PHYSICAL REVIEW D VOLUME 51, NUMBER 12 15 JUNE 1995

Why aren't black holes infinitely produced?

Steven B. Giddings*
Department of Physics, University of California, Santa Barbara, California 93106-9530
(Received 28 December 1994)

Unitarity and locality imply a remnant solution to the information problem, and also imply that Reissner-Nordström black holes have an infinite number of internal states. Pair production of such black holes is reexamined including the contribution of these states. It is argued that the rate is proportional to the thermodynamic quantity $\mathrm{Tr}e^{-\beta H}$, where the trace is over the internal states of a black hole; this is in agreement with estimates from an effective field theory for black holes. This quantity, and the rate, is apparently infinite due to the infinite number of states. One obvious out is if the number of internal states of a black hole is finite.

PACS number(s): 04.70.Dy, 04.50.+h

I. INTRODUCTION

Despite much recent effort the problem of what happens to quantum-mechanical information thrown into a black hole remains a puzzling problem. A variety of detailed scenarios can be boiled down to three basic pictures:[1] information is destroyed, information is returned in the Hawking radiation, or information is left behind in a black hole remnant. As is by now well known, if one attempts to describe black hole formation and evaporation from a low-energy, effective point of view each of these possibilities encounters serious conflicts with basic low-energy principles such as energy conservation, locality, and crossing symmetry.

Those advocating a remnant scenario [6–8] have attempted to evade the problem of infinite production by hypothesizing that black hole remnants are not correctly described by low-energy effective field theory and/or that crossing somehow fails [9–11]. Fertile ground for the investigation of these possibilities is provided by the phenomenon of pair production of charged black holes in background electromagnetic fields. If one hypothesizes that information is neither destroyed nor reemitted, then there should be an infinite number of internal states of such a black hole: one can feed in arbitrarily large amounts of information-rich matter, then allow evaporation to extremality [12]. Furthermore, instantons for such processes are described in [13–16]. Remnant advocates have hoped that a reliable calculation of the resulting rate for pair creation would be finite, and that charged black holes would therefore serve as a guide to formulation of theories of infinitely degenerate remnants with finite production, which might also extend to neutral remnants.

Indeed, this testing ground is critical. Charged black holes might show us the way to a theory of remnants, but if they do not, then the remnant hypothesis apparently dies with them. The reason is that the basic postulates of the remnant hypothesis, namely, unitarity (no information destruction) and locality (no reradiation of information in Hawking information) imply an infinite degeneracy of charged black holes, and if this implies that charged black holes are infinitely pair produced then these postulates are not correct, removing the *raison d'être* for neutral remnants. Pair production of charged black holes is thus a litmus test for the theory of remnants.

This paper will at the outset accept these postulates and attempt to investigate their viability through a more careful investigation of the pair production problem for Reissner-Nordström black holes. The essential features of these arguments extend as well to pair production of dilatonic black holes [15,16]. It begins by reviewing some of the basic features of remnant theories and the argument for infinite pair production that follows from an effective description, as well as issues that remnants raise for black hole thermodynamics. Next the role of Reissner-Nordström black holes as remnants of the Hawking process in the charged sector is reviewed, and a description of the infinite states appropriate to an outside observer is outlined. The following section contains a reinvestigation of the Schwinger process for Reissner-Nordström black holes. It is argued that the contribution of the infinite number of states is contained in the fluctuation determinant around the instanton, and that this cannot be computed without full knowledge of Planck scale physics. However, the calculation is nearly identical to that of $\mathrm{Tr}e^{-\beta H}$ for a black hole in contact with a heat bath, and if there are an infinite number of nearly degenerate black hole states, then this appears infinite independent of our inability to describe them explicitly. The emergence of such a factor agrees with the rate computation done in the effective approach. In closing, possibilities for avoiding infinite production of Reissner-Nordström black holes are outlined.

II. BASICS OF REMNANTS

As stated in the Introduction, the postulates of unitarity and locality imply that the information lost in the Hawking evaporation of a black hole is left behind in a

*Electronic address: giddings@denali.physics.ucsb.edu
[1]For reviews see [1–5].

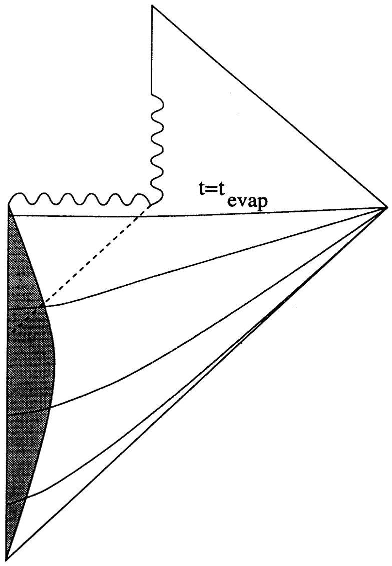

FIG. 1. The Penrose diagram for an evaporating neutral black hole, together with a time slicing.

