Quantum fluctuation-dissipation theorem for general relativity

Emil Mottola
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106
(Received 12 April 1985; revised manuscript received 25 November 1985)

By applying linear-response methods long familiar in many-body theory to quantum fluctuations in curved spacetime, the vacuum is shown to behave like a *dissipative* as well as a fluctuating medium. The implications of this quantum vacuum "viscosity" for cosmology and the cosmological constant are discussed. Since the de Sitter phase is *unstable* to particle creation processes, this vacuum dissipative mechanism may be responsible for the approximate flatness of the observed Universe.

I. INTRODUCTION

Ever since Hawking first discovered that a black hole in the presence of quantum matter fields radiates at a temperature equal to $1/8\pi M$ (Ref. 1), it has become increasingly clear that the methods of statistical mechanics may be extended and fruitfully reapplied to general relativity. Macroscopic thermodynamic concepts, such as temperature and entropy, make their appearance in the four laws of black-hole thermodynamics.[2] The microscopic quantities which underlie these concepts are the Green's functions of finite-temperature field theory. In ordinary, nonrelativistic, statistical mechanics the real-time Green's functions are used to study both the fluctuations of the system about its equilibrium configuration *and* the dissipative transport properties of the system under the influence of external fields. The response of a physical quantity (such as the electric current) to first order in a perturbation (such as an applied electric field) describes the dissipative effects of the medium (Ohm's law). The fluctuations in equilibrium and the dissipation in an applied field are closely related, being the symmetric and antisymmetric parts, respectively, of an appropriate polarization tensor. Because of the analytic properties of the Green's functions, specifically their periodicity in imaginary time the symmetric and antisymmetric parts of the polarization tensor are related in a simple way. This relation is one form of the fluctuation-dissipation theorem.[3]

The primary purpose of this paper is to show that such an analysis may be carried out in general relativity. Specifically, quantum matter fields in curved spaces which possess a timelike Killing field are considered (Sec. II). The existence of the Killing field $\partial/\partial t$ is necessary in order that some concept of time independence and equilibrium remains intact in the full general relativistic setting. The thermal Green's functions are uniquely defined by their regularity on the Euclidean section, $t \to it$ (see Fig. 1), and vanishing at spacelike infinity (if the space is not compact). Corresponding to this thermal Green's function is a state of the quantum field(s) or, more precisely, a thermal density matrix. We may now ask how the expectation value of a physical quantity in this state, such as the energy-momentum tensor T_{ab}, changes as the background metric is changed. That is, we consider the (linear) response of $\langle T_{ab}(x)\rangle$ to an external perturbation $\delta g_{cd}(x')$, to first order in the perturbation. The function (of x and x') that relates the two is the polarization tensor $\Pi_{ab}{}^{cd}(x,x')$. Because of the analyticity properties of the thermal Green's functions, this polarization tensor enjoys all of the same properties as its nonrelativistic analogs. The relation between the fluctuations about equilibrium in the thermal state and the dissipation under the influence of the external perturbation then follows immediately (Sec. III). Just as in nonrelativistic examples, of which the prototype is Brownian motion, this means that the random fluctuations about equilibrium and the systematic damping effect must derive from the *same* underlying physics.

The dissipation is associated with the possibility of real particle production by the background metric, when time-asymmetric boundary conditions are imposed on the matter field(s). Thus, the close connection between quantum fluctuations in the matter field (or thermal fluctuations in the Hawking temperature of the horizon) on the one hand, and black-hole radiance and decay on the other, becomes manifest. These are general features of quantum fields in curved space of which the black hole is only one familiar prototype. The quantum vacuum in an external field (gravitational or not) generally behaves as a *dissipative* as well as a fluctuating medium.

Furthermore, the spectral functions appearing in the linear response analysis provide a very powerful tool for analyzing the *stability* of the spacetime under the quantum (thermal) matter fluctuations. In Sec. IV we consider the fluctuation $\delta\langle T_{ab}\rangle$ as a *source* for the linearized semi-

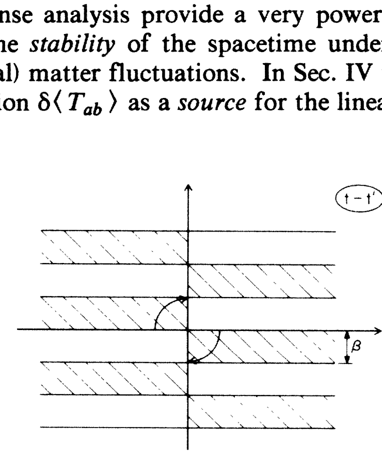

FIG. 1. The complex t-t' plane. The shaded strips of width β are the regions in which the real-time thermal Green's functions and polarization operator Π are analytic.

classical Einstein equations. Since this source is nonlocal, i.e., $\delta\langle T_{ab}(x)\rangle$ depends on $\delta g_{cd}(x')$ for $x\neq x'$, we obtain a linear integro-differential equation for the dynamic metric fluctuations. This equation automatically incorporates the correct boundary conditions on the metric perturbations away from the equilibrium state; so, we expect no unphysical "runaway" modes. From the previous fluctuation-dissipation theorem any unstable fluctuation uncovered by the linear response analysis must be connected with the possibility of real particle creation. That this is indeed the case is also shown in Sec. IV.

The second purpose of this paper is to apply this general stability analysis to de Sitter space. In two previous papers,[4,5] two very different arguments were given demonstrating that de Sitter space is unstable in the presence of quantized matter fields. In the first paper (I) $|\,\text{in}\rangle$ and $|\,\text{out}\rangle$ states were defined, by which the Bogoliubov transformation and associated particle creation effects could be calculated. Just as in the case of black-hole radiance, the effect of this process is to decrease the background curvature: the *effective* cosmological "constant" (i.e., the scalar curvature) decreases monotonically as the coherent vacuum energy is radiated into matter modes. In paper II, the Euclidean or thermal vacuum state was also shown to be unstable. In this case the instability manifests itself in the macroscopic fluctuations of the Hawking—de Sitter temperature. This quantum-thermal instability of de Sitter space was suggested as a physical mechanism by which the cosmological constant could be screened by quantum matter, i.e., by any positive vacuum energy being dissipated into matter-field modes.

Detailed analysis of the vacuum polarization tensor $\Pi^{ab}{}_{cd}(x,x')$ in de Sitter space shows that the unstable mode does indeed exist. (See Fig. 2.) The time scale for the instability to develop is estimated and agrees with paper I. This confirms and, together with the fluctuation-dissipation theorem, clarifies the relationship between the previous calculations in papers I and II: de Sitter space is unstable to fluctuations which, as they grow, dissipate more and more vacuum energy into matter-field modes.

If the very early Universe ever entered a de Sitter phase, it would have exited from it by particle creation and entropy generation on an enormous scale. This means that an inflationary scenario is possible which suffers from none of the usual objections: there are no gross inhomogeneities in the form of bubble walls, monopoles are produced in negligible numbers if at all and the vacuum energy relaxes to zero without any fine-tuning. These implications of the present work are discussed in the final section.

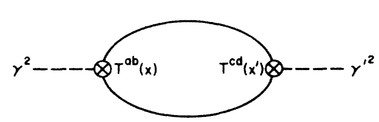

FIG. 2. The one-loop vacuum polarization graph which represents Π in a noninteracting scalar field theory.

