Quantum black hole entropy and Newton constant renormalization

J. L. F. Barbón*
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544

R. Emparan†
Departamento de Física de la Materia Condensada, Universidad del País Vasco, Apartado 644, 48080 Bilbao, Spain
(Received 1 March 1995)

We discuss the status of the black hole entropy formula $S_{\rm BH} = A_H/4G$ in low-energy effective field theory. The low-energy expansion of the black hole entropy is studied in a nonequilibrium situation: the semiclassical decay of hot flat space by black hole nucleation. In this context the entropy can be defined as an enhancement factor in the semiclassical decay rate, which is dominated by a sphaleronlike saddle point. We find that all perturbative divergences appearing in Euclidean calculations of the entropy can be renormalized in low-energy couplings. We also discuss some formal aspects of the relation between the Euclidean and Hamiltonian approaches to the one-loop corrections to black hole entropy and geometric entropy, and we emphasize the virtues of the use of covariant regularization prescriptions. In fact, the definition of black hole entropy in terms of decay rates *requires* the use of covariant measures and, accordingly, covariant regularizations in path integrals. Finally, we speculate on the possibility that low-energy effective field theory could be sufficient to understand the microscopic degrees of freedom underlying black hole entropy. We propose a qualitative physical picture in which black hole entropy refers to a space of quasicoherent states of infalling matter, together with its gravitational field. We stress that this scenario might provide a low-energy explanation of both the black hole entropy and the information puzzle.

PACS number(s): 04.70.Dy, 11.10.Gh, 97.60.Lf

I. INTRODUCTION

Black hole entropy has a neat phenomenological meaning. During the late stages of the collapse process in which a large black hole radiates thermally (i.e., according to Hawking's calculation [1]), the interaction of the black hole with the rest of the world occurs as if it had an effective density of states $\rho \sim \exp(A_H/4G)$, where A_H is the horizon area and G is the low-energy (renormalized) Newton constant. This is obtained from the Hawking temperature formula $T_H = (8\pi GM)^{-1}$ and the equilibrium equation $\partial S_{\rm BH}/\partial M = T_H^{-1}$.

A seemingly equivalent phenomenological derivation of black hole entropy due to 't Hooft [2] does not assume any thermal equilibrium boundary conditions (which are probably unphysical for real collapsing black holes). Instead, this derivation is based on a comparison of semiclassical absorption and emission rates and, in contrast, the main assumptions are those of unitarity and CPT invariance. The absorption cross section of the black hole would be proportional to the horizon area:

$$\sigma_{\rm in} \sim |H_I({\rm in})|^2 \rho_{\rm BH}(M) \sim A_H \sim (GM)^2 \,, \qquad (1)$$

where the interaction is described by a Hamiltonian H_I in some underlying quantum description. According to Hawking, the emission rate for particles of mass δM has a thermal profile:

$$\Gamma_{\rm out} \sim |H_I({\rm out})|^2 \rho_{\rm BH}(M-\delta M) \sim A_H e^{-\beta_H \delta M} \sim (GM)^2 e^{-8\pi GM\delta M} \,. \qquad (2)$$

Now, from CPT invariance and unitarity, $|H_I({\rm in})| = |H_I({\rm out})|$, and one obtains the semiclassical relation

$$\frac{\rho_{\rm BH}(M-\delta M)}{\rho_{\rm BH}(M)} \sim e^{-8\pi GM\delta M} \,, \qquad (3)$$

which is satisfied by $\rho_{\rm BH}(M) = e^{S_{\rm BH}} = \exp(4\pi GM^2 + C)$.

This phenomenological derivation relies on the Hawking emission formula, which is a low-energy result. Hence, one would expect that a microscopic description of the degrees of freedom leading to $S_{\rm BH}$ could be achieved in low-energy effective field theory, at least for large enough black holes, which accurately follow the Hawking radiation formula. Strikingly, the only such calculation of $S_{\rm BH}$ giving the correct result $A_H/4G$ regards the entropy as a purely classical entity, without any statistical interpretation [3]. This "intrinsic" entropy is at odds with the phenomenological notion presented above. Several proposals for an *ab initio* quantum construction of the Bekenstein-Hawking entropy have been put forward over the years (for a review see [4]). Notably, geometric (or entanglement) entropy has been proposed as responsible for all or part of $S_{\rm BH}$ [2,5].

Entanglement entropy [6,7] arises when the support of physical operators is conventionally restricted to a proper

*Electronic address: barbon@puhep1.Princeton.edu
†Electronic address: wmbemgar@lg.ehu.es

region of space, and finds its origin in quantum correlations across the boundary between both regions. It is a fully quantum object naturally scaling as the area, but it is ultraviolet divergent in field theory. Moreover, this divergence poses subtle conceptual questions regarding its physical interpretation. Perhaps one should operationally choose a physical cutoff such that all the entropy of large black holes comes from entanglement and its value is precisely $A_H/4G$, with G the low-energy Newton constant. Concrete pictures of this kind include estimates of the cutoff based on horizon fluctuations [5]. Other suggestions involving nontrivial quantum gravity physics are, for example, the hypothesis of a "holographic" description of the state of collapsed matter [8], and the idea that black hole entropy looks classical because it lives in a Hilbert space of states which cannot be realized in field theory, such as special (perturbative) string configurations [9] (the question of one-loop corrections in string theory is a subtle one; see Refs. [10]). A striking feature of all these scenarios, in which one invokes subtle effects of quantum gravity to target the value $A_H/4G$, is that automatically Hawking's semiclassical calculation becomes suspect. One could wonder that not only black hole entropy, but also Hawking's temperature, would arise as "miraculous" successes of the semiclassical approximation.

A different alternative, which is compatible with some of the ideas above, has been advocated by Susskind and Uglum in [9]. According to this proposal geometric entropy is just a correction to a classical entropy, and the divergences organize in such a way that they renormalize the Newton constant, such that $S_{\rm BH} = A_H/4G$ to all orders. In this context, it is important to note that the problem of understanding black hole entropy can be addressed independently of the details of the end point of black hole evaporation. Phenomenologically, one assigns entropy to a black hole only during the period in which it is radiating thermally. For this reason, a concrete low-energy picture should be attainable to the extent that Hawking's emission formula can be considered as accurate at least for some period of time.

Many of the discussions of black hole entropy are carried out in the Euclidean formalism for the canonical ensemble where, as stated before, there is a classical contribution whose explanation in physical terms is not clear. In addition, the one-loop fluctuation determinant in the gravitational sector has a negative eigenvalue which leads formally to a complex one-loop entropy. While this may be sufficient to call into question the relevance of the whole approach, it is likely that the imaginary part of the effective action still has a physical meaning interpreted as a decay rate for black hole nucleation. This is the interpretation of Gross, Perry, and Yaffe in [11] of the original Hawking-Gibbons calculation. In the light of these comments, it becomes interesting to recast the ideas of [9] about entropy renormalization in the language of decay rates, rather than in the static definitions of entropy.

There is an elegant operational definition of black hole entropy appropriate for nonequilibrium situations, based on the fact that semiclassical decay processes can be computed by means of Euclidean instanton methods. For example, when charged black holes can be pair created in a background electromagnetic field, the total rate may be written as

$$\Gamma \sim |\text{amplitude}|^2 \rho_{\rm eff}\,, \tag{4}$$

where $\rho_{\rm eff}$ is the effective density of final states. The classical instanton contribution is such that $\rho_{\rm eff}$ is enhanced for nonextremal black holes with respect to the extremal ones (and to monopole pair production) by precisely the Bekenstein-Hawking degeneracy factor $\rho_{\rm BH} = e^{A_H/4G}$ [13].

It was pointed out in Ref. [14] that the quantum corrections (fluctuation determinant) to Eq. (4) diverge uncontrollably because the Gaussian fluctuation determinant contains the factor

$$\mathrm{Tr}_{\mathcal{H}_{ph}} e^{-\beta_H H^{(0)}}\,, \tag{5}$$

where β_H is the Hawking inverse temperature (the black holes are created in equilibrium with the Hawking radiation), and $H^{(0)}$ is the free Hamiltonian for the physical (transverse) fluctuations around the instanton. This quantity diverges because of the continuous spectrum of field excitations in the black hole background [2], in a way which makes the corresponding renormalization a subtle question. The reason is that the divergence can be traced back to the presence of the horizon as an infinite redshift surface. One may then argue that, after low-energy couplings have been renormalized according to, e.g., graviton scattering far from the black hole, the horizon is still in place and the spectrum in a finite box is still continuous. According to this argument, it would appear that the calculation of the quantum corrections to $S_{\rm BH}$ requires explicit knowledge of Planckian physics and therefore the phenomenological formula $S_{\rm BH} = A_H/4G$ could not be substantiated in a low-energy effective description. This state of affairs would contradict the low-energy theorem of Susskind and Uglum [9], which states that an inambiguous low-energy expansion exists for $S_{\rm BH}$. Whereas a breakdown of the effective description could be expected in a remnant scenario, one would regard it as unreasonable if the classical theory failed to provide an adequate description of large enough black holes.

