PHYSICAL REVIEW D VOLUME 40, NUMBER 4 15 AUGUST 1989

Unimodular theory of canonical quantum gravity

W. G. Unruh
Cosmology Program, Canadian Institute for Advanced Research, and Department of Physics, University of British Columbia, Vancouver, British Columbia, Canada V6T 2A6
and Institute for Theoretical Physics, University of California, Santa Barbara, Santa Barbara, California 93106
(Received 4 November 1988)

Einstein's theory of gravity is reformulated so that the cosmological constant becomes an integration constant of the theory, rather than a "coupling" constant. However, in the Hamiltonian form of the theory, the Hamiltonian constraint is missing, while the usual momentum constraints are still present. Replacing the Hamiltonian constraint is a secondary constraint, which introduces the cosmological constant. The quantum version has a normal "Schrödinger" form of time development, and the wave function does not obey the usual "Wheeler-DeWitt" equation, making the interpretation of the theory much simpler. The small value of the cosmological constant in the Universe at present becomes a genuine question of initial conditions, rather than a question of why one of the coupling constants has a particular value. The key "weakness" of this formulation is that one must introduce a nondynamic background spacetime volume element.

INTRODUCTION

It has long been a "folk theorem" that any nonlinear "second-order" theory which gives a free massless spin-2 field at linearized order must be Einstein's theory.[1-3] Recently it has been shown that there exist more general theories which obey this requirement.[4] One of the simplest forms is most succinctly formulated as a metric, non-coordinate-invariant theory of a symmetric unit-determinant field $g^{\mu\nu}$. The action is given by

$$S=\int R(g^{\mu\nu})d^4x \ , \tag{1}$$

where the field variables are defined by the requirement that $g^{\mu\nu}$ be symmetric and obey $\det(g^{\mu\nu})=1$. This theory is *not* coordinate invariant, but does have a gauge freedom: namely,

$$g_{\mu\nu}\to g_{\mu\nu}+K^{\alpha\beta}{}_{,\beta}g_{\mu\nu,\alpha}+g_{\mu\alpha}K^{\alpha\beta}{}_{,\beta,\nu}+g_{\nu\alpha}K^{\alpha\beta}{}_{,\beta,\mu} \ , \tag{2}$$

where $K^{\mu\nu}=-K^{\nu\mu}$ is an arbitrary antisymmetric tensor. Note that this gauge transformation could be interpreted as a coordinate transformation with the infinitesimal coordinate change given by

$$\xi^\alpha=K^{\alpha\beta}{}_{,\beta} \ . \tag{3}$$

This infinitesimal coordinate change obeys

$$\xi^\alpha{}_{;\alpha}=0 \tag{4}$$

and thus leaves the determinant of $g^{\mu\nu}$ invariant. Viewed in this way it is obvious that the above Lagrangian does have this gauge transformation as a symmetry. Furthermore, if one assumes that the coupling between any other matter fields and gravity is via this $g_{\mu\nu}$ used as the metric, the matter actions also have this gauge freedom.

One might ask what the difference is between this theory and the usual Einstein theory restricted to a coordinate system in which $\det(g^{\mu\nu})=1$. The difference lies in what one feels the fundamental variables of the theory are and the variations of the Lagrangian which are allowed in deriving the field equations. In this theory, unlike in Einstein's, the variations in $g^{\mu\nu}$ must all obey $g_{\mu\nu}\delta g^{\mu\nu}=0$. It is this restriction on the dynamical variables which leads to the differences in the theories. Restrictions such as this are used all the time in formulating theories. For example, one could extend Einstein's theory by imagining that the "metric" is a completely general nonsymmetric matrix. In doing so one could obtain additional and different field equations for the "metric" than if one assumed that the field variables were symmetric from the start.

The action of Eq. (1) is not generally coordinate invariant, but general coordinate invariance could be reinstituted in the usual way by introducing a background volume form on the spacetime $\eta_{\mu\nu\rho\sigma}=\eta_{[\mu\nu\rho\sigma]}$, and writing the action as

$$S=\int \sqrt{\eta}R((\eta/g)^{1/4}g_{\mu\nu})d^4x \ , \tag{5}$$

where $\sqrt{\eta}=\eta_{0123}$. Under a coordinate transformation, η/g is a scalar, while the presence of the matrix $g_{\mu\nu}$ (now regarded as a fully arbitrary symmetry matrix) only in the combination $g_{\mu\nu}/g^{1/4}$ ensures that the variation $g^{\mu\nu}\delta g_{\mu\nu}$ is zero. In the following, I will assume that η is 1.