II. LINEAR RESPONSE IN CURVED SPACETIME

In this section the necessary preliminaries of the linear-response method are defined and discussed. Let $\bar{g}_{ab}$ be some solution to Einstein's equations, with or without a cosmological constant. If all curvature and mass scales are well below the Planck scale, then a semiclassical approximation should be possible. That is, the metric may be treated as a classical variable responding to the quantum fluctuations of all other fields through the semiclassical Einstein equation:

$$G_{ab}+\Lambda g_{ab}=R_{ab}-\tfrac{1}{2}g_{ab}R+\Lambda g_{ab}=8\pi\langle T_{ab}\rangle \ . \tag{2.1}$$

Actually, the spin-2 fluctuations of the metric itself contribute just as much as any other matter field. However, it is convenient to consider a scalar matter field as a test field and source of the gravitational field. Formally, this can be justified by taking the large-N limit of a theory with N identical scalar fields. This imposes no essential restriction, as fields of any spin may also be considered.

The matter stress-energy operator appearing in (2.1) is defined by

$$T_{ab}=\frac{2}{\sqrt{-g}}\frac{\delta}{\delta g^{ab}}S^{\text{ren}} \ , \tag{2.2}$$

where S^{ren} is the *fully renormalized* quantum effective action operator of the matter field(s) in the curved background under consideration. In addition to renormalization of the masses and coupling constants, specifying S^{ren} generally requires further subtractions to define the composite operator insertions (such as Φ^2 or Φ^4 in a scalar-field theory) which appear in T_{ab}. Any convenient covariant scheme may be employed to isolate the divergences. The dimensional regularization method has been discussed by several authors.[6] Because the divergences have a purely local structure, they may be accounted for by a finite set of terms involving the background metric and its derivatives. The Schwinger-DeWitt short-distance expansion is most useful for characterizing these local divergences and bringing out the physical meaning of the renormalization counterterms.[7] Since T_{ab} and S^{ren} are dimension-four operators, these counterterms will involve up to four metric derivatives:

$$c_0\Lambda+c_1R+c_2R^2+c_3C_{abcd}C^{abcd} \ . \tag{2.3}$$

The first two terms correspond to renormalization of the cosmological constant and Newtonian constant [set equal to unity in Eq. (1)]. The fourth-order terms, proportional to the squares of the Ricci scalar and Weyl tensor C_{abcd}, should also appear in the effective action for the matter field with arbitrary dimensionless coefficients. Taking Eq. (2.1) as the equation of motion with *no* fourth-order kinetic terms amounts to setting the renormalized coefficients, c_2 and c_3, to zero. Equivalently, we could restrict ourselves to configurations where the higher-order *kinetic* terms vanish. As soon as time-dependent dynamics is allowed to develop, matter excitations are produced and the higher derivatives in these fourth-order terms have the potential for totally changing the character of the evolution of the spacetime. As the aim in this paper is to approach this radiation reaction effect in a systematic way, but without tackling the full back-reaction problem in a quantum theory of gravity, we will restrict our attention to *static* metric backgrounds, $\bar{g}_{ab}$. Specifically, we will assume that the background spacetime possesses a timelike Killing field $\partial/\partial t$. Then the notion of a time-independent, time-reversal-invariant "vacuum" state can be defined in a precise way.[8] No particle production or decay in the usual sense can occur in this state. $\langle T_{ab}\rangle$ may be evaluated and the size of the order R^2 terms (with finite coefficients of order unity) in (2.3) estimated. As long as the curvature is small compared to the Planck scale, $\langle T_{ab}\rangle$ contains *only time-independent* vacuum polarization effects which are *numerically small* compared to the left-hand side of (2.1). In this special situation of a timelike Killing field, then, it makes no essential difference whether one takes $\bar{g}_{ab}$ to be a solution to the full semiclassical equations (with c_2 and c_3 nonzero) or simply the *vacuum* Einstein equations with $\langle T_{ab}\rangle=0$. In the absence of a Killing time, this certainly would not be the case.

Also, since the system is assumed static (in the mean) with respect to the Killing-time coordinate, the possible competing suggestions for defining a finite $\langle T_{ab}\rangle$ satisfying Wald's axioms[9] all yield the same result: the $|\text{in}\rangle$ and $|\text{out}\rangle$ states are one and the same in static equilibrium, where no particle production is possible.

The only other point about the renormalization procedure that bears mentioning here is that the strictly massless conformal limit of the field theory is delicate. This is clear from the renormalization-group equations and/or trace anomaly derived by means of dimensional continuation methods and is even more obvious from the Schwinger-DeWitt expansion which involves *inverse* powers of the mass.[7]

Consider now a small perturbation in the static metric:

$$g_{ab}=\bar{g}_{ab}+\delta g_{ab}\ . \tag{2.4}$$

For the moment the perturbation (which need *not* be static) may be taken to be one induced by an external agency, i.e., a nondynamic source term in the Lagrangian. Later, our interest will be in the dynamic fluctuations induced by the matter itself. In either case, the variation in the stress tensor may be expressed in the form

$$\delta\langle T_{ab}(x)\rangle=-\tfrac{1}{4}M_{ab}{}^{cd}\delta g_{cd}(x)+\tfrac{1}{2}\int d^4x'(-g')^{1/2}\Pi_{ab}{}^{cd}(x,x')\delta g_{cd}(x')\ . \tag{2.5}$$

The above notation has been used in writing

$$\langle\delta T_{ab}(x)\rangle=\int d^4x'\left\langle\frac{\delta T_{ab}(x)}{\delta g_{cd}(x')}\right\rangle\frac{\delta g_{cd}(x')}{2}=-\tfrac{1}{4}M_{ab}{}^{cd}\delta g_{cd}(x) \tag{2.6}$$

for the purely local part of the variation, which records the explicit dependence of T_{ab} on the metric, and

$$\Pi_{ab}{}^{cd}(x,x')=i\langle\mathcal{T}T_{ab}(x)T^{cd}(x')\rangle_{\text{con}} \tag{2.7}$$

for the nonlocal quantum polarization tensor. The definition of T_{ab} in (2), in terms of the renormalized operator insertion S^{ren}, is sufficient to define a renormalized polarization tensor in (2.7) whose local divergences as $x\to x'$ have been removed by precisely the same set of counterterms (2.3) needed to define S^{ren} itself.[10]

The time-ordered product of $\Pi_{ab}{}^{cd}$ deserves some comment. Formally, it results from the definition of the Schwinger variational principle, corresponding to the boundary conditions of the Feynman propagator. That is, it is an "in-out" matrix element which contains both advanced and retarded (particle and antiparticle) effects. Since we have in mind a time-independent state, $|\text{in}\rangle$ is the same as $|\text{out}\rangle$ (up to a trivial overall phase) and these Feynman boundary conditions are precisely the correct ones for describing the dynamical fluctuations about the time-independent equilibrium state. As soon as we wish to study dissipation, however, the time symmetry must be broken by choosing, for example, retarded Green's functions in place of the Feynman propagator. The perturbation's effects are felt only in the past and $|\text{in}\rangle$ is not necessarily the same as $|\text{out}\rangle$. This will be discussed in detail in the next section.

The local term of (2.6) contains all of the local renormalization counterterm contributions from (2.3). It is most conveniently dealt with by relating it to the real part of the polarization tensor via[18]

$$\begin{aligned}M_{ab}{}^{cd}\delta g_{cd}(x)=2(&-g_{ab}g^{ce}g^{df}+\delta^c_b g^{de}\delta^f_a\\&+\delta^d_b g^{cf}\delta^e_a)\langle T_{ef}\rangle\delta g_{cd}(x)\\&+\text{Re}\int d^4x'(-g')^{1/2}\Pi_{ab}{}^{cd}(x,x')\\&\quad\times\delta g_{cd}(x')\ .\end{aligned} \tag{2.8}$$

Because of the assumption of a timelike Killing field $\partial/\partial t$, $\Pi_{ab}{}^{cd}(x,x')$ depends on the corresponding static time coordinate only through the difference t-t'.

