The cosmological constant problem*

Steven Weinberg

Theory Group, Department of Physics, University of Texas, Austin, Texas 78712

Astronomical observations indicate that the cosmological constant is many orders of magnitude smaller than estimated in modern theories of elementary particles. After a brief review of the history of this problem, five different approaches to its solution are described.

CONTENTS

As I was going up the stair,
I met a man who wasn't there.
He wasn't there again today,
I wish, I wish he'd stay away.

Hughes Mearns

I. INTRODUCTION

Physics thrives on crisis. We all recall the great progress made while finding a way out of various crises of the past: the failure to detect a motion of the Earth through the ether, the discovery of the continuous spectrum of beta decay, the τ-θ problem, the ultraviolet divergences in electromagnetic and then weak interactions, and so on. Unfortunately, we have run short of crises lately. The "standard model" of electroweak and strong interactions currently faces neither internal inconsistencies nor conflicts with experiment. It has plenty of loose ends; we know no reason why the quarks and leptons should have the masses they have, but then we know no reason why they should not.

Perhaps it is for want of other crises to worry about that interest is increasingly centered on one veritable crisis: theoretical expectations for the cosmological constant exceed observational limits by some 120 orders of magnitude.[1] In these lectures I will first review the history of this problem and then survey the various attempts that have been made at a solution.

II. EARLY HISTORY

After completing his formulation of general relativity in 1915–1916, Einstein (1917) attempted to apply his new theory to the whole universe. His guiding principle was that the universe is static: "The most important fact that we draw from experience is that the relative velocities of the stars are very small as compared with the velocity of light." No such static solution of his original equations could be found (any more than for Newtonian gravitation), so he modified them by adding a new term involving a free parameter λ, the cosmological constant:[2]

$$R_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}R-\lambda g_{\mu\nu}=-8\pi G T_{\mu\nu}\ . \tag{2.1}$$

Now, for $\lambda>0$, there was a static solution for a universe filled with dust of zero pressure and mass density

$$\rho=\frac{\lambda}{8\pi G}\ . \tag{2.2}$$

Its geometry was that of a sphere S_3, with proper circumference $2\pi r$, where

$$r=1/\sqrt{8\pi\rho G}\ , \tag{2.3}$$

so the mass of the universe was

$$M=2\pi^2 r^3\rho=\frac{\pi}{4}\lambda^{-1/2}G^{-1}\ . \tag{2.4}$$

In some popular history accounts, it was Hubble's discovery of the expansion of the universe that led Einstein to retract his proposal of a cosmological constant. The real story is more complicated, and more interesting.

One disappointment came almost immediately. Einstein had been pleased at the connection in his model between the mass density of the universe and its geometry, because, following Mach's lead, he expected that the mass distribution of the universe should set inertial frames. It was therefore unpleasant when his friend de Sitter, with whom Einstein remained in touch during the war, in 1917 proposed another apparently static cosmological model with no matter at all. (See de Sitter, 1917.) Its line element (using the same coordinate system as de Sitter, but in a different notation) was

$$d\tau^2=\frac{1}{\cosh^2 Hr}[dt^2-dr^2-H^{-2}\tanh^2 Hr(d\theta^2+\sin^2\theta\, d\varphi^2)]\ , \tag{2.5}$$

---

*Morris Loeb Lectures in Physics, Harvard University, May 2, 3, 5, and 10, 1988.

[1]For a good nonmathematical description of the cosmological constant problem, see Abbott (1988).

[2]The notation used here for metrics, curvatures, etc., is the same as in Weinberg (1972).

with H related to the cosmological constant by

$$H=\sqrt{\lambda/3} \quad (2.6)$$

and $\rho=p=0$. Clearly matter was not needed to produce inertia.

At about this time, the redshift of distant objects was being discovered by Slipher. Over the period from 1910 to the mid-1920s, Slipher (1924) observed that galaxies (or, as then known, spiral nebulae) have redshifts $z\equiv\Delta\lambda/\lambda$ ranging up to 6%, and only a few have blueshifts. Weyl pointed out in 1923 that de Sitter's model would exhibit such a redshift, increasing with distance, because although the metric in de Sitter's coordinate system is time independent, test bodies are not at rest; there is a nonvanishing component of the affine connection

$$\Gamma^r_{tt}=-H\sinh Hr\tanh Hr \quad (2.7)$$

giving a redshift proportional to distance

$$z\simeq Hr \quad \text{for } Hr\ll 1\ . \quad (2.8)$$

In his influential textbook, Eddington (1924) interpreted Slipher's redshifts in terms of de Sitter's "static" universe.

But of course, although the cosmological constant *was* needed for a static universe, it was not needed for an expanding one. Already in 1922, Friedmann (1924) had described a class of cosmological models, with line element (in modern notation)

$$d\tau^2=dt^2-R^2(t)\left[\frac{dr^2}{1-kr^2}+r^2(d\theta^2+\sin^2\theta\,d\varphi^2)\right]. \quad (2.9)$$

These are comoving coordinates; the universe expands or contracts as $R(t)$ increases or decreases, but the galaxies keep fixed coordinates r,θ,φ. The motion of the cosmic scale factor is governed by an energy-conservation equation

$$\left[\frac{dR}{dt}\right]^2=-k+\tfrac{1}{3}R^2(8\pi G\rho+\lambda)\ . \quad (2.10)$$

The de Sitter model is just the special case with $k=0$ and $\rho=0$; in order to put the line element (2.5) in the more general form (2.9), it is necessary to introduce new coordinates,

$$t'=t-H^{-1}\ln\cosh Hr\ ,$$
$$r'=H^{-1}\exp(-Ht)\sinh Hr\ , \quad (2.11)$$
$$\theta'=\theta,\quad \varphi'=\varphi\ ,$$

and then drop the primes. However, we can also easily find expanding solutions with $\lambda=0$ and $\rho>0$. Pais (1982) quotes a 1923 letter of Einstein to Weyl, giving his reaction to the discovery of the expansion of the universe: "If there is no quasi-static world, then away with the cosmological term!"

III. THE PROBLEM

Unfortunately, it was not so easy simply to drop the cosmological constant, because anything that contributes to the energy density of the vacuum acts just like a cosmological constant. Lorentz invariance tells us that in the vacuum the energy-momentum tensor must take the form

$$\langle T_{\mu\nu}\rangle=-\langle\rho\rangle g_{\mu\nu}\ . \quad (3.1)$$

(A minus sign appears here because we use a metric which for flat space-time has $g_{00}=-1$.) Inspection of Eq. (2.1) shows that this has the same effect as adding a term $8\pi G\langle\rho\rangle$ to the effective cosmological constant

$$\lambda_{\text{eff}}=\lambda+8\pi G\langle\rho\rangle\ . \quad (3.2)$$

Equivalently we can say that the Einstein cosmological constant contributes a term $\lambda/8\pi G$ to the total effective vacuum energy

$$\rho_V=\langle\rho\rangle+\lambda/8\pi G=\lambda_{\text{eff}}/8\pi G\ . \quad (3.3)$$

A crude experimental upper bound on λ_{eff} or ρ_V is provided by measurements of cosmological redshifts as a function of distance, the program begun by Hubble in the late 1920s. The present expansion rate is today estimated as

$$\left[\frac{1}{R}\frac{dR}{dt}\right]_{\text{now}}\equiv H_0\simeq 50-100\ \text{km/sec Mpc}$$
$$\simeq(\tfrac{1}{2}-1)\times 10^{-10}/\text{yr}\ .$$

Furthermore, we do not gross effects of the curvature of the universe, so very roughly

$$|k|/R^2_{\text{now}}\lesssim H_0^2\ .$$

Finally, the ordinary nonvacuum mass density of the universe is not much greater than its critical value

$$|\rho-\langle\rho\rangle|\lesssim 3H_0^2/8\pi G\ .$$

Hence (2.10) shows that

$$|\lambda_{\text{eff}}|\lesssim H_0^2$$

or, in physicists' units,

$$|\rho_V|\lesssim 10^{-29}\ \text{g/cm}^3\approx 10^{-47}\ \text{GeV}^4\ . \quad (3.4)$$

A more precise observational bound will be discussed in Sec. V, but this one will be good enough for our present purposes.

As everyone knows, the trouble with this is that the energy density $\langle\rho\rangle$ of empty space is likely to be enormously larger than 10^{-47} GeV^4. For one thing, summing the zero-point energies of all normal modes of some field of mass m up to a wave number cutoff $\Lambda\gg m$ yields a vacuum energy density (with $\hbar=c=1$)

$$\langle\rho\rangle=\int_0^{\Lambda}\frac{4\pi k^2dk}{(2\pi)^3}\frac{1}{2}\sqrt{k^2+m^2}\simeq\frac{\Lambda^4}{16\pi^2}\ . \quad (3.5)$$

If we believe general relativity up to the Planck scale, then we might take $\Lambda \simeq (8\pi G)^{-1/2}$, which would give

$$\langle \rho \rangle \approx 2^{-10}\pi^{-4}G^{-2} = 2\times 10^{71}\ \mathrm{GeV}^4 \ . \tag{3.6}$$

But we saw that $|\langle \rho \rangle + \lambda/8\pi G|$ is less than about 10^{-47} GeV^4, so the two terms here must cancel to better than 118 decimal places. Even if we only worry about zero-point energies in quantum chromodynamics, we would expect $\langle \rho \rangle$ to be of order $\Lambda_{\mathrm{QCD}}^4/16\pi^2$, or 10^{-6} GeV^4, requiring $\lambda/8\pi G$ to cancel this term to about 41 decimal places.

Perhaps surprisingly, it was a long time before particle physicists began seriously to worry about this problem, despite the demonstration in the Casimir effect of the reality of zero-point energies.[3] Since the cosmological upper bound on $|\langle \rho \rangle + \lambda/8\pi G|$ was vastly less than any value expected from particle theory, most particle theorists simply assumed that for some unknown reason this quantity was zero. But cosmologists generally continued to keep an open mind, analyzing cosmological data in terms of models with a possibly nonvanishing cosmological constant.

In fact, as far as I know, the first published discussion of the contribution of quantum fluctuations to the effective cosmological constant was triggered by astronomical observations. In the late 1960s it seemed that an excessively large number of quasars were being observed with redshifts clustered about $z = 1.95$. Since $1+z$ is the ratio of the cosmic scale factor $R(t)$ at present to its value at the time the light now observed was emitted, this could be explained if the universe loitered for a while at a value of $R(t)$ equal to $1/2.95$ times the present value. A number of authors [Petrosian, Salpeter, and Szekeres (1967); Shklovsky (1967); Rowan-Robinson (1968)] proposed that such a loitering could be accounted for in a model proposed by Lemaître (1927, 1931). In this model there is a positive cosmological constant λ_{eff} and positive curvature $k = +1$, just as in the static Einstein model, while the mass of the universe is taken close to the Einstein value (2.4). The scale factor $R(t)$ starts at $R = 0$ and then increases; however, when the mass density drops to near the Einstein value (2.2), the universe behaves for a while like a static Einstein universe, until the instability of this model takes over and the universe starts expanding again. In order for this idea to explain a preponderance of redshifts at $z = 1.95$, the vacuum energy density ρ_V would have to be $(2.95)^3$ times the present nonvacuum mass density ρ_0.

These considerations led Zeldovich (1967) to attempt to account for a nonzero vacuum energy density in terms of quantum fluctuations. As we have seen, the zero-point energies themselves gave far too large a value for $\langle \rho \rangle$, so Zeldovich assumed that these were canceled by $\lambda/8\pi G$, leaving only higher-order effects: in particular, the gravitational force between the particles in the vacuum fluctuations. (In Feynman diagram terms, this corresponds to throwing away the one-loop vacuum graphs, but keeping those with two loops.) Taking Λ^3 particles of energy Λ per unit volume gives the gravitational self-energy density of order

$$\langle \rho \rangle \approx (G\Lambda^2/\Lambda^{-1})\Lambda^3 = G\Lambda^6 \ . \tag{3.7}$$

For no clear reason, Zeldovich took the cutoff Λ as 1 GeV, which yields a density $\langle \rho \rangle \approx 10^{-38}$ GeV^4, much smaller than from zero-point energies themselves, but still larger than the observational bound (3.4) on $|\langle \rho \rangle + \lambda/8\pi G|$ by some 9 orders of magnitude. Neither Zeldovich nor anyone else felt encouraged to pursue these ideas.

The real beginning of serious worry about the vacuum energy seems to date from the success of the idea of spontaneous symmetry breaking in the electroweak theory. In this theory, the scalar field potential takes the form (with $\mu^2 > 0$, $g > 0$)

$$V = V_0 - \mu^2\phi^\dagger\phi + g(\phi^\dagger\phi)^2 \ . \tag{3.8}$$

At its minimum this takes the value

$$\langle \rho \rangle = V_{\mathrm{min}} = V_0 - \frac{\mu^4}{4g} \ . \tag{3.9}$$

Apparently some theorists felt that V should vanish at $\phi = 0$, which would give $V_0 = 0$, so that $\langle \rho \rangle$ would be negative definite.[4] In the electroweak theory this would give $\langle \rho \rangle \simeq -g(300\ \mathrm{GeV})^4$, which even for g as small as α^2 would yield $|\langle \rho \rangle| \simeq 10^6$ GeV^4, larger than the bound on ρ_V by a factor 10^{53}. Of course we know of no reason why V_0 or λ must vanish, and it is entirely possible that V_0 or λ cancels the term $-\mu^4/4g$ (and higher-order corrections), but this example shows vividly how unnatural it is to get a reasonably small effective cosmological constant. Moreover, at early times the effective temperature-dependent potential has a positive coefficient for $\phi^\dagger\phi$, so the minimum then is at $\phi = 0$, where $V(\phi) = V_0$. Thus, in order to get a zero cosmological constant today, we have to put up with an enormous cosmological constant at times before the electroweak phase transition. [This is *not* in conflict with experiment; in fact, the phase transition occurs at a temperature T of order $\mu/\sqrt{g}$, so the black-body radiation present at that

[3]Casimir (1948) showed that quantum fluctuations in the space between two flat conducting plates with separation d would produce a force per unit area equal to $\hbar c\pi^2/240d^4$, or 1.30×10^{-18} dyn cm$^2/d^4$. This was measured by Sparnaay (1957), who found a force per area of $(1\text{–}4)\times 10^{-18}$ dyn cm$^2/d^4$, when d was varied between 2 and 10 μm.

[4]Veltman (1975) attributes this view to Linde (1974), himself (quoted as "to be published"), and Dreitlein (1974). However, Linde's paper does not seem to me to take this position. Dreitlein's paper proposed that Eq. (3.9) could give an acceptably small value of $\langle \rho \rangle$, with $\mu/\sqrt{g}$ fixed by the Fermi coupling constant of weak interactions, if μ is very small, of order 10^{-27} MeV. Veltman's paper gives experimental arguments against this possibility.

time has an energy density of order μ^4/g^2, larger than the vacuum energy by a factor $1/g$ (Bludman and Ruderman, 1977).] At even earlier times there were other transitions, implying an even larger early value for the effective cosmological constant. This is currently regarded as a good thing; the large early cosmological constant would drive cosmic inflation, solving several of the long-standing problems of cosmological theory (Guth, 1981; Albrecht and Steinhardt, 1982; Linde, 1982). We want to explain why the effective cosmological constant is small now, *not* why it was always small.

Before closing this section, I want to take up a peculiar aspect of the problem of the cosmological constant. The appearance of an effective cosmological constant makes it impossible to find any solutions of the Einstein field equations in which $g_{\mu\nu}$ is the constant Minkowski term $\eta_{\mu\nu}$. That is, the original symmetry of general covariance, which is always broken by the appearance of any given metric $g_{\mu\nu}$, cannot, without fine-tuning, be broken in such a way as to preserve the subgroup of space-time translations.

