Black-hole thermodynamics and the Euclidean Einstein action

James W. York, Jr.
Institute of Field Physics, Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27514
(Received 11 October 1985)

Using the approach to black-hole thermodynamics initiated by Gibbons and Hawking, in terms of the Euclidean Einstein action, I show that the canonical ensemble with elements of radius r and temperature $T(r)$ for hot gravity with black holes is well defined. This follows from the double valuedness of solutions of the Euclidean Einstein equation with canonical boundary conditions. One of the solutions is a locally stable hole. Its partition function is well defined and implies the entropy $S=4\pi M^2$ as well as a generalized version of black-hole thermodynamics that reduces to the usual form if $rM^{-1}\rightarrow\infty$. The density of states of the locally stable hole is real and nonpathological. The free energy of this hole can be negative, while that of the other (unstable) solution is always positive. Consequently, the direct nucleation of black holes from hot flat space, as proposed by Gross, Perry, and Yaffe, can be given a thermodynamically consistent description. The scaling laws for hot gravity are obtained and applied to phase transitions between hot flat space and locally stable holes. The free energy of the unstable solution forms the effective potential barrier between these phases. The ground state of the canonical ensemble is always locally stable in the semiclassical approximation. If N is the effective number of massless fields of helicity zero in hot flat space, then when *either* $r \lesssim N^{1/2}$ *or* $T \gtrsim N^{-1/2}$, hot flat space is the most probable ground state. Independently of N, if $rT<(27)^{1/2}(8\pi)^{-1}$ there can be no real black hole in the canonical ensemble.

I. INTRODUCTION

In the original demonstration that a Schwarzschild black hole of mass M has a temperature $T_\infty=(8\pi M)^{-1}$ measured at a large distance from the hole ($r \gg M$), Hawking used techniques of quantum field theory on a given classical background spacetime.[1] Subsequently, several studies that attempt to relate the thermal properties of black holes to quantum gravity have appeared.[2-4] The case of a hole in thermal equilibrium with its surroundings is particularly important in making the connection to quantum gravity. For this purpose, it is useful to describe the hole in terms of the "real Euclidean section" of its geometry, in which the time is rotated to a purely imaginary value and is given the period $\beta_* = T_\infty{}^{-1}$; this is the so-called "Hartle-Hawking-Israel" state.[5] Quantum gravity enters through the interpretation of the contribution of the Euclidean Schwarzschild geometry to the canonical partition function for hot gravity written as a Euclidean functional integral.

Two very different conclusions have been reached using the Euclidean approach. Gibbons and Hawking,[2] and Hawking,[3] deduced the thermodynamic properties of black holes that had already been found from the temperature formula and a reinterpretation of the classical laws of black-hole mechanics in a manner suggested by the work of Bekenstein.[6] In particular, they found that the black-hole entropy $S=4\pi M^2$ arises from the contribution of the *classical* first-order Euclidean Einstein action $I_{\rm BH}$ of the Schwarzschild geometry to the canonical partition function. This conclusion was important because ordinarily the value of the classical action is not considered to be physically significant; it can simply be absorbed into the normalization of the functional integral. However, Hawking[3] argued forcibly that the classical action of a black hole does contribute to the entropy and is therefore not to be ignored. Nevertheless, the relation $T_\infty=(8\pi M)^{-1}$ implies negative heat capacity for a black hole in a large cavity and, as a result, the root-mean-square energy fluctuations are imaginary. Hence, in this approach, the canonical ensemble seems not to be well defined and appeal was made to a microcanonical picture.[3] However, I shall not follow that path here.

On the other hand, Gross, Perry, and Yaffe[4] reached a different conclusion. They demonstrated explicitly that the action of a Schwarzschild geometry of mass M in a large cavity ($r \gg M$) corresponds to a saddle point, rather than to a minimum, with respect to small perturbations of the background geometry. They found in the perturbation spectrum of the Schwarzschild geometry a single "negative mode" that decreases its action in the second variation.[7] The negative mode gives rise from a field-theoretic argument to an imaginary free energy for the black hole. This was interpreted as indicating an instability of "hot flat space," that is, an instability of gravitons and other massless fields on a flat background geometry. The idea is that activation of this instability by thermal (quantum) fluctuations could spontaneously excite hot flat space over an effective potential barrier, resulting in the nucleation of a black hole. This would be a purely quantum-gravitational phenomenon having no classical counterpart.