The Killing-time coordinate also implies the existence of a conserved Hermitian generator of time translations on the fields, i.e., a Hamiltonian H. For concreteness a free scalar field Φ may be considered, although our results are easily generalized to fields of any spin. The Hamiltonian which satisfies

$$\frac{\partial}{\partial t}\Phi=i[H,\Phi] \tag{2.9}$$

may be presented in the form

$$H=\int_\Sigma d\Sigma^b K^a T_{ab}\ , \tag{2.10}$$

where K^a is the timelike Killing field and Σ is a spacelike surface, everywhere normal to K^a. In the interesting case of spacetimes with horizons, the integral in (2.10) extends only over the region where K^a is timelike and future directed. Since this surface does *not* generally represent a complete Cauchy surface for the fields, any state specified on Σ alone must be a mixed state. If a complete set of modes on Σ is introduced (with *zero flux* through the horizon), as in Ref. 11, operating on which H has a real eigenspectrum, then a thermal density matrix $e^{-\beta H}$ may be introduced and the thermal ensemble average of the bilinear operator $\mathscr{T}\Phi(x)\Phi(x')$ defined in the usual way:

$$\langle \mathscr{T}\Phi(x)\Phi(x')\rangle_\beta = \frac{\mathrm{Tr}[\mathscr{T}\Phi(x)\Phi(x')e^{-\beta H}]}{\mathrm{Tr}(e^{-\beta H})} . \tag{2.11}$$

Gibbons and Perry[11] show that the Green's functions defined on the Euclidean section, $t \to it$, correspond precisely to Eq. (11) provided that β is identified with the periodicity of the Euclidean metric. Back in real Lorentzian time, the Green's functions so defined are equivalent to filling the stationary states of the field, $H|\omega\rangle = \omega|\omega\rangle$, with a Bose-Einstein (Fermi-Dirac) distribution at temperature β^{-1},

$$n_{B(F)} = \frac{1}{e^{\omega\beta} \mp 1} , \tag{2.12}$$

if the fields have integral (half-integral) spin.

The analytic properties of these thermal Green's functions have also been discussed in Ref. 11. They are analytic in the shaded strips of the complex t-t' plane illustrated in Fig. 1. Because of the (real) time-ordering operation in both Eqs. (2.7) and (2.11), these *same* analyticity properties are shared by the polarization tensor $\Pi_{ab}{}^{cd}(x,x')$. We could prove this claim directly, by working with the definition of $\Pi_{ab}{}^{cd}(x,x')$ through Eqs. (2.7) and (2.11) in the thermal state, and then use the cyclic property of the trace in the standard way.[3,12] This analyticity of $\Pi_{ab}{}^{cd}$ in the complex t-t' plane is the single most important technical ingredient to our subsequent analysis.

As in the nonrelativistic case[12] it is convenient to separate $\Pi_{ab}{}^{cd}$ into two pieces, depending on the time ordering:

$$\Pi_{ab}{}^{cd}(x,x') = \theta(t-t')\Pi_{ab}^{>cd}(x,x') + \theta(t'-t)\Pi_{ab}^{<cd}(x,x') . \tag{2.13}$$

Viewed as matrices in (ab,x) and (cd,x') these functions are anti-Hermitian

$$\Pi_{ab}^{>,<cd}(x,x') = -\Pi^{>,<cd}{}_{ab}(x',x)^* \tag{2.14}$$

and pure imaginary positive definite

$$\frac{1}{i}\Pi_{ab}^{>,<cd}(x,x') \geq 0 \tag{2.15}$$

in the sense that

$$\int d^4x(-g)^{1/2} \int d^4x'(-g')^{1/2} \times f^{ab}(x)\left[\frac{1}{i}\Pi_{ab}^{>,<cd}(x,x')\right] f_{cd}(x') \geq 0$$

for any smooth normalizable function $f_{ab}(x)$. Since the two functions $\Pi^<$ and $\Pi^>$ are really the same function evaluated at different arguments, i.e.,

$$\Pi_{ab}^{>cd}(x,x') = \frac{i}{Z}\mathrm{Tr}[e^{-\beta H}T_{ab}(x)T^{cd}(x')], \quad t > t' ,$$

$$\Pi_{ab}^{<cd}(x,x') = \frac{i}{Z}\mathrm{Tr}[e^{-\beta H}T^{cd}(x')T_{ab}(x)], \quad t < t' , \tag{2.16}$$

$$Z = \mathrm{Tr}(e^{-\beta H}) ,$$

we have

$$\Pi_{ab}^{<cd}(x,x') = \Pi^{>cd}{}_{ab}(x',x) \tag{2.17}$$

which in virtue (2.14) becomes

$$\Pi_{ab}^{<cd}(x,x') = -\Pi_{ab}^{>cd}(x,x')^* . \tag{2.18}$$

The Fourier transforms are now introduced in the usual way:

$$\Pi_{ab}^{>,<cd}(x,x') = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} F_{ab}^{>,<cd}(\mathbf{r},\mathbf{r}';\omega) , \tag{2.19}$$

where we have used $\mathbf{r},\mathbf{r}'$ for the spatial components of x,x'. Properties (2.14)–(2.18) translate into the corresponding statements for the Fourier transforms:

$$F_{ab}^{>,<cd}(\mathbf{r},\mathbf{r}';\omega) = -F^{>,<cd}{}_{ab}(\mathbf{r}',\mathbf{r};\omega)^* , \tag{2.20}$$

$$\frac{1}{i}F_{ab}^{>,<cd}(\mathbf{r},\mathbf{r}';\omega) \geq 0 , \tag{2.21}$$

$$F_{ab}^{<cd}(\mathbf{r},\mathbf{r}';\omega) = F^{>cd}{}_{ab}(\mathbf{r}',\mathbf{r};-\omega) = -F_{ab}^{>cd}(\mathbf{r},\mathbf{r}';-\omega)^* . \tag{2.22}$$

Finally, the relation exhibiting the periodicity in imaginary time, which follows from the trace structure of (2.16) is

$$\Pi_{ab}^{<cd}(\mathbf{r},\mathbf{r}';t-t') = \Pi_{ab}^{>cd}(\mathbf{r},\mathbf{r}';t-t'-i\beta) \tag{2.23}$$

or

$$F_{ab}^{<cd}(\mathbf{r},\mathbf{r}';\omega) = e^{-\omega\beta}F_{ab}^{>cd}(\mathbf{r},\mathbf{r}';\omega) . \tag{2.24}$$

Another interesting quantity is the anticommutator function

$$\langle [T_{ab}(x), T^{cd}(x')]_+ \rangle_\beta = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} e^{-i\omega(t-t')} [F_{ab}^{<cd}(\mathbf{r},\mathbf{r}';\omega) + F_{ab}^{>cd}(\mathbf{r},\mathbf{r}';\omega)] . \tag{2.25}$$

Its Fourier transform $S_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)$ obeys

$$S_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega) = S_{ab}{}^{cd}(\mathbf{r}',\mathbf{r};-\omega) \geq 0 \tag{2.26}$$

and

$$S_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega) = \frac{1}{i}(1+e^{-\omega\beta})F_{ab}^{>cd}(\mathbf{r},\mathbf{r}';\omega) = \frac{1}{i}(1+e^{\omega\beta})F^{<cd}_{ab}(\mathbf{r},\mathbf{r}';\omega) . \tag{2.27}$$

