace. Indeed, if minisuperspace possessed a timelike conformal Killing field which scaled the "potential term" in the Wheeler-DeWitt equation at the same rate as the metric, and if this potential term were of the correct sign on all of minisuperspace, then this structure would provide such a natural choice of μ by the procedures used for quantum fields in stationary spacetimes.[12] However, although full superspace does possess a timelike conformal Killing field, it does not properly scale the potential term,[13] and, in any case, in general models the potential term is not everywhere of the correct sign. Thus, there is a serious problem in this approach with the construction of the Hilbert space of states. Furthermore, even if this obstacle could be overcome, it is not clear that sensible analogs of the Newton-Wigner operators will exist. Finally, it is far from clear that these ideas can be generalized to spatially inhomogeneous models, since the metric on superspace will no longer have Lorentz signature when there is more than one conformal degree of freedom; i.e., the resulting superspace metric signature has both multiple plus and minus signs. Thus, at present, this approach does not appear to be viable, although it certainly would appear to be worthy of further investigation.

The basic difficulty with the conditional probability interpretation can be restated in a language which indicates its direct connection to the issue of time is quantum gravity. The basic property that a variable C should satisfy in order to be appropriate for "setting the conditions" is that, for each fixed value of C, a measurement of any of the other dynamical variables must yield one and only one value; it is only under this circumstance that one could expect to meaningfully talk about probability distributions for these other variables. Thus, for example, in ordinary Schrödinger quantum mechanics, one should expect the particle position variable X to be inappropriate for setting the conditions, because when $X=X_0$, the dynamical variable Y could take on many values (since the particle could be at $X=X_0$ at many different times) or no value (since the particle might never be at $X=X_0$). As discussed further by one of us elsewhere,[14] this property for a variable needed to properly set the conditions for the other variables is a key feature of our intuitive notion of "time," as is well expressed in the aphorism "time is that which allows contradictory things to occur." Following Ref. 14 we will refer to this as the "Heraclitian property" of time, on account of Heraclitus' view of the flow of time as a "war of opposites." Thus a variable which is suitable for setting the conditions for the remaining variables could reasonably be referred to as a "time variable." The central difficulty with the conditional probability interpretation is that in quantum gravity the time variable (or various allowed possible choices of time variable) has not been specified. (Such a specification is given in the approach described in the previous two paragraphs, but as concluded above, this approach does not appear to be viable.)

In most theories the notion of "time" that is present at the classical level can be taken over directly in the formulation of the quantum theory. In classical general relativity a spacelike hypersurface in spacetime provides an appropriate realization of the notion of an "instant of time." Thus, at the classical level, the specification of a foliation of spacetime by spacelike hypersurfaces defines a satisfactory notion of time. The problem with this notion of time is that it is closely analogous to the notion of time in a so-called "parametrized version" of particle mechanics, and this notion of time is unsuitable for use in quantum theory in the same manner as the time of a "deparametrized" theory. We now shall briefly review parametrized particle mechanics and the quantum theory based upon it. In doing so we shall also elucidate some of the points raised above regarding the above "naive" and conditional probability interpretations of canonical quantum gravity.

Consider the theory of a nonrelativistic particle moving in one dimension, with dynamical variable X and with t denoting the ordinary (Galilean) time parameter. Let the action S be given by

$$S=\int \tilde{L}\, dt \tag{1.16}$$

with Lagrangian $\tilde{L}$ of the form

$$\tilde{L}=\tilde{L}(X,dX/dt)\ , \tag{1.17}$$

where it is assumed that $\tilde{L}$ is nondegenerate in the sense that the relation between momentum and velocity [see Eq. (1.22) below] is one to one. We rewrite this theory in the following manner. We introduce a new "parameter time" τ given in terms of t by an arbitrary monotonic function $\tau=\tau(t)$. We now view t as a dynamical variable, which may be thought of as the readings of a perfect clock. To emphasize this new viewpoint we shall use the new symbol T to denote this quantity. We rewrite the action S as

$$S=\int L(X,T,\dot{X},\dot{T})d\tau , \tag{1.18}$$

where the overdot denotes derivatives with respect to τ and where

$$L(X,T,\dot{X},\dot{T})\equiv\dot{T}\tilde{L}(X,\dot{X}/\dot{T}) . \tag{1.19}$$

For the Lagrangian L, the momenta canonically conjugate to X and T are

$$P_X=\partial L/\partial\dot{X}=\tilde{P}_X , \tag{1.20}$$

$$P_T=\partial L/\partial\dot{T}=\tilde{L}-(\dot{X}/\dot{T})\tilde{P}_X , \tag{1.21}$$

where $\tilde{P}_X$ denotes the momentum conjugate to X for the original Lagrangian $\tilde{L}$:

$$\tilde{P}_X=\partial\tilde{L}/\partial(dX/dt) . \tag{1.22}$$

The Euler-Lagrange equations for L are

$$\dot{P}_X=\frac{\partial L}{\partial X}=\dot{T}\frac{\partial\tilde{L}}{\partial X} , \tag{1.23}$$

$$\dot{P}_T=\frac{\partial L}{\partial T}=0 . \tag{1.24}$$

It is not difficult to verify that Eq. (1.24) is redundant; i.e., it follows from Eqs. (1.21) and (1.23). It also is easily seen, using $P_X=\tilde{P}_X$ [see Eq. (1.20)], that Eq. (1.23) is equivalent to the original Euler-Lagrange equations for $\tilde{L}$,

$$\frac{d\tilde{P}_X}{dt}=\frac{\partial\tilde{L}}{\partial X} , \tag{1.25}$$

in the sense that the relation $X=X(T)$ between the dynamical variables X and T implied by Eq. (1.23) is the same as the relation $X=X(t)$ between the dynamical variable X and the time parameter t implied by Eq. (1.25). [That this must be the case can be seen from the fact that, with the identification $t=T=\int\dot{T}\,d\tau$, both Euler-Lagrange equations (1.23) and (1.25) arise from the same variational problem for S.] In this sense, at the classical level, the theory described by the original Lagrangian $\tilde{L}$ (with dynamical variable X and time parameter t) is equivalent to the theory described by L (with dynamical variables X, T, and time parameter τ).

The above parametrized theory can be given a Hamiltonian formulation which is closely analogous to the Hamiltonian formulation of general relativity. The canonical momenta P_X and P_T were already obtained above [see Eqs. (1.20) and (1.21)]. Normally, one would proceed by inverting these formulas to eliminate $\dot{X}$ and $\dot{T}$ in favor of P_X and P_T. However, this cannot be done here (for any choice of the original Lagrangian $\tilde{L}$), since the Jacobian matrix of second partial derivatives of L with respect to "velocities" $\dot{X}$ and $\dot{T}$ is degenerate, as a consequence of the fact that, apart from an overall factor of $\dot{T}$, L depends on $\dot{X}$ and $\dot{T}$ only in the combination $\dot{X}/\dot{T}$. Indeed, if the original Lagrangian $\tilde{L}$ is nondegenerate as we have assumed, then we can use Eq. (1.20) to eliminate $\dot{X}/\dot{T}$ in favor of $\tilde{P}_X=P_X$. By Eq. (1.21) we then have

$$P_T=\tilde{L}(X,\dot{X}/\dot{T})-(\dot{X}/\dot{T})P_X , \tag{1.26}$$

thus showing that P_T can be expressed as a function of X and P_X and thus cannot be an independent variable.