One of the aims of this paper is to show that a prescription can be given so that such a breakdown does not occur. This will require disentangling some of the subtleties involved in different ways of considering the quantum black hole entropy. We choose to discuss these issues in the context of thermal nucleation of black holes: this provides a scenario rich in both technical features and appealing interpretations in physical terms. The final outcome of our analysis will be a low-energy characterization of black hole entropy.

The paper is planned as follows. In the next section we cast the low-energy theorem of Ref. [9] into the language of Eq. (4) and find that the continuous spectrum problem is absent in the Euclidean formalism, where all ultraviolet divergences can be renormalized in a standard way. In Sec. III we turn to a more detailed study of Eq. (5). We derive a Euclidean prescription which is markedly different from the conical singularity formalism, and makes manifest the problem of the continuous spectrum. It also explains why this problem does not appear in the treatment of Sec. II. Roughly speaking, what happens is that the divergent Hamiltonian partition function equation (5) admits several formal representations as a determinant. One of them uses an operator with continuous spectrum and the natural regularization is noncovariant (a brick wall). The other representation uses a fully covariant operator with discrete spectrum, which is the one appearing in the calculation of decay rates following the formalism of [11].

In Sec. IV we offer some heuristic arguments showing that the continuous spectrum problem is also absent in semiclassical models for black hole collapse. It becomes an artifact of the eternal black hole geometry as an asymptotic approximation, and the structure of ultraviolet divergences should be again renormalizable in low-energy couplings in the standard way. Finally, in the last section we speculate on a physical picture for the quantum origin of black hole entropy. We point out that the space of quasicoherent states of the infalling matter and gravitational field could be used to parametrize the microscopic degrees of freedom of black hole entropy. This point of view does not necessarily rely on Planckian physics. One of the appendices contains a calculation that would otherwise disrupt the main line of the text. The other develops the subject of the thermal instabilities of the vacuum of two-dimensional dilaton gravity.

II. THERMAL NUCLEATION OF BLACK HOLES AND ENTROPY

Black hole entropy as a classical enhancement factor of the final state degeneracy may be studied in a technically simple situation, which nevertheless retains many physical features of the charged pair production, at least for large black holes. This is the thermal nucleation of neutral black holes in hot flat space, as studied in Ref. [11], where it was computed to one-loop order in the dilute instanton gas approximation. The corresponding instanton is simply the Euclidean section of the Schwarzschild geometry, which mediates the nucleation of black holes of critical mass $M = \beta/8\pi G$ inside a thermal bath of gravitons in flat space at temperature $T = 1/\beta$. The rate per unit volume is given by [11]

$$\Gamma = \frac{\omega_0}{2\pi\beta}\,\frac{m_{\rm Pl}^3}{64\pi^3}\,(\mu\beta)^{212/45}\,\exp\left(-\frac{m_{\rm Pl}^2}{16\pi T^2}\right). \tag{6}$$

The relation between the imaginary part of the free energy and the nucleation rate in this expression is the one corresponding to semiclassical excitation over the barrier (see Ref. [15]):

$$\Gamma = \frac{\omega_0\beta}{\pi}{\rm Im}\left(\frac{F}{V}\right). \tag{7}$$

In other words, the Euclidean instanton is better thought of as a sphaleron. Indeed, it is time independent (the Schwarzschild metric is static) and the fluctuation determinant in the physical sector has exactly one negative eigenvalue $\lambda = -\omega_0^2 \simeq -0.19/(GM)^2$, responsible for the appearance of an imaginary part of the free energy. The term proportional to $m_{\rm Pl}^3$ in Eq. (6) comes from the integration over the collective coordinates of the sphaleron, and the term $(\mu\beta)^{212/45}$ is because of the anomalies associated with the Euler number counterterm, which is nonvanishing in the Euclidean Schwarzschild section. The mass scale μ appears as a dimensional transmutation (analogous to $\Lambda_{\rm QCD}$) of the dimensionless Euler number coupling, which then becomes a running coupling, and may be phenomenologically determined (a natural value in this context is $\mu \simeq m_{\rm Pl}$). Finally, the exponential suppression factor comes entirely from the classical gravitational action $e^{-I_{\rm cl}}$, which is given to leading order by the Hilbert action

$$I_H = -\frac{1}{16\pi G}\int_{\mathcal{M}} R - \frac{1}{8\pi G}\oint_{\partial\mathcal{M}}(K - K^{(0)}). \tag{8}$$

The interpretation of Eq. (6) according to the Fermi rule Eq. (4) is based on the fact that, due to the nontrivial topology of the Euclidean Schwarzschild section, the classical suppression factor is not exactly the Boltzmann factor $e^{-\beta M}$. Indeed, on a manifold with cylindrical topology (a usual thermal manifold), the Hilbert action equals the canonical action and $I_{\rm cl}({\rm cylinder} \times S^2) = \beta M_{\rm ADM}$. On the other hand, on the Schwarzschild manifold, with topology $\mathcal{M} = {\rm disk} \times S^2$ there is one boundary missing, which produces $I_{\rm cl}(\mathcal{M}) = \frac{1}{2}\beta M_{\rm ADM}$. The Boltzmann factor is in excess exactly by the value of the classical black hole entropy (recall $\beta = 8\pi GM$ for the nucleated black holes):

$$e^{-\beta^2/16\pi G} = e^{-\beta M}\, e^{4\pi G M^2} = e^{-\beta M}\rho_{\rm BH}. \tag{9}$$

Hence, an operational definition of black hole entropy in this context would be the excess of the classical action over the Boltzmann factor

$$S_{\rm BH} \equiv \beta M - I_{\rm cl}(\mathcal{M}). \tag{10}$$

Since the gravitational sector may be regarded as a low-energy effective theory, quantum corrections require the introduction in $I_{\rm cl}$ of the whole tower of counterterms, leading to a low-energy expansion

$$S_{\rm BH} = \frac{A_H}{4G} + \sum_n \frac{\lambda_n}{m_{\rm Pl}^{d_n-4}}\int_{\mathcal{M}}\mathcal{O}_n. \tag{11}$$

Here we have suppressed the cosmological constant counterterm, since asymptotic flatness is a condition of the problem. The leading term absorbs the renormalization of the Newton constant, while the others are also phenomenologically determined. It is important to note that the definition of $S_{\rm BH}$ given in Eq. (10) is not in general equivalent to others based on the thermodynamic formula

$$S = (\beta\partial_\beta - 1)I_{\rm cl}. \tag{12}$$

In the original analysis of Gibbons and Hawking [3], the derivative was taken on the space of classical solutions of the field equation. Alternatively, in the conical singularity method [16,17], one holds M fixed while varying β, thus going off shell due to the conical singularity at the horizon.

Actually, one may regard the sphaleronlike interpretation of the Euclidean Schwarzschild instanton [11] as more natural than the thermostatic interpretation behind formula (12), because the negative mode at one loop signals an instability of the canonical ensemble for thermal gravitons, in addition to the infrared Jeans instability. Yet, the low-energy effective theory predicts a value for the decay rate which may have physical meaning. The definitions based on Eqs. (10) and (12) will give, in general, a slightly different low-energy expansion, although the leading term seems to be universal (in a fashion similar to the first two terms of the β function in gauge theories, which are independent of the definition of physical coupling).

A very striking feature of Eq. (6) is the absence of the potentially troublesome partition function in Eq. (5). In fact, it is easy to see that it cancels against the flat space normalization, up to an ultraviolet finite boundary or "surface tension" term. To be more specific, let us consider the total partition function given as a dilute multi-instanton sum:

$$Z = \exp(-\beta \mathrm{Re} F + \frac{i\pi}{\omega_0} V\Gamma) = \sum_{N=1}^{\infty} \frac{1}{N!} Z_{(N)} e^{-I_{\mathrm{cl}}^{(N)}}\ , \quad (13)$$

where $I_{\mathrm{cl}}^{(N)} \simeq N I_{\mathrm{cl}}^{(1)} = N\beta M/2$ and $Z_{(N)}$ denotes the perturbative partition function around the N-multiblackhole solution. To one-loop order one finds

$$Z_{(N)} = (i/2)^N C^N \det\nolimits_{+}^{-1/2}(I''_{(N)})\,, \quad (14)$$

where C stands for the contribution of the collective coordinates (zero modes) and anomalous scaling. The factor $i/2$ comes from the usual half-contour rotation for the N negative modes and $I''_{(N)}$ is a combination of second order elliptic differential operators which includes fluctuation kernels for the physical as well as unphysical graviton polarizations, and the corresponding ghost terms [19]. Roughly speaking, the ghost determinant cancels the longitudinal and trace parts of the graviton excitations, leaving the physical (transverse-traceless) fluctuations.