The Fourier transform of the commutator is then given by

$$D_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega) = \frac{1}{i}[F_{ab}^{<cd}(\mathbf{r},\mathbf{r}';\omega) - F_{ab}^{>cd}(\mathbf{r},\mathbf{r}';\omega)] = \tanh\left[\frac{\beta\omega}{2}\right] S_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega) . \tag{2.28}$$

In terms of $S_{ab}{}^{cd}$ the polarization tensor becomes

$$\Pi_{ab}{}^{cd}(x,x')=\frac{i}{2}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega(t-t')}\times\left[\epsilon(t-t')\tanh\left[\frac{\beta\omega}{2}\right]+1\right]\times S_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)\ . \tag{2.29}$$

The $\epsilon(t-t')$ discontinuous function may be represented as

$$\epsilon(t-t')=\frac{1}{i}\int\frac{d\omega'}{2\pi}e^{i\omega'(t-t')}\left[\frac{1}{\omega'-i\eta}+\frac{1}{\omega'+i\eta}\right]=\frac{2}{i}P\int\frac{d\omega'}{2\pi}\frac{e^{i\omega'(t-t')}}{\omega'} \tag{2.30}$$

as $\eta\to 0$, so that (29) yields the Fourier transform

$$\Pi_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)=\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\left[P\frac{1}{\omega'-\omega}+i\pi\coth\left[\frac{\beta\omega}{2}\right]\times\delta(\omega'-\omega)\right]\times\tanh\left[\frac{\beta\omega'}{2}\right]S_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega')\ . \tag{2.31}$$

If instead of the Feynman time-ordered boundary conditions we have considered purely retarded boundary conditions, then $\epsilon(t-t')$ is replaced by $2\theta(t-t')$ and there is no purely symmetric piece in the polarizability tensor. For example,

$$\Pi_{ab}^{(\mathrm{ret})cd}(\mathbf{r},\mathbf{r}';\omega)=\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\left[P\frac{1}{\omega'-\omega}+\pi i\delta(\omega'-\omega)\right]\times D_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega')\ . \tag{2.32}$$

This completes the development of the formal machinery of linear response theory and the properties of the polarization tensor along precisely the same lines of nonrelativistic many-body theory.

III. THE FLUCTUATION-DISSIPATION THEOREM

The thermal equilibrium state defined by continuation from the Euclidean section of the metric is automatically time symmetric. The fluctuations in the energy-momentum tensor in this state are given directly by the symmetric product correlation function defined in Eqs. (2.25)–(2.27). This function describes the lowest-order deviations of T_{ab} from its average value in the equilibrium state. Thus, it contains information about the fluctuations away from the mean as given by Eqs. (1.1), in this state.

In the general case this state of thermal equilibrium may not be the one in which we are most interested. For instance, initial plus spatial boundary conditions may be specified instead, by which we would like to calculate the real-time evolution of the system. In such cases the full geometry cannot be truly static. Nevertheless, an *approximate* timelike Killing field may exist over a large part of the spacetime. An example is that of a collapsing star leaving behind the exterior Schwarzschild (or Kerr-Newman) geometry. The gravitational background is time independent to lowest order in $\hbar$ but the quantum fields obey certain time-reversal-*non*invariant boundary conditions. The effects of quantum fluctuations can now include dissipation as well, in the sense that real particle creation can occur through Bogoliubov mixing of positive and negative frequencies at late times, if no particles are present at early times. It is this dissipative side of the quantum fluctuations we now wish to consider.

In nonrelativistic statistical mechanics it is the antisymmetric or commutator function $D_{ab}{}^{cd}$ that describes the dissipative effects. To show that this is the case here as well requires that we relate the rate of increase of matter entropy due to the particle production processes to this function $D_{ab}{}^{cd}$. This is most easily accomplished by again considering the linear response equation (2.5), this time with time-*asymmetric* boundary conditions. Physically, we wish to consider the situation in which the background field $\bar{g}_{ab}$ has been turned on adiabatically from flat space in the infinite past by a sequence of small perturbations δg_{ab}. It is only by introducing time asymmetry in this way that particle production (i.e., instability of the adiabatic vacuum to pair creation) and the accompanying generation of entropy is possible in quantum field theory.

With these retarded boundary conditions the response of the energy-momentum tensor to a change in the metric is again given by Eq. (2.5) with $\Pi_{ab}^{(\mathrm{ret})cd}$ replacing $\Pi_{ab}{}^{cd}$. Introducing the Fourier transform (2.32) and differentiating with respect to the static time t gives

$$\frac{\partial}{\partial t}\delta\langle T_b^a(x)\rangle=\tfrac{1}{2}\int d^4x'(-g')^{1/2}\int\frac{d\omega}{2\pi}(-i\omega)e^{-i\omega(t-t')}\int\frac{d\omega'}{2\pi}\frac{D^a{}_b{}^{cd}(\mathbf{r},\mathbf{r}';\omega')}{\omega'-\omega-i\eta}\delta g_{cd}(\mathbf{r}',t')\ . \tag{3.1}$$

The contact term $-\frac{1}{4}M^a{}_b{}^{cd}\delta g_{cd}(x)$ in fact, will not contribute to this variation if we consider adiabatic changes in the background, for which δg_{cd} is essentially independent of t:

$$\int dt' e^{i\omega t'}\delta g_{cd}(\mathbf{r}')=2\pi\delta(\omega)\delta g_{cd}(\mathbf{r}')\ . \tag{3.2}$$

Then Eq. (3.1) simplifies enormously, since

$$\tfrac{1}{2}\int \frac{d\omega}{2\pi}(-i\omega)\delta(\omega) \times \int d\omega' \int_{\Sigma} K'^{e} d\Sigma'_{e} D^{a}{}_{b}{}^{cd}(\mathbf{r},\mathbf{r}';\omega') \times \left[P\left[\frac{1}{\omega'-\omega}\right] + i\pi\delta(\omega'-\omega)\right]\delta g_{cd}(\mathbf{r}')$$

receives a contribution only from the $i\pi\delta(\omega'-\omega)$ term and (3.1) becomes

$$\delta\frac{\partial}{\partial t}\langle T^{a}_{b}(x)\rangle = \tfrac{1}{4}\lim_{\omega\to 0}\int_{\Sigma} K'^{e} d\Sigma'_{e}[\omega D^{a}{}_{b}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)] \times \delta g_{cd}(\mathbf{r}') \ . \tag{3.3}$$

Setting $a=b=0$ and

$$\langle T^{0}_{0}\rangle = \rho \tag{3.4}$$

gives

$$\delta\left[\frac{\partial}{\partial t}\rho\right] = \tfrac{1}{4}\lim_{\omega\to 0}\int_{\Sigma} K^{b'} d\Sigma'_{a}[\omega D^{a}{}_{b}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)] \times \delta g_{cd}(\mathbf{r}') \ . \tag{3.5}$$

This equation shows that in the adiabatic static field limit matter creation takes place if and only if $\omega D^{a}{}_{b}{}^{cd}(\omega)$ is finite as $\omega \to 0$. This is a general criterion for particle creation in a static background. Thus as long as the creation rate is slow enough to justify the relevance of the adiabatic switching on of the background field, the boundary conditions do not need to be rederived for each different background configuration. We simply need to check whether or not $D^{a}{}_{b}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)$ develops a singular $1/\omega$ behavior as $\omega \to 0$. In this adiabatic method the actual creation rate in the static background is then given by Eq. (3.5) by integrating up the linear response from flat space to the desired background; if the metric is related to flat space by variation of a one-parameter sequence of metrics, labeled λ ($\lambda = M$ in the Schwarzschild case), then

$$\frac{\partial\rho}{\partial t} = \int_0^{\lambda} d\lambda \frac{\partial}{\partial\lambda}\left[\frac{\partial\rho}{\partial t}\right] = \int_0^{\lambda}\frac{d\lambda}{4}\lim_{\omega\to 0}\int_{\Sigma} K^{b'} d\Sigma'_{a}[\omega D^{a}{}_{b}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)]\frac{\delta g_{cd}(\mathbf{r}')}{\delta\lambda} \tag{3.6}$$

gives the full particle creation rate in adiabatic approximation.