Nevertheless, we can give a Hamiltonian formulation wherein Eq. (1.26) is treated as a constraint. To do so we define the quantity $\mathcal{H}$ by the usual formula for a Hamiltonian

$$\mathcal{H}=\dot{T}P_T+\dot{X}P_X-L=\dot{T}[P_T+\tilde{H}(X,P_X)] , \tag{1.27}$$

where $\tilde{H}(X,P_X)$ is the Hamiltonian associated with the original Lagrangian $\tilde{L}$:

$$\tilde{H}(X,P_X)=\frac{dX}{dt}\tilde{P}_X-\tilde{L}=(\dot{X}/\dot{T})P_X-\tilde{L} . \tag{1.28}$$

Note that the constraint (1.26) then is equivalent to

$$\mathcal{H}=0 . \tag{1.29}$$

The inability to eliminate $(\dot{T},\dot{X})$ in favor of (P_T,P_X) comes into play in Eq. (1.27) in that $\dot{T}$ cannot be related to P_T and P_X. However, we can proceed by viewing $\dot{T}$ as a "Lagrangian multiplier" which enforces the constraint (1.29). To emphasize this viewpoint we write $N(\tau)=\dot{T}$ and view N as an arbitrary, unspecified function. Thus, we write

$$\mathcal{H}(N;T,X,P_T,P_X)=N[P_T+\tilde{H}(X,P_X)] . \tag{1.30}$$

Hamilton's equation of motion then yield

$$\dot{T}=\frac{\partial\mathcal{H}}{\partial P_T}=N , \tag{1.31}$$

$$\dot{P}_T=-\frac{\partial\mathcal{H}}{\partial T}=0 , \tag{1.32}$$

$$\dot{X}=\frac{\partial\mathcal{H}}{\partial P_X}=N\frac{\partial\tilde{H}}{\partial P_X} , \tag{1.33}$$

$$\dot{P}_X=-\frac{\partial\mathcal{H}}{\partial X}=-N\frac{\partial\tilde{H}}{\partial X} \tag{1.34}$$

which, when supplemented by the constraint (1.29) obtained by variation of $\mathcal{H}$ with respect to N, are easily verified to be equivalent to the previously derived equations of motion. Thus, we have succeeded in obtaining a Hamiltonian formulation of the parametrized particle theory. The very close similarity of this formulation to the Hamiltonian formulation of general relativity is not accidental, since the presence of the constraint $\mathcal{H}=0$ in both theories is directly related to the fact that in both theories "time evolution" is equivalent to a "gauge transformation" (i.e., a spacetime diffeomorphism or a reparametrization of τ).

We turn now to the quantum theory of the parametrized particle. By the standard canonical quantization rules for the "coordinate representation" already mentioned above, states of the system are described by wave functions $\Psi(\tau;T,X)$. The operators T and X are represented by multiplication whereas P_T and P_X are represented as $-i\partial/\partial T$ and $-i\partial/\partial X$. Time evolution is governed by the Schrödinger equation

$$i\frac{\partial\Psi}{\partial\tau}=\mathcal{H}\Psi \ . \tag{1.35}$$

However, the constraint (1.29) is imposed by the condition

$$\mathcal{H}\Psi=0 \tag{1.36}$$

thus showing that Ψ is independent of τ. Thus, this system possesses many of the features of canonical quantum gravity, but, of course, it has the advantage that the correct quantum theory is known, namely, that obtained from the original Lagrangian $\tilde{L}$ and Hamiltonian $\tilde{H}$.

Note that the constraint (1.36), which is the analog of the Wheeler-DeWitt equation (1.6), is just the ordinary Schrödinger equation for $\Psi(T,X)$:

$$-i\frac{\partial\Psi}{\partial T}+\tilde{H}(X,P_X)\Psi=0 \ . \tag{1.37}$$

Thus, the above rules for the canonical quantization applied to our parametrized theory yield the correct equation for Ψ. The above "naive interpretation" applied to this case would demand that $\Psi(T,X)$ be square integrable with respect to T and X, and would then attempt to interpret Ψ as giving the (time-independent) amplitude for measuring T and X. This clearly does not make sense; indeed, there presumably do not even exist any solutions of (1.37) which are square integrable with respect to T and X. The "WKB interpretation" yields results consistent with the standard quantum theory, but, of course, has a severely limited range of applicability. Finally, the "conditional probability interpretation" with T chosen as the variable which "sets the conditions" would assert that, given that the time is T, $\Psi(T,X)$ yields the amplitude for finding the particle at X. If the measure is taken to be $d\mu_X=dX$, this corresponds precisely to the standard quantum theory of a particle. However, the "conditional probability interpretation" with X chosen as the variable which "sets the conditions" would assert that at given particle position X, $\Psi(T,X)$ should give the amplitude that the time is T. This does not make sense, and thus illustrates the point already mentioned above that only certain variables are suitable for "setting the conditions."

The analogy with the quantum theory of parametrized particle mechanics may be viewed as supporting the "conditional probability approach" to the interpretation of the wave function of the Universe in quantum gravity, provided that a suitable "time variable" is identified from among the collection of dynamical variables. In parametrized particle mechanics, the dynamical variable T can be distinguished from the other dynamical variables by the fact that its canonical momentum P_T appears only linearly in the Hamiltonian $\mathcal{H}$. However, in canonical quantum gravity, although there has been considerable research effort on this issue, no suitable "time variable" has yet been identified. In particular, the Wheeler-DeWitt constraint (1.6) is quadratic in all the momenta. We refer the reader to the reviews of Kuchař[15] and references cited therein for extensive further discussion of this issue.

Thus, the conditional probability interpretation presently remains in the unsatisfactory state described above. Clearly, one approach to the interpretation of the wave function of the Universe would be to attempt to solve the problem of identifying a suitable time variable. However, we are skeptical that such an approach will succeed. One reason for our skepticism is that no solution has yet emerged after over 20 years of effort. A more fundamental reason is that the selection of a preferred time variable (or class of time variables) would appear to be in conflict with the generally covariant nature of general relativity. The ability to "deparametrize" the above particle theory by choosing T as a preferred time variable appears to be directly related to the fact that the underlying classical theory possesses a preferred time parametrization; however, there does not appear to be any analogous preferred time slicing of generic spacetimes in general relativity. Finally, perhaps the strongest reason for our skepticism is that we do not believe that any realistic dynamical variable can satisfy the "Heraclitian property" required for a time variable. The main support for this view arises from the fact that in the context of ordinary Schrödinger quantum mechanics, no dynamical variable in a system with Hamiltonian bounded from below can act as a perfect clock in the sense that there is always a nonvanishing amplitude for any realistic dynamical variable to "run backwards." (Note that, by construction, the dynamical variable T in our parametrized particle model does act as a perfect clock; however, its Hamiltonian $H_T=P_T$ is unbounded from below.) In particular, any realistic dynamical variable may take on the same value at two distinct Schrödinger parameter times. Hence, it would appear that in Schrödinger quantum mechanics, all other dynamical variables can be multivalued at a given value of any realistic dynamical variable which is selected to be a time variable. We proceed, now, to state and prove this "no perfect clock" theorem.[16]

Consider any arbitrary system in ordinary Schrödinger quantum mechanics, restricted only by the requirement that its Hamiltonian H be bounded from below. We seek an observable (i.e., operator) T which can serve as a "monotonically perfect clock" in the sense that, for some choice of initial state, its observed values increase monotonically with Schrödinger time t. Since T may have continuous spectrum we formulate a minimal condition on such an operator T as follows. Break up the spectrum of T into nonoverlapping intervals of finite size. We require of T that there exist an infinite sequence of states $|T_0\rangle$, $|T_1\rangle$, $|T_2\rangle, \ldots$ having the following properties. (i) Each $|T_n\rangle$ is an eigenstate of the projection operator onto the spectral interval centered around the value T_n, with $T_0<T_1<T_2<\ldots$, (ii) for each n there exists an $m>n$ and a $t>0$ such that the amplitude to go from $|T_n\rangle$ to $|T_m\rangle$ in time t is nonvanishing (i.e., the "clock" has a nonzero probability of running forward in time); (iii) for each m and for all $t>0$, the amplitude to go in time t from $|T_m\rangle$ to any $|T_n\rangle$ with $n<m$ vanishes (i.e., the

"clock" cannot run backward in time). We then have the following theorem.

Theorem If the Hamiltonian H is bounded from below, then there does not exist an operator T satisfying properties (i)–(iii) above.

Proof. Let n,m be as in condition (ii). Consider the quantity

$$f(t)=\langle T_n|\exp(-iHt)|T_m\rangle \tag{1.38}$$

for $t\in\mathbb{C}$. Since H is bounded from below, f is holomorphic in the lower half t-plane and hence f cannot vanish on any open interval of real t unless it vanishes identically for all t with imaginary part ≤ 0 (see, e.g., Ref. 17). Since by (iii) we have $f(t)=0$ for $t>0$, it follows that $f(t)=0$ for all real t. However, for all $t>0$ we have

$$\begin{aligned}\langle T_m|\exp(-iHt)|T_n\rangle &= \langle T_n|\exp(+iHt)|T_m\rangle^* \\ &= f^*(-t) \\ &= 0\end{aligned} \tag{1.39}$$

which contradicts property (ii). □

The above theorem can be interpreted as saying that any realistic "clock" in Schrödinger quantum mechanics which can run forward in time must have a nonvanishing probability to run backward in time. Of course, the theorem allows this probability to be very small. Nevertheless, it shows that if we attempt to replace the Schrödinger time parameter t by a dynamical variable T, we cannot expect to obtain more than a crude, approximate interpretation of the theory. Indeed, since clocks typically degrade (so that they are "good" only for a finite period of Schrödinger time) one should expect severe problems to arise if one attempts to replace the condition "at Schrödinger time t" with the condition "when the clock is measured to read T"; the latter condition may include occurrences in the distant future (in Schrödinger time t), when the clock is working very poorly and has taken a large jump backward, or has fallen apart completely, leaving the hand pointed at some random digit.