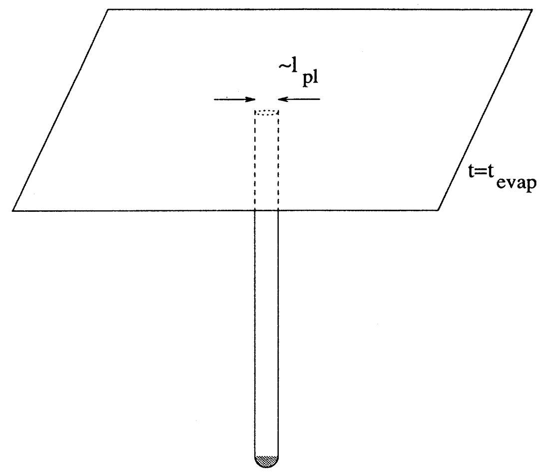

FIG. 2. The spatial geometry of a late time slice through Fig. 1. As the radius of the black hole decreases through evaporation, the slice becomes a thin, and in the limit, Planckian fiber attached to the asymptotic geometry.

black hole remnant. Consider an initial black hole of mass M that leaves behind a remnant. This should have a mass $m \sim M_{\rm Pl}$, and due to the difficulty in emitting its large information $I \sim M^2$ with its small available energy, it will have a very long lifetime [17,2]:

$$\tau \sim M^4. \tag{2.1}$$

With an arbitrary initial black hole, an arbitrarily large amount of information can be stored in such a remnant, and so there must be an infinite number of internal states or species of such an object.

At first sight two obvious issues leap to mind. First, the infinite number of remnant species would appear to lead to infinite total remnant production rates in various physical processes. Second, it seems rather strange to have absolutely stable remnants, and it is not obvious what physics would give a remnant decay time as in (2.1).

Answers to the second question may be provided by understanding how the internal physics of remnants returns the information and respects the constraints placed by the relationship between information and energy. One approach to this physics has been recently proposed by Polchinski and Strominger [18,19]. They discuss the proper treatment of a scenario where the black hole interior branches off a baby universe in an attempt to carry information off. As in the case of baby universes, there is not a repeatable loss of information [20,21,5], and the couplings self-adjust so that the interior takes a long time to split off and a long-lived remnant results. Alternatively, note that if we consider a late time slice through a plausible Penrose diagram of an evaporating black hole, Figs. 1 and 2, this slice consists of a Planckian fiber attached to a flat geometry. With this picture of a remnant in mind, it is quite plausible that the appropriate behavior follows from the necessarily Planckian physics of the fiber. Thus solutions to the second problem are easily imagined.

The first issue is much more difficult. Since remnants are localized massive objects, we expect that the only description of them that is local and/or causal, Lorentz invariant, and quantum mechanical is in terms of an effective field ϕ_a. Here a is a species label. The couplings of the remnant field to other fields may be quite complicated, but near zero momentum transfer their approximate form would appear to be dictated simply by the mass and charge of the remnant. If, for simplicity, we think of electrically charged scalar remnants (the magnetic case follows via electromagnetic duality), they should therefore be described by an effective action

$$S_{\rm eff} = \int d^4x \sum_a (-|D_\mu\phi_a|^2 - m_a^2|\phi_a|^2) + \cdots, \tag{2.2}$$

with $D_\mu = \partial_\mu + iQA_\mu$, and where higher dimension terms are not written. At low momenta transfers the latter terms are expected to be negligible.

Such a coupling will allow Schwinger production of pairs of remnants. The decay rate of a background electric field A^0_μ is given by the imaginary part of the Euclidean vacuum-to-vacuum amplitude:

$$V_4\Gamma = 2\,\mathrm{Im}\,\ln\left(\int \mathcal{D}A_\mu\mathcal{D}\phi e^{-S[A^0+A]-S_{\rm eff}[\phi_a]} \Big/ \int \mathcal{D}A\mathcal{D}\phi e^{-S[A]-S_{\rm eff}[\phi_a]}\right), \tag{2.3}$$

where V_4 is the four-volume and $S[A]$ is the Maxwell action. To lowest order in the coupling electromagnetic fluctuations are neglected, and one finds

$$V_4\Gamma = \text{Im ln det}\left\{\left[-(\partial_\mu + iQA^0_\mu)^2 + m_a^2\right]/\left[-\partial_\mu^2 + m_a^2\right]\right\}. \tag{2.4}$$

Then ln det=Tr ln, and the operator traces can be rewritten in terms of single particle amplitudes, giving

$$\text{Im}\ln\det[-(\partial_\mu + iQA^0_\mu)^2 + m_a^2] = 2\,\text{Im}\int_0^\infty \frac{dT}{T}\int_{X(0)=X(T)} \mathcal{D}X \exp\left\{-\int_0^T d\tau\left(\frac{\dot{X}^2}{2} + iQA_\mu \dot{X}^\mu\right)\right\}\text{tr}_a \exp\left\{-\frac{T}{2}m_a^2\right\}. \tag{2.5}$$

Each term in the sum over a is well approximated by a Schwinger saddle point corresponding to circular Euclidean motion, and the decay rate is then given by

$$\Gamma \sim \sum_a e^{-\pi m_a^2/QE}. \tag{2.6}$$

If there are an infinite number of species, the sum diverges and the total production rate is infinite. If we furthermore suppose that the remnant spectrum consists of nearly degenerate states, $m_a = M + \Delta m_a$, with $\Delta m \ll M$, then (2.6) can be written

$$\Gamma \sim e^{-\pi M^2/QE}\text{Tr}_a e^{-\beta\Delta m}, \tag{2.7}$$

with $\beta = 2\pi M/QE$. Thus it is proportional to the partition function for the nearly degenerate states.

One might attempt to find remnant effective theories that avoid this problem, perhaps by avoiding minimal couplings altogether.[2] However, no such theories have been formulated. Furthermore, as will be shown in Sec. IV, a remarkably similar result holds for Schwinger production of black holes.

In closing this general discussion, it is also worth emphasizing the conflict between theories of remnants and black hole thermodynamics [22]. In particular, in a remnant scenario the Bekenstein-Hawking entropy of a black hole is not related in any obvious way to the number of its internal states. The Bekenstein bound is directly violated by remnants. In a viable remnant scenario it may be that the only role of black hole thermodynamics is to furnish a macroscopic description of properties of the black hole: for example, the temperature versus mass relation is independent of the number of black hole internal states.