We have yet to relate this particle creation rate to entropy production. If the matter is produced with a Planck spectrum, however, this relation is immediate. The appearance of a local temperature θ for the matter is simply a reflection of the noncompleteness of the three-surface Σ, viewed as a Cauchy surface for the time evolution of the fields. Boulware[13] had shown previously that corresponding to the Hawking temperature of emission from a black hole there is also a matter entropy as soon as the pure quantum states of the full system are converted to mixed states by summing over inaccessible degrees of freedom behind the event horizon. Recently, the general requirements for a Planck spectrum and mixed state density matrix formulation have been investigated.[14] The noncompleteness of Σ and thermal aspects of the emission process should be expected whenever a timelike Killing field K^a exists whose Euclidean orbits are periodic. The periodicity β is the inverse Hawking temperature of the emission process in adiabatic approximation. The usual Tolman red-shift formula would then be expected to relate the local temperature to the Hawking temperature. The local entropy production in this near thermal equilibrium state is then related to (3.6) by

$$\frac{\partial\rho}{\partial t} = \theta\frac{\partial s}{\partial t} \tag{3.7}$$

which is just the first law of thermodynamics for the matter fluid.

Thus, the commutator function $D^{a}{}_{b}{}^{cd}$ does indeed describe the dissipative effects of particle creation through Eqs. (3.6) and (3.7) while the symmetric function $S^{a}{}_{b}{}^{cd}$ describes the fluctuations about the time-symmetric state. Thus, the fluctuations in the time-symmetric equilibrium state, described by the symmetric function $S_{ab}{}^{cd}$ and the dissipation of gravitational field energy into matter, described by Eqs. (3.6) and (3.7) are really two different aspects of the *same* quantum-thermal matter fluctuations.

The relationship between the two is summarized in Eqs. (2.25)–(2.31). Specifically,

$$\tfrac{1}{2}S_{ab}{}^{cd}(x,x') = \tfrac{1}{2}\langle\{T_{ab}(x),T^{cd}(x')\}\rangle_{\beta} = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega(t-t')}\left[\frac{1}{2}+\frac{1}{e^{\beta\omega}-1}\right] \times D_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega) \tag{3.8}$$

is the desired fluctuation-dissipation theorem for quantum matter in curved spaces possessing a timelike Killing symmetry. Moreover, the positivity properties of the fluctuations, Eq. (2.26), can be used to derive results about the *direction* in which the particle creation drives the geometry. The tendency of the quantum particle creation to *decrease* the background curvature parameter, i.e., dissipate the coherent gravitational field into particle degrees of freedom (as in the Schwarzschild case), is actually a much more general phenomenon. However, the analysis involves the decomposition of $\Pi_{ab}{}^{cd}$ into scalar and tensor parts which must be treated independently, and will be reserved for a future publication.

IV. COSMOLOGICAL VACUUM INSTABILITY AND DECAY

The formalism developed in the preceding sections describes the Gaussian stochastic fluctuations of quantum matter in curved background fields. The polarization tensor gives the linear response of the energy-momentum tensor to a gravitational perturbation δg_{ab} through Eqs.

(2.4)–(2.7). This perturbation was assumed nondynamic and external in Sec. II. However, because of the semiclassical Einstein equations (2.1), fluctuations in the matter stress tensor must induce fluctuations in the metric as well. It should be emphasized that these metric fluctuations are purely slaves to the matter fluctuations in the semiclassical theory through (2.1), whereas in a full quantum theory of gravitation, as independent degrees of freedom they would also have to be allowed to fluctuate independently. So, we may view this semiclassical analysis as coming to the threshold of quantum gravity without actually broaching it.

Once matter and/or metric fluctuations are admitted into the discussion, we can use our linear response formalism to analyze the stability of the thermal equilibrium state to these quantum disturbances. We simply regard the linear response $\delta\langle T_{ab}\rangle$ (2.5) as a source for the semiclassical Einstein equations linearized about the given solution $\bar{g}_{ab}$. The resulting *linear* equation for δg_{ab},

$$\delta G_{ab}+\Lambda\delta g_{ab}=8\pi\delta\langle T_{ab}\rangle\ , \tag{4.1}$$

can now be analyzed by standard methods. Because of the timelike Killing field of $\bar{g}_{ab}$ it is useful to introduce a Fourier decomposition of the fluctuations and write

$$\delta g_{ab}(x)=\mathrm{Re}[e^{-i\omega t}h_{ab}(\mathbf{r})]\ . \tag{4.2}$$

Stability is tested by demanding that there be no solutions of (1) with $\omega^2<0$. On the other hand, when particle production and dissipation is possible, i.e., when the residue of $D_{ab}{}^{cd}(\mathbf{r},\mathbf{r}';\omega)$ at $\omega=0$ is nonzero, from Eq. (3.5), we expect an unstable mode to exist.

The polarization tensor is given in terms of the spectral functions by Eq. (2.31), which has both a real and an imaginary part. The integral over ω' in Eq. (2.31) may also be converted into an integral over ω'^2 from 0 to ∞, by using the symmetry property of the spectral function, Eq. (2.26). Then it is clear that the imaginary part of $\Pi_{ab}{}^{cd}$ is present only for $\omega^2>0$. We are interested primarily in any unstable modes with $\omega^2<0$. There the $i\pi\delta(\omega'-\omega)$ term in Eq. (2.31) does not contribute, the principal part prescription becomes superfluous and $\Pi_{ab}{}^{cd}$ is purely real. It is this real part which must be substituted into the right-hand side of Eq. (4.1) in order to test for the existence of unstable modes. This is the variation in T_{ab} away from its value in the equilibrium or Euclidean state in which all the analysis of Sec. II applies.

To illustrate the general method and discuss the relevance to the cosmological constant question, consider the case of de Sitter space in the coordinates:

$$ds^2=-dt^2(1-H^2r^2)+\frac{dr^2}{1-H^2r^2}+r^2d\Omega^2\ . \tag{4.3}$$

de Sitter space clearly has a timelike Killing field $\partial/\partial t$ in the region $0\le r\le H^{-1}$. The scalar curvature of the background is rigidly fixed classically, i.e., $R=4\Lambda=\text{const}$. With the quantum matter source in Eq. (4.1) this is no longer the case. Thus, the conformal mode, for which h_{ab} is proportional to g_{ab} itself is the interesting one. Tracing over the semiclassical Einstein equations (2.1) with the background de Sitter metric $\bar{g}_{ab}$ gives

$$\mathrm{Re}\delta\langle T_a{}^a(x)\rangle=\langle\delta T_a{}^a(x)\rangle+\tfrac{1}{2}\mathrm{Re}\int d^4x'(-g')^{1/2}\Pi_a{}^{acd}(x,x')\delta g_{cd}(x')\ . \tag{4.4}$$