The Hamiltonian, Eq. (1.4), of general relativity is unbounded from below (or above), so the above theorem is not directly applicable to it. Nevertheless, we see no reason to expect the presence of a dynamical variable which satisfies the "Heraclitian property."

Our belief that a "Heraclitian time variable" is needed in quantum theory but that dynamical variables are unsuitable for this role leads us to seek a formulation of canonical quantum gravity where a suitable parameter time is explicitly present in the theory. We shall proceed to do so in the following manner. In the next section we reformulate Schrödinger quantum mechanics in a manner entirely equivalent to the usual formulation, but using an arbitrary time parameter τ rather than the usual Schrödinger time t, in order to help distinguish between the roles of t as a "Heraclitian time variable" and as a variable containing dynamical information. (This procedure of introducing an arbitrary parameter time *after* quantization should be distinguished clearly from the above procedure of parametrization *prior* to quantization where the original time t is treated as a dynamical variable T.) In the reformulation, τ itself is unmeasurable, but the usual predictions can be expressed in terms of correlations between dynamical variables at (ordered) sequences of τ times. Furthermore, if one of the dynamical variables T is a "good clock" (i.e, at each τ its possible values are sharply peaked about a value which varies monotonically with τ), then one can pass (approximately) from the Schrödinger wave function, ψ to an "effective wave function" Ψ which depends only on the dynamical variables and (approximately) satisfies Eq. (1.37). This effective wave function conveniently encodes the correlations between clock readings T and the values of the other dynamical variables; however, one must return to the original wave function ψ in order to obtain a sensible, precise interpretation of the theory.

The above results motivate our proposal for canonical quantum gravity given in Sec. III. The "effective wave function" Ψ of Sec. II, is closely analogous to the wave function of the Universe, and it satisfies Eq. (1.37), which is closely analogous to the Wheeler-DeWitt equation (1.6). Thus, we seek a theory possessing a nondynamical "Heraclitian time variable" τ, which, in the presence of a "good clock" dynamical variable T, gives rise to an "effective wave function" Ψ satisfying the Wheeler-DeWitt equation. In Sec. III we present a proposal for such a theory in our minisuperspace model based directly upon analogy with the analysis of Sec. II. This theory has the desired qualitative properties, but it turns out that the "effective wave function" Ψ satisfies an equation which differs from the Wheeler-DeWitt equation (unless T is a "perfect clock," with Hamiltonian $H_T=P_T$). We study a slightly modified version of this proposal which is seen to correspond classically to Einstein's equation with an arbitrary, unspecified cosmological constant. This modified proposal yields a mathematically consistent, interpretable quantum theory in minisuperspace models, but for general spacetimes it may suffer from a shortage of measurable observables similar to (though possibly not as severe as) the situation for the "naive interpretation" of canonical quantum general relativity discussed above.

We conclude this section with a discussion of the issue of obtaining an interpretation of the wave function of the Universe by explicitly incorporating observers into the theory. It often is suggested that the interpretive difficulties of canonical quantum gravity arise from the failure to include dynamical variables representing an observer in the wave function of the Universe.[18] What these discussions appear to have in mind is the following. Let Y denote the relevant dynamical variables associated with an observer (which we may view, perhaps, as representing his "state of consciousness"). Let X denote all of the remaining dynamical variables (e.g., in our minisuperspace model a and ϕ). Then the wave function of the Universe is a function of Y and X, $\Psi=\Psi(Y,X)$. We view Ψ as describing the correlations between Y and X. Specifically, given that the observer is in state Y_0, $\Psi(Y_0,X)$ gives the amplitude for the various possible values of the remaining dynamical variables.

It is easily seen that this interpretation is just the above "conditional probability interpretation," with the "preferred time variable" chosen to be Y. However, if one takes seriously the treatment of an observer as an ordinary dyanmical quantum system (as is done in this viewpoint), we see no reason why Y should have this special property. Indeed, presumably at best, Y could be expected to be a "Heraclitian time variable" only while the observer is alive. How is this limitation to be implemented in this viewpoint? It may well be that the proper incorporation of observers into the theory will have many important (and, perhaps, radical) implications for quantum theory in general and for quantum gravity in particular. However, it seems to us highly unlikely that such an incorporation can be attained simply by means of treating observers as ordinary dynamical systems. In any case, unless some argument, presumably based on a detailed dynamical model of observers, can be given showing that the "observer variables" Y have the desired property of a "preferred time variable," the above proposed interpretation based upon observers amounts to nothing more than asserting that some (unspecified) "time variable" can be selected so that the conditional probability interpretation is valid. Further indication that explicit incorporation of observers should not be relevant to the resolution of the interpretive issues under consideration here comes from the fact that the same interpretive difficulties occur for the above quantum theory of a parametrized particle, but the accepted version of the theory which resolves these difficulties (namely, ordinary Schrödinger quantum mechanics) does not rely upon the incorporation of observers into the theory.

Note that when observers are incorporated into the theory, although the Wheeler-DeWitt equation (1.6) is imposed upon the wave function of the Universe $\Psi(Y,X)$, the quantity $\Psi(Y_0,X)$ [or, more precisely, the projection of $\Psi(Y,X)$ onto the perceived "eigenstate of observer variables"] will, in general, fail to satisfy the Wheeler-DeWitt equation.[19] Thus, it is far from clear that an observer would perceive that the Wheeler-DeWitt equation is satisfied and that Einstein's equation holds in the classical limit. Actually, as we shall now discuss, the situation is considerably worse, since an observer presumably has access only to Y and cannot directly determine anything at all about the "universe variables" X.

The difficulty just alluded to is, of course, nothing more than the age-old problem of skepticism, but it arises in a particularly virulent form when one examines the logical consequences of taking seriously the incorporation of observers into any (classical or quantum) theory so that they are treated as ordinary dynamical systems. The key point is that, even if the dynamics of observers is completely specified (which, of course, realistically is far from the case), the theory loses all predictive power concerning the Universe unless some new principle of physics governing the correlations of observers with the Universe (presumably formulated in terms of initial conditions) can be invoked. To illustrate this point explicitly, consider the simple example of the theory of classical mechanics of point particles. We describe the state of our dynamical system by dynamical variables (X_i,P_{X_i}) and we model our observer as similarly described by dynamical variables (Y_i,P_{Y_i}). The various possible perceptions (including memories) of the observer would then be assumed to correspond to particular regions of observer phase space. The difficulty is that unless severe restrictions are imposed upon initial conditions, then all kinematically possible observer states are dynamically possible. In particular, the observer may "perceive" or "remember" things about the (X_i,P_{X_i}) system which did not occur or even were dynamically impossible. Thus, what may have started out as a theory of the dynamics of the (X_i,P_{X_i}) system would become transformed, by the incorporation of observers, into a theory dealing primarily with the kinematical restrictions on observer phase space and with the allowed initial states for observers. Indeed, the requirement that these initial states be such that the observers' perceptions and memory correlate properly with what actually occurred in the (X_i,P_{X_i}) system is not experimentally testable, nor is even the existence of an (X_i,P_{X_i}) system, and would have to be viewed as an entirely *ad hoc* hypothesis. Exactly the same difficulties occur, of course, in quantum theory. The refusal to incorporate observers into the theory as ordinary dynamical systems does not solve the problem of skepticism, but it does, at least, remove it from one's doorstep.

Thus, in view of all the above discussion, it does not appear to us to be fruitful to attempt to resolve interpretive issues of quantum gravity by invoking the explicit incorporation into the theory of observers, treated as ordinary dynamical systems. (We also do not have any serious proposals for treating observers as extraordinary systems.) Thus, for the remainder of this paper we will continue to treat observers in the same "phenomenological" way that they have been treated in all prior physical theories.