The equations of motion for this theory (in the absence of matter) are

$$R^{\mu\nu}-\tfrac{1}{4}Rg^{\mu\nu}=0 \ , \tag{6}$$

where $R^{\mu\nu}$ is the usual Ricci tensor for $g^{\mu\nu}$ regarded as a metric. With matter fields, the right-hand side becomes $T^{\mu\nu}-\frac{1}{4}Tg^{\mu\nu}$, where $T^{\mu\nu}$ is the usual energy-momentum tensor for the matter fields.

Defining

$$R_T^{\mu\nu}=R^{\mu\nu}-\tfrac{1}{4}Rg^{\mu\nu}, \quad T_T^{\mu\nu}=T^{\mu\nu}-\tfrac{1}{4}Tg^{\mu\nu} \ , \tag{7}$$

we have by the Bianchi identity for R and from the equations of motion for T that

$$R_{T\mu}{}^{\nu}{}_{;\nu}=\tfrac{1}{2}R_{,\mu} \quad (8)$$

and

$$T_{T\mu}{}^{\nu}{}_{;\nu}=-\tfrac{1}{2}T_{,\mu} , \quad (9)$$

giving

$$(R+T)_{,\mu}=0 \quad (10)$$

by the equations of motion, or

$$R+T=-\Lambda \quad (11)$$

for some constant Λ. Thus the field equations can be written as

$$G^{\mu\nu}=T^{\mu\nu}+\Lambda g^{\mu\nu} , \quad (12)$$

where Λ is now an arbitrary integration constant, rather than a coupling constant as in Einstein's theory with cosmological constant.

Since any solution of Einstein's equations with cosmological constant can always be written (at least over any topologically R^4 open subset of the spacetime) in a coordinate system with $\det(g_{\mu\nu})=1$, we find that if we identify the matrix $g_{\mu\nu}$ of this theory with the metric components of Einstein's theory in that special coordinate system, then any solution of this theory is a solution of Einstein's theory with some cosmological constant, and any solution of Einstein's equations with any cosmological constant, in an appropriate coordinate system gives a matrix which is a solution of this theory. The physical predictions of this classical theory are thus identical to those of Einstein's theory with some cosmological constant.

Because of the introduction in this theory of a cosmological constant as an integration constant rather than as an explicit coupling constant, theories such as this have been examined by others as a possible way of attacking the cosmological-constant problem.[5] This is not my purpose here. Rather my purpose will be to examine the quantization of this theory. In particular, the ultimate hope is that this theory, or one like it, could help to resolve some of the interpretive problems[6,7] of canonical quantum gravity, especially with respect to the role of time in such an interpretation.

CANONICAL QUANTIZATION

Let us now look at the Hamiltonian form of this theory. We will follow the usual procedure used in the case of Einstein's thoery.[8] Define variables such that

$$\gamma_{ij}=g_{ij}, \quad N_i=N^j\gamma_{ij}=g_{0i} .$$

I will also introduce the variable N defined by

$$N^2=-g_{00}+N^iN_i .$$

Unlike the case in general relativity, N is not an independent variable, but is just given by

$$N^2=1/\det(\gamma_{ij}) .$$

This relation arises from the definition of $g_{\mu\nu}$: namely, that $\det(g_{\mu\nu})=1$. In the usual way, we define conjugate momenta π^{ij} to the dynamic variables γ_{ij} by

$$\pi^{ij}=\delta S/\delta\dot{\gamma}_{ij} .$$

In the above the indices i,j,k, etc., run from 1 to 3, while 0 refers to the "timelike" direction. The overdot denotes a time derivative. Now, going through the usual procedure we finally end up with the super-Hamiltonian density

$$\tilde{H}=NH_0+N^iH_i=H_0\sqrt{\gamma}+N^iH_i$$

and super Hamiltonian

$$\tilde{\mathcal{H}}=\int\tilde{H}\,d^3x ,$$

where H_0 and H_i are the usual Hamiltonian and momentum "constraint" functions given by