In any covariant regularization, the ultraviolet divergences in the perturbative effective action $W_{\mathrm{eff}} = \frac{1}{2}\ln\det(I''_{(N)})$ can be absorbed in the counterterm series of I_{cl}. A very convenient one-loop prescription is given by ζ function regularization, which only requires the spectrum of I'' to be discrete. This is always the case at finite volume, since all operators are elliptic and the Euclidean manifold is compact with boundary. If we write the ultraviolet finite part of W_{eff} as a volume integral it is clear that, within the dilute gas approximation and in the large volume limit, W_{eff} is dominated by the free energy of gravitons in flat space. Let us separate the contribution of the asymptotic thermal gravitons from those close to the horizon (up to, say, a radius $r \sim 3GM$). Then one finds

$$W_{\mathrm{eff}} \simeq N W_{\mathrm{hor}} + \beta f_g (V - N V_{\mathrm{BH}})\,, \quad (15)$$

where $f_g = -\pi^2/45\beta^4$ denotes the free energy density of gravitons in flat space, and V_{BH} is the excluded volume per black hole. If we multiply and divide by the flat space partition function $Z_{(0)} = \exp(V\pi^2/45\beta^3)$ we get an overall factor in the N-instanton term

$$Z_{(0)} e^{-N\beta F_B}\,, \quad (16)$$

where F_B is a boundary-free energy given by the contribution to W_{eff} coming from the horizon region minus the graviton-free energy in the same volume of flat space

$$\beta F_B = W_{\mathrm{hor}} - \beta f_g V_{\mathrm{BH}}\,. \quad (17)$$

Notice that $\beta f_g V_{\mathrm{BH}}$ is a pure number, independent of M. In fact, βF_B appears as a constant term in the $1/V$ expansion of W_{eff}:

$$W_{\mathrm{eff}} = \beta f_g V + N\beta F_B + O(V^{-1})\,. \quad (18)$$

For small black holes (corresponding to high temperature) this term should approach zero, whereas for large black holes (i.e., low temperature), W_{hor} scales like the vacuum energy of Euclidean Rindler space, $W_{\mathrm{hor}} \sim (M/m_{\mathrm{Pl}})^4 + \mathrm{const}$. Hence one concludes that

$$\beta F_B \sim \left(\frac{M}{m_{\mathrm{Pl}}}\right)^4 \left[1 + O\left(\frac{m_{\mathrm{Pl}}}{M}\right)\right]\,. \quad (19)$$

When the multi-instanton sum is performed the term βF_B exponentiates and it contributes to the imaginary part of the free energy. The real part is given by the flat space free energy as it should be, and the corrected low-energy expansion for the rate reads

$$\Gamma_{\mathrm{corr}} = \frac{\omega_0}{2\pi\beta} \frac{m_{\mathrm{Pl}}^3}{64\pi^3} f(\mu\beta) e^{-\beta M} e^{S_{\mathrm{BH}}} e^{-\beta F_B}\,, \quad (20)$$

where S_{BH} is defined in Eq. (11) and $f(\mu\beta) = (\mu\beta)^{212/45} + \cdots$ is obtained from the perturbative expansion of the β function associated to the dimensionless coupling to the Euler number.

The boundary term $\exp(-\beta F_B) \simeq \exp[-C(M/m_{\mathrm{Pl}})^4]$ dominates the suppression factor at low temperatures. This agrees with the fact that surface effects become increasingly important: in the large M limit the Rindler region covers all of (Euclidean) space. The boundary partition function is thus related to the part of Eq. (5) coming from the vicinity of the horizon.

In the construction presented above, black hole entropy is fundamentally a classical object with no microscopic interpretation, and quantum corrections organize in a low-energy expansion. Furthermore, the renormalization of the Newton constant implied in the definition of S_{BH} is the same that one would obtain from graviton scattering far from the black hole, as long as a covariant procedure, such as ζ-function regularization, is employed everywhere. This, in turn, is ensured by the fact that the finite volume Euclidean manifold is compact and smooth and at the equilibrium temperature there is no global distinction between finite temperature free energy and vacuum energy.

An important point to stress is that, at least for nonextremal black holes, the problem of continuous spectrum is absent from the previous discussion. All the operators involved are elliptic, and have discrete spectrum at finite volume. An explicitly covariant regularization is possible and there is no obstruction to the low-energy theorem of [9]. For example, having discrete spectrum one can use ζ function regularization at one loop, in which there are no divergences at all and the total black hole entropy comes out clearly as $A_H/4G$.

III. CONTINUOUS VERSUS DISCRETE FLUCTUATION SPECTRUM

In this section we study some aspects of the Hamiltonian partition function, Eq. (5), which following Unruh [20], is related to the entanglement density matrix in the vacuum of the extended eternal black hole geometry. We start by reviewing the disease caused by continuous black hole spectrum, as first pointed out by 't Hooft in Ref. [2], and work backwards to derive an Euclidean formulation which makes manifest the differences between Eq. (5) and the term $\exp(-\beta F_B)$ that we have found in the previous section.

A. Statistical mechanics of the fluctuation degrees of freedom

Although we have in mind the physical situation studied in Sec. II (thermal gravitons) the discussion may be generalized to different matter contents. In general, let the Gaussian (Lorentzian) action for quadratic fluctuations around the black hole be

$$S^{(2)} = \frac{1}{2}\int_{\mathcal{M}} \varphi\, \mathcal{L}\, \varphi\,, \tag{21}$$

where $\mathcal{L} = -\nabla^2 + V(g)$ and φ represents the physical (transverse) excitations. For example, for a scalar field we have $\mathcal{L} = -\nabla^2 + m^2 + \xi R + \cdots$, while for transverse-traceless gravitons, the case relevant to the previous section, we must consider $\mathcal{L}_{TT} h_{\alpha\beta} = -\nabla^2 h_{\alpha\beta} - 2R_{\alpha\beta\gamma\delta}h^{\gamma\delta}$ (we are focusing on bosonic fields for simplicity).

Choosing a time slicing adapted to the Killing vector $\partial/\partial t$, where t is the asymptotically Minkowskian time, we may express the free canonical Hamiltonian associated to Eq. (21) in terms of the eigenfrequencies as

$$H^{(0)} = \sum_\omega \omega a^\dagger_\omega a_\omega + \Lambda_B\,, \tag{22}$$

where $\Lambda_B = \frac{1}{2}\sum_\omega \omega$ is the (Boulware) vacuum energy, formally infinite, and $a^\dagger_\omega$ generate the physical Fock space. The one-loop free energy takes then the well known form

$$\beta F - \beta\Lambda_B = \sum_\omega \ln(1-e^{-\beta\omega})\,. \tag{23}$$

At finite volume, this quantity is ill defined even with an ultraviolet cutoff in the frequency sum: $\omega \leq \Lambda$. The reason is that the spectrum of a black hole in a box is still continuous because the horizon behaves as a noncompact boundary. The eigenvalue problem for the frequencies is

$$(-g_{00})(-\vec{\nabla}^2 + V(g))\psi_\omega(\mathbf{x}) = \omega^2\psi_\omega(\mathbf{x}). \tag{24}$$

In tortoise coordinates $r_* = r + 2GM\ln|r/2GM - 1|$ this is a Schrödinger problem for radial excitations with L^2 metric, and with an effective potential $V_{\text{eff}} \propto -g_{00} \sim \exp(4\pi T_H r_*)$ as we approach the horizon $(r_* \to -\infty)$. As a result, the spectrum is continuous unless a horizon regulator (brick wall) is imposed. This all looks very different from the discussion in the preceding section, where all operators would present discrete spectrum after standard infrared regularization (large volume cutoff).

In particular, as pointed out in Ref. [14], the problem of continuous spectrum seems to remain even after the Newton constant has been renormalized according to graviton scattering far from the black hole, because it only depends on the existence of the horizon as an infinite redshift surface. Heuristically, the brick wall boundary condition is a local ultraviolet cutoff, because the condition $\omega \leq \Lambda$ is not a uniform cutoff for local static observers, who measure local frequencies $\omega_{\text{loc}} = \omega/\sqrt{-g_{00}}$. Thus, the brick wall cuts off unphysical static frames. It is, however, very disturbing that this interpretation of the cutoff is frame dependent. This is a first indication of the fact that the continuous spectrum cannot be easily cut off in a covariant way.

In order to bring the discussion to the terms of Sec. II, it is necessary to rewrite the free energy equation (23) in Euclidean form. This can be done directly, as in flat space, by means of the ζ function identity (we follow Ref. [21])

$$\prod_{n=1}^{\infty}(A + n^2/B) = \frac{2}{\sqrt{A}}\sinh(\pi\sqrt{AB})\,. \tag{25}$$

We get

$$\beta F = \frac{1}{2}\ln\prod_{n\in Z}\prod_\omega(4\pi^2n^2/\beta^2 + \omega^2) \equiv -\ln\det{}^{-1/2}(\widehat{\mathcal{L}})\,. \tag{26}$$

This defines $\widehat{\mathcal{L}}$ as the operator with eigenvalues $4\pi^2n^2/\beta^2 + \omega^2$. From Eq. (24) we conclude that $\widehat{\mathcal{L}}$ is given by

$$\widehat{\mathcal{L}} = -\partial_\theta^2 + (-g_{00})(-\vec{\nabla}^2 + V(g))\,, \tag{27}$$

acting on periodic functions of the Euclidean time θ of the form $\hat{\psi}_{n,\omega} = e^{2\pi i n\theta/\beta}\psi_\omega(\mathbf{x})$, where $\psi_\omega(\mathbf{x})$ are the spatial harmonics in Eq. (24).