II. SCHRÖDINGER QUANTUM MECHANICS IN AN ARBITRARY TIME PARAMETRIZATION

The parameter t which appears in Schrödinger quantum mechanics has a rather unusual status. It is not an "observable" in any normal sense; in particular, there is no operator on the Hilbert space of states which corresponds to measuring "time." Rather, t primarily plays the role of a nondynamical, "Heraclitian time variable" (see Sec. I) which "sets the conditions" for measurements of the dynamical variables. It is not true that one "measures" t by examining a clock, since by examining a clock one is simply measuring some ordinary dynamical variable (e.g., the position of the hands of the clock). Indeed, we showed in the previous section that no realistic dynamical variable can even monotonically correlate with t with certainty. Nevertheless, one can make inferences about t by performing measurements of dynamical variables. In this section we shall attempt to clarify the role of t in ordinary Schrödinger quantum mechanics by reformulating the theory in terms of an arbitrary time parameter τ. In this way, the roles of t as a "Heraclitian time variable" and a quantity which is, in some sense, "observable," can be clearly distinguished.

Our main purpose in giving this reformulation is that it will aid us in explaining how information concerning correlations between measurements of a "good clock" dynamical variable T and another dynamical variable X can be expressed in terms of an "effective wave function" $\Psi(T,X)$ which depends only upon observable dynamical variables and which (approximately) satisfies the Schrödinger equation. However, unless the clock is "perfect," the effective wave function $\Psi(T,X)$ will possess an interpretation only in a crude, approximate sense; one must return to the exact Schrödinger wave function $\psi(\tau;T,X)$ in order to formulate a precise interpretation of the theory. Our proposal for formulating an interpretable quantum theory of gravity, to be given in the next section, consists, in essence, of simply paralleling for the wave function of the Universe, $\Psi(a,\phi)$, the steps which lead backward from $\Psi(T,X)$ to $\psi(\tau;T,X)$.

Our reformulation of Schrödinger quantum theory consists simply of replacing the parameter t in the Schrödinger equation by a parameter τ related to t by an arbitrary monotonic function $\tau=\tau(t)$. (Note that, as already pointed out in the previous section, this is not equivalent to formally quantizing a classical parametrized system with "time" treated as a dynamical variable.) Thus, we take the Hilbert space to be as in standard quantum theory; i.e., if the collection of all the dynamical (configuration) variables are denoted as Z, states are again represented by wave functions $\psi(Z)$, which vary with time τ. However, the Schrödinger equation now is replaced by the following condition. There exists a function $N(\tau)$ not known or specified in advance, such that

$$i\frac{\partial\psi}{\partial\tau}(\tau;Z)=N(\tau)H\psi(\tau;Z)\ , \qquad (2.1)$$

where H is the Hamiltonian operator of the usual formulation. (We assume that H has no explicit time dependence, so that the operator H is independent of τ.) The interpretation of ψ is the usual one. At a fixed value of parameter time τ, $\psi(\tau;Z)$ gives the amplitude for the values of the dynamical variables to be Z.

The above formulation easily can be seen to be equivalent to the usual formulation via the simple substitution $t=\int N(\tau)d\tau$. However, the above formulation suggest a viewpoint which we feel corresponds much more closely to the observable structure actually present in ordinary quantum theory. In this viewpoint, an observer has access to time orderings of his observations given by the label τ whose numerical values are of no significance except for the ordering they provide. (In keeping with the "phenomenological" approach to observers taken at the end of the previous section, we do not attempt to explain the mechanism by which the observer obtains access to this ordering defined by τ.) However, an observer does not have any "innate" knowledge of the "metrical" properties of time, as given by t; as already mentioned above t can be determined only by inference from measurements of dynamical variables. [As will be discussed further below, this is easily done if a "good clock" dynamical variable is present; otherwise much more sophisticated methods would have to be employed. The probability distribution for Z can be expressed in terms of the unknown parameter t, and techniques of statistical inference (such as the maximum-likelihood method) could be used to infer t.] Thus, the arbitrary (except for monotonicity) label time τ provides the essential "background structure" of quantum mechanics. To avoid confusion, we emphasize here that we have in mind the measurements of a single observer (who, of course, may send out probes, or other observers, to spatially distant regions and later measure the readings of these probes). The interpretation is formulated for this single observer. If a family of observers is present, we do *not* assume that different members of this family have access to the same τ; i.e., we do not assume that a family of observers has any innate knowledge of simultaneity.

Since $N(\tau)$ is unknown, the Schrödinger equation (2.1) does not yield unambiguous predictions for the wave function $\psi(\tau;Z)$ at time τ. (This, of course, corresponds to the fact that τ is an arbitrary "label time.") Predictions primarily consist of statements concerning the possible values and correlations of the dynamical variables Z. Thus, the situation is much like that occurring in general relativity, formulated in terms of arbitrary coordinates x^μ. Einstein's equation does not unambiguously determine the value of the metric components $g_{\mu\nu}$ evaluated at the point labeled by x^μ, and typical statements about the theory expressed in this manner usually are meaningless because they refer to the arbitrary, unmeasurable labels x^μ.

The meaningful statements in general relativity are typically formulated in the following manner: "There exists a point (i.e., event) in spacetime such that the measurements of . . . made at that point by the specific procedure . . . would have the outcome . . . ." Similarly, in the above reformulation of quantum mechanics, the meaningful statements typically would take the form, "There exists a time τ such that the probability of measuring the dynamical variables to be Z is . . . ." Statements concerning whether this time τ came "before" or "after" a time at which certain other measurements were made also would be meaningful.

In general, these meaningful statements may be rather cumbersome to formulate. However, as we shall now explain, if one of the dynamical variables is a "good clock variable," then meaningful statements can be formulated in a very simple manner (although these statements will have only approximate validity). We say that a dynamical variable, denoted T, is a *good clock variable* over the interval $I=[a,b]$ of τ time if the state vector ψ and Hamiltonian H satisfy the following two conditions: (i) For all $\tau\in I$, T (nearly) decouples from all the other dynamical variables, which we denote as X, in the sense that ψ (nearly) takes the product form

$$\psi(\tau;T,X)\approx\chi(\tau;T)\phi(\tau;X) \qquad (2.2)$$

and we have

$$H\psi\approx(H_T+H_X)\psi\ , \qquad (2.3)$$

where H_T is independent of X, and H_X is independent of T; (ii) at each $\tau\in I$, $\psi(\tau;T,X)$ is sharply peaked in T about the value $f(\tau)$, where f is a monotonic function of τ. Although we proved in the previous section that if H is bounded from below, then no dynamical variable T can serve as a "perfect monotonic clock," there is no obstacle to the existence of a "good clock variable" in the above sense.

In condition (i) it should be noted that it is not required that the Hamiltonian operator be generally expressible as $H = H_T + H_X$; i.e., there may be states (other than ψ) for which Eq. (2.3) fails. In condition (ii), since H is independent of τ, it follows from the Schrödinger equation (2.1) that f can depend on τ only in the combination $\int N(\tau)d\tau$. By redefining T, if necessary, we will assume, without loss of generality, that, in fact, $f(\tau) = \int N(\tau)d\tau$. Thus, Eq. (2.2) and assumption (ii) imply that ψ takes the form

$$\psi(\tau;T,X) \approx \lambda\left[\tau;T - \int N(\tau)d\tau\right]\phi(\tau;X)\ , \tag{2.4}$$

where for each τ, the function $\lambda(\tau;\cdot)$ is sharply peaked around zero. Furthermore, Eq. (2.3) further implies that λ and ϕ satisfy

$$i\frac{\partial\lambda}{\partial\tau} \approx N(\tau)(H_T + C)\lambda\ , \tag{2.5}$$

$$i\frac{\partial\phi}{\partial\tau} \approx N(\tau)(H_X - C)\phi\ , \tag{2.6}$$

where C is constant. By redefining λ and ϕ by multiplying and dividing, respectively, by the phase factor $\exp[iC\int N(\tau)d\tau]$, we may assume that $C = 0$. We also may normalize λ and ϕ so that they each have unit norm in the T and X Hilbert spaces, respectively.