This situation is unusual. Usually if a theory is invariant under some group G, we would not expect to have to fine-tune the parameters of the theory in order to find vacuum solutions that preserve any given subgroup $H \subset G$. For instance, in the electroweak theory, there is a finite range of parameters in which any number of doublet scalars will get vacuum expectation values that preserve a U(1) subgroup of SU(2)$\times$U(1). So why will this not work for the translational subgroup of the group of general coordinate transformations? Suppose we look for a solution of the field equations that preserves translational invariance. With all fields constant, the field equations for matter and gravity are

$$\frac{\partial \mathcal{L}}{\partial \psi_i}=0 , \tag{3.10}$$

$$\frac{\partial \mathcal{L}}{\partial g_{\mu\nu}}=0 . \tag{3.11}$$

With N ψ's, these are $N+6$ equations for $N+6$ unknowns, so one might expect a solution without fine-tuning. The problem is that when (3.10) is satisfied, the dependence of $\mathcal{L}$ on $g_{\mu\nu}$ is too simple to allow a solution of (3.11). There is a GL(4) symmetry that survives as a vestige of general covariance even when we constrain the fields to be constants: under the GL(4) transformation

$$g_{\mu\nu} \to A^{\rho}{}_{\mu} A^{\sigma}{}_{\nu} g_{\rho\sigma} , \tag{3.12}$$

$$\psi_i \to D_{ij}(A)\psi_j ; \tag{3.13}$$

the Lagrangian transforms as a density,

$$\mathcal{L} \to \mathrm{Det}\, A \mathcal{L} . \tag{3.14}$$

When Eq. (3.10) is satisfied, this implies that $\mathcal{L}$ transforms as in (3.14) under (3.12) *alone*. This has the unique solution

$$\mathcal{L}=c(\mathrm{Det} g)^{1/2} , \tag{3.15}$$

with c independent of $g_{\mu\nu}$. With this $\mathcal{L}$, there are no solutions of Eq. (3.11), unless for some reason the coefficient c vanishes when (3.10) is satisfied.

Now that the problem has been posed, we turn to its possible solution. The next five sections will describe five directions that have been taken in trying to solve the problem of the cosmological constant.

IV. SUPERSYMMETRY, SUPERGRAVITY, SUPERSTRINGS

Shortly after the development of four-dimensional globally supersymmetric field theories, Zumino (1975) pointed out that supersymmetry in these theories would, if unbroken, imply a vanishing vacuum energy. The argument is very simple: the supersymmetry generators Q_α satisfy an anticommutation relation

$$\{Q_\alpha, Q_\beta^\dagger\}=(\sigma_\mu)_{\alpha\beta}P^\mu , \tag{4.1}$$

where α and β are two-component spin indices; σ_1, σ_2, and σ_3 are the Pauli matrices; $\sigma_0=1$; and P^μ is the energy-momentum 4-vector operator. If supersymmetry is unbroken, then the vacuum state $|0\rangle$ satisfies

$$Q_\alpha|0\rangle = Q_\alpha^\dagger|0\rangle = 0 , \tag{4.2}$$

and from (4.1) and (4.2) we infer that the vacuum has vanishing energy and momentum

$$\langle 0|P^\mu|0\rangle = 0 .$$

This result can also be obtained by considering the potential $V(\phi,\phi^*)$ for the chiral scalar fields ϕ^i of a globally supersymmetric theory:

$$V(\phi,\phi^*)=\sum_i \left|\frac{\partial W(\phi)}{\partial \phi^i}\right|^2 , \tag{4.3}$$

where $W(\phi)$ is the so-called superpotential. (Gauge degrees of freedom are ignored here, but they would not change the argument.) The condition for unbroken supersymmetry is that W be stationary in ϕ, which would imply that V take its minimum value,

$$\langle \rho \rangle = V_{\min} = 0 . \tag{4.4}$$

Quantum effects do not change this conclusion, because with boson-fermion symmetry, the fermion loops cancel the boson ones.

The trouble with this result is that supersymmetry is broken in the real world, and in this case either (4.1) or (4.3) shows that the vacuum energy is positive-definite. *If* this vacuum energy were the sole contribution to the effective cosmological constant, then the effect of supersymmetry would be to convert the problem of the cosmological constant from a crisis into a disaster.

Fortunately this is not the whole story. It is not possible to decide the value of the effective cosmological constant unless we explicitly introduce gravitation into the theory. Any globally supersymmetric theory that involves gravity is inevitably a locally supersymmetric supergravity theory. In such a theory the effective cosmological constant *is* given by the expectation value of the potential, but the potential is now given by (Cremmer *et al.*, 1978, 1979; Barbieri *et al.*, 1982; Witten and Bagger, 1982)

$$V(\phi,\phi^*)=\exp(8\pi GK)[D_i W(\mathcal{G}^{-1})^i{}_j(D_j W)^* -24\pi G|W|^2] , \tag{4.5}$$

where $K(\phi,\phi^*)$ is a real function of both ϕ and ϕ^* known as the Kahler potential, $D_i W$ is a sort of covariant derivative

$$D_i W \equiv \frac{\partial W}{\partial \phi^i}+8\pi G\frac{\partial K}{\partial \phi^i} , \tag{4.6}$$

and $(\mathcal{G}^{-1})^i{}_j$ is the inverse of a metric

$$\mathcal{G}^i{}_j \equiv \frac{\partial^2 K}{\partial \phi^{i*}\partial \phi^j} . \tag{4.7}$$

The condition for unbroken supersymmetry is now $D_i W=0$. This again yields a stationary point of the potential, but now it is one at which V is generally negative. In fact, even if we fine-tuned W so that there were a supersymmetric stationary point at which $W=0$ and hence $V=0$, such a solution would not, in general, be the state of lowest energy, though it would be stable [Coleman and de Luccia (1980), Weinberg (1982)]. It should, however, be mentioned that if there is a set of field values at which $W=0$ and $D_i W=0$ for all i in lowest order of perturbation theory, then the theory has a supersymmetric equilibrium configuration with $V=0$ to all orders of perturbation theory, though not necessarily beyond perturbation theory (Grisaru, Siegel, and Rocek, 1979). The same is believed to be true in superstring perturbation theory (Dine and Seiberg, 1986; Friedan, Martinec, and Shenker, 1986; Martinec, 1986; Attick, Moore, and Sen, 1987; Morozov and Perelomov, 1987).

Without fine-tuning, we can generally find a nonsupersymmetric set of scalar field values at which $V=0$ and $D_i W\neq 0$, but this would not normally be a stationary point of V. Thus in supergravity the problem of the cosmological constant is no more a disaster, but just as much a crisis, as in nonsupersymmetric theories.

On the other hand, supergravity theories offer opportunities for changing the context of the cosmological constant problem, if not yet for solving it. Cremmer *et al.* (1983) have noted that there is a class of Kahler potentials and superpotentials that, for a broad range of most parameters, automatically yield an equilibrium scalar field configuration in which $V=0$, even though supersymmetry is broken. Here is a somewhat generalized version: the Kahler potential is

$$K=-3\ln|T+T^*-h(C^a,C^{a*})|/8\pi G +\tilde{K}(S^n,S^{n*}) \tag{4.8}$$

while the superpotential is

$$W=W_1(C^a)+W_2(S^n) , \tag{4.9}$$

and T,C^a,S^n are all chiral scalar fields. No constraints are placed on the functions $h(C^a,C^{a*})$, $\tilde{K}(S^n,S^{n*})$, $W_1(C^a)$, or $W_2(S^n)$, except that h and $\tilde{K}$ are real, and functions all depend only on the fields indicated; in particular, the superpotential must be independent of the single chiral scalar T.

With these conditions the potential (4.5) takes the form

$$V=\exp(8\pi G\tilde{K})\left[\frac{1}{3(T+T^*+h)^3}\left[\frac{\partial W}{\partial C^a}\right](\mathcal{N}^{-1})^a{}_b \times\left[\frac{\partial W}{\partial C^b}\right]^* +(D_n W)(\mathcal{G}^{-1})^n{}_m(D_m W)^*\right] , \tag{4.10}$$

where $(\mathcal{N}^{-1})^a{}_b$ is the reciprocal of the matrix

$$\mathcal{N}^a{}_b=\frac{\partial^2 h}{\partial C^{a*}\partial C^b} . \tag{4.11}$$

The matrices $\mathcal{N}^a{}_b$ and $\mathcal{G}^n{}_m$ are necessarily positive-definite, because of their role in the kinetic part of the scalar Lagrangian

$$\mathcal{L}_{\rm kin}=-\mathcal{G}^i{}_j\frac{\partial\phi^{i*}}{\partial x^\mu}\frac{\partial\phi^j}{\partial x_\mu} =-\frac{3}{(T+T^*+h)^2}\left[\frac{\partial T}{\partial x^\mu}+\frac{\partial h}{\partial C^a}\frac{\partial C^a}{\partial x^\mu}\right]\left[\frac{\partial T}{\partial x_\mu}+\frac{\partial h}{\partial C^b}\frac{\partial C^b}{\partial x_\mu}\right]^* -\frac{3}{|T+T^*-h|}\mathcal{N}^a{}_b\frac{\partial C^{a*}}{\partial x^\mu}\frac{\partial C^b}{\partial x_\mu}-\mathcal{G}^n{}_m\frac{\partial S^{n*}}{\partial x^\mu}\frac{\partial S^m}{\partial x_\mu} . \tag{4.12}$$

Hence Eq. (4.10) is positive and therefore, without further fine-tuning, may be expected to have a stationary point with $V=0$, specified by the conditions

$$\frac{\partial W}{\partial C^a}=D_n W=0 . \tag{4.13}$$

But this is not necessarily a supersymmetric configuration, because here

$$D_a W\equiv\frac{\partial W}{\partial C^a}+8\pi G\frac{\partial K}{\partial C^a}W =-\frac{3}{|T+T^*+h|}\frac{\partial h}{\partial C^a}W , \tag{4.14}$$

and this does not necessarily vanish. (However, to have supersymmetry broken, it is essential that the superpotential actually depend on all of the chiral scalars S^n, because otherwise the conditions $D_n W=0$ would require $W=0$ and hence $D_a W=0$.)

The superpotential W depends on C^a and S^n, but not on T, so the conditions (4.13) will generally fix the values of C^a and S^n at the minimum of V, while leaving T undetermined. The field T enters the potential only in the overall scale of the part that depends on the C^a, so such theories are called "no-scale" models. An intensive phenomenological study of these models was carried out at CERN for several years following 1983 (Ellis, Lahanas, *et al.*, 1984; Ellis, Kounnas, *et al.*, 1984; Barbieri *et al.*, 1985).

Of course, these models do not solve the cosmological constant problem, because neither Eq. (4.8) nor Eq. (4.9) is dictated by any known physical principle. In particular, in order to cancel the second term in Eq. (4.5), it is essential that the coefficient of the logarithm in the first term in (4.8) be given the apparently arbitrary value $-3/8\pi G$.

It was therefore exciting when, in some of the first work on the physical implications of superstring theory, it was found that compactification of six of the ten original dimensions yielded a four-dimensional supergravity theory with Kahler potential and superpotential of the form (4.8) and (4.9). Specifically, Witten (1985) found a Kahler potential of the form (4.8), with h quadratic in the C's and $\tilde{K}=-\ln(S+S^*)/8\pi G$, but with a superpotential that depended solely on the C's. By including nonperturbative gaugino condensation effects, Dine *et al.* (1985) were able to give the superpotential a dependence on S (though they did not treat the dependence of the Kahler potential or superpotential on the C^a fields). In this work, the S field is a complex function (now often called Y) of four-dimensional dilaton and axion fields, while the T field represents the scale of the compactified six-dimensional manifold. The factor 3 in Eq. (4.8) arises in these models because one compactifies on a complex manifold with $(10-4)/2=3$ complex dimensions (Chang *et al.*, 1988).

Intriguing as these results are, they have not been taken seriously (even by the original authors) as a solution of the cosmological constant problem. The trouble is that no one expects the simple structures (4.8) and (4.9) to survive beyond the lowest order of perturbation theory, because they are not protected by any symmetry that survives down to accessible energies.

Recently Moore (1987a, 1987b) has attempted a more specifically "stringy" attack on the problem. Early work by Rohm (1984) and Polchinski (1986) had shown that in the calculation of the vacuum energy density, the sum over zero-point energies can be converted into an integral over a complex "modular parameter" τ. (In string theories, two-dimensional conformal symmetry makes the tree-level vacuum energy vanish.) Last year Moore pointed out that for some special compactifications there is a discrete symmetry of modular space, known as Atkin-Lehner symmetry, that makes the integral over τ vanish despite the absence of space-time supersymmetry. So far, the only examples where this occurs entail a compactification to two rather than four space-time dimensions, but it does not seem unlikely that four-dimensional examples could be found. A more serious obstacle is that the Atkin-Lehner symmetry seems irretrievably tied to one-loop order.

Indeed, it is very hard to see how any property of supergravity or superstring theory could make the effective cosmological constant sufficiently small. It is not enough that the vacuum energy density cancel in lowest order, or to all finite orders of perturbative theory; even nonperturbative effects like ordinary QCD instantons would give far too large a contribution to the effective cosmological constant if not canceled by something else. According to our modern theories, properties of elementary particles, like approximate baryon and lepton conservation, are dictated by gauge symmetries of the standard model, which survive down to accessible energies. We know of no such symmetry (aside from the unrealistic example of unbroken supersymmetry) that could keep the effective cosmological constant sufficiently small. It is conceivable that in supergravity the property of having zero effective cosmological constant does survive to low energies without any symmetry to guard it, but this would run counter to all our experience in physics.

V. ANTHROPIC CONSIDERATIONS

I now turn to a very different approach to the cosmological constant, based on what Carter (1974) has called the anthropic principle.[5] Briefly stated, the anthropic principle has it that the world is the way it is, at least in part, because otherwise there would be no one to ask why it is the way it is. There are a number of different versions of this principle, ranging from those that are so weak as to be trivial to those that are so strong as to be absurd. Three of these versions seem worth distinguishing here.

(i) In one very weak version, the anthropic principle amounts simply to the use of the fact that we are here as one more experimental datum. For instance, recall M. Goldhaber's joke that "we know in our bones" that the lifetime of the proton must be greater than about 10^{16} yr, because otherwise we would not survive the ionizing particles produced by proton decay in our own bodies. No one can argue with this version, but it does not help us to explain anything, such as *why* the proton lives so long. Nor does it give very useful experimental information; certainly experimental physicists (including Goldhaber) have provided us with better limits on the proton lifetime.

[5]Recent discussions of the anthropic principle are given in the books by Davies (1982) and Barrow and Tipler (1986), and in articles by Carter (1983), Page (1987), and Rees (1987).

(ii) In one rather strong version, the anthropic principle states that the laws of nature, which are otherwise incomplete, are completed by the requirement that conditions must allow intelligent life to arise, the reason being that science (and quantum mechanics in particular) is meaningless without observers. I do not know how to reach a decision about such matters and will simply state my own view, that although science is clearly impossible without scientists, it is not clear that the universe is impossible without science.

(iii) A moderate version of the anthropic principle, sometimes known as the "weak anthropic principle," amounts to an explanation of which of the various possible eras or parts of the universe we inhabit, by calculating which eras or parts of the universe we *could* inhabit. An example is provided by what I think is the first use of anthropic arguments in modern physics, by Dicke (1961), in response to a problem posed by Dirac (1937). In effect, Dirac had noted that a combination of fundamental constants with the dimensions of a time turns out to be roughly of the order of the present age of the universe:

$$\hbar/Gcm_\pi^3 = 4.5\times10^{10}\ \mathrm{yr}\ . \tag{5.1}$$

[There are various other ways of writing this relation, such as replacing m_π with various combinations of particle masses and introducing powers of $e^2/\hbar c$. Dirac's original "large-number" coincidence is equivalent to using Eq. (5.1) as a formula for the age of the universe, with m_π replaced by $(137 m_p m_e^2)^{1/3} = 183$ MeV. In fact, there are so many different possibilities that one may doubt whether there is any coincidence that needs explaining.] Dirac reasoned that if this connection were a real one, then, since the age of the universe increases (linearly) with time, some of the constants on the left side of (5.1) must change with time. He guessed that it is G that changes, like $1/t$. [Zee (1985) has applied similar arguments to the cosmological constant itself.] In response to Dirac, Dicke pointed out that the question of the age of the universe could only arise when the conditions are right for the existence of life. Specifically, the universe must be old enough so that some stars will have completed their time on the main sequence to produce the heavy elements necessary for life, and it must be young enough so that some stars would still be providing energy through nuclear reactions. Both the upper and lower bounds on the ages of the universe at which life can exist turn out to be roughly (very roughly) given by just the quantity (5.1). Hence there is no need to suppose that any of the fundamental constants vary with time to account for the rough agreement of the quantity (5.1) with the present age of the universe.