Curiously enough, this is not the covariant fluctuation operator, but rather a local multiple:

$$\widehat{\mathcal{L}} = (-g_{00})\mathcal{L}\,. \tag{28}$$

The inner product for $\widehat{\mathcal{L}}$, as inherited from the L^2 inner product in tortoise coordinates (or the Klein-Gordon metric in Lorentzian signature) is

$$\langle\hat{\psi}|\hat{\psi}'\rangle = \int_{\mathcal{M}} d^4x \frac{\sqrt{-g}}{(-g_{00})}\, \hat{\psi}^* \hat{\psi}' \,. \tag{29}$$

From this we can derive a path integral formula. In general, given an inner product

$$\langle\psi,\psi'\rangle_\rho = \int d^4x\; \rho(x)\psi^*(x)\psi'(x)\,, \tag{30}$$

the determinant of an operator $\mathcal{L}_\rho$ admits the representation

$$\det{}^{-1/2}(\mathcal{L}_\rho) = \int \mathcal{D}_\rho\varphi \exp\left(-\frac{1}{2}\langle\varphi, \mathcal{L}_\rho\varphi\rangle\right)\,, \tag{31}$$

where the measure is formally given by

$$\mathcal{D}_\rho\varphi = \prod_n \frac{dc_n}{\sqrt{2\pi}} = \prod_x \frac{d\varphi_x}{\sqrt{2\pi}}\rho_x^{1/2}\,. \tag{32}$$

Here c_n are the Fourier coefficients of the field φ in a basis orthonormal with respect to the product (30).

It is interesting that the inner product equation (29) precisely gives the action $S^{(2)}$ in the exponent:

$$S^{(2)} = \frac{1}{2}\int d^4x\sqrt{-g}\;\varphi\mathcal{L}\varphi = \langle\varphi|\widehat{\mathcal{L}}|\varphi\rangle\,. \tag{33}$$

So, the operator $\widehat{\mathcal{L}}$ with the inner product equation (29) is *classically* equivalent to the operator $\mathcal{L}$ with the covariant inner product. However, quantum mechanically, there is a difference in the path integral measure.

We have then established

$$\begin{aligned}\mathrm{Tr}_{\mathcal{H}_{ph}} e^{-\beta H^{(0)}} &= \det{}^{-1/2}(\widehat{\mathcal{L}}) \\ &= \int \prod_x \frac{d\varphi_x}{\sqrt{2\pi}} \left(\frac{\sqrt{-g}}{-g_{00}}\right)_x^{1/2} e^{-S^{(2)}[\varphi]}\,. \end{aligned}\tag{34}$$

This result was also obtained in Ref. [22] using the canonical derivation of the path integral. It is remarkable because it shows that the canonical partition function equation (5) is *not* formally equal to $\det^{-1/2}(\mathcal{L})$. Rather, it equals the determinant of a related operator which is singular at the horizon where $g_{00} = 0$. Accordingly, the operator $\widehat{\mathcal{L}}$ has continuous spectrum at finite volume, and does not admit ζ-function regularization unless we provide some kind of brick wall cutoff. Therefore, the topology of the Euclidean manifold appropriate to $\widehat{\mathcal{L}}$ is cylindrical: $\widehat{\mathcal{M}} = \mathcal{M} - \{\text{horizon}\}$ is noncompact in the vicinity of the horizon. If we would use this as the physical thermal manifold, the classical contribution to the entropy would vanish.

The peculiar topology associated to Eq. (34) can be traced back to its origin as geometric or entanglement entropy, at least when it is calculated as a thermal sum. For example, if we consider the entanglement entropy generated by performing a trace over half of Minkowski space [6], the formal procedure to expose the thermal nature of the density matrix uses a trick due to Unruh [20] (see also Refs. [23,12]).

One decomposes the total Cauchy surface into two noncompact left and right components by an appropriate coordinate mapping, which in this case is equivalent to the Rindler acceleration. Since the two components are noncompact, in fact the origin (the position of the boundary) is not part of the mapping. In other words, one writes $\mathcal{H}' = \mathcal{H}_L \otimes \mathcal{H}_R$, where $\mathcal{H}_{L,R}$ are the left and right Hilbert spaces, and $\mathcal{H}'$ is the total Hilbert space minus the field oscillator at the boundary. This is the formal origin of the missing point in the Euclidean manifold $\widehat{\mathcal{M}}$.

The density matrix for the vacuum obtained by tracing out degrees of freedom in the left half-space can be found to be [23–25]

$$\langle\varphi|\rho|\varphi'\rangle = \langle\varphi|(\mathrm{Tr}_{\mathcal{H}_L}|0\rangle\langle 0|)|\varphi'\rangle = \prod_\omega \left(\frac{\omega}{\sinh 2\pi\omega}\right)^{1/2} \exp\left(-\frac{\omega}{2}\left\{\coth 2\pi\omega(\varphi_\omega^2 + {\varphi'}_\omega^2) - \frac{2\varphi_\omega {\varphi'}_\omega}{\sinh 2\pi\omega}\right\}\right)\,, \tag{35}$$

where φ_ω, ${\varphi'}_\omega$ are the Fourier components of the spatial fields in the right half-space, analyzed in the basis of spatial eigenfunctions $\psi_\omega(\mathbf{x})$ orthonormal with respect to the spatial section of the inner product (29). The exponential term in Eq. (35) corresponds to the classical action $S^{(2)}$ between configurations φ, φ', whereas the prefactor comes from the fluctuation determinant around the classical path. It is easy to check that to obtain it from the four-dimensional Euclidean path integral one must use the noncovariant measure in Eq. (34), and introduce φ, φ' as the values of the field at each side of the cut along $\theta = 0$ [23,12].

These results may seem disturbing at first, because they indicate that the Euclidean construction for entanglement entropy is formally defined in terms of $\widehat{\mathcal{L}}$ instead of the covariant operator $\mathcal{L}$. On the other hand, we know that the Hartle-Hawking Green's function defined without boundary condition on the Euclidean section $\mathcal{M}$ is the correct thermal Green's function for static observers. In fact, both Green's functions are equal: $\widehat{G}(x,x') = G_{\mathrm{HH}}(x,x')$ and there is no contradiction. Again, this follows easily from the freedom to choose different operators provided the inner product is changed accordingly. The Green's function of an operator $\mathcal{L}_\rho$ defined as

$$G_\rho(x,x') = \langle x|\mathcal{L}_\rho^{-1}|x'\rangle \tag{36}$$

satisfies the equation

$$\mathcal{L}_\rho(x)G_\rho(x,x') = \delta_\rho(x,x') = \frac{\delta(x-x')}{\rho_x}. \tag{37}$$

Using the expression for $\widehat{\mathcal{L}}$ and $\hat{\rho}$ in terms of $\mathcal{L}$ and ρ, it is trivial to realize that $G_{\rm HH}$ and $\widehat{G}$ satisfy the same equation

$$\mathcal{L}_x\widehat{G}(x,x') = \frac{\delta(x-x')}{\sqrt{-g_x}}. \tag{38}$$

Therefore, $G_{\rm HH}$ and $\widehat{G}$ are obviously identical when the boundary conditions are the same, such as, in a brick wall model. For the no-boundary case, the equality is not obvious, because $\widehat{G}(x,x')$ cannot be extended to the horizon in terms of the eigenfunctions of $\widehat{\mathcal{L}}$. However, an explicit computation in Rindler space can be done (see Appendix A) which ensures $G_{\rm HH} = \widehat{G}$ also in the no-boundary case.

Thus, for local physics, the difference between $\widehat{\mathcal{M}}$ and $\mathcal{M}$ is just the way in which the no-boundary condition of Hartle and Hawking is introduced. However, the difference is important for the issue of the total number of states of the black hole in low-energy field theory. This is due to the fact that, in going from the Green's function to the extensive free energy, one has to give sense to the expression

$$\frac{1}{2}{\rm Tr}_{\{x\}}\ln G(x,x). \tag{39}$$

Different prescriptions for the spatial trace and the coincidence limit turn into the different determinants above. The disease of continuous black hole spectrum arises when one works with $\widehat{G}$, which leads to considering $\det^{-1/2}(\widehat{\mathcal{L}})$. As explained above, this is naturally regularized by means of a brick wall cutoff. On the other hand, use of $G_{\rm HH}$ in Eq. (39) is concomitant to the computation of $\det^{-1/2}(\mathcal{L})$, which is free of the continuous spectrum problem. In this case the regularization procedure is fully covariant and we obtain the results of Sec. II. In fact, both prescriptions are formally related by a conformal transformation. To see this, we recall that, according to Eq. (26), there are many path integral versions of the same determinant, because we can change the operator at the price of rescaling the inner product (thereby changing the functional measure). A particularly nice variation is given by the "optical" inner product, which is covariant with respect to the conformally-related ("optical") metric $\bar{g}_{\alpha\beta} = g_{\alpha\beta}/(-g_{00})$ and has weight $\bar{\rho} = \sqrt{-g}/(-g_{00})^{d/2} = (-\bar{g})^{-1/2}$ in d spacetime dimensions. Then the operator

$$\bar{\mathcal{L}} \equiv (-g_{00})^{\frac{d+2}{4}}\,\mathcal{L}\,(-g_{00})^{\frac{2-d}{4}} \tag{40}$$

has the same eigenvalues as $\widehat{\mathcal{L}}$ and $\det^{-1/2}(\widehat{\mathcal{L}}) = \det^{-1/2}(\bar{\mathcal{L}})$.