$$H_0=(\gamma_{ik}\gamma_{jl}-\tfrac{1}{2}\gamma_{ij}\gamma_{kl})\pi^{ij}\pi^{kl}/\sqrt{\gamma}+\sqrt{\gamma}\,{}^{(3)}R ,$$

$$H_i=\gamma_{ij}\pi^{ik}{}_{|k} .$$

Here the vertical bar denotes a covariant three-derivative with respect to the metric γ_{ij}, remembering that π^{ij} is a three-tensor density. Note that the fact that in this theory $N=1/\sqrt{\gamma}$ does not alter the formal manipulations which lead to this Hamiltonian since N does not depend on $\dot{\gamma}_{ij}$. However, since N is no longer a free variable in this theory, only the momentum constraints

$$H_i=0$$

survive as constraints. The super Hamiltonian $\tilde{\mathcal{H}}$ is thus not zero.

If we evaluate the Poisson brackets of the constraints, or rather of $\int\xi^iH_id^3x$ with ξ^i arbitrary, with $\tilde{\mathcal{H}}$ we find that we need

$$\int\xi^i(H_0/\sqrt{\gamma})_{,i}d^3x=0$$

as a secondary constraint.[9] This implies that

$$H_0=\sqrt{\gamma}\Lambda ,$$

where Λ is a spatial constant. Furthermore, the equations of motion imply that Λ is also a constant of the motion. It thus enters into the equations in just the way a cosmological constant does.

We can formally quantize this theory. The wave function $\Psi(\gamma^{ij})$ would obey the momentum constraint

$$\int\xi^k{}_{|j}\gamma_{ij}(\delta\Psi/\delta\gamma_{ij})d^3x=0 ,$$

the secondary constraint

$$\int\xi^i(H_0/\sqrt{\gamma})_{,i}d^3x\,\Psi=0 ,$$

and the Schrödinger equation

$$i\frac{\partial\Psi}{\partial t}=\int(\gamma)^{-1/2}H_0d^3x\;\Psi .$$

Since the wave function now contains an explicit time dependence, the interpretation of this wave function becomes much easier than it is in the usual quantum theory where the Hamiltonian constraint explicitly sets the super-Hamiltonian to zero, and thus makes the wave function time independent. It is important to point out here that the time t is a time just like that in ordinary quantum mechanics. It is not, as in the usual Wheeler-DeWitt approach to canonical quantum gravity, one of the dynamical (unconstrained) phase-space variables. It is external—an extra variable over and above the phase-space variables—rather than internal—one of the phase-space variables.

The eigenstates of "cosmological constant" will be given by wave functions which obey the Wheeler-DeWitt-type equation

$$(\gamma)^{-1/2}H_0\Psi_\Lambda=\Lambda\Psi_\Lambda \ .$$

Thus the general solution to the time-dependent Schrödinger equation is

$$\Psi(t)=\int \alpha(\Lambda)\exp\left[-i\int \Lambda\, d^3x\, dt\right]\Psi_\Lambda d^3x$$

for some function α and for some choice of the Ψ_Λ. Note that since $g=1$, and since the volume element of the spacetime is not a quantum variable but is a given classical volume element, the phase factor in the exponential is just the cosmological constant times the four-volume of the spacetime between the initial time slice and the time slice of interest.

INTERPRETATION

Although the interpretation has been made somewhat simpler because of the explicit dependence of the wave function on time, there are certainly still problems that remain. Because of the existence of the constraints, one would like one's observables to be defined only on the Hilbert subspace which obeys the constraint equations. Thus, because of the presence of the momentum constraints, we can demand that the observables of the theory commute with these constraints. Thus the only quantities which are to be regarded as measurable are those which are invariant under any spatial coordinate transformations. For example, $\gamma_{ij}(x)$ would not be a measurable quantity, but $\int \sqrt{\gamma}\phi(x)\delta({}^{(3)}R(x)-r)d^3x$ might be. [This is just the operator ϕ restricted to those points in space where ${}^{(3)}R(x)=r$].