It is this "weak anthropic principle" that will be applied here. Its relevance arises from the fact that, in some modern cosmological models, the universe does have parts or eras in which the effective cosmological constant takes a wide variety of values. Here are some examples.

(1) The vacuum energy may depend on a scalar field vacuum expectation value that changes slowly as the universe expands, as in a model of Banks (1985).

(2) In a model of Linde (1986, 1987, 1988b), fluctuations in scalar fields produce exponentially expanding regions of the universe, within which further fluctuations produce further subuniverses, and so on. Since these subuniverses arise from fluctuations in the fields, they have differing values of various "constants" of nature.

(3) The universe may go through a very large number of first-order phase transitions in which bubbles of smaller vacuum energy form; within these bubbles there form further bubbles of even smaller vacuum energy, and so on. This can happen if the potential for some scalar field has a large number of small bumps, as in a model of Abbott (1985). Alternatively, the bubble walls may be elementary membranes coupled to a 3-form gauge potential $A_{\mu\nu\lambda}$, as in the work of Brown and Teitelboim (1987a, 1987b).

(4) The universe may start in a quantum state in which the cosmological constant does not have a precise value. Any "measurement" of the properties of the universe yields a variety of possible values for the cosmological constant, with *a priori* probabilities determined by the initial state (Hawking, 1987a). We will see examples of this in Secs. VII and VIII.

In models of these types, it is perfectly sensible to apply anthropic considerations to decide which era or part of the universe we could inhabit, and hence which values of the cosmological constant we could observe.

A large cosmological constant would interfere with the appearance of life in different ways, depending on the sign of λ_{eff}. For a large *positive* λ_{eff}, the universe very early enters an exponentially expanding de Sitter phase, which then lasts forever. The exponential expansion interferes with the formation of gravitational condensations, but once a clump of matter becomes gravitationally bound, its subsequent evolution is unaffected by the cosmological constant. Now, we do not know what weird forms life may take, but it is hard to imagine that it could develop at all without gravitational condensations out of an initially smooth universe. Therefore the anthropic principle makes a rather crisp prediction: λ_{eff} must be small enough to allow the formation of sufficiently large gravitational condensations (Weinberg, 1987).

This has been worked out quantitatively, but we can easily understand the main result without detailed calculations. We know that in our universe gravitational condensation had already begun at a redshift $z_c \geq 4$. At this time, the energy density was greater than the present mass density ρ_{M_0} by a factor $(1+z_c)^3 \geq 125$. A cosmological constant has little effect as long as the nonvacuum energy density is larger than ρ_V, so one can conclude that a vacuum energy density ρ_V no larger than, say $100\rho_{M_0}$ would not be large enough to prevent gravitational condensations. [The quantitative analysis of Weinberg

(1987) shows that for $k=0$, a vacuum energy density no greater than $\pi^2(1+z_c)^3\rho_{M_0}/3$ would not prevent gravitational condensation at a redshift z_c; this is $410\rho_{M_0}$ for $z_c=4$.]

This result suggests strongly that if it is the anthropic principle that accounts for the smallness of the cosmological constant, then we would expect a vacuum energy density $\rho_V \sim (10-100)\rho_{M_0}$, because there is no anthropic reason for it to be any smaller.

Is such a large vacuum energy density observationally allowed? There are a number of different types of astronomical data that indicate differing answers to this question.

A. Mass density

If, as often assumed, the universe now has negligible spatial curvature, then

$$\Omega_V+\Omega_{M_0}=1 , \tag{5.2}$$

where Ω_V and Ω_{M_0} are the ratios of the vacuum energy density and the present mass density to the critical density

$$\Omega_V \equiv \frac{8\pi G\rho_V}{3H_0^2}=\frac{\lambda_{\text{eff}}}{3H_0^2} , \qquad \Omega_{M_0}=\frac{8\pi G\rho_{M_0}}{3H_0^2} . \tag{5.3}$$

The dynamics of clusters of galaxies seems to indicate that Ω_{M_0} is in the range 0.1–0.2 (Knapp and Kormendy, 1987), which with these assumptions would indicate a value for ρ_V/ρ_{M_0} in the range 4–9. If we discount the evidence from the dynamics of clusters of galaxies, then Ω_{M_0} could be as small as 0.02 (Knapp and Kormendy, 1987), corresponding to a value of $\rho_V/\rho_{M_0}\simeq 50$. [See also Bahcall *et al.* (1987).]

B. Ages

In a dust-dominated universe with $k=0$ and $\rho_V=0$, the age of the universe is $t_0=2/3H_0$. For $H_0=100$ km/sec Mpc, this is 7×10^9 yr, considerably less than the ages usually estimated for globular clusters (Renzini, 1986). On the other hand, for a dust-dominated universe with $k=0$ and $\rho_V\neq 0$, the present age of an object that formed at a redshift z_c is

$$t_0(z_c)=\frac{2}{3}\left[1+\frac{\rho_{M_0}}{\rho_V}\right]^{1/2} H_0^{-1}\left\{\sinh^{-1}\left[\frac{\rho_V}{\rho_{M_0}}\right]^{1/2}-\sinh^{-1}\left[\left[\frac{\rho_V}{\rho_{M_0}}\right]^{1/2}(1+z_c)^{-3/2}\right]\right\} . \tag{5.4}$$

For instance, for $z_c=4$ and $\rho_V/\rho_{M_0}=9$ (i.e., $\Omega_{M_0}=0.1$), this gives an age $1.1H_0^{-1}$ in place of $\frac{2}{3}H_0^{-1}$. This is not in conflict with globular cluster ages even for Hubble constants near 100 km/sec Mpc.

These considerations of cosmic age and density have led a number of astronomers to suggest a fairly large positive cosmological constant, with $\rho_V>\rho_{M_0}$ [de Vaucouleurs (1982, 1983); Peebles (1984, 1987a, 1987b); Turner, Steigman, and Krauss (1984)]. However, there recently has appeared a strong argument against this view, which we shall now consider.

C. Number counts

Loh and Spillar (1986) have carried out a survey of numbers of galaxies as a function of redshift, subsequently analyzed by Loh (1986). For a uniformly distributed class of objects that are all bright enough to be detectable at redshifts $\le z_{\max}$, the number of objects observed at redshift less than $z\le z_{\max}$ in a dust-dominated universe with $k=0$ is

$$N(<z)\propto \int_{(1+z)^{-3/2}}^{1} ds\, s^{4/3}(1+\rho_V s^2/\rho_{M_0})^{-1/2}\left[\int_0^1 ds' s'^{-2/3}(1+\rho_V s'^2/\rho_{M_0})^{-1/2}\right]^2 . \tag{5.5}$$

Of course, in the real world there are always some objects too dim to be seen. Loh's analysis allowed for an unknown luminosity distribution, assuming only that its shape does not evolve with time. Under these assumptions, he found that the vacuum energy must be quite small: specifically,

$$\rho_V/\rho_{M_0}=0.1^{-0.4}_{+0.2} .$$

This is more than 3 orders of magnitude below the anthropic upper bound discussed earlier. If the effective cosmological constant is really this small, then we would have to conclude that the anthropic principle does *not* explain why it is so small. [However, there are reasons to be cautious in reaching this conclusion. Bahcall and Tremaine (1988) have recently reanalyzed the data of Loh and Spillar, using a plausible model of galaxy evolution in which the shape of the luminosity distribution does change with time. They considered only the case $\rho_V=0$, leaving Ω_{M_0} undetermined, and found that evolution in this model could increase or decrease the inferred value of Ω_{M_0} by as much as unity. Presumably it would also have a similarly large effect on the inferred value of ρ_V/ρ_{M_0} when $\Omega_{M_0}+\Omega_V$ is constrained to be unity. In addition, the redshifts of Loh and Spillar are photometric and therefore less certain than those obtained from shifts of individual spectral lines.]

Now let us consider a cosmological constant of the other sign, $\lambda_{\rm eff}<0$. Here the cosmological constant does not interfere with the formation of gravitational condensations. Instead (for $k=0$ or $k=+1$), the whole universe collapses to a singularity in a finite time T. The anthropic constraint here is simply that the universe last long enough for the appearance of life (Barrow and Tipler, 1986), say, $T\gtrsim 0.5H_0^{-1}$, where H_0^{-1} is the Hubble time in *our* universe. For a dust-dominated universe with $k=0$, we have

$$T=\pi(8\pi G|\rho_V|)^{1/2}\ ,\qquad H_0^{-1}=(8\pi G\rho_{M_0}/3)^{1/2}\ , \tag{5.6}$$

so the anthropic constraint here is just

$$|\rho_V|\lesssim\rho_{M_0}\ . \tag{5.7}$$

In this case the anthropic principle can explain why the cosmological constant is as small as found by Loh (1986), but not much smaller. On the other hand, a negative cosmological constant would not help with the cosmic mass and age problems.

Before closing this section, let me take up one possibility that may confront us in a few years. Suppose it really is confirmed that, as suggested by cosmic ages and densities, there is a cosmological constant with ρ_V of order ρ_{M_0}. Would we then have any alternative to an anthropic explanation for this value of ρ_V? The mass density ρ_M changes with time, so without anthropic considerations it is very hard to explain why a constant ρ_V should equal the value that ρ_M happens to have at present. But perhaps ρ_V really is not constant. For instance, Peebles and Ratra (1988) and Ratra and Peebles (1988) have considered a model in which the vacuum energy depends on a scalar field that changes as the universe expands. In order to qualify as a vacuum energy, it is only necessary for ρ_V to be accompanied with a pressure $p_V=-\rho_V$; the value of ρ_V can change if the vacuum exchanges energy with matter and radiation. The conservation of energy then relates the change of ρ_V to the change in the densities of matter (with $p_M=0$) and radiation (with $p_R=\rho_R/3$):

$$\frac{d}{dt}\rho_V+R^{-3}\frac{d}{dt}(R^3\rho_M)+R^{-4}\frac{d}{dt}(R^4\rho_R)=0\ . \tag{5.8}$$

Freese *et al.* (1987) have considered the possibility that energy is exchanged only between the vacuum and matter, or the vacuum and radiation, in such a way that either ρ_V/ρ_M or ρ_V/ρ_R remain constant, respectively (see also Reuter and Wetterich, 1987). In order for the vacuum to transfer energy to ordinary matter in such a way that ρ_V/ρ_M remains fixed, and if baryon number is conserved, then it would be necessary to create baryon-antibaryon pairs at a sufficient rate to produce a troublesome γ-ray background. Alternatively, if the vacuum transfers energy to radiation in such a way that ρ_V/ρ_R remains constant, and if ρ_V is comparable with the present mass density ρ_{M_0}, then ρ_V/ρ_R must be rather large, completely changing the results of cosmological nucleosynthesis.

One more possibility that was not considered by Freese *et al.* is that the vacuum transfers energy to radiation, avoiding the problems of baryon-antibaryon annihilation, but in such a way as to keep a fixed ratio ρ_V/ρ_M rather than ρ_V/ρ_R. However, this also does not work. With $\rho_V=c\rho_M$ and $R^3\rho_M$ constant, Eq. (5.8) yields

$$\rho_R=\rho_{R_0}\left[\frac{R_0}{R}\right]^4+3c\rho_{M_0}\left[\left[\frac{R_0}{R}\right]^3-\left[\frac{R_0}{R}\right]^4\right] \underset{R\ll R_0}{\longrightarrow}(\rho_{R_0}-3c\rho_{M_0})\left[\frac{R_0}{R}\right]^4 . \tag{5.9}$$

So that there is no interference with calculations of cosmological nucleosynthesis, we need

$$\rho_R\approx\rho_{R_0}\left[\frac{R_0}{R}\right]^4 ,$$

and therefore

$$|\rho_V|/\rho_M\equiv|c|\lesssim\frac{\rho_{R_0}}{3\rho_{M_0}}\ll 1\ . \tag{5.10}$$

Thus, even if we are willing to suppose that the vacuum energy changes with time, a vacuum energy density comparable with the present mass density seems very difficult to explain on other than anthropic grounds.

VI. ADJUSTMENT MECHANISMS

I now turn to an idea that has been tried by virtually everyone who has worried about the cosmological constant [see, e.g., Dolgov (1982); Wilczek and Zee (1983); Wilczek (1984, 1985); Peccei, Solá, and Wetterich (1987); Barr and Hochberg (1988)]. Suppose there is some scalar ϕ whose source is proportional to the trace of the energy-momentum tensor

$$\Box^2\phi\propto T^\mu{}_\mu\propto R\ . \tag{6.1}$$

(Here $T^{\mu\nu}$ is the total energy-momentum tensor that includes a possible cosmological constant term $-\lambda g^{\mu\nu}/8\pi G$.) Suppose also that $T^\mu{}_\mu$ depends on ϕ and vanishes at some field value ϕ_0. Then ϕ will evolve until it reaches an equilibrium value ϕ_0, where $T^\mu{}_\mu=0$, and the

Einstein field equations have a flat-space solution.

Of course, we do not observe such a scalar field, but for these purposes it can couple as weakly as we like; a weak coupling simply implies that the equilibrium value ϕ_0 is very large. In this respect the scalar ϕ is analogous to the axion, especially in its later "invisible" version [Kim (1979); Dine, Fischler, and Srednicki (1981)].

Even very weakly coupled, it is possible that the ϕ field could have interesting effects, because it must have very small mass. If it has any nonzero mass M_ϕ, then at energies below m_ϕ we can work with an effective Lagrangian in which ϕ has been "integrated out," and so does not appear explicitly. But massless fields like the gravitational and electromagnetic field will still appear in this effective Lagrangian, and their vacuum fluctuations will contribute to the effective cosmological constant. In order to keep $\rho_V < 10^{-48}$ GeV^4, we need the scalar field adjustment to cancel the effect of gravitational and electromagnetic field fluctuations down to frequencies 10^{-12} GeV; for this purpose we must have $m_\phi < 10^{-12}$ GeV. A field this light will have a macroscopic range: $\hbar/m_\phi c \gtrsim 0.01$ cm.

Unfortunately it seems to be impossible to construct a theory with one or more scalar fields having the assumed properties. This can be seen in very general terms. What we want is to find an equilibrium solution of the field equations in which $g_{\mu\nu}$ and all matter fields ψ_n (perhaps tensors as well as scalars) are constant in space-time. For such constant fields the Euler-Lagrange equations are simply

$$\frac{\partial \mathcal{L}}{\partial g_{\mu\nu}}=0 , \tag{6.2}$$

$$\frac{\partial \mathcal{L}}{\partial \psi_n}=0 . \tag{6.3}$$

As we saw in Sec. III, the problem is in satisfying the trace of the gravitational field equation. To make a solution natural, we would like this trace to be a linear combination of the ψ_n field equations; that is, we want

$$g_{\lambda\nu}\frac{\partial \mathcal{L}(g,\psi)}{\partial g_{\lambda\nu}}=\sum_n \frac{\partial \mathcal{L}(g,\psi)}{\partial \psi_n}f_n(\psi) \tag{6.4}$$

for all constant $g_{\mu\nu}$ and ψ_n. This can be restated as a symmetry condition: for constant fields the Lagrangian must be invariant under the transformation

$$\delta g_{\lambda\nu}=2\varepsilon g_{\lambda\nu}, \quad \delta\psi_n=-\varepsilon f_n(\psi) . \tag{6.5}$$

With this condition, if we find a solution $\psi^{(0)}$ of the Euler-Lagrange equations for ψ_n,

$$\frac{\partial \mathcal{L}}{\partial \psi_n}=0 \quad \text{at } \psi_n=\psi_n^{(0)} , \tag{6.6}$$

then the trace of the field equation for $g_{\mu\nu}$ is automatically satisfied.