III. EXTREMAL REISSNER-NORDSTRÖM STATES

In studying the possibility of a remnant solution to the information problem it is useful to investigate a situation where the Hawking process is guaranteed to leave a remnant: the evaporation of a charged black hole.

Indeed, following the preceding logic, suppose that we begin with a charged[3] black hole. There are several ways that one could be obtained; it could come from collapse of charged matter, or be one end of a Wheeler wormhole created in Schwinger production, or be an extremal black hole either of primordial origin or created in the Schwinger process. In each case the global geometry of the solution differs. However, each shares the important feature that the end point of the Hawking process leaves an infinite number of internal black holes states, and that these states are practically indistinguishable independent of the origin of the black hole.

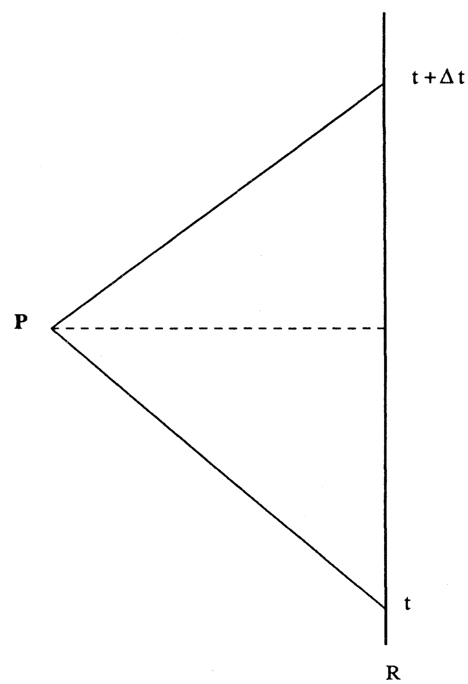

FIG. 3. Shown in the geometry motivating the definition of the radiolocation coordinates (r_*, t) of an arbitrary point P.

[2] It is conceivable that there are theories with no minimal couplings to remnants, in which minimal couplings are mimicked by more complicated couplings. For example, it may be that there is no sense in which we can scatter a charged particle off an extremal charged black hole without exciting it.

[3] This discussion will consider either electrically or magnetically charged black holes; issues connected with Schwinger charge loss of an electric black hole will be postponed to a subsequent section.

To see this, notice that we can throw an arbitrary matter configuration into our black hole over an arbitrarily long time. We then allow it to evaporate; in each case the end point of the evaporation process is a black hole with the extremal value of the mass, $M = Q$. If the information from the infalling matter neither escapes the black hole nor is annihilated, then we have managed to construct an infinite variety of internal black hole quantum states depending on the configuration of the infalling matter.

It is useful to be more explicit in describing these states. To do so we introduce a particular choice of coordinates. Suppose that the metric is asymptotically flat; we base the coordinates on an asymptotic observer fixed at large radius R from the black hole. Let the observer carry a clock. The coordinates of an arbitrary event P are defined by radiolocation; see Fig. 3. If an inwardly directed light signal emitted from R at time t is reflected from P and returns to R at time $t + \Delta t$, P is assigned coordinates

$$(r_*, t) = \left(R - \frac{\Delta t}{2}, t + \frac{\Delta t}{2}\right).$$

In the case of a static black hole geometry, this prescription gives the usual tortoise coordinates. Notice in particular that the interior of the black hole is not covered.

The states of the black hole can be described using data specified on these time slices. The time scale for evaporation of a black hole to extremality [12] is $\tau \sim Q^3$. Consider the configuration on a slice in the far future, long after the mass excess $M - Q$ has been radiated past radius R. Thus in terms of the data on the slice inside R, the states have energy Q and the external appearance

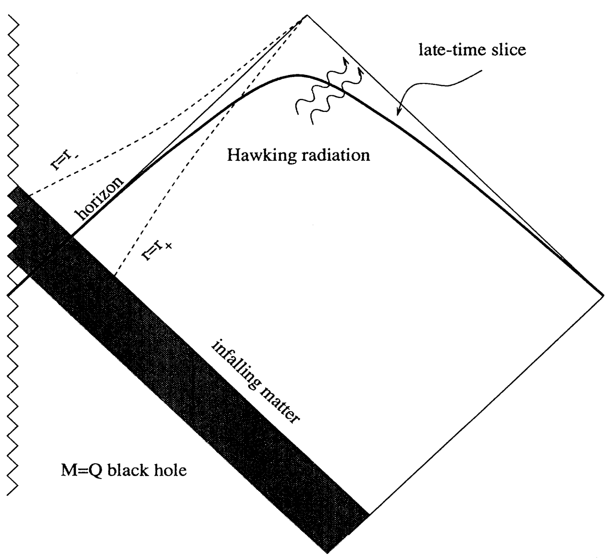

FIG. 4. The Penrose diagram for an initially extremal black hole, into which some matter is thrown, and which subsequently evaporates back to extremality.

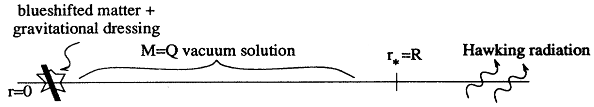

FIG. 5. Shown is a schematic description of the state of the black hole of Fig. 4 on a late time slice that stays outside the true horizon.

of a black hole. However, differences between states are seen if one investigates near the horizon. The slice crosses the infalling matter, whose different configurations imply different quantum states of the black hole. These features are illustrated in Figs. 4 and 5.