In addition to these momentum constraints, we must also demand that one's observables also commute with the secondary constraint of Eq. (7). This just leads to the conclusion that if the observable is O, we must have not only that O commutes with H_i but that $\dot{O}$ does as well. The observables are then those quantities which are coordinate invariant, and whose time derivative is also coordinate invariant. Constants of the motion will, of course, satisfy this requirement, but the allowed set of observables is larger than just constants of the motion. For example, in the usual homogeneous and isotropic minisuperspace models, all of the operators are observable since homogeneity and isotropy imply that the constraints are trivially satisfied. This contrasts with the usual case in which both the "momentum" constraints as well as the Hamiltonian constraints are imposed on the theory. In that case, following the above philosophy, the observables must commute with the constraints, which leads to the conclusion that the only observables are constants of the motion.

How does one go about interpreting measurements in this theory?[6] Measurements are assumed to be made at some time t. The time t is, of course, not directly observable or measurable. So, although the wave function describes the development in t time of the probabilities of various outcomes, how can one use these predictions since t cannot be directly determined? The answer is, as usual, by the use of "clocks." For certain wave functions Ψ, some dynamic variable may correlate well with the value of t—i.e., over some limited range of values of t, the measured value of that observable may allow one to infer (in a statistical sense) what the value of t is. Having made the measurement of that observable, one can then use the information one has about t to make predictions about the values of other observables.

How does this differ from the usual approach to the "wave function of the Universe?" There one has a wave function Ψ which is independent of time. One argues that there one must use some of one's variables as "clocks" and that the wave function predicts correlations between the clock and the other observables of the theory, which seems similar to the procedure I have sketched in the time-dependent theory. The difference is that there the observables are only the sets of constants of the motion, constants not for some limited "time" but forever. Furthermore, one does not have there the concept of an approximate clock, good for only a limited time, and of limited accuracy.

There is one rather novel feature in the theory presented here. In the course of a measurement on any physical quantity, the outcome of the measurement would, in general, not leave the system with its "energy," or rather, its cosmological constant unchanged.

Let us look at a minisuperspace theory as an example. Let us take the metric to be

$$ds^2=dt^2/a(t)^6-a(t)^2(dx^2+dy^2+dz^2)$$

which has unit determinant as required. I will take the closed model by identifying the x, y, and z coordinates with period 1. Let us minimally couple the metric to a homogeneous scalar field ϕ with potential $V(\phi)$. The resulting Hamiltonian is given by

$$\mathcal{H}=-\pi_a^2/a^6+\pi_\phi^2/a^6+V(\phi) \ .$$

Since the model is homogeneous, the constraints are automatically satisfied. The Schrödinger equation then becomes

$$i\frac{\partial\Psi(a,\phi,t)}{\partial t}=\frac{\partial^2\Psi}{\partial(a^3)^2}-\frac{1}{a^6}\frac{\partial^2\Psi}{\partial\phi^2}+V(\phi)\Psi$$

or

$$i\frac{\partial\Psi}{\partial t}=\frac{3}{8\bar{z}}\frac{\partial}{\partial z}z\frac{\partial\Psi}{\partial\bar{z}}-\frac{1}{2z^2}\frac{\partial^2\Psi}{\partial\phi^2}+V(\phi)\Psi \ ,$$

where I have chosen the "Laplacian" ordering for the momentum terms. Here $z=a^3$. I will take the simplest model with $V(\phi)=0$.

The general solution for the equation is given by

$$\Psi(z,\phi,t)=\int A(p,\Lambda)e^{-i(\Lambda t-p\phi)} \times J_{i\eta}((8\Lambda/3)^{1/2}z)dp\,d\Lambda\ ,$$

where $\eta=(\frac{4}{3})^{1/2}p$.

Note that at any time t, we have

$$\int|\Psi|^2 d\phi\, a^5 da=\int|A|^2 dp\, d\Lambda$$

and we can thus find square integrable solutions to the wave equation. However, if $A(p,\Lambda)\propto\delta(\Lambda)$, the solutions are not square integrable. Since the equation for Ψ with some definite value Λ is just the Wheeler-DeWitt equation with that definite cosmological constant Λ, we see how this approach avoids the persistant problem besetting the usual approach, namely, that the wave function there is usually not square integrable, making any probablistic interpretation of the wave function difficult.

For large a, the wave function behaves like a free particle with coordinate $z=a^3$, mass$=\frac{3}{16}$, and energy$=\Lambda$. We have $\langle a^3\rangle\propto t$, which would imply an exponential expansion of the space with proper time. [Remember that we have chosen $N=1/a^3$ in order to impose the $\det(g_{\mu\nu})=1$ condition.]