The problem is that under these assumptions, it is impossible (without fine-tuning $\mathcal{L}$) to find a solution to the field equations (6.3) for the ψ_n. To see this, we replace the N fields ψ_n with $N-1$ fields σ_a (not necessarily scalars) and one scalar ϕ, in such a way that the symmetry transformation (6.5) takes the form

$$\delta g_{\lambda\nu}=2\varepsilon g_{\lambda\nu}, \quad \delta\sigma_a=0, \quad \delta\phi=-\varepsilon . \tag{6.7}$$

[To do this, we first define a "transverse" surface S in field space by an equation $T(\psi)=0$, where $T(\psi)$ is any function on which $\sum_n(\partial T/\partial\psi_n)f_n(\psi)$ does not vanish. We take σ_a as any set of coordinates on this $(N-1)$-dimensional surface, and define $\psi_n(\sigma;\phi)$ as the solution of the ordinary differential equation $d\psi_n/d\phi=f_n(\psi)$ subject to the condition that at $\phi=0$, ψ_n is at the point on S with coordinates σ. The condition that S be a transverse surface ensures that, at least within a finite region of field space, any point ψ_n is on just one of these trajectories.] This symmetry simply ensures that for constant fields the Lagrangian can depend on $g_{\lambda\nu}$ and ϕ only in the combination $e^{2\phi}g_{\lambda\nu}$. The general arguments of Sec. III then show that when the field equations for σ are satisfied, the Lagrangian must take the form

$$\mathcal{L}=e^{4\phi}(\mathrm{Det}g)^{1/2}\mathcal{L}_0(\sigma) . \tag{6.8}$$

We see that the source of ϕ is the trace of the energy-momentum tensor

$$\frac{\partial \mathcal{L}}{\partial \phi}=T^{\mu}{}_{\mu}(\mathrm{Det}g)^{1/2} , \tag{6.9}$$

$$T^{\mu\nu}=g^{\mu\nu}e^{4\phi}\mathcal{L}_0(\sigma) . \tag{6.10}$$

It is true that if there were a value of ϕ where $\mathcal{L}$ is stationary in ϕ, then the trace of the Einstein field equations would automatically be satisfied at this point, but clearly there is no such stationary field value (unless, of course, we fine-tune $\mathcal{L}_0$ so that it vanishes at its stationary point). To put this another way, since $\mathcal{L}$ depends only on ϕ and $g_{\mu\nu}$ only in the combination $\hat{g}_{\mu\nu}\equiv e^{2\phi}g_{\mu\nu}$ (and derivatives of ϕ and $g_{\mu\nu}$), we might as well redefine the metric as $\hat{g}_{\mu\nu}$ instead of $g_{\mu\nu}$. Then ϕ is just a scalar with only derivative couplings and clearly cannot help with our problem.[6]

As one example of many failed attempts along this line, let us consider a proposal of Peccei, Solà, and Wetterich (1987). They observed that the symmetry (6.5) or (6.7) may be broken by conformal anomalies, such as those that produce the β function of quantum chromodynamics, in such a way that the effective Lagrangian becomes[7]

$$\mathcal{L}_{\mathrm{eff}}=(\mathrm{Det}g)^{1/2}[e^{4\phi}\mathcal{L}_0(\sigma)+\phi\Theta^{\mu}{}_{\mu}] , \tag{6.11}$$

where $\Theta^{\mu}{}_{\mu}$ represents the effect of the conformal anoma-

[6]This remark is due to Polchinski (1987).

[7]An equation essentially equivalent to (6.11) appeared in the preprint version of the paper by Peccei, Solá, and Wetterich (1987). In the published version this equation was removed, and it was acknowledged that fine-tuning is still needed to make the cosmological constant vanish. However, this equation was quoted in the meantime in a paper by Ellis, Tsamis, and Voloshin (1987), which mostly deals with the observable consequences of the light scalar particle in this model.

ly. The source of the ϕ field is now

$$\frac{\partial\mathcal{L}}{\partial\phi}=(T^{\mu}{}_{\mu}+\Theta^{\mu}{}_{\mu})(\mathrm{Det}g)^{1/2}\ , \tag{6.12}$$

with $T^{\mu\nu}$ the previous energy-momentum tensor (6.10). Now we can find an equilibrium solution for the ϕ field, at a value ϕ_0 such that

$$4e^{4\phi_0}\mathcal{L}_0+\Theta^{\mu}{}_{\mu}=0\ . \tag{6.13}$$

The trouble is that this is *not* the condition for a flat-space solution; the Einstein equation for a constant metric is

$$0=\frac{\partial\mathcal{L}_{\text{eff}}}{\partial g_{\mu\nu}}\propto e^{4\phi}\mathcal{L}_0+\phi\Theta^{\mu}{}_{\mu}\ , \tag{6.14}$$

which is not the same as (6.13). The point is that just calling the anomalous term in (6.11) $\Theta^{\mu}{}_{\mu}$ does not make it a term in the trace of the energy-momentum tensor to which $g_{\mu\nu}$ is coupled. This result is not surprising, since (6.11) does not obey the symmetry (6.7). One cannot have it both ways: either we preserve the symmetry, in which case there is no equilibrium solution for ϕ, or we break the symmetry, in which case such an equilibrium solution does not imply a solution of the field equations for a constant metric. (Also see Coughlan *et al.*, 1988; Wetterich, 1988.)

In a slightly different version of this general class of models, we can try coupling a scalar field so that it is the curvature scalar R rather than the trace of the energy-momentum tensor that directly serves as the source of the scalar field. [See, e.g., Dolgov (1982); Barr (1987); Ford (1987).] For instance, we might take the Lagrangian as

$$\mathcal{L}=\sqrt{g}\left[-\tfrac{1}{2}\partial_{\mu}\phi\partial^{\mu}\phi-\frac{\lambda}{8\pi G}-\frac{1}{16\pi G}R-U(\phi)R\right]\ . \tag{6.15}$$

This has a flat-space solution with $g_{\mu\nu}=\eta_{\mu\nu}$ and $\phi=\phi_0$ (a constant), provided

$$U(\phi_0)=\infty\ . \tag{6.16}$$

However, as the above authors observed, the effective gravitational coupling in this theory is given by

$$G_{\text{eff}}=\frac{G}{1+16\pi GU(\phi_0)}=0\ . \tag{6.17}$$

This is not much progress; we always knew that a nonzero vacuum energy does not prevent a flat-space solution if the gravitational constant is zero.

The "no-go" theorem proved in this section should not be regarded as closing off all hope in this direction. No-go theorems have a way of relying on apparently technical assumptions[8] that later turn out to have exceptions of great physical interest. (A famous example is the Coleman-Mandula theorem.) More discouraging than any theorem is the fact that many theorists have tried to invent adjustment mechanisms to cancel the cosmological constant, but without any success so far.

VII. CHANGING GRAVITY

A number of authors have suggested changing the rules of classical general relativity in such a way that the cosmological constant appears as a constant of integration, unrelated to any parameters in the action [Van der Bij *et al.* (1982); Weinberg (1983); Wilczek and Zee (1983); Buchmüller and Dragon (1988a, 1988b)]. This does not solve the cosmological constant problem, but it does change it in a suggestive way.

I will describe one version of this idea, in which one maintains general covariance, but reinterprets the formalism so that the determinant of the metric is not a dynamical field. *Any* theory can be written in a way that is formally generally covariant, so by the usual arguments we can take the action for gravity and matter as

$$I[\psi,g]=\frac{-1}{16\pi G}\int d^4x\sqrt{g}R+I_M[\psi,g]\ , \tag{7.1}$$

where ψ are a set of matter fields appearing in the matter action I_M. (I_M includes a possible cosmological constant term $-\lambda\int\sqrt{g}\,d^4x/8\pi G$.) The variational derivative of Eq. (7.1) with respect to the metric is

$$\frac{\delta I}{\delta g_{\mu\nu}}=\frac{1}{8\pi G}(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R)+T^{\mu\nu}\ , \tag{7.2}$$

where, as usual, $T^{\mu\nu}$ is the variational derivative of I_M with respect to $g_{\mu\nu}$. In ordinary general relativity all components of the metric are dynamical fields, so Eq. (7.2) vanishes for all μ,ν, yielding the usual Einstein field equations. However, just because we use a generally covariant formalism does not mean that we are committed to treating all components of the metric as dynamical fields. For instance, we all learn in childhood how to write the equations of Newtonian mechanics in general curvilinear spatial coordinate systems, without supposing that the 3-metric has to obey any field equations at all.

In particular, if the determinant g is not dynamical, then the action only has to be stationary with respect to variations in the metric that keep the determinant fixed,

[8]For instance, we assumed that in the solution for flat space all fields are constant, but it might be that this solution preserves only some combination of translation and gauge invariance, in which case some gauge-noninvariant fields might vary with space-time position. (This is the case for the 3-form gauge field model discussed at the end of Sec. VII and in Sec. VIII.) Furthermore, it is possible that the foliation of field space, which allows us to replace the ψ_n with σ_a and ϕ, does not work throughout the whole of field space.

i.e., for which $g^{\mu\nu}\delta g_{\mu\nu}=0$; hence only the traceless part of (7.2) needs to vanish, yielding the field equation

$$R^{\mu\nu}-\tfrac{1}{4}g^{\mu\nu}R=-8\pi G(T^{\mu\nu}-\tfrac{1}{4}g^{\mu\nu}T^{\lambda}{}_{\lambda}) . \quad (7.3)$$

This is just the traceless part of the Einstein field equations; these equations evidently contain less information than Einstein's, but as we shall see, not much less. Because the whole formalism is generally covariant, the energy-momentum tensor satisfies the usual conservation law

$$T^{\mu\nu}{}_{;\mu}=0 , \quad (7.4)$$

and of course the Bianchi identities still hold,

$$(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R)_{;\mu}=0 . \quad (7.5)$$

The full Einstein field equations are automatically consistent with (7.4) and (7.5), but for the traceless part we get a nontrivial consistency condition. Taking the covariant derivative of Eq. (7.3) with respect to x^{μ} yields

$$\tfrac{1}{4}\partial_{\mu}R=8\pi G\tfrac{1}{4}\partial_{\mu}T^{\lambda}{}_{\lambda} ,$$

or, in other words, $R-8\pi GT^{\lambda}{}_{\lambda}$ is a constant, which we will call -4Λ:

$$R-8\pi GT^{\lambda}{}_{\lambda}=-4\Lambda \quad (\text{constant}) . \quad (7.6)$$

From (7.3) and (7.6), we obtain

$$R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R-\Lambda g^{\mu\nu}=-8\pi GT^{\mu\nu} . \quad (7.7)$$

Thus we recover the Einstein field equations, but with a cosmological constant that has nothing to do with any terms in the action or vacuum fluctuations, arising, instead, as a mere integration constant. To put this another way, Eq. (7.3) does not involve a cosmological constant; the contribution of vacuum fluctuations automatically cancel on the right-hand side of Eq. (7.3), so this equation does have flat-space solutions in the absence of matter and radiation. The remaining problem in this formulation is: why should we choose the flat-space solutions?

Before proceeding with this theory, I should pause to mention that it is closely related[9] to a proposal made long ago by Einstein (1919). After his formulation of general relativity and its application to cosmology, Einstein turned to the old problem of a field theory of matter. In a paper titled "Do Gravitational Fields Play an Essential Part in the Structure of the Elementary Particles of Matter?" he proposed to replace the original gravitational field equation with the equation

$$R_{\mu\nu}-\tfrac{1}{4}g_{\mu\nu}R=-8\pi Gt_{\mu\nu} . \quad (7.8)$$

This is consistent only if $t_{\mu\nu}$ is traceless; however, Einstein took for $t_{\mu\nu}$ not the full energy-momentum tensor of matter and radiation, but just the traceless tensor of radiation alone. This is, of course, conserved only outside matter. In such regions there is no difference between Eqs. (7.8) and (7.3), so by the same calculation as shown here, Einstein was able to recover Eq. (7.7), with Λ a constant of integration. However, inside matter, Eq. (7.8) is different from (7.3), the difference being that the right-hand side of Eq. (7.3) includes the traceless part of the energy-momentum tensor of matter. A consequence of this difference is that in charged matter R is an undetermined function, except that it is constant along world lines.

I will also take the opportunity of this pause to comment on the connection between the formulation described here and that of Zee (1985) and Buchmüller and Dragon (1988a, 1988b). These authors take as their starting point the assumption that the action is invariant not under the group of all coordinate transformations, but only under the subgroup of transformations $x^{\mu}\rightarrow x'^{\mu}$ with $\text{Det}(\partial x'^{\mu}/\partial x^{\nu})=1$. This is not really in conflict with the formulation presented here; the general covariance of Eq. (7.1) is achieved at the cost of introducing a metric that is partly nondynamical (just as we can make Newtonian mechanics formally Lorentz invariant by introducing a nondynamical quantity, the velocity of the reference frame). However, in giving up general covariance, one may be led to a theory with unnecessary elements. Under transformations with $\text{Det}(\partial x'/\partial x)=1$, the determinant of the metric g behaves just like any scalar field, so one can introduce arbitrary functions of g here and there in the action. There is nothing wrong with this, but it is not necessary, no different from inserting a new scalar field into the theory.

Now let us return to the theory described by the field equations (7.3). In my view, the key question in deciding whether this is a plausible classical theory of gravitation is whether it can be obtained as the classical limit of any physically satisfactory quantum theory of gravitation. To help in answering this, and also to illuminate the points raised in the previous paragraph, let us look at a simple model (Teitelboim, 1982) that shares several features with the theory of gravitation studied here.

Consider a free relativistic particle, with space-time trajectory $x^{\mu}(s)$ parametrized by a variable s. In order for the action to be invariant under arbitrary reparametrizations $s\rightarrow s'(s)$, we must introduce an "einbein" $g(s)$, with transformation rule

$$g(s)\rightarrow g'(s')=g(s)\left[\frac{ds'}{ds}\right]^{-1} . \quad (7.9)$$

The action may then be taken as

$$I[x,g]=\tfrac{1}{2}\int ds\, g^{-1}(s)\frac{dx^{\mu}(s)}{ds}\frac{dx_{\mu}(s)}{ds} -\frac{m^2}{2}\int ds\, g(s) . \quad (7.10)$$

[9]This was pointed out to me by someone in the audience of the lectures at Harvard. I thank my informant for this interesting historical reference.

The conditions that I be stationary with respect to variations in $x^\mu(s)$ and $g(s)$ are, respectively,

$$\frac{dp^\mu}{ds}=0 , \qquad (7.11)$$

$$p^\mu p_\mu=-m^2 , \qquad (7.12)$$

where p_μ is the canonical conjugate to x^μ:

$$p_\mu(s)=g^{-1}(s)\frac{dx_\mu(s)}{ds} . \qquad (7.13)$$

However, just because we choose to write the action in a reparametrization-invariant way does not necessarily mean that we must treat the einbein $g(s)$ as a dynamical quantity. If we treat $x^\mu(s)$, but not $g(s)$, as dynamical variables, then we obtain Eq. (7.11), but not (7.12). Of course, Eq. (7.11) implies that $p_\mu p^\mu$ is a constant [just as Eq. (7.3) implies that $R-8\pi G T^\lambda{}_\lambda$ is constant]. If we like, we can call this constant $-m^2$, but this is now a mere integration constant, unrelated to anything in the original action.