Now, since $\lambda(\tau;\cdot)$ is peaked sharply about zero, we will make little error in evaluating $\psi(\tau;T,X)$ if, in Eq. (2.2), we replace $\phi(\tau;X)$ by the "effective wave function" $\Psi(T;X)$ defined by

$$\Psi(T,X) \equiv \phi(\tau(T);X)\ , \tag{2.7}$$

where $\tau(T)$ is determined by the peak in λ, i.e., by

$$T = \int N(\tau)d\tau\ . \tag{2.8}$$

Thus, if T is a "good clock variable" we have

$$\psi(\tau;T,X) \approx \lambda\left[\tau;T - \int N(\tau)d\tau\right]\Psi(T,X)\ . \tag{2.9}$$

The factor λ contains the information concerning the probability for obtaining "false readings" from measuring the "clock variable" T. To the extent that the probability for the measured value of T to deviate significantly from $\int N(\tau)d\tau$ is negligibly small (i.e., to the extent that T is a "good clock" variable), all the remaining information concerning the system is usefully encoded in $\Psi(T,X)$. Indeed, to the extent that T is a "good clock variable," we have the following simple "interpretation" of Ψ. "If the clock variable is measured to have the value T, then the amplitude for measuring the remaining dynamical variables to have the value X is $\Psi(T,X)$." This interpretation, of course, is only an approximate one, and if T is a "poor clock" variable, then one must return to the full wave function $\psi(\tau;T,X)$ in order to formulate the predictions of the theory, which would have to be stated in the manner discussed above.

It follows immediately from Eqs. (2.6), (2.7), and (2.8), that if T is a good clock variable, then $\Psi(T,X)$ satisfies

$$-i\frac{\partial\Psi}{\partial T} + H_X\Psi \approx 0\ . \tag{2.10}$$

This equation is, of course, formally the same as the usual Schrödinger equation, except that the Schrödinger parameter time t has been replaced by the dynamical variable T. More importantly for our purposes, Eq. (2.10) is formally identical to the constraint equation (1.37) which arose in the quantization of a parametrized classical theory in Sec. I. Thus, we may summarize the results of this section as follows. We began with Schrödinger quantum mechanics—a well-defined and interpretable quantum theory, possessing no constraints. We gave an entirely equivalent reformulation of this theory in terms of an arbitrary label time τ. We then showed that if one of the dynamical variables in the theory is a "good clock," one can pass to an "effective wave function" $\Psi(T,X)$ which depends only upon the dynamical variables and satisfies Eq. (2.10), which is formally equivalent to the constraint equation (1.37). Furthermore, one has an interpretation of Ψ only in the approximate sense described above; the "exact" interpretation of the theory can only be formulated in terms of the original wave function $\psi(\tau;T,X)$.

The relevance of the above discussion for the interpretation of the wave function of the Universe should now be clear. The wave function of the Universe $\Psi(a,\phi)$ is closely analogous to the "effective wave function" $\Psi(T,X)$, with the Wheeler-DeWitt equation playing the role of Eq. (2.10). Thus, we propose to view $\Psi(a,\phi)$ as an "effective wave function"—an approximate concept, valid only when a "good clock" dynamical variable is present. We propose that it arises from an exact theory which possesses a "label time" τ and is well defined and interpretable in all circumstances. We shall attempt to implement this proposal in the next section by reversing the steps above which led from $\psi(\tau;T,X)$ to $\Psi(T,X)$.

III. A PROPOSAL FOR CANONICAL QUANTUM GRAVITY

In this section we shall explore the possibility that the "exact" quantum theory of gravity is such that a "Heraclitian time parameter" τ is explicitly present. In such a theory with time parameter τ, if a "good clock variable" is present, we can pass to an "effective wave function" Ψ, which depends only on dynamical variables, in the manner described in the previous section. Our goal is to obtain a theory whereby this effective wave function satisfies the Wheeler-DeWitt equation, since presumably this would be necessary (and, presumably, also sufficient) for the theory to reduce to general relativity in the classical limit.

We shall proceed by making a straightforward proposal for the desired theory in the simple context of a minisuperspace model (see Sec. I). We shall then explain why this proposal fails to yield the Wheeler-DeWitt equation for the effective wave function unless the "good clock" actually is a "perfect clock" in the sense that its Hamiltonian is $H_T = -i\partial/\partial T$. Nevertheless, we then will proceed to describe a general (i.e., not restricted to minisuperspace models) formulation of a class of proposals having the character of our minisuperspace proposal.

We will focus attention on one of them and show that it corresponds classically to general relativity with an arbitrary, unspecified cosmological constant. Thus, although this proposal fails to do what was originally intended (since it does not correspond classically to ordinary general relativity) and although (as discussed further below) it may suffer from serious difficulties in the context of general spacetimes, we hope that both the nature of the attempt and the resulting theory will be of some interest.

Our proposal in the minisuperspace case is based directly upon analogy with the discussion of the previous section. It consists, in essence, of simply omitting the constraint (1.12) from the theory. Thus, the wave function of the Universe is taken to be a function of a "Heraclitian time parameter" τ and the dynamical variables, which, for simplicity and definiteness, we take to be a and ϕ as in the model of Sec. I. Thus, $\psi=\psi(\tau;a,\phi)$ is directly analogous to an ordinary Schrödinger wave function. It satisfies the Schrödinger equation

$$i\frac{\partial\psi}{\partial\tau}=\mathcal{H}\psi \tag{3.1}$$

which, for our model takes the explicit form

$$i\frac{\partial\psi}{\partial\tau}=N\left[+\frac{1}{12a^2}\frac{\partial}{\partial a}\left[a\frac{\partial\psi}{\partial a}\right]-\frac{1}{2a^3}\frac{\partial^2\psi}{\partial\phi^2}+a^3V(\phi)\psi\right] \tag{3.2}$$

[see Eq. (1.10) above] but now the constraint equation $\mathcal{H}\psi=0$ is *not* imposed. As a consequence, ψ has a nontrivial time development. Indeed, if we normalize ψ in a and ϕ and use the natural (metric) volume element $a^2da\,d\phi$ on minisuperspace, we may consistently give the straightforward interpretation of $\psi(\tau;a,\phi)$ as yielding the amplitude for an observer to measure the values a and ϕ at time τ.

Thus, our proposal for the minisuperspace case can be made to yield a mathematically sensible quantum theory which can be interpreted in a straightforward way. However, it remains to be seen whether it corresponds to general relativity (or, some other physically viable theory) in the classical limit. To investigate this issue we assume that one of our dynamical variables, denoted T, is a "good clock variable" in the sense of the previous section. Then, as discussed in the previous section, the wave function ψ will approximately factor as

$$\psi(\tau;T,X)=\chi(\tau;T)\Psi(T,X)\ , \tag{3.3}$$

where X denotes the remaining dynamical variable(s). (In our simple model there would be only one such additional dynamical variable, but our discussion does not depend upon the details of this model and would apply if many more dynamical degrees of freedom were present.) As in the previous section, the Schrödinger equation (3.2) together with our "good clock" assumptions will imply an equation for our effective wave function $\Psi(T,X)$. Indeed, by precisely the same derivation as led to Eq. (2.10) we obtain

$$-i\frac{\partial\Psi}{\partial T}+H_X\Psi=0\ . \tag{3.4}$$

However, except in the case $H_T=-i\partial/\partial T$ (corresponding to the Hamiltonian of a perfect clock), this is *not* the same as the Wheeler-DeWitt equation which, in the present circumstances, takes the form

$$H_T\Psi+H_X\Psi=0\ . \tag{3.5}$$

Thus, the correlations among the dynamical variables implied by Eq. (3.4) will not, in general, be the same as those implied by Eq. (3.5). Unless the dynamical variable T is a "perfect clock," we will not obtain Einstein's equation in the classical limit.

The reason why Eqs. (3.4) and (3.5) are different can be understood as follows. The only property of the "good clock variable" that enters the derivation of Eq. (3.4) is the behavior of the clock, i.e., the correlation between T and τ. On the other hand, the relevant feature of the clock with regard to the Wheeler-DeWitt equation (3.5) is its energy, i.e., its contribution as a source of gravitation in the Hamiltonian constraint equation. But it is easy to build two clocks which have the same behavior (i.e., essentially the same correlation between T and τ) but vastly different energies. For example, one could use the position of a free particle as a clock.[1] By using particles of different masses, one can obtain a simple example of two "clock variables" which would make the same contribution to Eq. (3.4) but very different contributions to Eq. (3.5). Only in the case of a clock whose energy is related to its behavior in the same way as for a "perfect clock" (i.e., a dynamical system whose true Hamiltonian is $H_T=-i\partial/\partial T$) are the contributions of the clock to Eqs. (3.4) and (3.5) the same. Note that one can have an "extremely good clock" such that Eqs. (3.4) and (3.5) differ drastically; the difference between Eqs. (3.4) and (3.5) has to do with the "active gravitational mass" of the clock, not how well it runs.