At large times the Hawking radiation turns off and the black hole must asymptote to a superposition of its exact energy eigenstates. Semiclassically these states all have the same appearance. The matter distribution asymptotes to $r_* = -\infty$, and outside the solution should be the extremal vacuum geometry. In this geometry the proper distance to the matter also becomes infinite: the extremal Reissner-Nordström black hole has an infinitely long throat. In the long-time limit, the only difference between solutions is in the matter configuration at the end of this throat.

Of course, the semiclassical description based on our slices eventually fails. One way of estimating where this happens is to inquire when observers traveling on world lines of fixed r_* see the proper frequencies or wave numbers at the Planck scale. For example, an outgoing s wave of the Hawking radiation for a massless field behaves like

$$e^{-i\omega(t-r_*)};$$

at infinity the typical frequency is $\omega \sim T_H$. Planck physics becomes relevant at the radius where the proper frequency becomes Planckian; if n^μ is the unit normal to the time slices, this occurs where $n^\mu k_\mu \sim M_{\rm Pl}$. Likewise, if the infalling matter is followed in, the description fails when it becomes Planckian. Notice that while one ordinarily expects to have a description of the infalling state that does not require Planck scale physics if one uses the frame of the infalling observer, it is the translation of this description to the frame of the outside observer that requires Planckian physics. Therefore it seems that the differences between the infinite number of states are not discernible without a full theory of quantum gravity.

As stated above, there are several distinct types of black hole, depending on whether one began with a truly extremal black hole, with one end of a wormhole, or with a nonextremal black hole. However, in each case the final state of the Hawking process is a solution with $M = Q$, and in the semiclassical description these ground-state solutions differ only in the configuration at $r_* = -\infty$. It is not known if this statement is modified in the full quantum theory.

In our later discussion black holes that are thermally excited will play a central role. Suppose we take an $M = Q$ black hole, and place it in a box of blackbody radiation. The black hole will then absorb radiation until the accretion rate and the Hawking radiation rate match;

this should happen when the temperature of the radiation and black hole are equal. For neutral black holes this equilibrium can be arranged to be stable, despite the negative specific heat, by taking the radius of the box to be sufficiently small, $r \sim M^{5/3}$. For charged black holes sufficiently close to extremality the specific heat is positive, so this is even easier to achieve. If we compute the partition function, $\mathrm{Tr}e^{-\beta H}$, for states in the box, the infinite number of ground states will contribute and the partition function will therefore diverge.

As first pointed out by Gibbons and Hawking [23], an elegant path integral derivation of the partition function also exists. The evolution operator $e^{-\beta H}$ can be turned into a Euclidean path integral by the usual steps, and the trace corresponds to the periodic identification. Thus one is instructed to sum over asymptotically flat geometries with period β at infinity. One ordinarily assumes that these geometries should be regular in the vicinity of the horizon. However, in accord with the above arguments, doing so would discard the contribution of the infinite number of states: these correspond to configurations that do not behave smoothly at the horizon. Thus it seems that the instruction to sum over regular geometries only captures a finite subset of the states, and does not give the correct result for $\mathrm{Tr}e^{-\beta H}$. The infinite states only appear to be accounted for if one allows singular behavior in the vicinity of the horizon.[4] Although the trace must contain an infinite factor from the infinite number of states, the prescription for a detailed calculation cannot be given in the absence of a quantum theory of gravity.

It should be emphasized that the infinite number of states contributing to the trace can be explicitly counted. For example, one could imagine forming black holes from diffuse collapsing matter. The initial state of this matter can be taken to be in finite volume, and can be specified in the presence of a short-distance cutoff. The number of such states that collapse to form a black hole is therefore enumerable and finite. If one assumes unitary evolution (and in particular conservation of the norm of a state) and that information does not escape from black holes, then the infinite volume limit gives an infinite number of final states containing a black hole and in which the only difference is in the internal state of the black hole. This infinity in the number of states should also appear in the quantity $\mathrm{Tr}e^{-\beta H}$. Other authors [25,26] have advocated calculations that give a finite result for the latter quantity. It would seem that this is only possible either if these calculations are not including all black hole states or if our assumptions are wrong and black holes only have finitely many states.

[4]It seems quite likely to the present author that the contribution of the infinite number of states and the conformal factor problem are closely connected. Indeed, the divergent integral over the conformal factor quite likely is connected to the required infinity in the partition function, and the unstable behavior seems connected to the irregularity of the geometry. (A possibly related comment has been made by Carlip and Teitelboim [24].)

IV. PAIR PRODUCTION VIA TUNNELING

A good starting point for the description of Schwinger production is the Wheeler-DeWitt equation, or its completion in a full theory of quantum gravity. This equation acts on wave functionals

$$\Psi[^3g, f(x), A(x), T]$$

of the three geometry, the matter fields f, the gauge field A, and asymptotic time T. The solutions of this equation are given by the Lorentzian path integral. Where the semiclassical approximation is valid and in classically allowed domains, leading order solutions of this equation are simply given in terms of classical solutions of the coupled Einstein-Maxwell-matter equations according to the standard WKB formalism.

One can likewise consider tunneling through classically forbidden regions, as in the Schwinger process. Again where the semiclassical approximation is valid, the leading semiclassical wave functions are given by classical solutions, in this case of the Euclidean continuation of the equations of motion.[5] This gives the tunneling rate to a configuration that is a classical turning point. The system can also tunnel to nearby configurations via paths near the Euclidean solution. The contribution to the tunneling rate of these nearby configurations is well approximated by the usual fluctuation determinant [28], as in standard instanton calculations. Alternatively, these results can be obtained directly from the Euclidean functional integral [29].