In the conformally-invariant case, a nice relation between the determinants of $\mathcal{L}$ and $\widehat{\mathcal{L}}$ can be written using the optical operator as an intermediate step. A conformally-invariant fluctuation operator for scalars is given by

$$\mathcal{L}_c(g) = -\nabla^2 + \frac{d-2}{4(d-1)}R. \tag{41}$$

A simple computation shows that $\bar{\mathcal{L}}_c = \mathcal{L}_c(\bar{g})$ and we may write

$$\begin{aligned}\det{}^{-1/2}(\bar{\mathcal{L}}_c) &= \det{}^{-1/2}(\mathcal{L}_c(\bar{g}))\\ &= \int\prod_x \frac{d\varphi_x}{\sqrt{2\pi}}(-\bar{g}_x)^{1/4}\\ &\quad\times\exp-\left[\frac{1}{2}\int\sqrt{-\bar{g}}\varphi\mathcal{L}_c(\bar{g})\varphi\right].\end{aligned} \tag{42}$$

But the last path integral is conformally related to the covariantly regularized path integral for the operator in the physical metric. Then we obtain

$$\det{}^{-1/2}(\widehat{\mathcal{L}}_c) = \det{}^{-1/2}(\bar{\mathcal{L}}_c) = e^{-I_L[\ln g_{00}]}\det{}^{-1/2}(\mathcal{L}_c) \tag{43}$$

(see also Ref. [22]). In two dimensions I_L is the standard Liouville functional, while in four dimensions it is in general a nonlocal action [26].

B. Brick wall regularization and renormalization

Equations (42,43) only make sense with a brick wall in place, because otherwise the noncompact operators have no well-defined determinant. In such a situation $I_L \sim \beta \times$ finite; i.e., it contributes only to the vacuum energy (not to the entropy). This means that one can compute the entropy in the presence of the brick wall directly from the ultraviolet finite part of $\det^{-1/2}(\mathcal{L})$. It is important to recognize that, in the absence of a brick wall, Eq. (43) has a formal status, because it relates infinite quantities.

The leading brick wall divergence is in fact independent of the particular potential term occurring in $\mathcal{L} = -\nabla^2 + V(g)$, provided $V(g)$ is regular in the horizon region. This is due to the fact that the leading divergence depends only on the effective potential $-g_{00}V(g)$, which vanishes exponentially in the horizon region. The potentially troublesome angular degrees of freedom [27], which may spoil the accuracy of the WKB approximation, sum up such that the WKB result is surprisingly correct [this is easy to check by using Eq. (43) and computing in the optical metric [18,22]].

The final answer for the leading divergence per degree of freedom in $d > 2$ dimension is

$$S_{\rm div} = -(\beta_H F - \beta_H \Lambda_B)_{\rm div} d$$
$$= \frac{d\,\Gamma(d/2)\,\zeta(d)}{(d-2)\pi^{3d/2-1}2^{d-1}} \frac{A_H}{\epsilon_{\rm BW}^{d-2}}\,, \qquad (44)$$

where $\epsilon_{\rm BW}$ is the brick wall cutoff. In two dimensions $S_{\rm div} = 1/6 \ln \epsilon_{\rm BW}^{-1}$. Also, for fermions one obtains the usual statistical correction factor $S_{\rm Fermi} = (1-2^{1-d}) S_{\rm Bose}$.

At this point one can adopt different attitudes. If black hole entropy is primarily regarded as a quantum object and Eq. (44) considered at least part of it, then the entropy is clearly cutoff dependent. We cannot predict its value using low-energy quantum gravity nor understand what degrees of freedom $S_{\rm BH}$ accounts for. In this view, the final result $S_{\rm BH} = A_H/4G$ with G the long distance Newton constant, would seem to come out in a rather "miraculous" way from Planckian dynamics in quantum gravity. Variations of this idea have been put forward by various authors [2,5].

Another possibility is to consider a classical entropy, and take Eq. (44) as a counterterm renormalizing Newton constant. However, there is some arbitrariness here since the renormalization conventions appropriate for graviton scattering far from the black hole and for Eq. (44) do not agree in general. For example, the counterterms induced by a scalar field on the vacuum energy are (in Schwinger proper time regularization) readily found from the heat kernel expansion

$$\Lambda_{\rm counter} = -\frac{{\rm Vol}(\mathcal{M})}{d(4\pi)^{d/2}} \frac{1}{\epsilon^d} + \frac{\sqrt{\pi}{\rm Vol}(\partial\mathcal{M})}{2(d-1)(4\pi)^{d/2}} \frac{1}{\epsilon^{d-1}}$$
$$-\frac{\epsilon^{2-d}}{(d-2)(4\pi)^{d/2}} \left(\frac{1}{6}\int_{\mathcal{M}} R + \frac{1}{3}\oint_{\partial\mathcal{M}} K \right) + \cdots\,. \qquad (45)$$

The last term induces a renormalization (in four dimensions)

$$G_{\rm bare}^{-1} \to G_{\rm bare}^{-1} + \frac{1}{12\pi\epsilon^2}\,. \qquad (46)$$

Now, in order to compare Eqs. (44) and (45) we would need an invariant relation between both cut offs. It is unlikely that such a relation exists because, as we pointed out before, the physical interpretation of the brick wall as an ultraviolet cutoff is fundamentally frame dependent. If one insisted on comparing the results, the only possible "natural" relation should be based on the fact that the Schwinger proper time cutoff is a length cutoff for paths in the first quantized path integral representation of determinants. Then, one could declare that the Schwinger cutoff is set by the minimum length noncontractible path in the brick wall manifold,

$$\epsilon \simeq \epsilon_{\rm BW} \frac{2\pi\beta}{\beta_H}\,, \qquad (47)$$

and this would lead to $S_{\rm div} = (\pi/90)(A_H/\epsilon^2)$ in four dimensions, which could not be absorbed with the renormalization above. As a consequence, if we wanted the renormalization to work along the lines of Sec. II, we would be led to *ad hoc* choices of brick wall cutoff.

This situation may be summarized by saying that the use of brick wall regulators has a heuristic value but, if we assume that there is an inambiguous classical entropy, a systematic treatment of the renormalizaton procedure in low-energy theory requires the use of covariant schemes, based on the Hartle-Hawking regular manifold $\mathcal{M}$, as in Sec. II, or a conical deformation of it (this agrees with remarks made in recent papers [28]). In fact, in the context of the black hole nucleation approach, we can say that the covariant method is the only possible choice. This is due to the fact that the continuous spectrum operator $\widehat{\mathcal{L}}$ has positive spectrum by construction. There is no way we could get a negative eigenvalue from this operator and thus no imaginary part for the free energy. As a result, this operator cannot appear if we want to maintain the physical picture of hot space decay.

IV. REDSHIFT ARGUMENTS IN MIRROR MODELS

In this section we argue on a heuristic basis that, in semiclassical collapse models, the continuous spectrum problem seems to be spurious. In the WKB approximation one basically gets the results of naive redshift calculations, i.e., formula (44) can be obtained from [21]

$$S_{\rm div} \simeq \frac{d\,\Gamma(d/2)\,\zeta(d)}{\pi^{d/2}} \int d({\rm Vol})(\beta_H\sqrt{-g_{00}})^{1-d}\,. \qquad (48)$$

This suggests that the divergence in Eq. (44) should be properly related to the unphysical observers close to the horizon. Any quantity computed from $\widehat{\mathcal{L}}$ refers to a family of static observers which become singular at the horizon—a physical static frame at the horizon has infinite energy. Yet, this is an artifact of the eternal black hole geometry as an effective approximation to a collapse solution. This point deserves further explanation.