The first term is purely real and local and may be handled exactly as in Eq. (2.8):

$$\langle\delta T_a{}^a(x)\rangle=-\tfrac{1}{4}\mathrm{Re}\int d^4x'(-g')^{1/2}\Pi_a{}^{acd}(x,x')\delta g_{cd}(x')\ . \tag{4.5}$$

Thus,

$$\mathrm{Re}\delta\langle T_a{}^a(x)\rangle=\tfrac{1}{4}\mathrm{Re}\int d^4x'(-g')^{1/2}\Pi_a{}^{acd}(x,x')\delta g_{cd}(x')\ . \tag{4.6}$$

Upon substituting (4.2), this can be conveniently rewritten in terms of the Fourier spectral functions introduced in Sec. II:

$$\mathrm{Re}\delta\langle T_a{}^a(x)\rangle=\frac{\mathrm{Re}}{4}\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}P\frac{1}{\omega'-\omega}e^{-i\omega t}\times\int_\Sigma(-d\Sigma_b'K^{b'})\times D_a{}^{acd}(\mathbf{r},\mathbf{r}';\omega)h_{cd}(\mathbf{r}')\ . \tag{4.7}$$

For the conformal mode, we may take $h_{cd}(\mathbf{r})=g_{cd}(\mathbf{r})$. Introducing the notation

$$\sigma(\omega^2)=\frac{1}{2\omega}\int_\Sigma(-d\Sigma_b'K^{b'})D_a{}^a{}_c{}^c(\mathbf{r},\mathbf{r}';\omega)\,|_{\mathbf{r}=0} \tag{4.8}$$

justified by the fact that $D_{ab}{}^{cd}$ is an odd function of ω, leads to

$$\begin{aligned}\mathrm{Re}\delta\langle T_a{}^a(x)\rangle\,|_{\mathbf{r}=0}&=\frac{\mathrm{Re}}{2}\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}P\frac{1}{\omega'-\omega}\omega'\sigma(\omega'^2)e^{-i\omega t}\\&=\mathrm{Re}P\int_0^\infty\frac{d\omega'}{2\pi}\frac{\omega'^2\sigma(\omega'^2)}{\omega'^2-\omega^2}e^{-i\omega t}\ .\end{aligned} \tag{4.9}$$

The left-hand side of the linearized Einstein equations evaluated at $\mathbf{r}=0$ is trivial: $\mathrm{Re}(\omega^2+4H^2)e^{-i\omega t}$, where $H^2=\Lambda/3$. Canceling the now irrelevant time-dependent phase factor results in the scalar linear response equation in Fourier space:

$$\omega^2+4H^2=P\int_0^\infty d\omega'\frac{\omega'^2\sigma(\omega'^2)}{\omega'^2-\omega^2}\equiv F(\omega^2)\ . \tag{4.10}$$

This is the desired linear response equation for spatially homogeneous scalar fluctuations in de Sitter space. Several aspects of Eq. (4.10) deserve comment. First, from the definition of σ in Eq. (4.8) and the relations (2.26) and (2.28) we see that σ is also a positive spectral function:

$$\sigma(\omega^2)\ge 0\ . \tag{4.11}$$

Thus, for $\omega^2<0$, where the principle part prescription in

(4.10) is unnecessary, the integral is always positive.

Second, ω^2 appears in the denominator of the integrand. This means that the integral *cannot* be written as a local expression of a finite number of time derivatives operating on the metric fluctuation (4.2). It is completely nonlocal, as also must have been evident from its form in position space, Eq. (4.6). The fact that Eq. (4.10) is the Fourier transform of an integro-differential equation for the dynamic metric fluctuations means that it already incorporates boundary conditions. These are just the time-symmetric boundary conditions of the time-ordered product which defines the polarization tensor (2.5). As in discussions of radiation reaction in electromagnetism the present treatment with boundary conditions included eliminates the many unphysical "runaway" modes that would be obtained if we tried explaining the denominator of Eq. (4.10) in powers of ω^2 and truncating at some finite order, for example.

Indeed, Eq. (4.10) makes it quite clear that when $H^2=0$ (flat space), there are *no* unstable modes of Eq. (4.10), for the left-hand side of the equation is negative while the right-hand side is positive for $\omega^2<0$. Thus flat space is *stable* under conformal fluctuations of the metric. Spatially inhomogeneous modes can only be more stable since then a negative $-\mathbf{k}^2$ appears on the left-hand side of Eq. (4.10) in addition to ω^2.

Usually one thinks of the physics of the scalar part of Einstein's equations as uninteresting, because the conformal modes do not freely propagate as do the transverse, trace-free spin-2 excitations. The conformal modes are constrained by the matter sources, much as in electromagnetism, the longitudinal modes of the electric field are constrained by the charge sources through Gauss's law. If the sources are fixed (nondynamic) nothing can happen in these constrained modes. As soon as the sources are themselves allowed to respond dynamically to changes in the metric (or electric potentials) as in our linear response analysis then these previously constrained degrees of freedom are quite physical. They are really *matter* degrees of freedom whose dynamics has been reexpressed as conformal excitations of the metric through the constraints imposed by the Einstein equations themselves.

Returning to the de Sitter case $H^2>0$, we can verify these statements explicitly. It appears that the linear equations would have an unstable mode even in the complete absence of matter, since $\omega^2=-4H^2$ is a solution of Eq. (4.10) with $\delta\langle T_{ab}\rangle=0$. However, this is a pure gauge mode as may easily be seen by calculating the curvature to linear order in the perturbation away from pure de Sitter space. There is *no* unstable conformal mode in the absence of matter since $R=4\Lambda$ is completely fixed. When the vacuum polarization of the matter on the right-hand side of (4.10) is turned on this gauge mode still exists, albeit at a slightly shifted value of ω^2, corresponding to the slightly different value of the effective cosmological constant induced by the matter.

The existence or nonexistence of a second *nontrivial* solution for $\omega^2<0$ depends on the behavior of $F(\omega^2)$ for small ω^2. In particular, suppose that the residue of $D_{ab}{}^{cd}$ at $\omega=0$ is nonzero. From Eqs. (2.28) and (4.8) this implies that

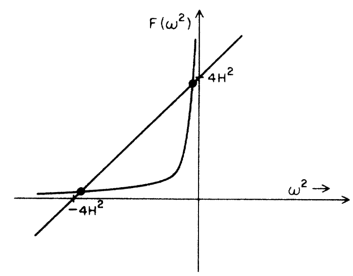

FIG. 3. The graphical solution of Eq. (4.10). The straight line is the left-hand side of (4.10) and the curve is the right-hand side, $F(\omega^2)$. The schematic form of $F(\omega^2)$ is inferred from the small-ω^2 behavior (4.13), and the positivity and the monotonic decrease of F as $\omega^2\rightarrow-\infty$. The solution at $\omega^2\cong-4H^2$ is a gauge mode as explained in the text, while the solution at small ω^2 is a physical unstable conformal mode of de Sitter space.

$$\sigma(\omega^2)\sim\int(-d\Sigma'_b K^{b'})S_a{}^a{}_c{}^c(\mathbf{r}=0,\mathbf{r}';\omega)\sim\frac{1}{\omega^2} \qquad (4.12)$$

as $\omega^2\rightarrow 0$.