There are two types of Euclidean solutions describing pair production of black holes.[6] The first [14,30] describes production of a pair of oppositely charged black holes connected by a Wheeler-wormhole throat. The black holes are consequently above extremality. The second [13,16] describes production of an oppositely charged pair of extremal black holes that are not connected. For simplicity we will henceforth focus on creation of magnetically charged black holes in a magnetic field. Then both of these solutions asymptote to the Melvin universe [31], which is the closest approximation to a uniform magnetic field in general relativity.

Necessary conditions for validity of the semiclassical approximation are that $Q \gg 1$ (super-Planckian black holes) and $QB \ll 1$ (weak magnetic fields). The leading semiclassical tunneling rate is given by the action

$$\Gamma \sim e^{-S} \sim e^{-\pi m^2/QB}. \tag{4.1}$$

[5]The connection between real-time and Euclidean solutions for other field theories is made explicit in [27,28].

[6]For a more complete discussion see [16].

However, this estimate clearly misses the contributions to the tunneling rate of the infinite number of states. The instanton describes tunneling to the classical turning point, which is a pair of vacuum Reissner-Nordström black holes. However, nearby configurations, reached by nearby paths, have nontrivial matter and gravitational excitations. Following our preceding discussion, these should include the infinite number of states of the black hole. Their contributions are therefore included to linear order by the fluctuation determinant around the instanton, or including interactions, by doing the full Euclidean functional integral in the vicinity of the saddle point.

Even the fluctuation determinant is difficult to compute directly. However, it is closely related to the result of another calculation, namely, that of the thermal partition function for a Reissner-Nordström black hole.

To see this, let us give a more detailed description of the Ernst instanton solutions [32]. For simplicity focus on the magnetic case. They are

$$\begin{aligned} ds^2 &= (x-y)^{-2}A^{-2}\Lambda^2[\{G(y)dt^2 - G^{-1}(y)dy^2\} \\ &\quad +G^{-1}(x)dx^2] \\ &\quad +(x-y)^{-2}A^{-2}\Lambda^{-2}G(x)d\varphi^2, \end{aligned} \tag{4.2}$$

$$A_\varphi = -\frac{2}{B\Lambda}\left[1+\frac{1}{2}Bqx\right]+k,$$

where the functions $\Lambda \equiv \Lambda(x,y)$ and $G(\xi)$ are given by

$$\lambda = \left[1+\frac{1}{2}Bqx\right]^2 + \frac{B^2}{4A^2(x-y)^2}G(x), \tag{4.3}$$

$$G(\xi) = (1-\xi^2 - r_+A\xi^3)(1+r_-A\xi),$$

and q is given by

$$q^2 = r_+r_-. \tag{4.4}$$

A and B are parameters, and the constant k in the expression for the gauge field is introduced so that the Dirac string of the magnetic field of a black hole is confined to one axis. Finally, it is useful to factorize G,

$$G = -r_+r_-A^2(\xi-\xi_1)(\xi-\xi_2)(\xi-\xi_3)(\xi-\xi_4), \tag{4.5}$$

with

$$\xi_1 = -\frac{1}{r_-A} \tag{4.6}$$

and ξ_2, ξ_2, ξ_2 the ordered roots of the remaining expression. For small acceleration, $r_+A \ll 1$, the zeros have expansions

$$\begin{aligned} \xi_2 &= -\frac{1}{r_+A} + r_+A + \cdots, \\ \xi_3 &= -1 - \frac{r_+A}{2} + \cdots, \\ \xi_4 &= 1 - \frac{r_+A}{2} + \cdots. \end{aligned} \tag{4.7}$$

The solution (4.2) describes a pair of black holes with opposite magnetic charge in a background magnetic field. The independent parameters of this solution are $r_\pm$, A, and B, to be thought of (roughly) as the inner and outer horizon radii, the acceleration, and the magnetic field strength. For general parameters the solution (4.2) is not regular. In particular, if the acceleration is not related to the charge, mass, and magnetic field, then there will by a physical string singularity attaching the two black holes. With these parameters matched, the Lorentzian geometry is regular outside the horizons, and it can readily be shown that the black holes follow approximately hyperbolic trajectories corresponding to uniform acceleration. The time t used in (4.2) corresponds to Rindler time asymptotically far from the black hole, as can be shown by investigating the limit $x \to y$.[7]

The instanton for pair production follows from substituting $\tau = it$. Now there is another condition that must be imposed on the parameters to have a regular solution. To see this, note that the point $y = \xi_3$ corresponds to the acceleration horizon and regularity there requires a specific periodicity for Euclidean Rindler time τ as in standard treatments of Rindler space. However, $y = \xi_2$ corresponds to the black hole horizon, and a periodic identification of τ is also required there as in standard treatments of the black hole. Equating these periods gives a second relation between the parameters. In a sense, this is a condition matching the acceleration and Hawking temperatures so that the black hole can be thought of as being in static equilibrium with the acceleration radiation.

Consider the case of small acceleration. There are actually two solutions to the temperature-matching condition. The first [14] is if the black hole is taken to be slightly above extremality in order to raise the temperature enough above zero to match the acceleration temperature. When matched onto Lorentzian solutions, these instantons are seen to create pairs of nonextremal black holes connected by a Wheeler wormhole. The second [13] is at first sight surprising: for truly extremal black holes, the horizon is at infinite proper distance and so any periodicity is allowed. This latter case corresponds to the limit $\xi_1 = \xi_2$ and pair produces extremal black holes.