Hawking radiation is dynamically generated by the time-dependent gravitational background in the vicinity of the collapsing matter. In the asymptotic regime, the time-dependent background can be eliminated in favor of a dynamical boundary condition by an appropriate choice of coordinates. This gives the mirror model description of black hole emission. Locally, for free field propagation in radial modes, the point $r = 0$ is a perfectly reflecting boundary which behaves as a time-dependent brick wall, following an asymptotic trajectory in tortoise coordinates:

$$r_*(r=0) \simeq -t - Ae^{-t/2GM} + B\,. \qquad (49)$$

In these models, the position of the infalling matter at late times stays asymptotically at a fixed tortoise distance from the origin, and provides a natural cut off for the static Cauchy surfaces. At any finite t, the spectrum of fields inside a large box is discrete, becoming continuous only in the mathematical limit $t = \infty$, which is totally unphysical because of the back reaction. We can rewrite Eq. (48) in terms of the optical volume $\bar{V}$ outside the infalling matter shell:

$$S_{\rm div} \simeq \frac{d\,\Gamma(d/2)\,\zeta(d)}{\pi^{d/2}}\,\frac{\bar{V}}{\beta_H^{d-1}}\,. \tag{50}$$

In two dimensions the optical volume diverges linearly with the tortoise coordinate (logarithmically in proper distance), whereas in four dimensions

$$\bar{V}_t \sim A_H GM e^{-r_*/2GM} \sim (GM)^3 e^{t/2GM}\,. \tag{51}$$

If we want to regard Eq. (50) as the geometric entropy outside the infalling matter we must get rid of the boundary divergence at the position of the outer shell. This can be done following Ref. [29], by subtracting the geometric entropy in the vacuum (the mirror remaining stationary). The result should be an extensive entropy with respect to the optical volume (this was explicitly checked in two dimensions in Ref. [29], and it is very plausible in four dimensions as well). In any case, as the tortoise position of the infalling matter recedes to $r_* \to -\infty$, the optical volume diverges exponentially and we find the divergence of 't Hooft. Notice that in Eq. (51) the Newton constant is the one entering in the mirror trajectory, i.e., the renormalized G.

Therefore, if we regularize an eternal black hole by a physical collapsing star, the continuous spectrum disease becomes an artifact of the time slicing used inside the collapsing star, or else it corresponds to the infinite entropy production at $t = \infty$.

The entropy source in these models is formally the mirror itself, although a more accurate interpretation would be that the time-dependent state of the infalling matter and gravitational fields decays with a thermal cross section. In this sense, the difficulties in locating the proper degrees of freedom of black hole entropy are naturally due to the classical treatment of the radiation source.

V. DISCUSSION

We have discussed several aspects, both technical and conceptual, of the black hole entropy problem. In Sec. II we have shown that classical "intrinsic" entropy makes sense in low-energy effective theory even in a nonequilibrium situation. The fact that it appears as a classical object could be due to the use of stationary saddle points to approximate the path integral. After all, in the sphaleron interpretation of black hole nucleation out of hot flat space one is talking about a classical process of excitation *over* the barrier; i.e., the nucleated black holes are formed by physical collapse of graviton "matter." But, of course, there are no temporal Killing vectors inside the collapsing matter, even asymptotically. The low-energy theorem of Susskind and Uglum can be applied to this situation provided a covariant regularization procedure is used throughout. We also found it useful to distinguish between ultraviolet divergences in determinants of operators with discrete spectrum, from others with continuous spectrum, such as the ones appearing in the brick wall model. Then we have argued in favor of fully covariant path integral prescriptions (leading to operators with discrete spectrum before ultraviolet regularization) in systematic discussions of entropy renormalization.

Some recent proposals for the solution of the black hole entropy problem and the information puzzle involve, in one way or another, a breakdown of low-energy effective field theory in the vicinity of the horizon, even for big black holes and early stages of the evaporation process. Since the discussions of black hole entropy renormalization (particularly those based on Euclidean methods) assume the validity of low-energy field theory, this question becomes very relevant to the matters discussed in this paper. Therefore, we would like to end with some speculations on the related question of a low-energy description of the microscopic degrees of freedom responsible for $S_{\rm BH}$.

From the point of view of mirror models, one would associate the quantum degrees of freedom of black hole entropy with the radiation source: the infalling matter and corresponding *time-dependent* gravitational field. The problem, of course, is that this Hilbert space has dimension $\sim A_H^{3/2}/\ell_{\rm Pl}^3$, instead of the required $A_H/\ell_{\rm Pl}^2$. Here is where exotic quantum gravity physics, such as the 'holographic' phase [8,30], seems unavoidable.

Actually, there is a natural notion of black hole entropy, closely related to the phenomenological derivation of 't Hooft given in Sec. I, which avoids explicit input from Planck-scale physics. It is based on the idea that a black hole radiates not because it is thermally excited in some way, but just because its cross section for decay happens to be thermal.

In Hawking's approximation one computes the decay rate by scattering the asymptotic vacuum off the time-dependent *classical* gravitational field. In a full quantum treatment the condition that the external field approach is a sensible approximation can be formalized by taking a coherent state for the infalling matter state (and the induced graviton condensate). By a coherent state we mean a minimum spread wave packet or, a state in which expectation values of operators are given as classical functions of the expectation values of the coordinates and momenta in the regularized theory (with a cut off in place). One would then work in a Hilbert space of the form

$$\mathcal{H}_{\rm Haw} = \mathcal{H}_{\rm coherent} \otimes \mathcal{H}_{\rm rad}\,, \tag{52}$$

where states are approximated by the product of a coherent time-dependent infalling state $|\Psi_{\rm coh}(t)\rangle$ and a dilute radiation state $|\omega_1 \cdots \omega_n\rangle$ of n Hawking quanta. The interaction Hamiltonian in the extreme coherent approximation would induce an effective time-dependent background field potential for Hawking quanta:

$$\begin{aligned} &\langle \Psi_{\rm coh}(t)| \otimes \langle \omega_1 \cdots \omega_n | H_{\rm int} | \Psi_{\rm coh}(t)\rangle \otimes |0\rangle \\ &\simeq \langle \omega_1 \cdots \omega_n | V_{\rm eff}(g_{\mu\nu}(t))|0\rangle\,. \end{aligned} \tag{53}$$

This yields Hawking's analysis. However, quantum back reaction changes this picture, since after each radiative transition the initial coherent state slightly decoheres. One has

$$|\Psi_{\rm coh}(t)\rangle \otimes |0\rangle \to |\Psi_\omega(t)\rangle \otimes |\omega\rangle\,, \qquad (54)$$

where $|\Psi_\omega\rangle$ is at best quasicoherent, and is distributed depending on ω (i.e., it is entangled with $|\omega\rangle$). If $|\Psi_{\rm coh}(t)\rangle$ has mass M, then $\hat{M}|\Psi_\omega(t)\rangle = (M-\omega)|\Psi_\omega(t)\rangle$.

It is clear that most of the $A_H^{3/2}/\ell_{\rm Pl}^3$ states of the infalling Hilbert space are not quasicoherent and, therefore, if excited they do not decay thermally at all. For example, if a super-Planckian Hawking quantum is generated with $\omega \sim M/2$, then, obviously, the entangled states $|\Psi_{M/2}(t)\rangle$ must be very far from being coherent. Of course, during the first stages of the evaporation process we know that, as long as Hawking's computation is accurate, most quanta have $\omega \sim (GM)^{-1}$ and, since $(GM)^{-1} \ll M$, then all the states $|\Psi_{1/GM}(t)\rangle$ should be quasicoherent. How many of these states are there? This is a difficult computation to do, but one can estimate their number by counting the number of ways to extract independent subsystems of energy $(GM)^{-1}$ from a system of energy M:

$$\dim\left\{|\Psi_{1/GM}(t)\rangle\right\} \sim \frac{M}{\langle\omega\rangle} \sim GM^2 \sim S_{\rm BH}\,. \qquad (55)$$

That is the correct order of magnitude. We think that this notion of quasicoherence as a basis for black hole entropy is the closest to the spirit of the phenomenological derivation of the entropy based on Eq. (3) and, most importantly, it does not necessarily rely on unknown quantum gravity effects, which could pollute Hawking's calculation even in the earliest stages of the evaporation process.

In any case, if important deviations from thermality should occur from the beginning, variants of this picture can be accommodated. For example, if we consider a set of infalling states where the decay cross section has a (not necessarily thermal) profile

$$\Gamma_{\rm out} \sim A_H e^{-f(M,\delta M)}\,, \qquad (56)$$

then, following the discussion in the introduction, the entropy associated to this subset of the Hilbert space is

$$S \sim \int dM \frac{\partial f}{\partial(\delta M)}(M,0)\,, \qquad (57)$$

which does not necessarily scale as the horizon area.

It would be very interesting to further study these notions in simplified models.

It is amusing to speculate what this picture implies for the late stages of the evaporation process. With the definition (55), $S_{\rm BH}$ is clearly decreasing in time, because the quasiclassical infalling state progressively decoheres. It is clear that, after a number of soft emissions of order GM^2, so that the remaining mass is, say $M/2$, then the infalling state is very poorly approximated by an external classical field. Therefore, further decay will not proceed with a thermal cross section; it seems that the infalling matter can become "fuzzy" still at macroscopic masses, thus spoiling Hawking's prediction long before higher derivative gravity counterterms become important. In contrast with other scenarios [8,31], this would be a purely "soft" resolution of the information puzzle. Of course, under these conditions the "operational" version of the equivalence principle is violated: any infalling observer trying to experience a smooth transition through the horizon would have lost its classical properties in a much earlier stage.

ACKNOWLEDGMENTS

It is a pleasure to thank C. Callan, D. Gross, F. Larsen, and A. Peet for useful discussions. The work of J.L.F.B. was supported by the National Science Foundation under Grant No. PHY90-21984, while R.E. was partially supported by a FPI grant from MEC (Spain) and projects UPV 063.320-EB119-92 and CICYT AEN93-0435. R.E. would also like to thank the High Energy Physics Group at Princeton University for hospitality during the last stages of this work.