Given that our proposal does not correspond to ordinary classical general relativity, one nevertheless may inquire further as to its physical viability by determining what classical theory (if any) it does correspond to. In the minisuperspace context in which it was formulated, it is clear that it corresponds classically to a theory in which the Hamiltonian constraint is not imposed, but the usual Einstein evolution equations are retained. However, to investigate its viability, it is worthwhile to generalize our proposal and obtain the corresponding classical theory for arbitrary spacetimes. [Note that there are many possible ways to generalize the proposal since, in particular, there are many inequivalent (on general spacetimes) theories which reduce to the same theory in minisuperspace models.] Since our proposals for a quantum theory in the minisuperspace case involved dropping the applicable constraint equation, one way of generalizing our proposal to arbitrary spacetimes would be simply to drop all of the constraint equations of general relativity (at both the classical and quantum levels). More precisely we could take for the classical theory the same Hamiltonian, Eq. (1.4), as occurs in ordinary general relativity, but now view N and N^i as fixed (though unspecified) functions, which are not to be varied in obtaining Hamilton's equations of motion. In quantum theory, the Schrödinger equation (1.5) would remain, but the constraints (1.6) and (1.7) would be absent. Such a theory would possess the desired "Heraclitian time parameter" but would correspond classically to a theory which differs sufficiently from ordinary general relativity (in that it allows many more solutions to the field equations) that it would not be physically viable.

However, it is not necessary to take such a drastic step in generalizing the proposal to arbitrary spacetimes. In particular, there is no need to drop the momentum constraints (1.7). Thus, we could again take the Hamiltonian to be given by Eq. (1.4), now treat N^i as a Lagrange multiplier (to be varied in obtaining the equations of motion) as in ordinary general relativity, but treat N as a fixed (though unspecified) function. Again, however, the classical theory obtained in this manner will admit many more solutions than ordinary general relativity. Indeed, it is not difficult to show that solutions of the vacuum field equations in this theory can be put in correspondence with solutions of the ordinary Einstein equation with arbitrary irrotational dust matter.

However, it should be noted that the Hamiltonian constraint equation (1.2) of ordinary general relativity actually represents infinitely many constraints (namely, one holding at each point of space), but only one constraint need be dropped in order to obtain a nontrivial Heraclitian time variable. Thus, we could obtain a classical theory which corresponds much more closely to ordinary general relativity by allowing all but "one degree of freedom" of N also to be varied in the Hamiltonian. For example, we could fix only the spatial average of N and thereby obtain a classical theory which corresponds to ordinary general relativity with irrotational dust of uniform density on the orthogonal hypersurfaces. This would represent only a "one-parameter generalization" of general relativity, and thus would not differ as greatly from ordinary general relativity as the previous proposals. However, the classical theory has the undesirable feature of possessing locally preferred Lorentz frames.

Rather than investigate the nature of this theory at the classical and quantum level, we will consider a slight modification of it, whereby N is required to be a fixed function of the dynamical variables (so that it still cannot be varied independently) rather than a fixed (or partially fixed) function on spacetime. As we shall see, by a particular such choice of N, we can obtain a "one-parameter generalization" or ordinary classical general relativity which preserves "local Lorentz invariance," and, indeed, corresponds classically to general relativity with an arbitrary cosmological constant. The canonical quantization of this theory will possess a Heraclitian time parameter of the type we have been seeking. We shall proceed by giving Lorentzian and Hamiltonian formulations of this theory at the classical level and then investigating the nature of the quantum theory in our minisuperspace model. This theory has been proposed by a number of authors,[20] primarily with regard to the cosmological-constant problem. It has been described elsewhere by one of us[21] with regard to the problem of time in quantum gravity; for completeness we shall repeat some of the discussion of Ref. 21 here. An equivalent proposal has been made recently by Sorkin,[22] who was motivated by considerations which were different from (though not entirely unrelated to) ours.

Our theory involves, as usual, a spacetime metric g_{ab} on a four-dimensional background manifold M. However, we now additionally require that M be orientable and we fix a volume element $\eta_{abcd}=\eta_{[abcd]}$ on M as part of the background structure. We require the metric to satisfy $g\equiv\det(g_{ab})=-1$, where the determinant is calculated using η_{abcd}, i.e., $g=g_{ab}g_{cd}g_{ef}g_{gh}\eta^{aceg}\eta^{bdfh}$ where $\eta^{abcd}=\eta^{[abcd]}$ is determined by $\eta^{abcd}\eta_{abcd}=4!$. In other words we require the natural volume element associated with g_{ab} to agree with η_{abcd}. Locally, this requirement places no restriction on the spacetime geometry in the sense that, locally (i.e., in any sufficiently small open set), any metric is related to one satisfying $g=-1$ by a diffeomorphism. Globally this need not always be true since, in particular, the total volume of M computed using g_{ab} and η_{abcd} would have to be the same in order for g_{ab} to be diffeomorphic to a metric with $g=-1$. Such global restrictions actually could be avoided in the formalism by requiring g to be any fixed (i.e., not to be varied as g_{ab} is varied) but unspecified function on spacetime (rather than -1). In any case, such global restrictions will not affect the derivation of the equations of motion in the classical and quantum cases, which we ultimately shall take as defining theory, so we shall ignore these possible global restrictions.

$$S=\int R\ , \tag{3.6}$$

where the volume element η_{abcd} (which is required to agree with the natural volume element of g_{ab}) is understood in the integral. It is clear that any solution of the usual Einstein's equation (in the "gauge" $g=-1$) will be an extremum of S, and, thus, will satisfy the new classical field equations. However, more solutions of the new field equations are possible because S need only be an extremum with respect to variations δg_{ab} which preserve $g=-1$, i.e., which satisfy $g^{ab}\delta g_{ab}=0$. Since by the usual calculation we have

$$\delta S=\int G^{ab}\delta g_{ab} \tag{3.7}$$

it is clear that the extrema of S are precisely the solutions of the *trace-free* Einstein's equation

$$G^{ab}-\tfrac{1}{4}Gg^{ab}=0\ . \tag{3.8}$$

More generally, if matter fields are present, the field equations become

$$G^{ab}-\tfrac{1}{4}Gg^{ab}=T^{ab}-\tfrac{1}{4}Tg^{ab}\ , \tag{3.9}$$

where T_{ab} is the usual stress-energy tensor of the matter fields (obtained by functional differentiation of the matter action with respect to the metric). As in the ordinary Einstein case, by the equations of motion of the matter fields, T_{ab} will satisfy $\nabla^a T_{ab}=0$. Hence, taking the divergence of Eq. (3.9) and using the Bianchi identity we obtain

$$\nabla_a(G-T)=0\ , \tag{3.10}$$

i.e.,

$$G-T=\text{const}\equiv 4\Lambda \ . \tag{3.11}$$

Thus, the equations of motion (3.9) are equivalent to

$$G^{ab}-\Lambda g^{ab}=T^{ab} \ , \tag{3.12}$$

where Λ is a constant. This is precisely the form of Einstein's equation with a cosmological constant. However, here Λ is not prescribed in advance but rather comprises part of the initial data; solutions with all possible values of Λ are allowed. Thus, as claimed above, the classical theory obtained from the action (3.6) with the restriction $g=-1$ is the same as general relativity with an arbitrary, unspecified cosmological constant.