In a quantum treatment the solutions will receive corrections from the back reaction of the Hawking and/or acceleration radiation on the geometry. Equilibrium with thermal acceleration radiation is quite analogous to equilibrium with a thermal bath. In particular, the extremal case is likely no longer a solution as the black hole is raised above extremality. Therefore we focus on the nonextremal wormhole solutions.

Possible contributions of the infinite states arise in computing the functional integral over configurations near the instanton. This is hard. However, let us investigate the instanton in the throat region, where in

[7] For a more detailed description of the features of this solution, see, for example [16].

accord with our earlier discussion the infinite states are expected to be located if they are present.

The vicinity of the black hole corresponds to $y \to \xi_2$ in (4.2). Using the periodicity matching condition

$$\xi_1 - \xi_2 - \xi_3 + \xi_4 = 0 \tag{4.8}$$

and the expansions (4.7), the metric takes the form

$$ds^2 \to q^2[-\sinh^2 w\, dt^2 + dw^2 + d\Omega_2^2] \tag{4.9}$$

after a change of variables. This agrees exactly with the form of the free near-extremal Reissner-Nordström solution near the horizon, as can be seen from the substitution

$$\frac{r - r_+}{r_+ - r_-} = \frac{1}{2}(\cosh w - 1) \tag{4.10}$$

and rescaling t. Subleading corrections to this expression vanish in the limit $qB \to 0$ and $w \to 0$. In particular, the leading correction to g_{tt} is of the form $qB(\cosh w - 1) \sim qBw^2$. These corrections are small and furthermore do not shift the location of the horizon or qualitatively change the form of the solution in the vicinity of the black hole. The corrections do become substantial, however, when $w \sim -\ln(qB)$, where the transition to the asymptotic solution takes place. The length of the black hole throat is therefore $l \sim -q \ln(qB)$. The corrections are exponentially small in the length of the black hole throat. Finally, note that to leading order in the qB expansion, the parameter q and the physical charge Q are equal.

The solution will also receive corrections from the backreaction of the Hawking/acceleration radiation. Since all of the effects of the acceleration, with the exception of the thermal fluxes, die near the horizon, the back-reaction-corrected solution should be of the same form as that of a free black hole in equilibrium with radiation, plus small corrections. Detailed descriptions of such solutions have not been given, although back-reaction-corrected solutions for extremal Reissner-Nordström black holes without the thermal flux have been investigation in [33] and dilatonic black holes in equilibrium with an inward flux were found numerically in [34,35]. Outside the black hole these are expected to preserve the general form of the near-extremal solution. It should be noted that as in the free case, there are an infinite number of solutions which in the far future differ only in their state at the horizon.

In pair creation, these states are accounted for in the Euclidean functional integral about the instanton. Once again, we do not know how to evaluate this integral without understanding quantum gravity. However, we have just argued that in the throat region, for

$$w \ll -\ln(qB), \tag{4.11}$$

the solution is identical to that of a free black hole in equilibrium with thermal radiation, up to small corrections. Indeed, a quick check shows that the local temperature at the end of the black hole throat is $T \sim B$, the expected value from the acceleration radiation. Although the semiclassical approximation fails, the contributions to the functional integral from Planck-scale dynamics should be essentially the same in either case. Indeed, using the composition property of the functional integral, it can be split along the dotted line in Fig. 6. The contribution from the bottom of the cup should be approximately the *same* as that from corresponding region in the computation of the Euclidean functional integral for free black hole in contact with a thermal bath. As explained in Sec. III we also cannot calculate the latter functional integral, but we know it gives the partition function. Thus the production rate contains a factor of the form

$$\mathrm{Tr} e^{-\beta H}. \tag{4.12}$$

There will of course be differences between this quantity and the functional integral for the instanton, arising from differences outside the throat region. However, if the black hole has infinitely many states down the throat, there should be contributions of the form (4.12) from these states to the pair creation rate. Note finally that this factor corresponds to the factor found in the low-energy effective calculation of the rate (2.7).

To summarize these arguments, although the calculation of the functional integral cannot be done without using details of quantum gravity, the calculation should be the same as that for the throat contribution in Tr $\exp\{-\beta H\}$. If a Reissner-Nordström black hole of charge Q has infinite numbers of nearly degenerate ground states, there is a corresponding infinity in both expressions; the pair creation rate is infinite.

Although the contributions of the infinite states come into the instanton calculation through singular geometries, note that there are also smooth geometries that contribute to the functional integral: these are precisely the original Wheeler wormhole configurations, with the

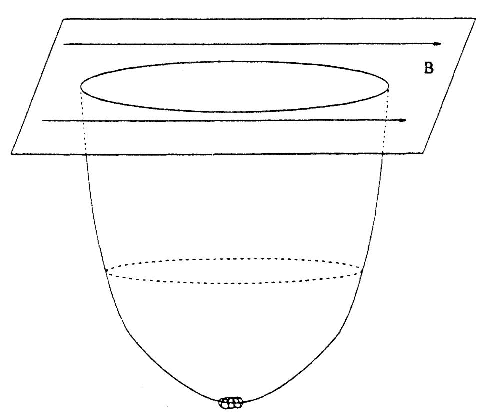

FIG. 6. A picture of the Euclidean instanton solution, for small QB. Below the dotted line, the solution is nearly identical to the Euclidean Reissner-Nordström solution. The contribution of the infinite number of states is expected to arise from configurations that rapidly oscillate near the would-be horizon.