Now to quantization. The Hamiltonian here is

$$H(s)=p_\mu\frac{dx^\mu}{ds}-L=\tfrac{1}{2}g(p^\mu p_\mu+m^2) , \qquad (7.14)$$

so in quantum mechanics we calculate amplitudes by the functional integral

$$A=\int[dx][dp][dg]\times\exp\left[i\int ds\left[p_\mu(s)\frac{dx^\mu(s)}{ds}-H(s)\right]\right] . \qquad (7.15)$$

The einbein $g(s)$ has no canonical conjugate, and so appears here only as a Lagrange multiplier, whose integral yields a factor

$$\prod_s\delta(p^\mu p_\mu+m^2) . \qquad (7.16)$$

Presumably the classical theory in which g is not dynamical would be obtained as the classical limit of a quantum theory in which we do not do a functional integral over $g(s)$, and hence do not get the factor (7.16). But then there would be nothing to keep p^μ timelike. This is such a trivial theory that it is hard to say that anything goes wrong physically; but we may anticipate that in less trivial theories, we need a field to serve as a Lagrange multiplier for every negative norm degree of freedom like p^0. This is the case, for instance, in string theories, where the integration over the world-sheet metric is needed to enforce the Virasoro conditions on physical states.

The quantum theory of gravitation can be put in similar terms. Using the Arnowitt-Deser-Misner (1962) formalism, we calculate amplitudes as functional integrals,

$$Z=\int[dh_{ij}][d\pi^{ij}][d\tilde{N}][dN^i]\times\exp\left[i\int\left[\pi^{ij}\frac{\partial h_{ij}}{\partial t}-(\tilde{\mathcal{H}}-2\lambda)\tilde{N}-\mathcal{H}_iN^i\right]d^4x\right] . \qquad (7.17)$$

Here h_{ij}, $\tilde{N}$, and N^i parametrize the 4-metric, with line element given by

$$d\tau^2=(h^{-1}\tilde{N}^2-N^iN^jh_{ij})dt^2-2h_{ij}N^idx^jdt-h_{ij}dx^idx^j , \qquad (7.18)$$

$$h\equiv\mathrm{Det}(h_{ij}) . \qquad (7.19)$$

Furthermore, π^{ij} is the canonical conjugate to h_{ij}, and $\tilde{\mathcal{H}}$ and $\mathcal{H}_i$ are functions of h_{ij} and π^{ij} and their space derivatives, given by

$$\tilde{\mathcal{H}}=\tfrac{1}{2}\mathcal{G}_{ij,kl}\pi^{ij}\pi^{kl}-{}^{(3)}R , \qquad (7.20)$$

$$\mathcal{H}_i=-2h_{ij}\nabla_k\pi^{jk} , \qquad (7.21)$$

where ${}^{(3)}R$ is the scalar curvature and ∇_k is the covariant derivative, both calculated using the 3-metric h_{ij}, and

$$\mathcal{G}_{ij,kl}\equiv h_{ik}h_{jl}+h_{il}h_{jk}-h_{ij}h_{kl} . \qquad (7.22)$$

We see that $\tilde{N}$ and N^i just act as Lagrange multipliers for $\tilde{\mathcal{H}}$ and $\mathcal{H}_i$, respectively. Moreover, from (7.18), we see that $\tilde{N}^2$ is just the quantity whose status is under question here, the determinant of the 4-metric[10]

$$\tilde{N}=(\mathrm{Det}g_{\mu\nu})^{1/2} . \qquad (7.23)$$

Thus, just as the integral over the einbein $g(s)$ enforced the constraint $p^\mu p_\mu=-m^2$, the integral over Detg enforces the constraint

$$\tilde{\mathcal{H}}=2\lambda . \qquad (7.24)$$

The two conditions are quite similar. Just as $\eta_{\mu\nu}$ has signature $+++-$, the quantity (7.22), viewed as a 6×6 matrix, has signature $+,+,+,+,+,-$. Hence the integration over $\mathrm{Det}g_{\mu\nu}$ has the effect of eliminating one negative norm degree of freedom for each $\mathbf{x}$, $\pi^{ij}\propto(h^{-1})^{ij}$, just as the integral over the einbein $g(s)$ allows one to eliminate the variable p^0. However, for gravity there is a "potential" term in $\tilde{\mathcal{H}}$, proportional to the 3-curvature, and it is not entirely clear to me that it really is necessary to constrain $\tilde{\mathcal{H}}$ to take a fixed value. For the present, the question of whether it is necessary to integrate over $\mathrm{Det}g_{\mu\nu}$ must be left open. [Recent work by Henneaux and Teitelboim (1988) shows that there is a sensible generally covariant quantum version of the classical theory described by Eq. (7.3).]

Before closing this section, I should note that several authors have made a rather different suggestion, which also has the effect of converting the cosmological constant from a function of parameters in the action into a constant of the motion (Aurilia *et al.*, 1980; Witten, 1983; Henneaux and Teitelboim, 1984). They proposed adding to the action a term

[10]In order to obtain this result, I have defined $\tilde{\mathcal{H}}$ and $\tilde{N}$ differently from the usual $\mathcal{H}$ and N, by moving a factor $h^{1/2}$ from N to $\mathcal{H}$.

$$I_F=-\tfrac{1}{48}\int d^4x\sqrt{g}\,F_{\mu\nu\rho\sigma}F^{\mu\nu\rho\sigma}\ , \qquad (7.25)$$

where $F_{\mu\nu\rho\sigma}$ is the exterior derivative of a 3-form gauge field $A_{\nu\rho\sigma}$,

$$F_{\mu\nu\rho\sigma}=\partial_{[\mu}A_{\nu\rho\sigma]}\ , \qquad (7.26)$$

and $g\equiv -\mathrm{Det}g_{\mu\nu}$. Since $F^{\mu\nu\rho\sigma}$ is totally antisymmetric, it can be expressed as

$$F^{\mu\nu\rho\sigma}=c\,\varepsilon^{\mu\nu\rho\sigma}/\sqrt{g}\ , \qquad (7.27)$$

where $\varepsilon^{\mu\nu\rho\sigma}$ is the Levi-Cività tensor density, with $\varepsilon^{0123}\equiv 1$, and c is a scalar field. The field equation for A is

$$F^{\mu\nu\rho\sigma}{}_{;\mu}=0\ , \qquad (7.28)$$

so, using (7.27)

$$\frac{\partial c}{\partial x^{\mu}}=0\ . \qquad (7.29)$$

But the action (7.25) then takes the form

$$I_F=+\tfrac{1}{2}c^2\int d^4x\sqrt{g}\ . \qquad (7.30)$$

In other words, whatever else contributes to the cosmological constant, there is one term that depends on the integration constant c,

$$\lambda_F=4\pi Gc^2\ . \qquad (7.31)$$

Again, this does not solve the cosmological constant problem, but it does change the way it arises.

If λ is a constant of integration, then in a quantum theory we expect the state vector of the universe to be a superposition of states with different values of λ, in which case the anthropic considerations of Sec. V would set a bound on th effective cosmological constant.

VIII. QUANTUM COSMOLOGY

The last approach to the cosmological constant problem that I shall describe here is based on the application of quantum mechanics to the whole universe. In 1984 Hawking (1984b) described how in quantum cosmology there could arise a distribution of values for the effective cosmological constant, with an enormous peak at $\lambda_{\text{eff}}=0$. Very recently, this approach has been revived in an exciting paper by Coleman (1988b), using a new mechanism for producing a distribution of values for the cosmological constant (that rests in part on other work of Hawking and Coleman) and finding an even sharper peak. Related ideas have also been recently discussed by Banks (1988). Before describing the work of Coleman and Hawking, I will have to say something about quantum cosmology in general.

Most treatments of quantum cosmology are based on the "wave function of the universe," a function $\Psi[h,\phi]$ of the 3-metric and matter fields on a spacelike surface. [The 3-metric h_{ij} can be conveniently defined by adapting the space-time coordinate system so that the spacelike surface has constant t, and then decomposing the 4-metric $g_{\mu\nu}$ as in Eq. (7.19).] This wave function satisfies a sort of Schrödinger equation, known as the Wheeler-DeWitt equation [DeWitt (1967); Wheeler (1968)]:

$$\left[\frac{1}{2h^{1/2}}\frac{\delta}{\delta h_{ij}}h^{1/2}g_{ij,kl}\frac{\delta}{\delta h_{kl}}-{}^{(3)}R-2\lambda \right.$$
$$\left. +8\pi GT_{00}\right]\Psi=0\ , \qquad (8.1)$$

with notation explained in Sec. VII (except that we now include a matter energy density T_{00}, in which the canonical conjugate of a matter field Φ is replaced with $\delta/\delta\Phi$). It will be very important in what follows that we express the solution as a *Euclidean* path integral

$$\Psi\propto\int[dg][d\Phi]\exp(-S[g,\Phi])\ , \qquad (8.2)$$

where we integrate over all Euclidean-signature 4-metrics $g_{\mu\nu}$ and matter fields Φ defined on a 4-manifold M_4, that have the 3-manifold $M_3[h,\phi]$ with 3-metric h_{ij} and matter fields ϕ as a boundary. (The Wheeler-DeWitt equation is the constraint obtained from integrating the Lagrange multiplier $\tilde{N}$ as discussed in Sec. VII.) Here S is the Euclidean action[11]

$$S=\frac{1}{16\pi G}\int_{M_4}\sqrt{g}\,(R+2\lambda)$$
$$+\text{matter terms}+\text{surface terms}\ . \qquad (8.3)$$

Since Eq. (8.1) is a differential equation in an infinite-dimensional space [the set of $h_{ij}(\mathbf{x})$ and $\phi(\mathbf{x})$ for all $\mathbf{x}$], it has an infinite variety of solutions, which can be specified by giving the 4-manifold in Eq. (8.2) other boundaries, besides the $M_3[h,\phi]$ on which the 3-metric and matter fields are specified. Hartle and Hawking (1983) proposed as a cosmological initial condition that the manifold M_4 should have no boundaries other than $M_3(h,\phi)$. We will see that Coleman's (1988b) approach does not depend critically on the choice of initial conditions.

There are technical problems associated with this formalism. One is an operator-ordering ambiguity: there are various ways of ordering[12] the h_{ij} fields and $\delta/\delta h_{ij}$ operators in (8.1), all of which have (8.2) as solution, but

[11]The Euclidean action S is opposite in sign to what we would get if we replaced the metric $g_{\mu\nu}$ in the action I in Eq. (7.1) with one of signature $+,+,+,+$. This sign of S is chosen so that ordinary matter makes a positive contribution to S.

[12]The insertion of factors $h^{-1/2}$ and $h^{1/2}$ in Eq. (8.1) represents one choice of operator ordering, which is made in order to allow the derivation of the conservation equation (8.8).

with different ways of calculating the measure $[dg][d\Phi]$ (Hawking and Page, 1986). Another problem, potentially more worrisome, is that for gravity the Euclidean action (8.3) is not bounded below. Gibbons, Hawking, and Perry (1978) have proposed rotating the contour of integration for the overall scale of the 4-metric so that it runs parallel to the imaginary axis. We will not need to go into these technicalities here, because it will turn out that we only need to deal with the effective action at its equilibrium point.

A problem that is more relevant to us here has to do with the probabilistic interpretation of the wave function Ψ and of Euclidean path integrals like (8.2). Hawking has proposed (1984a, 1984c) that $\exp(-S[g,\Phi])$ should be regarded as proportional to the probability of a particular metric and matter field history. It is not immediately clear what is meant by this—even supposing that we had the godlike ability to measure the gravitational and matter fields throughout space-time, it would be in a space-time of Lorentzian rather than Euclidean signature. However, since we can (sometimes) go from one signature to another by a complex coordinate transformation, it may be that a Euclidean history $g_{\mu\nu}(x)$, $\Phi(x)$ can be interpreted in terms of correlations of scalar quantities, just as if the space-time were Lorentzian. In much of Hawking's work (e.g., Hawking, 1979), these questions are avoided by using the formalism only to calculate the probability that, in the space-time history of the universe, there is a spacelike 3-surface with a given 3-metric $h_{ij}(\mathbf{x})$ and matter fields $\phi(\mathbf{x})$. For instance, with Hartle-Hawking (1983) initial conditions, we would integrate over all closed 4-manifolds that contain such a 3-surface. If this surface bisects the 4-manifold, then it can be regarded as the boundary of the two halves of the 4-manifold, and so the integral is (with some qualifications) just the square of the wave function (8.2). But questions still arise concerning the probabilistic interpretation of Ψ, particularly with regard to normalization. If $|\Psi[h,\phi]|^2$ is the probability density that there exists *some* 3-surface on which the 3-metric is $h_{ij}(\mathbf{x})$ and the matter fields are $\phi(\mathbf{x})$, then we would not simply want to set the functional integral of $|\Psi[h,\phi]|^2$ over $h_{ij}(\mathbf{x})$ and $\phi(\mathbf{x})$ equal to unity, because in this functional integral we are summing up possibilities that are not exclusive; if the universe has some $h_{ij}(\mathbf{x})$ and $\phi(\mathbf{x})$ on one 3-surface, then it may also have some other $h'_{ij}(\mathbf{x})$ and $\phi'(\mathbf{x})$ on some other 3-surface. After all, you would not expect that the probabilities that you ever in your life have flipped a coin and gotten heads, and that you ever in your life have flipped a coin and gotten tails, should add up to unity.

I would like to offer an interpretation of what is meant by treating $|\Psi[h,\Phi]|^2$ as a probability density, which seems to me implicit in Hawking's writings (and may already be stated explicitly somewhere in the literature). As everyone has recognized, the problem has to do with the role of time in quantum gravity. [See, e.g., Hartle (1987).] The problems raised here do not arise in asymptotically flat cosmologies, because in such theories there is a natural definition of time, and we generally ask for the probabilities that the fields have certain values at a definite time. However, here time is a coordinate with no objective significance, and this coordinate time is even imaginary. As Augustine (398) warned, "I must not allow my mind to insist that time is something objective."[13] Heeding this warning, suppose we choose some "time-keeping" field $a(\mathbf{x},t)$, for instance, the trace of the energy-momentum tensor, and use its value to define a local time α. Each value of α defines a 3-surface, on which the coordinate time t is a function $t(\mathbf{x},\alpha)$ defined implicitly by

$$a(\mathbf{x},t(\mathbf{x},\alpha))\equiv\alpha \ . \tag{8.4}$$

We are then interested in the probability that the tangential components of the metric and all matter fields other than $a(\mathbf{x},t)$ have specified values on this surface. Calling these quantities $b_n(\mathbf{x},t)$, we see that the probability density for the $b_n(\mathbf{x},t)$ to have the values $\beta_n(\mathbf{x})$ at local time α is

$$P_\alpha[\beta]=N\int_{M_4}[dg][d\Phi]\exp(-S[g,\Phi]) \times\prod_{n,\mathbf{x}}\delta(b_n(\mathbf{x},t(\mathbf{x},\alpha))-\beta_n(\mathbf{x})) \ , \tag{8.5}$$

with N a normalization factor, determined by the condition that the total probability of finding *any* value for the $b_n(\mathbf{x})$ at local time α should be unity:

$$1=\int P_\alpha[\beta][d\beta] =N\int_{M_4}[dg][d\Phi]\exp(-S[g,\Phi]) \ . \tag{8.6}$$

[This usually makes N a function of α, because in (8.5) and (8.6) we integrate only over matter and metric histories for which Eq. (8.4) is satisfied on some 3-surface. With some boundary conditions, this condition is automatically satisfied, and then N is α independent. For instance, if M_4 has two boundaries, on which $a(x)$ is required to take values α_1 and α_2, then there are 3-surfaces on which (8.4) is satisfied for all α in the range $\alpha_1<\alpha<\alpha_2$.] Where the surface of constant a bisects the 4-space, $P_\alpha[\beta]$ can be written as proportional to the square of the wave function $\Psi[\alpha,\beta]$, but with α constant in 3-space.

[13]This quote is not merely a display of useless erudition. Book XI of Augustine's *Confessions* contains a famous discussion of the nature of time, and it seems to have become a tradition to quote from this chapter in writing about quantum cosmology. Thus Hawking (1979) quotes "What did God do before He made Heaven and Earth? I do not answer as one did merrily: He was preparing a Hell for those that ask such questions. For at no time had God not made anything, for time itself was made by God." Coleman (1988a) quotes "The past is present memory." To this, I can add one more very relevant quote: "I confess to you, Lord, that I still do not know what time is. Yet I confess too that I do know that I am saying this in time, that I have been talking about time for a long time, ..."