The above solution has assumed that $\Lambda>0$. For $\Lambda<0$, we must replace $J_{i\eta}$ in the above solution by $H^1_{i\eta}$, in order that the solutions fall off at large a. In this case, the Universe will always "bounce" at some large value of a, with an exponentially decaying probability of finding larger values for a.

Since $\dot{a}$ is not equal to 0, the Hamiltonian and a do not commute. A measurement of a will, therefore, not preserve the "energy" or value of the cosmological constant.

This nonconstancy of the cosmological constant under quantum measurement of nonconstant variable leads to a puzzle for nonhomogeneous situations. How can my measurement here and now of some quantity change the cosmological constant for the whole Universe everywhere? One answer is in the speculation that the uncertainty in the cosmological constant induced by my measurement would be of the order of $\Delta\Lambda\sim\Delta E/V$, where ΔE is the uncertainty in the energy induced by the measurement, and V is the spatial volume of the Universe, a change which would be totally undetectable.

The other is that all observables in the theory must commute with the momentum-constraint functions: i.e., the generators of spatial coordinate transformations. All observables are thus in a sense homogeneous, and thus in some sense exist everywhere at once as far as the coordinates x are concerned. The change in the cosmological constant "everywhere at once" is thus not surprising since the measurement of the observable takes place "everywhere at once." This, however, evades the physical question, since we have a strong feeling that as far as we are concerned, causality is a law of nature, and our measurements made here on Earth should not affect the behavior of the cosmological constant at distant galaxies, at least not noncausally.

The question this raises is how does this theory look from the inside. In our description of the theory, I have employed concepts such as the coordinates x. These are, however, inaccessible to observers in the theory. Since we, as physical beings, are presumably objects describable within the theory, and since all of the aspects of our physical being correspond to observables in the theory, how do we describe our notions of causality, etc., and test to see whether or not the theory predicts that observers within the theory will experience a causal universe? How do observers within the theory experience the change in the cosmological constant caused by measurements made by those observers?

ACKNOWLEDGMENTS

I would especially like to thank R. Wald. It was in the course of conversations with him that the ideas for this theory arose. I would also like to thank the Natural Sciences and Engineering Research Council of Canada for research funds, and the Canadian Institute for Advanced Research and the LAC Minerals Corp. for financial support during the course of this work. I also thank the Institute for Theoretical Physics, University of California at Santa Barbara and NSF Grant No. PHY82-17853 for support in the preparation of this manuscript.

---

[1]R. P. Feynman, Lectures on Gravitation, California Institute of Technology Lecture Notes, 1962 (unpublished).

[2]S. Weinberg, Phys. Rev. **138**, 988 (1965).

[3]R. M. Wald, Phys. Rev. D **33**, 3613 (1986).

[4]K. Heiderich and W. G. Unruh (in preparation).

[5]The equations corresponding to this theory were first suggested by 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), who also pointed out the advantage of the theory in giving the cosmological constant as an integration constant rather than a coupling constant. The theory has been rediscovered and studied by a number of people recently in the context of the cosmological-constant problem. See, for example, J. J. van der Bij *et al.*, Physica **A116**, 307 (1982); F. Wilczek, Phys. Rep. **104**, 111 (1984); A. Zee, in *High Energy Physics,* proceedings of the 20th Annual Orbis Scientiae, Coral Gables, 1983, edited by B. Kursunoglu, S. C. Mintz, and A. Perlmutter (Plenum, New York, 1985); S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989). They are more negative about the theory than I am and do not really examine the quantum theory. In the context of the quantum theory, see also R. Sorkin, "Notes for the North Andover Seminar on Quantum Gravity" (unpublished), who uses the four-volume as a time coordinate, which leads to a theory with cosmological constant as a Lagrange multiplier. After this paper was submitted, I received a copy of a paper by M. Henneaux and C.

Teitelboim [University of Texas Physics report (unpublished)] which examines the classical theory with the view to reintroducing general covariance by using additional scalar and vector pure gauge fields.

[6]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), p. 252.

[7]W. G. Unruh and R. M. Wald, Report No. ITP88-190 (unpublished).

[8]See, for example, R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984), p. 459.

[9]See, for example, P. A. M. Dirac, *Quantum Mechanics* (Cambridge University Press, Cambridge, England, 1935).