APPENDIX A

In this appendix we check explicitly in Rindler space the equality of the Green's functions of the operators $\widehat{\mathcal{L}}$ and $\mathcal{L}$, when both are computed with no-boundary condition at the horizon.

The Green's function of $\widehat{\mathcal{L}}$ can be written as [see Eq. (36)]

$$\widehat{G}(x,x') = \langle x|\widehat{\mathcal{L}}^{-1}|x'\rangle = \sum_{n,\omega}\left(\frac{4\pi^2 n^2}{\beta^2}+\omega^2\right)^{-1} e^{2\pi i n\Delta\theta/\beta}\psi_\omega^*(\mathbf{x})\psi_\omega(\mathbf{x}')\,, \qquad \text{(A1)}$$

where $\psi_\omega(\mathbf{x})$ are eigenfunctions of Eq. (24) for the particular case of Rindler space. Clearly, this yields the solution to Eq. (38) in the text.

The metric of d-dimensional Euclidean Rindler space is

$$ds^2 = \xi^2 d\theta^2 + d\xi^2 + \sum_{j=1}^{d-2} dx_j^2\,, \qquad \text{(A2)}$$

where θ is the Euclidean time, ξ is the coordinate that labels constant acceleration trajectories, and x_j are flat transverse coordinates. In these coordinates, and after separation of transverse space variables, Eq. (24) takes the form of a Bessel equation. As stressed in the text, an important feature here is that in the absence of a cut off for small ξ, the spectrum of frequencies ω is continuous.

The following expression for the Green's function of a massive scalar can be readily obtained:

$$\widehat{G}^{\beta}(x,x') = \frac{1}{\pi^2}\int \prod_{j=1}^{d-2} \frac{dp_j}{2\pi} e^{ip_j \Delta x_j} \int_0^{\infty} d\omega \sinh \pi\omega \times K_{i\omega}(\mu\xi) K_{i\omega}(\mu\xi') \sum_{k=-\infty}^{+\infty} e^{-\omega|\Delta\theta + k\beta|}, \quad \text{(A3)}$$

where $\mu^2 \equiv m^2 + \sum_j p_j^2$, and $K_{i\omega}(\mu\xi)$ are modified Bessel functions.

Now, the sum over k can be performed as

$$\sum_{k=-\infty}^{+\infty} e^{-\omega|\Delta\theta+k\beta|} = \frac{\beta}{2\pi}\frac{\cosh\omega(\Delta\theta - \beta/2)}{\sinh \beta\omega/2}. \quad \text{(A4)}$$

We will be interested precisely in $\beta = \beta_H = 2\pi$.

Transverse momenta can also be integrated (details on similar manipulations can be found in Ref. [32]):

$$\int \prod_{j=1}^{d-2} \frac{dp_j}{2\pi} e^{ip_j\Delta x_j} K_{i\omega}(\mu\xi) K_{i\omega}(\mu\xi') = \frac{1}{2}\int_{-\infty}^{\infty} d\lambda \, e^{i\omega\lambda} \left(\frac{m}{2\pi\gamma}\right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(m\gamma), \quad \text{(A5)}$$

with $\gamma^2(\lambda) \equiv \xi^2 + \xi'^2 + 2\xi\xi' \cosh\lambda + \sum_j (\Delta x_j)^2$.

Therefore,

$$\widehat{G}^{2\pi}(x,x') = \frac{1}{2\pi^2}\int_0^{\infty} d\omega \cosh\omega(\Delta\theta - \pi) \int_{-\infty}^{\infty} d\lambda\, e^{i\omega\lambda} \times \left(\frac{m}{2\pi\gamma}\right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(m\gamma). \quad \text{(A6)}$$

At this moment we want to interchange the order of integrations. Convergence then requires $\mathrm{Im}\lambda > |\Delta\theta - \pi|$, so that after integrating ω we find

$$\widehat{G}^{2\pi}(x,x') = \frac{i}{2\pi^2}\int_C \frac{\lambda d\lambda}{\lambda^2 + (\Delta\theta-\pi)^2} \times \left(\frac{m}{2\pi\gamma}\right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(m\gamma), \quad \text{(A7)}$$

where the contour C runs from $-\infty$ to $+\infty$ passing above the pole at $\lambda = i|\Delta\theta - \pi|$. We can split the integration contour into a straight line along the real axis and a clockwise contour encircling the pole. The former contribution vanishes by antisymmetry of the integrand, whereas the latter yields

$$\widehat{G}^{2\pi}(x,x') = \frac{1}{2\pi}\left(\frac{m}{2\pi\sqrt{2\sigma}}\right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(m\sqrt{2\sigma}), \quad \text{(A8)}$$

where $2\sigma \equiv \xi^2 + \xi'^2 - 2\xi\xi'\cos\Delta\theta + \sum_j (\Delta x_j)^2$ is the geodesic separation between the points x, x' as written in Rindler coordinates. Then Eq. (A8) is precisely the Euclidean, zero-temperature, Green's function in Minkowski space, i.e., the Hartle-Hawking Green's function, with no-boundary condition placed at the horizon. It must be noted that Eq. (A8) admits an expansion into Bessel functions of integer order, corresponding to the standard solution of the Laplacian $\mathcal{L}$ in disk $\times R^{d-2}$, regular at the origin and with discrete frequency spectrum.

APPENDIX B

In this appendix we briefly study the possible thermal instabilities of the linear dilaton vacuum of two-dimensional dilaton gravity, along the lines of the four-dimensional analysis of Ref. [11]. This is an interesting exercise because Euclidean gravity is on a much firmer ground in two dimensions and there is a chance that all manipulations have a meaning in Lorentzian signature. For example, string theory in the light cone and the Euclidean covariant approach provides an example of such an equivalence.

The Euclidean action of two-dimensional dilaton gravity is [33]

$$I = -\frac{1}{2}\int_{\mathcal{M}} e^{-2\varphi}[R + 4(\nabla\varphi)^2 + 4\lambda^2] - \oint_{\partial\mathcal{M}} e^{-2\varphi}K + C_{\infty}, \quad \text{(B1)}$$

where C_{∞} is determined for our purposes by requiring that, on a Hamiltonian thermal manifold, $I(\text{cylinder}) = \beta M_{\text{ADM}}$.

In the conformal gauge $g_{\alpha\beta} = e^{2\rho}\delta_{\alpha\beta}$ we have

$$I = -\frac{1}{2}\int_{\text{Disk}} \left[-2e^{-2\varphi}\partial^2(\rho-\varphi) + 4\lambda^2 e^{2(\rho-\varphi)}\right] - \oint_{\partial\mathcal{M}} e^{-2\varphi}K + \lambda\oint_{\infty} e^{-2\varphi}. \quad \text{(B2)}$$

The (Euclidean) classical black holes are parametrized by the mass M,

$$ds^2 = \frac{d\sigma^2 + d\theta^2}{1 + e^{-2\lambda\sigma}M/\lambda}, \quad \text{(B3)}$$

and a dilaton $\varphi = -\frac{1}{2}\log[M/\lambda - \exp(2\lambda\sigma)]$ where σ is a tortoise coordinate (the horizon is at $\sigma = -\infty$). The solution with $M = 0$ is the linear dilaton vacuum: $g_{\alpha\beta} = \delta_{\alpha\beta}$, $\varphi = -\lambda\sigma$, which becomes strongly coupled at left infinity. In this case, unlike in four-dimensional black holes, the Hawking temperature is unrelated to the mass and only depends on the cosmological constant, $T_H = \lambda/2\pi$. This is an important difference, since it implies that all black holes have the same temperature and that the phenomenological entropy is proportional to the mass $S_{\text{BH}} = 2\pi M/\lambda$. The classical suppression factor for black hole nucleation vanishes in this case as

$$I_{\text{cl}}(M) = \beta M - \oint_{\infty} e^{-2\varphi}K = \beta M - \beta M = 0\,. \quad \text{(B4)}$$

Also, if we set $\beta \neq 2\pi/\lambda$, thus going offshell,

$$I_{\text{cl}}(M,\lambda,\beta) = \beta M - 2\pi e^{-2\varphi_H}\chi(\text{disk}) = \beta M - \frac{2\pi M}{\lambda}, \quad \text{(B5)}$$

where φ_H is the value of φ at the horizon and χ is the Euler-Poincaré characteristic. Therefore the conical singularity method yields the right answer for the entropy, as well as the classical method, Eq. (10), since, at the critical temperature,

$$S_{\rm BH} = \beta M - I_{\rm cl} = \beta M = \frac{2\pi M}{\lambda}\,. \tag{B6}$$

Now, the one-loop computation of the free energy around a particular instanton is similar to that in Ref. [11]. Here we have a renormalizable theory, but the position-dependent coupling $g_s = e^{\varphi}$ makes it very difficult the nonperturbative analysis of the path integral.