Coleman (1988b) short-circuits many of the problems that arise in giving a probabilistic interpretation to Euclidean path integrals by using such integrals only to calculate expectation values: the expectation value of an arbitrary scalar field $A_{g,\Phi}(x)$, which may depend on the metric and matter fields and their derivatives, is taken as

$$\langle A \rangle = \frac{\int [dg][d\Phi] A_{g,\Phi}(x) \exp(-S[g,\Phi])}{\int [dg][d\Phi] \exp(-S[g,\Phi])} . \quad (8.7)$$

The general covariance of the theory makes $\langle A \rangle$ independent of x. In fact, it should be emphasized that this sort of expectation values includes an average over the time in the history of the universe that A is measured. On the other hand, the probability $P_\alpha[\beta]$ discussed above is the expectation value of a nonlocal operator, the delta function in (8.5), and refers to a specific local time α.

(I should mention here that there is a very different and apparently unrelated approach to the problem of giving a probabilistic interpretation to the wave function Ψ. The Wheeler-DeWitt equation (8.1) is somewhat like the Klein-Gordon equation for a particle in a scalar potential and leads immediately to a somewhat similar conservation law (now given for pure gravity):

$$0 = \frac{\delta}{\delta h_{ij}(\mathbf{x})} \left[h^{1/2}(\mathbf{x}) \mathcal{G}_{ij,kl}(\mathbf{x}) \times \mathrm{Im} \left[\Psi^*[h] \frac{\delta}{\delta h_{kl}(\mathbf{x})} \Psi[h] \right] \right] . \quad (8.8)$$

Since the beginning, it was hoped that such a conservation law could be used to construct a suitable probability density (DeWitt, 1967). Usually (8.8) is stated in a minisuperspace context, where $h_{ij}(\mathbf{x})$ is constrained to depend on only a finite number of parameters. Since $\mathcal{G}_{ij,kl} h^{kl} = -h_{ij}$, it is natural to treat the overall scale of h_{ij} as a sort of global time coordinate, and take as a probability density the corresponding component of the conserved "current" in (8.8). I wish to point out here that such a construction is not limited to any particular minisuperspace formulation, but can be carried out in the general case. Take Ψ to depend on a "global time"

$$T[h] = \left[\int d^3x \, h^{1/2}(\mathbf{x}) \right]^{1/2} \quad (8.9)$$

and an arbitrary (in fact, infinite) number of other parameters $\zeta_n[h]$, all ζ_n independent of the overall scale of $h_{ij}(\mathbf{x})$:

$$0 = \int d^3x \, h^{1/2}(\mathbf{x}) h_{kl}(\mathbf{x}) \frac{\delta \zeta_n[h]}{\delta h_{kl}(\mathbf{x})} . \quad (8.10)$$

We also introduce a Jacobian $J(\zeta, T)$ and write the functional measure as

$$[dh] = J[d\zeta] dT . \quad (8.11)$$

Multiplying Eq. (8.8) with $\delta(T[h]-T)$ and doing an integral over $\mathbf{x}$ and a functional integral over $h_{ij}(\mathbf{x})$, we easily find a constancy condition

$$0 = \frac{d}{dT} \int J[d\zeta] \mathrm{Im} \left[\Psi^* \frac{\partial \Psi}{\partial T} \right] . \quad (8.12)$$

The trouble here is, of course, the same as that encountered in giving a probabilistic interpretation to the Klein-Gordon equation: the integrand in (8.12) is not, in general, positive. Banks, Fischler, and Susskind (1985) and Vilenkin (1986, 1988a), have considered minisuperspace models in which Ψ is complex, with increasing phase, for which the integrand of Eq. (8.12) is positive-definite; however, this is not the case in general, and, in particular, not for Hartle-Hawking boundary conditions. For a recent more general discussion, see Vilenkin (1988b).)

I now want to give a simplified description of Hawking's (1984b) proposed solution of the cosmological constant problem, using for this purpose parts of Coleman's (1988b) analysis. In order to make the cosmological constant into a dynamical variable, Hawking introduces a 3-form gauge field $A_{\mu\nu\lambda}$ of the sort described at the end of Sec. VII. According to the general ideas of Euclidean quantum cosmology, the probability distribution for the scalar $c(x)$ defined by Eq. (7.27) at any one point $x = x_1$ is

$$\begin{aligned} P(c) &= \langle \delta(c(x_1)-c) \rangle \\ &\propto \int [dA][dg][d\Phi] \delta(c(x_1)-c) \\ &\quad \times \exp(-S[A,g,\Phi]) . \end{aligned} \quad (8.13)$$

It is well known that such functional integrals can be expressed as exponentials of the effective action at its stationary point.[14] In the present case, we have

$$P(c) \propto \exp(-\Gamma[A_c, g_c, \Phi_c]) , \quad (8.14)$$

where $\Gamma[A,g,\Phi]$ is the total action (the sum of one-particle irreducible graphs with external lines replaced with fields A,g,Φ) and the subscript c indicates that this quantity is to be evaluated at a point where Γ is stationary with respect to any variations in $A_{\mu\nu\lambda}(x)$, $g_{\mu\nu}(x)$, or $\Phi(x)$ that leave $c(x_1)=c$ fixed. Now, among all the possible stationary points of Γ, there is one that can be found knowing only the effective action relevant to large

[14]The usual proof, for the case without a delta function in the integrand, proceeds by adding a term $\int J\Omega$ to the action, where Ω denotes the various fields, and J is a set of corresponding currents. The path integral is then $\exp[-W(J)] \equiv \int d\Omega \exp(-S-\int J\Omega)$. The effective action is defined by the Legendre transformation $\Gamma(\Omega) = W(J_\Omega) - \int J_\Omega \Omega$, where J_Ω is the current that produces a given expectation value $\Omega = \delta W/\delta J$. The condition for zero current is that $\Gamma(\Omega)$ be stationary with respect to Ω, and at this point $\Gamma(\Omega) = W(0)$. The delta function in (8.13) can be dealt with by writing it as an integral $\int d\omega \exp\{i\omega[c(x_1)-c]\}$. One can then use the above theorem to evaluate the functional integral before integrating over ω, now with no restriction on $c(x)$, and then doing the ω integral.

4-manifolds. In this case, it is convenient to set all fields except $A_{\mu\nu\lambda}$ and $g_{\mu\nu}$ equal to their (A- and g-dependent) stationary values, in which case the effective action can be expanded in inverse powers of the size of the manifold[15]

$$\Gamma_{\rm eff}[A,g]=\frac{\lambda}{8\pi G}\int\sqrt{g}\,d^4x+\frac{1}{16\pi G}\int\sqrt{g}R\;d^4x+\tfrac{1}{48}\int d^4x\sqrt{g}\,F_{\mu\nu\lambda\rho}F^{\mu\nu\lambda\rho}+\cdots\;, \qquad (8.15)$$

the omitted terms involving more than two derivatives of g and/or A. As we saw in Sec. VII, the condition that this be stationary in $A_{\mu\nu\lambda}$ [for variations that keep $c(x_1)$ fixed] is that $F_{\mu\nu\lambda\rho}$ have vanishing covariant divergence, from which it follows that c in Eq. (7.27) is constant; hence

$$\Gamma_{\rm eff}=\frac{\lambda(c)}{8\pi G}\int\sqrt{g}\,d^4x+\frac{1}{16\pi G}\int\sqrt{g}R\;d^4x+\cdots\;, \qquad (8.16)$$

where

$$\lambda(c)=\frac{c^2}{2}+\lambda\;. \qquad (8.17)$$

The condition that this be stationary in $g_{\mu\nu}$ is, of course, that $g_{\mu\nu}$ satisfy the Einstein field equations with cosmological constant $\lambda(c)$. For any such solution, $R=-4\lambda(c)$, so at the stationary point

$$\Gamma_{\rm eff}=-\frac{\lambda(c)}{8\pi G}\int\sqrt{g}\,d^4x\;. \qquad (8.18)$$

With Hartle-Hawking boundary conditions, the solution of the Einstein equations for $\lambda(c)>0$ is a 4-sphere of proper circumference $2\pi r$, where

$$r=\sqrt{3/\lambda(c)}\;, \qquad (8.19)$$

yielding a probability density proportional to

$$\exp(-\Gamma_{\rm eff})=\exp[3\pi/G\lambda(c)]\;. \qquad (8.20)$$

On the other hand, for $\lambda(c)<0$ the solutions can be made compact by imposing periodicity conditions, but they all have $\Gamma_{\rm eff}\geq 0$. Hawking's conclusion is that the probability density has an infinite peak for $\lambda(c)\to 0+$; hence, after normalizing P,

$$P(c)=\delta(c-c_0)\;, \qquad (8.21)$$

where c_0 is the value of c (assuming there is one), for which $\lambda(c)=0$.

It is important that the quantity $\lambda(c)$ is the true effective cosmological constant, previously called $\lambda_{\rm eff}$, that would be measured in gravitational phenomena at long ranges.[16] The constant λ in Eq. (8.15) includes all effects of fields other than $g_{\mu\nu}$ and $A_{\mu\nu\lambda}$, including all quantum fluctuations. Hence the result (8.21), if valid, really does solve the cosmological constant problem.

We can check that this result is not invalidated by the terms neglected in Eq. (8.16). For a large radius r, the exhibited terms in (8.16) are of order $\lambda r^4/G$ and r^2/G, respectively, while a term with $D\geq 4$ derivatives would yield a contribution to $\Gamma_{\rm eff}$ of order $(mr)^{4-D}$, where m is some combination of the Planck mass and elementary-particle masses. For $\lambda(c)\lesssim m^2$, this shifts the size of the manifold by

$$\delta r/r\approx G\lambda(c)[\lambda(c)/m^2]^{(D-4)/2}\ll 1\;.$$

The change in the stationary value of the action is then

$$\delta\Gamma_{\rm eff}\approx[\lambda(c)/m^2]^{(D-4)/2}\lesssim 1\;,$$

so these higher-derivative terms have no effect on the singularity (8.20).

Coleman (1988b) does not need to introduce a 3-form gauge field $A_{\mu\nu\lambda}$; rather, in order to make the cosmological constant into a dynamical variable, he considers the effect of topological fixtures known as wormholes.[17] An explicit example of a wormhole is provided by the metric (Hawking, 1987b, 1988)

$$ds^2=(1+b^2/x^\mu x^\mu)^2dx^\mu dx^\mu\;. \qquad (8.22)$$

This appears to have a singularity at $x^\mu=0$, but the line element is invariant under the transformation

$$x^\mu\to x'^\mu=x^\mu b^2/x^\nu x^\nu\;, \qquad (8.23)$$

so the region $x^\mu x^\mu<b^2$ actually has the same geometry as that with $x^\mu x^\mu>b^2$. The space described by Eq. (8.22) therefore consists of two asymptotically flat 4-spaces, joined together at the 3-surface with $x^\mu x^\mu=b^2$, a 3-sphere known as a "baby universe." This 4-metric is not a solution of the classical Einstein equations (though it does have $R=0$), but this is not very relevant; the action is

$$S=3\pi b^2/G\;, \qquad (8.24)$$

so the factor $\exp(-S)$ suppresses the effects of all

[15]Such an effective action may be used as the input for calculations in which we include quantum effects only from virtual massless particles with $|q^2|$ less than some cutoff Λ^2. Such effects are, of course, finite, and their Λ dependence is to be canceled by giving the coefficients in $\Gamma_{\rm eff}$ a suitable Λ dependence. (This point of view is described by Weinberg, 1979b.) In order to prevent these quantum effects from generating an unacceptable cosmological constant, the cutoff Λ must be taken very small.

[16]This property is shared by an imaginative solution to the cosmological constant problem proposed by Linde (1988a).

[17]The importance of quantum fluctuations in space-time topology at small scales has been emphasized for many years by Wheeler (e.g., 1964), and more recently by Hawking (1978) and Strominger (1984). Such "space-time foam" was considered as a mechanism for canceling a cosmological constant by Hawking (1983).

wormholes except those of Planck dimensions or less, for which quantum effects are surely important. [A model with classical wormhole solutions, based on a 2-form axion, has been presented by Giddings and Strominger (1988a).]

If Planck-sized wormholes can connect asymptotically flat 4-spaces, then they can connect any 4-spaces that are large compared to the Planck scale. We are therefore led to consider contributions to the Euclidean path integral from large 4-spaces [like the 4-sphere in Hawking's (1984b) theory] connected to themselves and each other with Planck-sized wormholes. Each wormhole can be regarded as the creation and subsequent destruction of a baby universe [like the 3-sphere of proper circumference $4\pi b$ in Hawking's (1987b, 1988) wormhole model], and such baby universes may also appear as part of the boundary of the 4-manifold.

What are the effects of these wormholes and baby universes? At scales large compared with the scale of the baby universe, the creation or destruction of a baby universe can only show up through the insertion of a local operator in the path integral. The various types of baby universes can be classified according to the form of these local operators. The effect of creating and destroying arbitrary numbers of baby universes of all types can thus be expressed by adding a suitable term in the action

$$\tilde{S}=S+\sum_i (a_i+a_i^\dagger)\int d^4x\, O_i(x) , \tag{8.25}$$

where a_i and $a_i^\dagger$ are the annihilation and creation operators for a baby universe of type i, and $O_i(x)$ is the corresponding local operator. [This was first stated by Hawking (1987b). Creation and annihilation operators for baby universes were earlier used by Strominger (1984). For a proof of Eq. (8.25), see Coleman (1988a) and Giddings and Strominger (1988b).] The path integral over all 4-manifolds with given boundary conditions is to be calculated as

$$\int [dg][d\Phi]e^{-S}=\int_{\rm No}[dg][d\Phi]\langle B|e^{-\tilde{S}}|B\rangle , \tag{8.26}$$

where No means that wormholes and baby universes are excluded, and $|B\rangle$ is a normalized baby-universe state depending on the boundary conditions. For instance, with Hartle-Hawking boundary conditions, $|B\rangle$ is the empty state

$$a_i|B\rangle=0 . \tag{8.27}$$

These baby universes have an important effect even if none of them appear as part of the boundary of the 4-manifold, as would be the case for Hartle-Hawking boundary conditions. Hawking (1987b, 1988) has suggested that since the baby universes are unobservable, their effect is an effective loss of quantum coherence. [See also Hawking (1982); Teitelboim (1982); Strominger (1984); Lavrelashvili, Rubakov, and Tinyakov (1987, 1988); Giddings and Strominger (1988b). A contrary view was taken by Gross (1984).] Recently Coleman (1988a) has argued (convincingly, in my view) for a different interpretation [see also Giddings and Strominger (1988b)]. The state $|B\rangle$ in Eq. (8.26) may always be expanded in eigenstates of the operators $a_i+a_i^\dagger$:

$$|B\rangle=\int f_B(\alpha)\prod_i d\alpha_i|\alpha\rangle , \tag{8.28}$$

$$(a_i+a_i^\dagger)|\alpha\rangle=\alpha_i|\alpha\rangle , \tag{8.29}$$

$$\langle\alpha'|\alpha\rangle=\prod_i \delta(\alpha_i'-\alpha_i) , \tag{8.30}$$

the function $f_B(\alpha)$ depending on the boundary conditions. For instance, for Hartle-Hawking conditions, $|B\rangle$ satisfies Eq. (8.27), and so

$$f_B(\alpha)=\prod_i \pi^{-1/4}\exp(-\alpha_i^2/2) . \tag{8.31}$$

(With n baby universes on the boundary of the 4-space, this would be multiplied with a Hermite polynomial of order n.) In the state $|\alpha\rangle$, the effect of the creation and annihilation of baby universes is to change the action S to

$$S_\alpha=S+\sum_i \alpha_i\int O_i(x)d^4x . \tag{8.32}$$

That is, the coupling constant multiplying each possible local term $\int O_i d^4x$ is changed by an amount α_i. As soon as we start to make any sort of measurements, the state of the universe breaks up into an incoherent superposition of these $|\alpha\rangle$'s, each appearing with *a priori* probability $|f_B(\alpha)|^2$; but for each term we have an ordinary wormhole-free quantum theory, with α-dependent action (8.32).