This was the condition we required to have nonzero dissipation at late times under time-asymmetric boundary conditions in Sec. III. It is also the condition that Eq. (4.10) have a (nongauge) negative ω^2 solution, for, with the behavior of σ indicated in (4.12),

$$F(\omega^2)\sim\frac{A}{|\omega^2|^{1/2}} \quad \text{as } \omega^2\rightarrow 0 \qquad (4.13)$$

by inspection of the integral in (4.10). If $|\omega^2|\ll 4H^2$ we then obtain the approximate solution

$$\omega^2\cong=-\left[\frac{A}{4H^2}\right]^2 . \qquad (4.14)$$

The general qualitative behavior of the left- and right-hand sides of Eq. (4.10) is illustrated in Fig. 3, in the case $A\neq 0$.

To show that the spectral functions do indeed behave as indicated in (4.12) requires a closer look at their specific forms in de Sitter space. Consider the concrete case of a scalar matter field. Then the wave equation can be separated in coordinates (4.3):

$$\phi_{\omega lm}(x)\sim e^{-i\omega t}f_{\omega l}(r)Y_{lm}(\Omega) \qquad (4.15)$$

with

$$\left[\frac{-\omega^2}{1-H^2r^2}-\frac{1}{r^2}\frac{d}{dr}\left[(1-H^2r^2)r^2\frac{d}{dr}\right]+\frac{l(l+1)}{r^2}+M^2\right]f_{\omega e}(r)=0 . \qquad (4.16)$$

The modes are normalized with respect to the inner product

$$(u,v)=i\int_Z d\Sigma^a u^*\partial_a v =i\int d\Omega\int_0^{H^{-1}}\frac{dr\,r^2}{1-H^2r^2}u^*\frac{\partial}{\partial t}v . \qquad (4.17)$$

For the modes (4.15) the normalization is

$$(\phi_{\omega_1 l_1 m_1}, \phi_{\omega_2 l_2 m_2}) = \delta(\omega_1 - \omega_2)\delta_{l_1 l_2}\delta_{m_1 m_2} \ . \qquad (4.18)$$

Because of the time derivative in (4.17) this implies that

$$f_{\omega l}(r) \sim \frac{1}{\sqrt{|\omega|}} \quad \text{for } |\omega| \to 0 \ . \qquad (4.19)$$

If we now calculate

$$D_a{}^a{}_c{}^c(0,\mathbf{r}';\omega) = \int_{-\infty}^{\infty} dt' \langle [T_a{}^a(0), T_c{}^c(\mathbf{r}',t')]_- \rangle e^{-i\omega t}$$
$$\sim \mathrm{Im} \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 e^{-i(\omega+\omega_2-\omega)t'} f_{\omega_{10}}(r') f_{\omega_{20}}(r') f^*_{\omega_{10}}(0) f^*_{\omega_{20}}(0) \ , \qquad (4.20)$$

where we have used the fact that $f_{\omega e}(r) \sim r^l$ so that only $l = m = 0$ survives, when the argument $\mathbf{r} = 0$, we find

$$\int_\Sigma (-K'^b d\Sigma'_b) D_a{}^a{}_c{}^c(\mathbf{r},\mathbf{r}';\omega)\,|_{\mathbf{r}=0}$$
$$\sim \mathrm{Im} \int_{-\infty}^{\infty} \frac{d\omega_1}{\omega_1} \frac{1}{(\omega-\omega_1)} \sim \frac{1}{\omega} \quad \text{as } \omega \to 0 \qquad (4.21)$$

due to the implied $i\epsilon$ prescription on the ω integrands which allows completion of the integration contour in the complex plane. Thus, since the residue in Eq. (3.4) is indeed finite in de Sitter space, nontrivial particle creation and dissipative effects can occur and $\sigma(\omega^2)$ has precisely the behavior advertised in (4.12) necessary for the existence of an unstable conformal mode solution to Eq. (4.10).

It might still appear that the same argument applied in the case $H = 0$. This is not correct, however, of the mass threshold for pair creation. That is, in flat space

$$\phi^{(0)}_{\omega lm} \sim e^{i\omega t} j_l(kr) Y_{lm}(\Omega)$$

with

$$\omega^2 = \mathbf{k}^2 + M^2 \geq M^2 \ . \qquad (4.22)$$

Thus, the integration ranges in (4.20) do not include $\omega = 0$, and the coefficient A in (4.13) and (4.14) is identically zero. A zero-energy conformal metric variation cannot produce a massive particle pair in flat space. In de Sitter space the horizon at $r = H^{-1}$ means that even an indefinitely small ω variation corresponds to a large local energy perturbation near $r = H^{-1}$, i.e., $\omega^2/(1 - H^2 r^2) > M^2$, even for $\omega \to 0$ as $r \to H^{-1}$. The particle production, dissipative effects and the instability of de Sitter spacetime are all traceable ultimately to the existence of this horizon.

For the massive field, the coefficient A can be estimated by considering the behavior of the normalized solutions to Eq. (4.16) near the horizon. The solutions may be written in terms of hypergeometric functions as

$$f_{\omega e}(r) \sim (Hr)^l (1 - H^2 r^2)^{-(a+b)/2} F\left[a,b,c;\frac{1}{1 - H^2 r^2}\right] , \qquad (4.23)$$

where

$$a = \tfrac{1}{2}(\tfrac{3}{2} \pm i\gamma + l - i\omega/H) \ ,$$
$$b = \tfrac{1}{2}(\tfrac{3}{2} \pm i\gamma + l + i\omega/H) \ ,$$
$$c = 1 \pm i\gamma \ , \qquad (4.24)$$
$$\gamma = \left[\frac{M^2}{H^2} - \frac{9}{4}\right]^{1/2} .$$

By using the inversion transformation for the hypergeometric function,[15] the solutions regular at $r = 0$ become a linear combination of

$$\frac{\Gamma(1 \pm i\gamma)\Gamma\left[\pm \frac{i\omega}{2H}\right]}{\Gamma\left[\frac{3}{4} \pm \frac{i\gamma}{2} \pm \frac{i\omega}{2H}\right]\Gamma\left[\frac{1}{4} \pm \frac{i\gamma}{2} \pm \frac{i\omega}{2H}\right]}(1 - Hr)^{\pm i\omega/2H} \qquad (4.25)$$

near $r = H^{-1}$. Since

$$|\Gamma(1 \pm i\gamma)|^2 \sim \mathrm{csch}^2 \pi\gamma \sim e^{-M/T_H} \qquad (4.26)$$

as $M/H \to \infty$ the magnitude of the residue at $\omega = 0$, A decreases exponentially as M/H becomes large:

$$A \sim e^{-M/T_H}, \quad \frac{M}{T_H} \to \infty \qquad (4.27)$$

with $T_H = H/2\pi$ the Hawking–de Sitter temperature. In this limit the approximation leading to Eq. (4.14) is valid and the time scale for the instability of de Sitter space to develop is seen to be exponentially long compared to the expansion time H^{-1}. This agrees with the earlier estimate in paper I.

Finally, it is worth commenting on the relation of the present work to previous analyses of the stability issue in flat space. In the linear response method we have explicitly neglected the time derivatives coming from the possible higher-order ($R^2, R_{abcd}R^{abcd}$) terms generated by the one-loop polarization tensor $\Pi_{ab}{}^{cd}$. These terms *do* generate unstable modes when expanded around flat space due to the ghost degrees of freedom present in the higher derivative theory. The unstable ghost modes have frequencies of order M_{Planck} and must therefore be called into question: Quantum gravity might eliminate them entirely. The mode found with ω^2 given by Eq. (4.14) is quite distinct from this. It is a *low*-frequency mode [cf. Eqs. (4.14) and (4.27)] well within the range of validity of the semiclassical analysis, which explicitly contains no higher derivative interactions and therefore no ghosts. This low-frequency instability of de Sitter space is physical, as is the statement that flat space is stable semiclassically—the ghost modes lying beyond the range of applicability of the analysis. These latter modes are rather like the Landau ghost modes of one-loop electromagnetism which are also presumed unphysical.