In order to proceed with the canonical quantization of this theory we must cast the classical theory in Hamiltonian form. This can be done by the same procedure as in ordinary general relativity. As in that case we identify the induced metric h_{ab} on a three-dimensional hypersurface Σ as the configuration variable. The canonically conjugate momentum variable

$$\pi^{ab}=\frac{\delta S}{\delta \dot{h}_{ab}} \tag{3.13}$$

again is given by

$$\pi^{ab}=h^{1/2}(K^{ab}-h^{ab}K) \ , \tag{3.14}$$

where K_{ab} is the extrinsic curvature of Σ and $h=\det(h_{ab})$. [Here the volume element ${}^{(3)}\eta_{abc}=\eta_{abcd}t^d$ on Σ is used to calculate h, where t^a is the "time flow vector field" (see, e.g., Ref. 2).] The Hamiltonian again takes the form

$$\mathcal{H}=\int_\Sigma (NH_0+N^aH_a) \ , \tag{3.15}$$

where N and N^a again have the intepretation, respectively, of lapse function and shift vector, and H_0 and H_a are given by the same expressions as in ordinary general relativity [see Eqs. (1.2) and (1.3) above]. (Here and in all integrals over Σ below, the volume element ${}^{(3)}\eta_{abc}$ is understood.) The major difference which occurs here is that, on account of the condition $g=-1$, the lapse function N no longer is an independent variable. Rather, it is given in terms of the dynamical variables h_{ab} by

$$N=h^{-1/2} \ . \tag{3.16}$$

As already indicated in our discussion above, the most important consequence of this change of status of N is that the Hamiltonian constraint equation, $H_0=0$, of ordinary general relativity, which is obtained by independent variation of N, no longer occurs as a constraint equation here and hence the Hamiltonian $\mathcal{H}$ does not vanish identically for solutions. The momentum constraints

$$H_a=0 \tag{3.17}$$

do remain present, since they are obtained by variation of the shift vector N^a, which remains an independent variable.

Although the Hamiltonian constraint is not present, it is almost recovered by the condition that dynamical evolution preserve the momentum constraints (3.17). Taking the Poisson brackets of the Hamiltonian $\mathcal{H}$ with the integrated momentum constraint $\int \xi^a H_a$ (where ξ^a is an arbitrary vector field on Σ) we obtain the additional constraint

$$0=\left\{\mathcal{H}, \int_\Sigma \xi^a H_a\right\}=\int_\Sigma \xi^a D_a(h^{-1/2}H_0) \ . \tag{3.18}$$

This implies that

$$H_0=h^{1/2}\Lambda \ , \tag{3.19}$$

where Λ is a spatial constant. The equations of motion then imply that Λ does not vary with time either. Thus, in essence, in the present theory the Hamiltonian constraint $H_0=0$, of ordinary general relativity, is replaced by Eq. (3.19). Note that the derivation of Eq. (3.19) corresponds in the Hamiltonian formulation to the derivation of Eq. (3.12) above.

The canonical quantization of the theory proceeds as in the case of ordinary general relativity as described in Sec. I. The state vector is taken to be a functional of the dynamical variable h_{ab} on Σ (as well as the function of time), $\psi=\psi(t;h_{ab})$. The momentum constraint corresponding to (3.17) is imposed as before by requiring ψ to satisfy

$$\int_\Sigma (D_a\xi_b)\frac{\delta\psi}{\delta h_{ab}}=0 \tag{3.20}$$

which, again, has the interpretation that ψ depends only on the three-geometry. However, in place of the Hamiltonian constraint (1.6), we now have the weaker condition arising from the constraint (3.18):

$$\int_\Sigma \xi^a D_a(h^{-1/2}H_0)\psi=0 \ . \tag{3.21}$$

Again, ψ will evolve via the Schrödinger equation

$$i\frac{\partial\psi}{\partial t}=\mathcal{H}\psi=\int_\Sigma h^{-1/2}H_0\psi \ . \tag{3.22}$$

However, since the right-hand side of Eq. (3.22) does not vanish, ψ will have a nontrivial time dependence, and t will play the role of a Heraclitian time parameter. Note that the theory does *not* have a "reparametrization invariance" $t\rightarrow\tau(t,x)$ on account of Eq. (3.16), which fixes the lapse function in terms of the dynamical variables. Thus t is analogous to the preferred time of Schrödinger theory. However, one could parametrize Eq. (3.22) directly (in the same manner as done in Sec. II) to write it in the form

$$i\frac{\partial\psi}{\partial\tau}=\int_\Sigma \mathcal{N}(\tau,x)h^{-1/2}H_0\psi \ , \tag{3.23}$$

where $\mathcal{N}$ is an arbitrary positive function, so that it appears more directly analogous to Eq. (2.1). As in ordinary Schrödinger quantum mechanics, neither t nor τ are directly measurable. The predictions of the theory must be formulated in terms of the correlations of the measurable dynamical variables. As discussed in Sec. II, such predictions are easiest to state when a "good clock" dynamical variable is present, although of course, the presence of such a variable is not necessary either for the formulation or interpretation of the theory.

Note that the solutions to Eq. (3.22) formally are superpositions of "eigenstates of the cosmological constant":

$$\mathcal{H}\psi_\Lambda = \int_\Sigma h^{-1/2} H_0 \Psi_\Lambda = \Lambda\psi_\Lambda \; ; \qquad (3.24)$$

i.e., formally, the general solution of the Schrödinger equation (3.22) is of the form

$$\psi(t) = \int d\Lambda\, \alpha(\Lambda) \exp\left[-i\left[\int \Lambda\right]\right]\psi_\Lambda \; , \qquad (3.25)$$

where ψ_Λ satisfies Eq. (3.24) and the integral in the phase factor is over the spacetime region between the initial slice and the time slice of interest, using the volume element η_{abcd}. (Thus, the phase factor has the interpretation of being simply the cosmological constant times the four-volume of this region of spacetime.) From Eq. (3.25) it can been seen that the nontrivial time dependence of ψ arises from the superposition of states corresponding to different values of the cosmological constant.

The new features of the wave function of the Universe in this theory are best elucidated by examining in concrete detail the nature of our minisuperspace model with dynamical variables a and ϕ. We again consider the cosmologically flat ("$k=0$") case but now write the metric in the form

$$ds^2 = -a^{-6}dt^2 + a^2(dx^2+dy^2+dz^2) \qquad (3.26)$$

so that $g=-1$. The Hamiltonian for this minisuperspace model in our new theory is

$$\mathcal{H} = -\frac{1}{12a^4}\pi_a^2 + \frac{1}{2a^6}\pi_\phi^2 + V(\phi) \; . \qquad (3.27)$$

Since the model is homogeneous, all of the constraints are automatically satisfied. Thus, any solution $\Psi(t;a,\phi)$ of the Schrödinger equation

$$i\frac{\partial\psi}{\partial t} = +\frac{1}{12a^5}\frac{\partial}{\partial a}\left[a\frac{\partial\psi}{\partial a}\right] - \frac{1}{2a^6}\frac{\partial^2\psi}{\partial\phi^2} + V(\phi)\psi \qquad (3.28)$$

is a possible wave function of the Universe. Here we have again chosen the "Laplacian" factor ordering for the momentum terms. Note that, even when written in the parametrized form (3.23), this equation is not the same as Eq. (3.2) above, as a consequence of the fact that here the lapse function is a function of dynamical variables and hence is an "operator" rather than a "c number." Nevertheless, Eqs. (3.2) and (3.28) are qualitatively similar, and the remarks made at the beginning of this section concerning the inequivalence of Eq. (3.2) to the Wheeler-DeWitt equation apply with equal validity to Eq. (3.28).

The Hamiltonian operator appearing on the right-hand side of Eq. (3.28) is formally self-adjoint; i.e., more precisely, it is symmetric on the domain of smooth functions of compact support in $L^2(a,\phi)$, where we now use the natural measure $a^5 da\, d\phi$ on minisuperspace [which arises from the metric on minisuperspace obtained from the kinetic terms in Eq. (3.27)]. Since this operator is "real," it always admits a self-adjoint extension (see Theorem X.3 of Ref. 23). Thus, the quantum dynamics can be made rigorously well defined, and both the mathematical structure and the interpretation of this model is exactly as in ordinary Schrödinger quantum mechanics.