"internal" states unexcited. In accordance with the arguments of [12,11,36] it is quite plausible that pair creation of these regular Wheeler wormholes is in fact finite because they are rather special states. More general states are found by throwing matter into a Wheeler wormhole and then letting it evolve back to equilibrium with the radiation.

Finally, an interesting question is what is the typical state of the Hawking radiation for the pair created black holes. Since the Euclidean section of the instanton closely approximates the Euclidean section of the unaccelerated black hole away from the horizon, the Green functions for excitations are computed according to the Hartle-Hawking prescription. This ensures that the state produced is essentially the Hartle-Hawking state [11,36].

V. CONCLUSIONS

If we assume the validity of quantum mechanics and also that information is not returned in Hawking radiation, this seems to inevitably lead to the statement that the Hawking process leaves behind an infinite number of "remnant" states in the evaporation of a black hole. We have argued that if this is the case, there is no obvious mechanism for suppressing the resulting infinite pair-creation rate for Reissner-Nordström black holes. This appears to be a catastrophe.

There are several ways to attempt to escape this conclusion. Let us consider them in turn.

One possibility is that extremal black holes do not exist as true ground states of any physical theory. A charged black hole itself sheds charge by Schwinger production, at a rate [37]

$$\frac{dQ}{dt} \simeq \frac{e^4 Q^3}{r_+} e^{-\pi m^2 r_+^2/eQ} \tag{5.1}$$

for quanta of mass m and charge e. This can, for example, be compared to rate of change of the mass through Hawking emission. In the electric case, black holes will rapidly discharge through electron emission unless $M \gtrsim 10^5 M_\odot$. The situation is improved in the magnetic case. If one, for example, considers a grand unified theory, production of magnetic monopoles by extremal black holes is highly suppressed for

$$M \gg g/M_{\rm mon}^2, \tag{5.2}$$

a much more reasonable constraint.[8] Furthermore, for

$$M \gtrsim gQ/M_{\rm mon} \tag{5.3}$$

monopole emission is forbidden. We can therefore easily create a charged black hole with an infinite number of internal states by beginning with a black hole satisfying (5.3) and feeding it information and energy at a sufficient rate to balance the Hawking radiation for as long as we please. If it then Hawking decays to extremality, one finds an infinite number of species of metastable extremal black holes. In [39] it was argued that, for remnant decay lifetimes larger than the Schwinger time,

$$\tau_S \sim l_S \sim M/QB, \tag{5.4}$$

finite decay rates do not substantially affect pair production. The exponential suppression of (5.1) makes this easy to achieve for moderate Q.

If one instead worked in a theory with no grand unified theory (GUT) monopoles, it is quite possible that magnetic black holes could be pair produced as Wheeler wormholes and then be absolutely stable to Schwinger emission. In any case, even if discharge instability were to make pair creation of Reissner-Nordström black holes finite, this would just shift the infinite production problem into the neutral remnant sector.

A second attempted out is to appeal that Schwinger production of black holes requires a very strong field that is uniform over extremely large scales. Indeed, for true Schwinger production the field should be uniform over at least the magnetic length of (5.4),

$$L \gtrsim \frac{10^{20}\ \text{cm}}{B(\text{tesla})}, \tag{5.5}$$

which is not likely to be realized. However, for much weaker fields that are nonuniform, one also expects there to be a finite but minuscule production rate for each species as long as there is sufficient energy available to make a pair of black holes, $E \gg 2M_{\rm Pl}$. Although this rate has enormous suppressions due to form factors, etc., these are overcompensated by the overall infinite number due to the infinite number of remnant species. If Schwinger production is not finite, it is probably not possible for these rates to be finite either.

A third possible escape is that although the difference between the instanton geometry and the geometry of the Euclidean asymptotically flat black hole vanishes far down the throat, this small difference conspires with Planck scale physics to make the calculation of the fluctuation determinant differ from that of the partition function by an infinite factor. In light of the fact that without this the calculation seems to be giving one a result in agreement with effective arguments, and in accord with the implications of crossing symmetry, this seems unlikely.

A fourth possibility is that despite the fact that black holes have an infinite number of states, there is a prescription to compute $\text{Tr}e^{-\beta H}$ for a black hole that gives a finite answer, and furthermore there is a reason that this is the correct prescription to use in calculating the production rate. One proposal is that the infinity can be absorbed through renormalization of Newton's constant [25,26].[9] However, this seems unlikely to work as

[8] Note that the instability of [38] is also absent for sufficiently large charge.

[9] I thank A. Strominger for conversations on this issue.

Newton's constant should be renormalized to give correct low-energy gravitational scattering amplitudes at low energies. Once this renormalization has been done, it is still apparently true that black holes have infinite numbers of states, and thus the trace over black hole states should still have a nonsubtracted infinity.

A fifth possibility is that $\mathrm{Tr}e^{-\beta H}$ is finite because Hawking was right: information is lost in quantum gravity, and this information loss causes black holes to have only a finite number of states. However the serious conflicts with energy conservation [40,41,11,5] that arise from this possibility remain; there is no known effective description of local information loss that conserves energy. This is a major problem.

The final possibility is that black holes have a finite number of states and information is conserved: it is emitted in the Hawking process. Despite the fact that this possibility has recently been vigorously pursued [42–44], there is as of yet little evidence of a concrete mechanism for string nonlocality or other physics to imprint information on the Hawking radiation. Nonetheless, given the results of this paper it is possible that the only way to avoid infinite production is if the information indeed comes out in the Hawking radiation.

ACKNOWLEDGMENTS

I wish to thank G. Horowitz, M. Srednicki, and A. Strominger for helpful discussions. This work was supported in part by DOE Grant No. DOE-91ER40618 and by NSF PYI Grant No. PHY-9157463.