V. THE COSMOLOGICAL CONSTANT PROBLEM

The cosmological constant in the present Universe is vanishingly small: $|\Lambda| < 10^{-122}$. The entropy, on the other hand, is enormous: $S > 10^{88}$. Inflationary models of the early Universe attempt to explain this large entropy by postulating a unified field theory of matter which undergoes a phase transition as the Universe cools through the transition temperature ($\sim 10^{15}$ GeV $\sim 10^{28}$ K).[16] The tiny cosmological term is not addressed at all in this kind of model.

The cosmological constant problem does not arise in flat-space field theory. It is only the gravitational field that weighs *all* energy, including cosmological vacuum energy that can give a nonzero Λ physical meaning through its curvature effect. The real problem, then, is *not* the value of Λ *per se*, but rather why the Universe we observe is so very nearly flat, even when quantum fluctuations and their associated zero-point energy are taken into account. That is, *why do vacuum fluctuations not curve spacetime?*

It might be said that this question cannot be tackled by present theory and that the issue must be laid at the door of quantum gravity—except that matter and/or radiation vacuum fluctuations are present even when spacetime is treated as a classical continuous manifold. So the issue of the gravitational effects of vacuum fluctuations is present long before Planck scales are reached. As long as the energy density is significantly smaller than the Planck scale it should then be feasible to approach this question within the context of the semiclassical theory, considered in this paper. Let us reconsider the cosmological constant problem, then, in the light of the semiclassical theory presented in the previous sections.

The effect of a positive cosmological energy density is to drive the Universe towards de Sitter space, classically.[17] However, the semiclassical methods of this and previous papers (I and II) show that de Sitter space is *unstable* in the presence of quantum matter. The analysis of Sec. IV relates the instability of the de Sitter vacuum to zero frequency matter fluctuations over one horizon volume. These are the *same* fluctuations which are responsible for the increase in matter energy and entropy density through Eq. (4.15).

The increase in ordinary matter energy density is clear enough. Particle excitations are created by the gravitational field—either by viewing it as a time-dependent scattering potential as in I, or because, in a time-independent gauge, the horizon acts as a potential barrier through which particle pairs may tunnel.[18] If the effect of this particle creation, to first order in $\hbar$, is to decrease the local scalar curvature, then a kind of gravitational Lenz law would be at work, in that particle creation in a gravitational field has the effect of decreasing the curvature, decreasing the creation rate still further. This effect has been noted previously.[19]

These observations have a profound implication for the cosmological constant problem. For they imply that when the quantum matter stresses in curved backgrounds are taken into account, there exists a dynamical mechanism for reducing the scalar curvature towards zero—*independently* of the value of Λ. de Sitter space is *not* a stable solution to the full semiclassical equations (2.1). Since this statement applies for all $\Lambda > 0$, it seems highly plausible that the only stable solution to Eqs. (2.1) is, in fact, flat space. Then the cosmological constant problem would "solve itself;" i.e., vacuum fluctuations do not curve spacetime, because if they did, vacuum dissipation would flatten it again, through particle creation effects. A definite arrow of time is distinguished by the imposition of physically reasonable time-asymmetric boundary conditions on otherwise time-reversal-invariant equations of motion. This gives a definite meaning to the ideas of "vacuum viscosity" which have from time to time appeared in the literature.[20]

Thus, one and the same mechanism, particle creation by gravitational fields may be responsible for *both* the enormous entropy of the present Universe *and* its approximate flatness. The cosmological constant may then be viewed as effectively screened by the quantum vacuum polarization of the matter fields. Its actual value is irrelevant since it does *not* imply a stable curved universe at the same scale. An extraordinarily large, flat universe such as the one we live in, with an *effectively zero* cosmological term, may then be the result of quantum vacuum dissipation actually taking place in a quasi–de Sitter phase of the very early Universe. Clearly what is needed is a detailed numerical investigation of the behavior of the semiclassical equations for long times, to determine whether quantitative agreement between an inflationary model based on these ideas and the observable Universe can be obtained.

ACKNOWLEDGMENTS

It is a pleasure to thank P. Mazur for useful discussions. This research was supported in part by the National Science Foundation under Grant No. PHY82-17853, supplemented by funds from the National Aeronautics and Space Administration.

[1]S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).
[2]J. M. Bardeen, B. Carter, and S. W. Hawking, Commun. Math. Phys. **31**, 181 (1973).
[3]H. B. Callen and T. A. Welton, Phys. Rev. **83**, 34 (1951); L. B. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1959).
[4]E. Mottola, Phys. Rev. D **31**, 754 (1985).
[5]E. Mottola, Phys. Rev. D **33**, 1616 (1986).
[6]L. S. Brown, Ann. Phys. (N.Y.) **126**, 135 (1980); L. S. Brown and J. C. Collins, *ibid.* **130**, 215 (1980); see also S. M. Christensen, Phys. Rev. D **14**, 2490 (1976).
[7]B. S. DeWitt, Phys. Rep. **19C**, 297 (1975); N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, 1982).
[8]G. W. Gibbons, Commun. Math. Phys. **44**, 245 (1975).
[9]R. M. Wald, Phys. Rev. D **17**, 1477 (1978).
[10]See, for example, O. Piguet and A. Rouet, Phys. Rep. **76**, 1 (1981). Section 1.3 is most relevant to the present work.

[11]G. W. Gibbons and M. J. Perry, Proc. R. Soc. London **A358**, 467 (1978).

[12]P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).

[13]D. G. Boulware, Phys. Rev. D **13**, 2169 (1976).

[14]T. D. Lee, Nucl. Phys. **B264**, 437 (1986).

[15]*Higher Transcendental Functions* (Bateman Manuscript Project), edited by A. Erdélyi (McGraw-Hill, New York, 1953), Vol. 1. See formula 2.14 (17), p. 63.

[16]A. Guth, Phys. Rev. D **23**, 374 (1981); R. Brout, F. Englert, J.-M. Frere, E. Gunzig, P. Nardone, and C. Truffin, Nucl. Phys. **B170**, 228 (1980). For a recent review see A. Linde, Rep. Prog. Phys. **47**, 925 (1984); R. Brandenberger, Rev. Mod. Phys. **57**, 1 (1985).

[17]W. Boucher and G. W. Gibbons, in *The Very Early Universe,* edited by G. W. Gibbons, S. W. Hawking, and S. T. C. Siklos (Cambridge University Press, Cambridge, 1983).

[18]E. Mottola (unpublished).

[19]B. L. Hu, Phys. Lett. **90A**, 375 (1982); **97A**, 368 (1983); L. Parker, Phys. Rev. **183**, 1057 (1969); Phys. Rev. Lett. **50**, 1009 (1983); in *Quantum Theory of Gravity,* edited by S. M. Christensen (Adam Hilger, Bristol, 1984).

[20]Ya. B. Zeldovich, Pis'ma Zh. Eksp. Teor. Fiz. **12**, 443 (1970) [JETP Lett. **12**, 307 (1970)]; D. W. Sciama, P. Candelas, and D. Deutsch, Adv. Phys. **30**, 327 (1981); J. B. Hartle, Phys. Rev. D **23**, 2121 (1983); L. Parker, Phys. Rev. Lett. **50**, 1009 (1983); B. L. Hu, in *Advances in Astrophysics,* edited by L. Z. Fang and R. Ruffin (World Scientific, Singapore, 1983), Vol. 1. Also see Ref. 13.