If all we want is to explain why the cosmological constant is not enormous, then our work is essentially done. The effective cosmological constant is a function of the α_i, because among the O_i there is a simple operator $O_1=\sqrt{g}$, whose coefficient contributes a term $8\pi G\alpha_1$ to λ, and also because the vacuum energy $\langle\rho\rangle$ depends on the couplings of all interactions, each of which has a term proportional to one of the α_i. Now, generic baby-universe states $|B\rangle$ will have components $|\alpha\rangle$ for which $\lambda_{\rm eff}(\alpha)$ is very small, as well as others for which it is enormous. The anthropic considerations of Sec. VI tell us that any scientist who asks about the value of the cosmological constants can only be living in components $|\alpha\rangle$ for which $\lambda_{\rm eff}$ is quite small, for otherwise galaxies and stars could never have formed (for $\lambda_{\rm eff}>0$), or else there would not be time for life to evolve (for $\lambda_{\rm eff}<0$).

However, it is of great interest to ask whether the effective cosmological constant is really zero, or just small enough to satisfy anthropic bounds, in which case it should show up observationally. The probability of getting any particular value of the α_i, and hence of finding a value $\lambda_{\rm eff}(\alpha)$, is not just given by the function $|f_B(\alpha)|^2$ arising from the boundary conditions, but is also affected by the functional integral itself.

In calculating this effect, Coleman (1988b) observed that although we are to integrate only over connected 4-manifolds, on a scale much large than the wormhole scale those manifolds that appear disconnected will really be connected by wormholes. Hence any sort of probability density or expectation value will contain as a factor a sum over disconnected manifolds consisting of arbitrary numbers of closed connected wormhole-free components.[18] Just as for Feynman diagrams, this sum is the exponential of the path integral for a single closed connected wormhole-free manifold

$$F(\alpha)=\exp\left[\int_{CC}[dg]\exp(-S_\alpha[g])\right] , \tag{8.33}$$

where CC indicates that we include only closed connected wormhole-free manifolds, and $S_\alpha[g]$ is the action (8.32) with all fields other than $g_{\mu\nu}(x)$ integrated out.

The path integral in (8.33) can be evaluated by precisely the same methods as described above in connection with Hawking's (1984b) model [and used for this purpose by Coleman (1988b)]. The result is that the probability density for $\lambda_{\rm eff}$ contains a factor (for $\lambda_{\rm eff}>0$)

$$F=\exp\left[\exp\left[\frac{3\pi}{G\lambda_{\rm eff}}+O(1)\right]\right] . \tag{8.34}$$

The fact that this is now an exponential of an exponential, instead of a mere exponential, is not essential in solving the cosmological constant problem (though it is important in fixing other constants, as described at the end of this section). Either way, the probability distribution has an infinite peak at $\lambda_{\rm eff}\rightarrow 0+$, which, after normalizing so that the total probability is unity, means that $P(\alpha)$ has a factor

$$P(\alpha)\propto\delta(\lambda_{\rm eff}(\alpha)) . \tag{8.35}$$

In addition, as in Hawking's case, from the way that F has been calculated it is clear that this $\lambda_{\rm eff}$ is the constant that appears in the effective action for pure gravity with all high-energy fluctuations integrated out; hence it is the cosmological constant relevant to astronomical observation.

Has the cosmological constant problem been solved? Perhaps so, but there are still some things to worry about in Coleman's approach, as also in the earlier work of Hawking. Here is a short list of qualms.

(1) Does Euclidean quantum cosmology have anything to do with the real world? It is essential to both Coleman and Hawking that the path integral be given by a stationary point of the Euclideanized action—the conclusion would be completely wiped out if in place of $\exp(3\pi/G\lambda_{\rm eff})$ we had found $\exp(3\pi i/G\lambda_{\rm eff})$. Some of the technical and conceptual difficulties of Euclidean quantum cosmology were discussed at the beginning of this section.

(2) What are the boundary conditions? It is always possible that the essential singularity in $\exp\{\exp[3\pi/G\lambda(\alpha)]\}$ is canceled by an essential zero in the *a priori* probability $|f_B(\alpha)|^2$. However, this is not the case for Hartle-Hawking boundary conditions, where $|f_B(\alpha)|^2$ is a simple Gaussian. Moreover, Coleman (1988b) has shown that in his theory such an essential zero would be destroyed by almost any perturbation of the boundary conditions; instead of its being unnatural to have zero cosmological constant, it would be highly unnatural not to. Still, the problem of boundary conditions is disturbing, because it reminds us that quantum cosmology is an incomplete theory.

(3) Are wormholes real? Coleman's calculation depends on there being a clear separation between the very large 4-manifolds, for which the long-range effective action is stationary (and large and negative), and very small wormholes, whose contribution to the action is of order unity (and generally positive). Furthermore, the wormholes have been assumed to be so well separated that we can ignore their interactions (the "dilute gas" approximation). It may be possible to construct a theory in which the wormhole scale [like b in Eq. (8.22)] is somewhat larger than the Planck scale, large enough to allow the wormhole metric to be calculated classically, but we would still have to ask whether this is actually the case. Hawking (1984b) does not need to worry about wormholes, but how do we know that the 3-form gauge field is real? A related question for both authors: even granting the existence of the stationary point of the action at which $\Gamma_{\rm eff}=-3\pi/\lambda G$, how do we know that this is the dominant stationary point?

(4) What about the other terms in the effective action? For instance, suppose we include the 6-derivative term[19]

$$\Gamma_{\rm eff}=\int d^4x\sqrt{g}\left[\frac{\lambda(\alpha)}{8\pi G(\alpha)}+\frac{R}{16\pi G(\alpha)}+\zeta(\alpha)G(\alpha)R_{\mu\nu}{}^{\lambda\rho}R_{\lambda\rho}{}^{\sigma\tau}R_{\sigma\tau}{}^{\mu\nu}\right] , \tag{8.36}$$

with $\zeta(\alpha)$ a dimensionless coefficient that, like λ and G, depends on the baby-universe parameters α_i. Hawking and Coleman found a stationary point of this action for which $\Gamma_{\rm eff}\rightarrow-\infty$ when $\lambda(\alpha)G(\alpha)\rightarrow 0$, but for this purpose it is essential that $\zeta(\alpha)$ remain bounded in this limit. (We recall that in our previous discussion of the higher-derivative terms in $\Gamma_{\rm eff}$, we assumed that the coefficient m^{4-D} of terms with $D\geq 4$ derivatives remains less than $\lambda^{(D-4)/2}$ as $\lambda\rightarrow 0$.) But if we can let $1/\lambda G$ go to infinity, then why not let ζ go to infinity also? In particular, why not use a dimensional factor $1/\lambda(\alpha)$ in place of $G(\alpha)$ in

[18]This sum actually includes manifolds that are truly not connected by wormholes or anything else, but their contribution is a harmless multiplicative factor, which will cancel out anyway in normalizing $P(\alpha)$.

[19]Terms involving the Ricci tensor $R_{\mu\nu}$ or its trace R are not included here, because they represent merely a redefinition of the metric; see, e.g., Weinberg (1979a). The 4-derivative term $R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho}$ is not included, because it can be combined with terms involving $R_{\mu\nu}$ or R to make a topological invariant and is therefore physically unmeasureable for fixed large-scale topology.

the last term of Eq. (8.36)? This would completely invalidate the analysis of the singularity in the probability density $P(\alpha)$, and could well wipe it out.

The last of these four qualms suggests some interesting possibilities. Suppose we do assume that for some reason constants like $\zeta(\alpha)$ in Eq. (8.36) are bounded. Then the effect of wormholes is not only to fix $\lambda(\alpha)$ at zero, but also to fix these other constants at their lower or upper bounds. [I think this is the correct interpretation of what Coleman (1988b) calls "the big fix."] For instance, for $\zeta(\alpha)$ bounded and $|\lambda(\alpha)G(\alpha)| \ll 1$, the action (8.36) is stationary for a sphere of proper circumference $2\pi r$, where

$$r^2 = \frac{3}{\lambda} - \frac{64\zeta\pi G^2\lambda}{3} , \tag{8.37}$$

for which the effective action takes the value

$$\Gamma_{\rm eff} = -\frac{3\pi}{G\lambda} - \frac{128\zeta G\lambda\pi^2}{3} . \tag{8.38}$$

Thus the probability distribution $\exp[\exp(-\Gamma_{\rm eff})]$ not only has an infinite peak at $\lambda(\alpha)=0$, but also contains a factor

$$\exp\left[\frac{128\zeta G\lambda\pi^2}{3}\exp\left[\frac{3\pi}{G\lambda}\right]\right] . \tag{8.39}$$

For $G\lambda \to 0$, the quantity $G\lambda \exp(3\pi/G\lambda)$ becomes infinite, so the normalized probability will have a delta function at the upper bound of $\zeta(\alpha)$. All constants in the effective action for gravitation, including terms with any numbers of derivatives, can be calculated in this way,[20] but they all have to be bounded as $\lambda(\alpha)G(\alpha) \to 0$ for any of this to make sense.

It may be that the bounds (if any) on parameters like $\zeta(\alpha)$ arise from the details of wormhole physics, in which case these remarks are not going to be useful numerically for some time. However, there is another more exciting possibility, that there are just unitarity bounds, which could be calculated working only with low-energy effective theory itself. Of course, we are not likely to be able to measure parameters like $\zeta(\alpha)$, but it would still be nice to be able to calculate them, because up to now the only really unsatisfactory feature of the quantum theory of gravitation has been the apparent arbitrariness of this infinite set of parameters.

[20]To the extent that it will become possible to calculate functions like $\lambda(\alpha)$, $G(\alpha)$, $\zeta(\alpha)$ etc., in terms of the parameters in an underlying fundamental theory, such as a string theory, the location of the delta functions in F may allow us to infer something about the values of the α_i and of the parameters in the underlying theory. However, without such an underlying theory, it is impossible to use calculations of λG, ζ, etc., to infer anything about the observed parameters of some intermediate theory like the standard model. This is because, in addition to charges, masses, etc., the standard model implicitly also involves parameters $\lambda_0, G_0, \zeta_0, \ldots$ appearing in the effective action for gravitation. When we integrate out the quarks, leptons, and gauge and Higgs bosons, we obtain new values for λ, G, ζ, etc.; but these new values depend on an equal number of unknowns λ_0, G_0, ζ_0, etc., as well as on charges and masses.

IX. OUTLOOK

All of the five approaches to the cosmological constant problem described in Secs. IV–VIII remain interesting. At present, the fifth, based on quantum cosmology, appears the most promising. However, if wormholes (or 3-form gauge fields) do produce a distribution of values for the cosmological constant, but without an infinite peak at $\lambda_{\rm eff}=0$, then we will have to fall back on the anthropic principle to explain why $\lambda_{\rm eff}$ is not enormously larger than allowed by observation. Alternatively, it may be some change in the theory of gravity, like that described here in Sec. VII, that produces the distribution in values for $\lambda_{\rm eff}$. The approaches based on supersymmetry and adjustment mechanisms described in Secs. IV and VI seem least promising at present, but this may change.

All five approaches have one other thing in common: They show that any solution of the cosmological constant problem is likely to have a much wider impact on other areas of physics or astronomy. One does not need to explain the potential importance of supergravity and superstrings. A light scalar like that needed for adjustment mechanisms could show up macroscopically, as a "fifth force." Changing gravity by making $\mathrm{Det}g_{\mu\nu}$ not dynamical would make us rethink our quantum theories of gravitation, and wormholes might force all the constants in these theories to their outer bounds. Finally, and of greatest interest to astronomy, if it is only anthropic constraints that keep the effective cosmological constant within empirical limits, then this constant should be rather large, large enough to show up before long in astronomical observations.

Note added in proof. As might have been expected, in the time since this report was submitted for publication there have appeared a large number of preprints that follow up on various aspects of the work of Coleman (1988b) and Banks (1988). Here is a partial list: Accetta *et al.* (1988); Adler (1988); Fischler and Susskind (1988); Giddings and Strominger (1988c); Gilbert (1988); Grinstein and Wise (1988); Gupta and Wise (1988); Hosoya (1988); Klebanov, Susskind, and Banks (1988); Myers and Periwal (1988); Polchinski (1988); Rubakov (1988). I am not able to review all of these papers here. However, I do want to mention two further qualms, regarding Coleman's proposed solution of the cosmological constant problem, that are raised by some of these papers. First, Fischler and Susskind (1988), partly on the basis of conversations with V. Kaplunovsky, have pointed out that the exponential damping of large wormholes may be overcome by Coleman's double exponential. If this were the case, we would be confronted with closely packed wormholes of macroscopic as well as Planck scales. This would be a disaster for Coleman's proposed solution of the cosmological constant problem, and would also indicate that we do not fully understand how to use Euclidean path integrals in quantum cosmology. Next, Polchinski (1988) has found that the Euclidean path integral over closed, connected, wormhole-free manifolds inside the exponential in (8.33) has a phase that might eliminate the peak in the probability distribution at zero cosmological constant. As pointed out here in footnote 15, when we use an effective action Γ_{eff} to evaluate such path integrals, the effective action must be taken as an input to calculations in which we include quantum fluctuations in massless particle fields with momenta up to some ultraviolet cutoff Λ. This cutoff must be taken as the same as the *infrared* cutoff that was used in calculating Γ_{eff}, so that all fluctuations are taken into account. It was remarked in footnote 15 that Λ must be taken very small, to avoid reintroducing a cosmological constant, but as Polchinski now remarks, no matter how small we take Λ, the integral over fluctuations in the gravitational field with momenta less than Λ produces a phase in the integral. Since this phase appears inside the exponential in Eq. (8.33), if its real part is not positive definite there would be no exponential peak at zero cosmological constant. On the other hand, in the absence of wormholes this phase would appear as an overall factor in front of a single exponential, so it would not affect the peaking at zero cosmological constant found by Hawking (1984b).

ACKNOWLEDGMENTS

I have been greatly helped in preparing this review by conversations with many colleagues. Here is a list of a few of those to whom my thanks are especially due. Section II: G. Holton; Sec. III: E. Witten; Sec. IV: S. de Alwis, J. Polchinski, E. Witten; Sec. V: P. J. E. Peebles, P. Shapiro, E. Vishniac; Sec. VI: J. Polchinski; Sec. VII: C. Teitelboim, F. Wilczek; Sec. VIII: L. Abbott, S. Coleman, B. DeWitt, W. Fischler, S. Giddings, L. Susskind, C. Teitelboim, F. Wilczek. Of course, they take no responsibility for anything that I may have gotten wrong. Research was supported in part by the Robert A. Welch Foundation and NSF Grant No. PHY 8605978.