---

[1] J. A. Harvey and A. Strominger, in Proceedings of the 1992 Trieste Spring School on String Theory and Quantum Gravity, Trieste, Italy, 1992 (unpublished).

[2] J. Preskill, in *Black Holes, Membranes, Wormholes, and Superstrings*, Proceedings of the International Symposium, Woodlands, Texas, 1992, edited by S. Kalara and D. V. Nanopoulos (World Scientific, Singapore, 1993).

[3] S. B. Giddings, in *Quantum Gravity*, Proceedings of the 7th Nishinomiya Yukawa Memorial Symposium, Nishinomiya, Japan, 1992, edited by K. Kikkawa and M. Ninomiya (World Scientific, Singapore, 1993).

[4] L. Thorlacius, presented at 1994 Trieste Spring School (unpublished).

[5] S. B. Giddings, presented at the Summer School in High Energy Physics and Cosmology, Trieste, Italy, 1994 (unpublished).

[6] C. G. Callan, S. B. Giddings, J. A. Harvey, and A. Strominger, Phys. Rev. D **45**, R1005 (1992).

[7] T. Banks, A. Dabholkar, M. R. Douglas, and M. O'Loughlin Phys. Rev. D **45**, 3607 (1992).

[8] S. B. Giddings and A. Strominger, Phys. Rev. D **46**, 627 (1992).

[9] T. Banks and M. O'Loughlin, Phys. Rev. D **47**, 540 (1993).

[10] T. Banks, M. O'Loughlin, and A. Strominger, Phys. Rev. D **47**, 4476 (1993).

[11] S. B. Giddings, Phys. Rev. D **49**, 4078 (1994).

[12] A. Strominger and S. Trivedi, Phys. Rev. D **48**, 5778 (1993).

[13] G. W. Gibbons, in *Fields and Geometry*, Proceedings of 22nd Karpacz Winter School of Theoretical Physics: Fields and Geometry, Karpacz, Poland, 1986, edited by A. Jadczyk (World Scientific, Singapore, 1986).

[14] D. Garfinkle and A. Strominger, Phys. Lett. B **256**, 146 (1991).

[15] F. Dowker, J. P. Gauntlett, D. A. Kastor, and J. Traschen, Phys. Rev. D **49**, 2209 (1994).

[16] F. Dowker, J. Gauntlett, S. B. Giddings, and G. T. Horowitz, Phys. Rev. D **50**, 2662 (1994).

[17] R. D. Carlitz and R. S. Willey, Phys. Rev. D **36**, 2327 (1987); **36**, 2336 (1987).

[18] J. Polchinski and A. Strominger, Phys. Rev. D **50**, 7403 (1994).

[19] A. Strominger, Report No. hep-th/9410187, UCSB UCSBTH-94-34 (unpublished).

[20] S. Coleman, Nucl. Phys. **B307**, 864 (1988).

[21] S. B. Giddings and A. Strominger, Nucl. Phys. **B307**, 854 (1988).

[22] J. D. Bekenstein, talk at Seventh Marcel Grossman meeting, Stanford, California, 1994 (unpublished).

[23] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2752 (1977).

[24] S. Carlip and C. Teitelboim, Report No. gr-qc/9312002, IAS/UC Davis IASSNS-HEP-93/84=UCD-93-34 (unpublished).

[25] L. Susskind and J. Uglum, Phys. Rev. D **50**, 2700 (1994).

[26] T. M. Fiola, J. Preskill, A. Strominger, and S. Trivedi, Phys. Rev. D **50**, 3987 (1994).

[27] H. J. DeVega, J. L. Gervais, and B. Sakita, Nucl. Phys. **B142**, 125 (1978); I. Bender and H. J. Rothe, *ibid.* **B142**, 177 (1978), and references therein.

[28] K. Bitar and S.-J. Chang, Phys. Rev. D **18**, 435 (1978).

[29] S. Coleman, in *Aspects of Symmetry: Selected Erice Lectures* (Cambridge University Press, London, England, 1985).

[30] D. Garfinkle, S. B. Giddings, and A. Strominger, Phys. Rev. D **49**, 958 (1994).

[31] M. A. Melvin, Phys. Lett. **8**, 65 (1964).

[32] F. J. Ernst, J. Math. Phys. **17**, 515 (1976).

[33] S. P. Trivedi, Phys. Rev. D **47**, 4233 (1993).

[34] B. Birnir, S. B. Giddings, J. A. Harvey, and A. Strominger, Phys. Rev. D **46**, 638 (1992).

[35] L. Susskind and L. Thorlacius, Nucl. Phys. **B382**, 123 (1992).

[36] P. Yi, Phys. Rev. D **51**, 2813 (1995).

[37] G. W. Gibbons, Commun. Math. Phys. **44**, 245 (1975).

[38] K. Lee, V. P. Nair, and E. J. Weinberg, Phys. Rev. Lett. **68**, 1100 (1992).

[39] S. B. Giddings, Phys. Rev. D **49**, 947 (1994).
[40] T. Banks, M. E. Peskin, and L. Susskind, Nucl. Phys. **B244**, 125 (1984).
[41] M. Srednicki, Nucl. Phys. **B410**, 143 (1993).
[42] G. 't Hooft, gr-qc/9310006, Utrecht Report No. THU-93/26, and references therein (unpublished).
[43] L. Susskind, L. Thorlacius, and J. Uglum, Phys. Rev. D **48**, 3743 (1993).
[44] L. Susskind, hep-th/9409089, Stanford Report No. SU-ITP-94-33, and references therein (unpublished).