At a perturbative level there is a potential instability coming from the fact that the dilaton field has the wrong metric. In this respect, it plays a role similar to the conformal factor of the metric in four-dimensional gravity, and should not be considered as a physical excitation. In fact, pure two-dimensional dilaton gravity has no propagating degrees of freedom. This is readily seen in the Lorentzian path integral with the action (B2). The functional integration over φ induces the condition that $\rho - \varphi$ be harmonic, so we can choose a (Kruskal) gauge in which $\rho = \varphi$. If we want to maintain this in the Euclidean path integral, we must integrate $\varphi = \varphi_{\rm cl} + i\delta\varphi$ over the imaginary axis, and this produces a functional δ function $\prod_x \delta(-\nabla^2(\rho - \varphi_{\rm cl})) = \det^{-1}(-\nabla^2)\prod_x \delta(\rho - \varphi_{\rm cl})$. This enforces $\rho = \varphi_{\rm cl}$ and the determinant is canceled by the ghost determinant.

The analysis goes through if one adds appropriate counterterms to take care of the one-loop conformal anomalies. Here one finds many variants of the same model. For example, the one-loop action studied in Refs. [34,35] is constructed such that the manipulations above make sense with $\exp(-2\varphi)$ replaced by $\Omega \equiv \exp(-2\varphi) + N\varphi/24$. In general, the effective action must preserve conformal invariance, and by means of nonlinear field redefinitions from ρ, φ to new fields X, Y, one can map the model to an open string theory [36]:

$$I = -\frac{1}{2}\int [-(\partial X)^2 + (\partial Y)^2 + 4\lambda^2 e^{C(X-Y)}] + I_{\rm boundary} + I^{(N)}_{\rm matter}\,. \tag{B7}$$

So we see that Y works like a target time. In Lorentzian quantization one must cancel X against Y, leaving the N "transverse" matter excitations. In Euclidean quantization one must rotate $Y \to iY$ as well, so that $\det^{-1}(-\nabla^2)$ from the X, Y integrals cancels against the ghost determinant.

As a result, for N scalar matter fields the perturbative partition function is proportional to $\det^{-N/2}(-\nabla^2)$, which is positive definite. No imaginary part of the free energy is generated, and consequently there is no black hole nucleation. In addition, the absence of propagating gravitons rules out any possible infrared Jeans instability.

This absence of tunneling barrier is compatible with the classical canonical thermodynamical analysis. The free energy for the combined system of two phases is (we neglect the boundary free energy)

$$F = F_{\rm rad} + F_{\rm BH} = -\frac{\pi}{6}NLT + M\left(1 - \frac{2\pi}{\lambda}T\right)\,. \tag{B8}$$

At the critical temperature $T_H = \lambda/2\pi$ there is a flat direction in M, and the canonical ensemble makes sense for two-dimensional black holes, at least within perturbation theory.

It is also interesting to analyze the classical microcanonical ensemble, where one maximizes the combined entropy $S = (\pi/3)NLT + 2\pi M/\lambda$ at fixed total energy $E = (\pi/6)NLT^2 + M$. The result in this case is very different from that in four dimensions [37]. If the energy density $\varepsilon = E/L$ is less than a critical value $\varepsilon_c = \lambda^2 N/(24\pi)$ then we have pure radiation with temperature $T = \sqrt{6\varepsilon/\pi N}$. Above this energy the temperature remains constant $T_H = \lambda/2\pi$ and the mass of the black hole grows linearly as $M = E - \varepsilon_c V$.

Regarding the one-loop divergences of the entropy, it is well known [9] that the logarithmic divergence $S_{\rm div} = N/6 \log \epsilon^{-1}$ from N matter fields contributes an infinite additive constant to S and cannot be renormalized in λ. In this respect, two-dimensional black holes follow a pattern different from their four-dimensional counterparts.

[1] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).
[2] G. 't Hooft, Nucl. Phys. **B256**, 727 (1985).
[3] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2752 (1977).
[4] J. Bekenstein, in *General Relativity*, Proceedings of the 7th Marcel Grossmann Meeting, Stanford, California, 1994, edited by M. Ruffini and M. Keiser (World Scientific, Singapore, 1995).
[5] V. Frolov and I. Novikov, Phys. Rev. D **48**, 4545 (1993).
[6] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373 (1986).
[7] M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993).
[8] G. 't Hooft, in *Salamfestschrift*, Proceedings of the Conference of Highlights of Particle and Condensed Matter Physics, Trieste, Italy, 1993, edited by A. Ali *et al.* (World Scientific, Singapore, 1993), p. 284.
[9] L. Susskind and J. Uglum, Phys. Rev. D **50**, 2700 (1994).
[10] J. L. F. Barbón, Phys. Lett. B **339**, 41 (1994); A. Dabholkar, Nucl. Phys. **B439**, 650 (1995); Phys. Lett. B **347**, 222 (1995); D. A. Lowe and A. Strominger, Phys. Rev. D **51**, 1793 (1995); R. Emparan, "Remarks on the Atick-Witten behavior and strings near black hole horizons," Report Nos. EHU-FT-94/10, hep-th/9412003 (unpublished).
[11] D. J. Gross, M. J. Perry, and L. G. Yaffe, Phys. Rev. D **25**, 330 (1982).

[12] C. Callan and F. Wilczek, Phys. Lett. B **333**, 55 (1994).
[13] D. Garfinkle, S. B. Giddings, and A. Strominger, Phys. Rev. D **49**, 958 (1994); S. W. Hawking, G. Horowitz, and S. F. Ross, *ibid.* **51**, 4302 (1995).
[14] S. B. Giddings, Phys. Rev. D **49**, 4078 (1994); **51**, 6860 (1995).
[15] I. K. Affleck, Phys. Rev. Lett. **46**, 388 (1981).
[16] M. Bañados, C. Teitelboim, and J. Zanelli, Phys. Rev. Lett. **72**, 957 (1992).
[17] L. Susskind, "Some speculations about black hole entropy in string theory," Report No. RU-93-44 (unpublished); Report No. hep-th/9309145 (unpublished).
[18] R. Emparan, Phys. Rev. D **51**, 5716 (1995).
[19] G. W. Gibbons, S. W. Hawking, and M. J. Perry, Nucl. Phys. **B138**, 141 (1978).
[20] W. G. Unruh, Phys. Rev. D **14**, 870 (1976).
[21] J. L. F. Barbón, Phys. Rev. D **50**, 2712 (1994).
[22] S. P. de Alwis and N. Ohta, "On the entropy of quantum fields in black hole backgrounds," Report Nos. COLO-HEP-347, OU-HET-207, hep-th/9412027 (unpublished).
[23] R. Laflamme, Nucl. Phys. **B324**, 233 (1989).
[24] A. O. Barvinsky, V. P. Frolov, and A. I. Zelnikov, Phys. Rev. D **51**, 1741 (1995).
[25] F. Larsen and F. Wilczek, "Geometric entropy, wave functionals and fermions," Report Nos. PUPT 1480, IASSNS 94/51, hep-th/9408089 (unpublished).
[26] M. R. Brown and A. C. Ottewill, Phys. Rev. D **31**, 2514 (1985); J. S. Dowker, *ibid.* **33**, 3150 (1986).
[27] D. Kabat and M. J. Strassler, Phys. Lett. B **329**, 46 (1994).
[28] J. G. Demers, R. Lafrance, and R. Myers, Phys. Rev. D **52**, 2245 (1995); D. Kabat, "Black hole entropy and entropy of entanglement," Report Nos. RU-95-06, hep-th/9503016 (unpublished); D. V. Fursaev and S. N. Solodukhin, "On one-loop renormalization of black hole entropy," Report Nos. E2-94-462, hep-th/9412020 (unpublished); F. Larsen and F. Wilczek, "Renormalization of black hole entropy and of the gravitational coupling constant," Report Nos. PUPT 1541, IASSNS 95/49, hep-th/9506066 (unpublished).
[29] C. Holzhey, F. Larsen, and F. Wilczek, Nucl. Phys. **B424**, 443 (1994).
[30] L. Susskind, "The world as a hologram," Report Nos. SU-ITP-94-33, hep-th/9409089 (unpublished).
[31] K. Schoutens, E. Verlinde, and H. Verlinde, in Proceedings of the 1993 Trieste Spring School (unpublished); Y. Kiem, H. Verlinde, and E. Verlinde, "Quantum horizons and complementarity," Report Nos. CERN-TH.7496/94, PUPT-1504, hep-th/9502074 (unpublished).
[32] P. Candelas and D. J. Raine, J. Math. Phys. **17**, 2101 (1976).
[33] C. Callan, S. B. Giddings, J. Harvey, and A. Strominger, Phys. Rev. D **45**, 1005 (1992).
[34] K. Schoutens, E. Verlinde, and H. Verlinde, Phys. Rev. D **48**, 2690 (1993).
[35] T. D. Chung and H. Verlinde, Nucl. Phys. **B418**, 305 (1994).
[36] A. Bilal and C. Callan, Nucl. Phys. **B394**, 73 (1993).
[37] G. W. Gibbons and M. J. Perry, Proc. R. Soc. London **A358**, 467 (1978).