In the case $V=0$, corresponding to a free, massless Klein-Gordon scalar field, Eq. (3.28) becomes simply

$$-i\frac{\partial\psi}{\partial t} = -\frac{3}{4\rho}\frac{\partial}{\partial\rho}\left[\rho\frac{\partial\psi}{\partial\rho}\right] + \frac{1}{2\rho^2}\frac{\partial^2\psi}{\partial\phi^2} \; , \qquad (3.29)$$

where $\rho\equiv a^3$, and the measure in the new variables (ρ,ϕ) is $\rho\, d\rho\, d\phi$. Remarkably, the Hamiltonian is identical (up to trivial factors) to that of the Laplacian operator in polar coordinates in a two-dimensional flat space, except for the minus sign occurring in the first term and the fact that here ϕ has the range $-\infty$ to ∞ rather than 0 to 2π. Thus, it can be seen immediately that the "eigenstates of cosmological constant" have the form

$$\psi_{\Lambda,p}(\rho,\phi) = \exp(-ip\phi) J_{i\eta}((4\Lambda/3)^{1/2}\rho) \; , \qquad (3.30)$$

where J_ν denotes the Bessel function of order ν and $\eta\equiv(\frac{2}{3})^{1/2}p$. These eigenstates correspond to solutions of the Wheeler-DeWitt equation for ordinary general relativity with a fixed value Λ of the cosmological constant; they fail to be square integrable on minisuperspace. However, there is, of course, no difficulty in obtaining superpositions of the form (3.25) which are square integrable. Thus, we see how a persistent problem for obtaining a probabilistic interpretation in the usual approach, namely, the failure of the wave function to be normalizable, is avoided here.

Typical dynamical variables, such as a, will fail to commute with the Hamiltonian. In contrast with the usual approach, this does not pose any difficulty with regard to the measurability of these quantities. However, it should be noted that such a measurement will influence the probability distribution for the value of the "energy," i.e., cosmological constant. Thus, in this theory, observers can, in effect, change the value of the cosmological constant by making measurements. However, one would expect that the uncertainty in the cosmological constant induced by a measurement would be of order $\Delta\Lambda=\Delta E/V$, where ΔE is the uncertainty induced in the energy and V is the spatial volume of the Universe (which would be finite for a three-torus model), in which case $\Delta\Lambda$ would be negligible for any physically realistic measurement.

In the context of minisuperspace models, the only difficulty with the above theory is its physical viability. It might appear that there is a serious problem in this regard, since it corresponds classically to general relativity with an arbitrary value of Λ. We know that the observed value of Λ in our Universe is extremely small ($<10^{-120}$), but there apparently is nothing in the theory to protect us from solutions with much larger values of Λ. However, it should be noted that a similar "cosmological-constant problem" occurs in the usual approach to quantum gravity. There Λ is a fixed parameter and one could argue that it is natural to set the "bare" value of Λ equal to zero. However, there is then apparently nothing to protect us from having the effective (renormalized) value of Λ take on much larger values. Thus, what we have is a modified version of the usual cosmological-constant problem. The "problem" now is shifted from the issue of why an effective coupling constant is so small to the issue of why the initial conditions of the Universe were such that a freely specificable quantity in the theory is so small.

Although in the context of minisuperspace models the above proposal provides a mathematically (and, perhaps, physically) viable quantum theory of gravitation, it is far from clear that it will continue to do so in the context of general spacetimes. Although an observable in the general theory no longer need commute with the Hamiltonian, it still must commute with the constraints (3.21) (which are trivial in the minisuperspace case on account of homogeneity). Thus, the theory may well possess the same type of difficulties that plague the "naive interpretation" of canonical quantum gravity discussed in Sec. I. It is clear that much more work will be required before one can determine whether any proposal of the type considered here (i.e., introducing an external parameter time) can provide a viable solution to the problem of time in quantum gravity. However, we hope that it may provide a step along the road to finding such a solution.

ACKNOWLEDGMENTS

We have benefited from numerous discussions with many colleagues, including Robert Geroch, James Hartle, Karel Kuchař, and Rafael Sorkin. This research was supported in part by the Canadian Institute for Advanced Research, by the Natural Sciences and Engineering Research Council of Canada, and by NSF Grant No. PHY 84-16691 to the University of Chicago. One of us (W.G.U.) wishes to thank the Institute for Theoretical Physics (supported by NSF Grant No. PHY 82-17853, supplemented by funds from NASA) for hospitality and support.

---

[1]For a discussion of the role of time in canonical quantum gravity which differs in some important respects from ours, see J. B. Hartle Phys. Rev. D **37**, 2818 (1988); and (to be published).

[2]R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).

[3]In fact, the Ashtekar representation may have significant technical advantages over the coordinate representation in general relativity. [See A. Ashtekar, *New Perspectives in Canonical Gravity* (Bibliopolis, Naples, in press).] In addition, in the coordinate representation for general relativity, the usual form of the canonical commutation relations imposed on h_{ab} and π^{ab} are not really appropriate on account of the "cone-space" structure of the configuration space [see C. J. Isham in *Relativity, Groups, and Topology II,* proceedings of the Les Houches Summer School, Les Houches, France, 1983, edited by R. Stora and B. S. DeWitt (Les Houches Summer School Proceedings Vol. 40) (Elsevier, Amsterdam, 1984).] However, neither of these points should affect the interpretive issues under consideration here.

[4]J. B. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983).

[5]A. Vilenkin, Phys. Rev. D **37**, 888 (1988).

[6]D. N. Page and W. K. Wootters, Phys. Rev. D **27**, 2885 (1983).

[7]A. Vilenkin, Phys. Rev. D **39**, 1116 (1989).

[8]T. Banks, W. Fischler, and L. Susskind, Nucl. Phys. **B262**, 159 (1985).

[9]J. J. Halliwell and S. W. Hawking, Phys. Rev. D **31**, 1777 (1985).

[10]C. W. Misner, in *Magic Without Magic,* edited by J. Klauder (Freeman, San Francisco, 1972).

[11]B. S. Kay and R. M. Wald (unpublished).

[12]A. Ashtekar and A. Magnon, Proc. R. Soc. London **A346**, 375 (1975); B. S. Kay, Commun. Math. Phys. **62**, 55 (1978).

[13]K. Kuchař, J. Math. Phys. **22**, 2640 (1981).

[14]W. G. Unruh, in *Proceedings of the Fourth Seminar on Quantum Gravity,* edited by M. A. Markov, V. A. Berezin, and V. P. Frolov (World Scientific, Singapore, 1988).

[15]K. Kuchař, in *Quantum Gravity 2,* edited by C. J. Isham, R. Penrose, and D. W. Sciama (Clarendon, Oxford, 1981); K. Kuchař, in the Proceedings of the Osgood Hill Conference, May, 1988 (unpublished).

[16]This theorem may be viewed as a strengthening of a result which appears in W. Pauli, *Die Allgemeinen Prinzipien der Wellenmechanik,* edited by S. Flugge (Handbuch d. Physik, V) (Springer, Berlin, 1958), p. 60. Note that our theorem conflicts with a statement in A. Peres, Am. J. Phys. **48**, 552 (1980).

[17]R. F. Streater and A. S. Wightman *PCT, Spin and Statistics, and All That* (Benjamin, New York, 1964).

[18]Note that in such discussions the importance of the Everett interpretation usually is stressed. However, we are concerned here solely with the specification of rules for what an observer "sees," not with extraneous language (such as "reduction of the wave packet"). Some discussions of the Everett interpretation appear to evade this issue of what an observer "sees," in which case they do not provide an "interpretation" in our sense. Those which do not evade this issue provide rules which are equivalent to those of other interpretations.

[19]This point also has been made by D. N. Page, in *Quantum Gravity,* proceedings of the Fourth Seminar, Moscow, USSR, 1987, edited by M. A. Markov, V. A. Berezin, and V. P. Frolov (World Scientific, Singapore, 1988).

[20]A. Einstein, Siz. Preuss. Acad. Scis. (1919) translated as "Do Gravitational Fields Play an Essential Role in the Structure of Elementary Particles of Matter" in *The Principle of Relativity,* by A. Einstein *et al.* (Dover, New York, 1952); J. J. van de Bij, H. van Dam, and Y. J. Ng. Phys. Acta A **116**, 307 (1982); F. Wilczek, Phys. Rep. **104**, 111 (1984); A. Zee, in *Gauge Interactions,* proceedings of the 20th Orbis Scientiae: Dedicated to P. A. M. Dirac's 80th year, Miami, Florida, 1983, edited by B. Kursunoglu and A. Perlmutter (Plenum, New York, 1984); S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989); M. Henneaux and C. Teitelboim, University of Texas report, 1988 (unpublished).

[21]W. G. Unruh, Phys. Rev. D **40**, 1048 (1989).

[22]R. Sorkin, in *History of Modern Gauge Theories,* edited by M. Dresaen and A. Rosenblum (Plenum, New York, in press).

[23]M. Reed and B. Simon, *Fourier Analysis, Self-Adjointness* (Academic, New York, 1975).