REFERENCES

Abbott, L., 1985, Phys. Lett. **B150**, 427.
Abbott, L., 1988, Sci. Am. **258** (No. 5), 106.
Accetta, F. S., A. Chodos, F. Cooper, and B. Shao, 1988, "Fun with the wormhole calculus," Yale University Preprint No. YCTP-P20-88.
Adler, S. L., 1988, "On the Banks-Coleman-Hawking argument for the vanishing of the cosmological constant," Institute for Advanced Study Preprint No. IASSNS-HEP-88/35.
Albrecht, A., and P. J. Steinhardt, 1982, Phys. Rev. Lett. **48**, 120.
Arnowitt, R., S. Deser, and C. W. Misner, 1962, in *Gravitation: An Introduction to Current Research,* edited by L. Witten (Wiley, New York) p. 227.
Attick, J., G. Moore, and A. Sen, 1987, Institute for Advanced Studies preprint.
Augustine, 398, *Confessions,* translated by R. S. Pine-Coffin (Penguin Books, Harmondsworth, Middlesex, 1961), Book XI.
Bahcall, J., T. Piran, and S. Weinberg, 1987, Eds., *Dark Matter in the Universe: 4th Jerusalem Winter School for Theoretical Physics* (World Scientific, Singapore).
Bahcall, S. R., and S. Tremaine, 1988, Astrophys. J. **326**, L1.
Banks, T., 1985, Nucl. Phys. B **249**, 332.
Banks, T., 1988, University of California, Santa Cruz, Preprint No. SCIPP 88/09.
Banks, T., W. Fischler, and L. Susskind, 1985, Nucl. Phys. B **262**, 159.
Barbieri, R., E. Cremmer, and S. Ferrara, 1985, Phys. Lett. B **163**, 143.
Barbieri, R., S. Ferrara, D. V. Nanopoulos, and K. S. Stelle, 1982, Phys. Lett. B **113**, 219.
Barr, S. M., 1987, Phys. Rev. D **36**, 1691.
Barr, S. M., and D. Hochberg, 1988, Phys. Lett. B **211**, 49.
Barrow, J. D., and F. J. Tipler, 1986, *The Anthropic Cosmological Principle* (Clarendon, Oxford).
Bernstein, J., and G. Feinberg, 1986, Eds., *Cosmological Constants* (Columbia University, New York).
Bludman, S. A., and M. Ruderman, 1977, Phys. Rev. Lett. **38**, 255.
Brown, J. D., and C. Teitelboim, 1987a, Phys. Lett. B **195**, 177.
Brown, J. D., and C. Teitelboim, 1987b, Nucl. Phys. B **297**, 787.
Buchmüller, W., and N. Dragon, 1988a, University of Hannover Preprint No. ITP-UH 1/88.
Buchmüller, W., and N. Dragon, 1988b, Phys. Lett. B **207**, 292.
Carter, B., 1974, in *International Astronomical Union Symposium 63: Confrontation of Cosmological Theories with Observational Data,* edited by M. S. Longair (Reidel, Dordrecht), p. 291.
Carter, B., 1983, in *The Constants of Physics, Proceedings of a Royal Society Discussion Meeting, 1983,* edited by W. H. McCrea and M. J. Rees (printed for The Royal Society, London, at the University Press, Cambridge), p. 137.
Casimir, H. B. G., 1948, Proc. K. Ned. Akad. Wet. **51**, 635.
Chang, N.-P., D.-X. Li, and J. Pérez-Mercader, 1988, Phys. Rev. Lett. **60**, 882.
Coleman, S., 1988a, Nucl. Phys. B **307**, 867.
Coleman, S., 1988b, "Why there is nothing rather than something: A theory of the cosmological constant," Harvard University Preprint No. HUTP-88/A022.
Coleman, S., and F. de Luccia, 1980, Phys. Rev. D **21**, 3305.
Coughlan, G. D., I. Kani, G. G. Ross, and G. Segré, 1988, CERN Preprint No. TH. 5014/88.
Cremmer, E., S. Ferrara, C. Kounnas, and D. V. Nanopoulos, 1983, Phys. Lett. B **133**, 61.
Cremmer, E., B. Julia, J. Scherk, S. Ferrara, L. Girardello, and P. van Nieuwenhuizen, 1978, Phys. Lett. B **79**, 231.
Cremmer, E., B. Julia, J. Scherk, S. Ferrara, L. Girardello, and P. van Nieuwenhuizen, 1979, Nucl. Phys. B **147**, 105.
Davies, P. C. W., 1982, *The Accidental Universe* (Cambridge University, Cambridge).
de Sitter, W., 1917, Mon. Not. R. Astron. Soc. **78**, 3 (reprinted in Bernstein and Feinberg, 1986).
de Vaucouleurs, G., 1982, Nature (London) **299**, 303.
de Vaucouleurs, G., 1983, Astrophys. J. **268**, 468, Appendix B.
DeWitt, B., 1967, Phys. Rev. **160**, 1113.
Dicke, R. H., 1961, Nature (London) **192**, 440.
Dine, M., W. Fischler, and M. Srednicki, 1981, Phys. Lett. B **104**, 199.

Dine, M., R. Rohm, N. Seiberg, and E. Witten, 1985, Phys. Lett. B **156**, 55.
Dine, M., and N. Seiberg, 1986, Phys. Rev. Lett. **57**, 2625.
Dirac, P. A. M., 1937, Nature (London) **139**, 323.
Dolgov, A. D., 1982, in *The Very Early Universe: Proceedings of the 1982 Nuffield Workshop at Cambridge,* edited by G. W. Gibbons, S. W. Hawking, and S. T. C. Siklos (Cambridge University, Cambridge), p. 449.
Dreitlein, J., 1974, Phys. Rev. Lett. **34**, 777.
Eddington, A. S., 1924, *The Mathematical Theory of Relativity,* 2nd Ed. (Cambridge University, London).
Einstein, A., 1917, Sitzungsber. Preuss. Akad. Wiss. Phys.-Math. Kl. 142 [English Translation in *The Principle of Relativity* (Methuen, 1923, reprinted by Dover Publications), p. 177; and in Bernstein and Feinberg, 1986].
Einstein, A., 1919, Sitzungsber. Preuss. Akad. Wiss., Phys.-Math. Kl. [English translation in *The Principle of Relativity* (Methuen, 1923, reprinted by Dover Publications), p. 191].
Ellis, J., C. Kounnas, and D. V. Nanopoulos, 1984, Nucl. Phys. B **247**, 373.
Ellis, J., A. B. Lahanas, D. V. Nanopoulos, and K. Tamvakis, 1984, Phys. Lett. B **134**, 429.
Ellis, J., N. C. Tsamis, and M. Voloshin, 1987, Phys. Lett. B **194**, A 291.
Fischler, W., and L. Susskind, 1988, "A wormhole catastrophe," Texas Preprint No. UTTG-26-88.
Ford, L. H., 1987, Phys. Rev. D **35**, 2339.
Freese, K., F. C. Adams, J. A. Frieman, and E. Mottola, 1987, Nucl. Phys. B **287**, 797.
Friedan, D., E. Martinec, and S. Shenker, 1986, Nucl. Phys. B **271**, 93.
Friedmann, A., 1924, Z. Phys. **21**, 326 [English translation in Bernstein and Feinberg, 1986, Eds., *Cosmological Constants* (Columbia University, New York)].
Gibbons, G. W., S. W. Hawking, and M. J. Perry, 1978, Nucl. Phys. B **138**, 141.
Giddings, S. B., and A. Strominger, 1988a, Nucl. Phys. B **306**, 890.
Giddings, S. B., and A. Strominger, 1988b, Nucl. Phys. B **307**, 854.
Giddings, S. B., and A. Strominger, 1988c, "Baby universes, third quantization, and the cosmological constant," Harvard Preprint No. HUTP-88/A036.
Gilbert, G., 1988, "Wormhole induced proton decay," Caltech Preprint No. CALT-68-1524.
Grinstein, B., and M. B. Wise, 1988, "Light scalars in quantum gravity," Caltech Preprint No. CALT-68-1505.
Grisaru, M. T., W. Siegel, and M. Rocek, 1979, Nucl. Phys. B **159**, 429.
Gross, D. J., 1984, Nucl. Phys. B **236**, 349.
Gupta, A. K., and M. B. Wise, 1988, "Comment on wormhole correlations," Caltech Preprint No. CALT-68-1520.
Guth, A. H., 1981, Phys. Rev. D **23**, 347.
Hartle, J. B., 1987, in *Gravitation in Astrophysics: Cargese 1986,* edited by B. Carter and J. B. Hartle (Plenum, New York), p. 329.
Hartle, J. B., and S. W. Hawking, 1983, Phys. Rev. D **28**, 2960.
Hawking, S. W., 1978, Nucl. Phys. B **144**, 349.
Hawking, S. W., 1979, in *Three Hundred Years of Gravitation,* edited by S. W. Hawking and W. Israel (Cambridge University, Cambridge).
Hawking, S. W., 1982, Commun. Math. Phys. **87**, 395.
Hawking, S. W., 1983, Philos. Trans. R. Soc. London, Ser. A **310**, 303.
Hawking, S. W., 1984a, Nucl. Phys. B **239**, 257.
Hawking, S. W., 1984b, Phys. Lett. B **134**, 403.
Hawking, S. W., 1984c, in *Relativity, Groups and Topology II,* NATO Advanced Study Institute Session XL . . . Les Houches 1983, edited by B. S. DeWitt and Raymond Stora (Elsevier, Amsterdam), p. 336.
Hawking, S. W., 1987a, remarks quoted by M. Gell-Mann, Phys. Scr. **T15**, 202 (1987).
Hawking, S. W., 1987b, Phys. Lett. B **195**, 337.
Hawking, S. W., 1988, Phys. Rev. D **37**, 904.
Hawking, S., and D. Page, 1986, Nucl. Phys. B **264**, 185.
Henneaux, M., and C. Teitelboim, 1984, Phys. Lett. B **143**, 415.
Henneaux, M., and C. Teitelboim, 1988, "The cosmological constant and general covariance," University of Texas preprint.
Hosoya, A., 1988, "A diagrammatic derivation of Coleman's vanishing cosmological constant," Hiroshima Preprint No. RRK-88-28.
Kim, J., 1979, Phys. Rev. Lett. **43**, 103.
Klebanov, I., L. Susskind, and T. Banks, 1988, "Wormholes and cosmological constant," SLAC Preprint No. SLAC-Pub.-4705.
Knapp, G. R., and J. Kormendy, 1987, Eds., *Dark Matter in the Universe: I.A.U. Symposium No. 117* (Reidel, Dordrecht).
Lavrelashvili, G. V., V. A. Rubakov, and P. G. Tinyakov, 1987, Pis'ma Zh. Eksp. Teor. Fiz. **46**, 134 [JETP Lett. **46**, 167 (1987)].
Lavrelashvili, G. V., V. A. Rubakov, and P. G. Tinyakov, 1988, Nucl. Phys. B **299**, 757.
Lemaître, G., 1927, Ann. Soc. Sci. Bruxelles, Ser. 1 **47**, 49.
Lemaître, G., 1931, Mon. Not. R. Astron. Soc. **91**, 483.
Linde, A. D., 1974, Pis'ma Zh. Eksp. Teor. Fiz. **19**, 320 [JETP Lett. **19**, 183 (1974)].
Linde, A. D., 1982, Phys. Lett. B **129**, 389.
Linde, A. D., 1986, Phys. Lett. B **175**, 395.
Linde, A. D., 1987, Phys. Scri. **T15**, 169.
Linde, A. D., 1988a, Phys. Lett. B **200**, 272.
Linde, A. D., 1988b, Phys. Lett. B **202**, 194.
Loh, E. D., 1986, Phys. Rev. Lett. **57**, 2865.
Loh, E. D., and E. J. Spillar, 1986, Astrophys. J. **303**, 154.
Martinec, E., 1986, Phys. Lett. B **171**, 189.
Moore, G., 1987a, Nucl. Phys. B **293**, 139.
Moore, G., 1987b, Institute for Advanced Study Preprint No. IASSNS-HEP-87/59, to be published in the proceedings of the Cargèse School on Nonperturbative Quantum Field Theory.
Morozov, A., and A. Perelomov, 1987, Phys. Lett. B **199**, 209.
Myers, R. C., and V. Periwal, 1988, "Constants and correlations in the Coleman calculus," Santa Barbara Preprint No. NSF-ITP-88-151.
Page, D., 1987, in *The World and I* (in press).
Pais, A., 1982, *'Subtle is the Lord . . .': The Science and the Life of Albert Einstein* (Oxford University, New York).
Peccei, R. D., J. Solà, and C. Wetterich, 1987, Phys. Lett. B **195**, 183.
Peebles, P. J. E., 1984, Astrophys. J. **284**, 439.
Peebles, P. J. E., 1987a, in *Proceedings of the Summer Study on the Physics of the Superconducting Super Collider,* edited by R. Donaldson and J. Marx (Division of Particles and Fields of the APS, New York).
Peebles, P. J. E., 1978b, Publ. Astron. Soc. Pac., in press.
Peebles, P. J. E., and B. Ratra, 1988, Astrophys. J. Lett. **325**, L17.
Petrosian, V., E. E. Salpeter, and P. Szekeres, 1967, Astrophys. J. **147**, 1222.
Polchinski, J., 1986, Commun. Math. Phys. **104**, 37.

Polchinski, J., 1987, private communication.
Polchinski, J., 1988, in preparation.
Ratra, B., and P. J. E. Peebles, 1988, Phys. Rev. D **37**, 3406.
Rees, M. J., 1987, New Sci. August 6, 1987, p. 43.
Renzini, A., 1986, in *Galaxy Distances and Deviations from Universal Expansion,* edited by B. F. Madore and R. B. Tully (Reidel, Dordrecht), p. 177.
Reuter, M., and C. Wetterich, 1987, Phys. Lett. **188**, 38.
Rohm, R., 1984, Nucl. Phys. B **237**, 553.
Rowan-Robinson, M., 1968, Mon. Not. R. Astron. Soc. **141**, 445.
Rubakov, V. A., 1988, "On the third quantization and the cosmological constant," DESY preprint.
Shklovsky, I., 1967, Astrophys. J. **150**, L1.
Slipher, V. M., 1924, table in Eddington (1924), *The Mathematical Theory of Relativity,* 2nd Ed. (Cambridge University, London), p. 162
Sparnaay, M. J., 1957, Nature (London) **180**, 334.
Strominger, A., 1984, Phys. Rev. Lett. **52**, 1733.
Teitelboim, C., 1982, Phys. Rev. D **25**, 3159.
Turner, M. S., G. Steigman, and L. M. Krauss, 1984, Phys. Rev. Lett. **52**, 2090.
Van der Bij, J. J., H. Van Dam, and Y. J. Ng, 1982, Physica **116A**, 307.
Veltman, M., 1975, Phys. Rev. Lett. **34**, 777.
Vilenkin, A., 1986, Phys. Rev. D **33**, 3560.
Vilenkin, A., 1988a, Phys. Rev. D **37**, 888.
Vilenkin, A., 1988b, Tufts Preprint No. TUTP-88-3.
Weinberg, S., 1972, *Gravitation and Cosmology* (Wiley, New York).
Weinberg, S., 1979a, in *General Relativity: An Einstein Centenary Survey,* edited by S. W. Hawking and W. Israel (Cambridge University, Cambridge), p. 800.
Weinberg, S., 1979b, Physica **96A**, 327.
Weinberg, S., 1982, Phys. Rev. Lett. **48**, 1776.
Weinberg, S., 1983, unpublished remarks at the workshop on "Problems in Unification and Supergravity," La Jolla Institute, 1983.
Weinberg, S., 1987, Phys. Rev. Lett. **59**, 2607.
Wetterich, C., 1988, Nucl. Phys. B **302**, 668.
Wheeler, J. A., 1964, in *Relativity, Groups and Topology,* Lectures Delivered at Les Houches, 1963 . . . , edited by B. DeWitt and C. DeWitt (Gordon and Breach, New York), p. 317.
Wheeler, J. A., 1968, in *Battelle Rencontres,* edited by C. DeWitt and J. A. Wheeler (Benjamin, New York).
Wilczek, F., 1984, Phys. Rep. **104**, 143.
Wilczek, F., 1985, in *How Far Are We from the Gauge Forces: Proceedings of the 1983 Erice Conference,* edited by A. Zichichi (Plenum, New York), p. 208.
Wilczek, F., and A. Zee, 1983, unpublished work quote by Zee (1985), in *High Energy Physics: Proceedings of the Annual Orbis Scientiae,* edited by S. L. Mintz and A. Perlmutter (Plenum, New York).
Witten, E., 1983, in *Proceedings of the 1983 Shelter Island Conference on Quantum Field Theory and the Fundamental Problems of Physics,* edited by R. Jackiw, N. N. Khuri, S. Weinberg, and E. Witten (MIT, Cambridge, Massachusetts), p. 273.
Witten, E., 1985, Phys. Lett. B **155**, 151.
Witten, E., and J. Bagger, 1982, Phys. Lett. B **115**, 202.
Zee, A., 1985, in *High Energy Physics: Proceedings of the 20th Annual Orbis Scientiae, 1983,* edited by S. L. Mintz and A. Perlmutter (Plenum, New York).
Zeldovich, Ya., B., 1967, Pis'ma Zh. Eksp. Teor. Fiz. **6**, 883 [JETP Lett. **6**, 316 (1967)].
Zumino, B., 1975, Nucl. Phys. B **89**, 535.