It is not the same process as "ordinary" gravitational collapse of hot flat space by a relativistic Jeans instability, as the authors pointed out.

One can view nucleation of a black hole (BH) from hot flat space (HFS) thermodynamically as a phase transition. There is no conservation law to forbid it because the chemical potentials of both phases (BH and HFS) are zero. Moreover, energy need not be conserved because the system is assumed to be coupled to a large heat reservoir that keeps it at a fixed temperature. Phase transitions are ordinarily treated under conditions of fixed temperature and pressure (and chemical potential), in which case the appropriate thermodynamic potential is the Gibbs free energy $G=E-TS+pV$, E=thermal energy or total internal energy. However, in the canonical ensemble temperature and volume ("size") are fixed and the relevant thermodynamic potential is therefore the Helmholz free energy $F=E-TS$, which is related formally to the Euclidean action I by $I=\beta F$, $\beta\equiv T^{-1}$. A spontaneous process, such as a phase transition, in the canonical ensemble should *never* imply an increase of F, in the same sense that a spontaneous process in a microcanonical ensemble (fixed energy and volume) *never* implies a decrease of the entropy, which would violate the second law of thermodynamics.

The above well-known points suggest a difficulty in the description of nucleation proposed by Gross, Perry, and Yaffe. The free energy $F_{\rm HFS}$ of hot flat space is zero classically and is negative quantum mechanically, where $F_{\rm HFS}=-3^{-1}aT^4V$, a=Stephan's constant. For a black hole in a large cavity, the thermodynamic identity $dM=T\,dS$ and the Euler relation $M=2TS$ imply, as is well known, that $F_{\rm BH}=(0.5)M$, which is positive. Thus, nucleation of a black hole of mass M from hot flat space would seem to imply a spontaneous increase of F. Nevertheless, Gross, Perry, and Yaffe found that the rate Γ of nucleation events could be quite significant at high temperatures. They found $\Gamma=D\exp(-B)$, where D is a determinant and $B=(16\pi T^2)^{-1}$. They assumed that the critical black-hole mass nucleated would be $M=(8\pi T)^{-1}$.

Clearly there are problems to be resolved in assessing the physical significance of the black-hole action in its contribution to the Euclidean functional integral. Two profoundly different interpretations have been given, but each seems at best incomplete.

The purpose of this paper is to resolve the difficulties described above. I deal with the following points.

(1) The boundary conditions that define a canonical ensemble whose elements are spherical cavities of radius r and temperature T at r are such as to admit either no black hole (when $rT<$ a critical value) or two physically distinct ones, of which one is always unstable and the other is always at least locally stable. This holds even for arbitrarily large r.

(2) The locally stable (larger mass) hole can have a negative free energy under certain conditions, which enables one to see that it could form by nucleation from hot flat space. Its mass, however, is not $(8\pi T)^{-1}$. The unstable hole always has a positive free energy (or action) that forms the "effective potential" barrier between hot flat space and the locally stable hole. Its action is fairly well approximated by $B=(16\pi T^2)^{-1}$ if back-reaction effects are ignored. The determinant D is not zero because the free energy of the light hole corresponds to a saddle point of the action, implying the existence of a negative mode. The free energy of the heavy hole is a local minimum devoid of negative modes.

(3) The canonical ensemble for hot gravity possesses a locally stable state of lowest free energy (ground state). This state can be either hot flat space, or a "large" hole (that does not engulf the cavity), or a superposition of these phases. Which candidate prevails is readily determined from the given r and T.

(4) From the classical Euclidean action of the locally stable solution, one can deduce black-hole thermodynamics by generalizing slightly the approach of Gibbons and Hawking. One finds that the entropy is one-fourth of the area of the event horizon ($4\pi M^2$), but that there is a distinction between the mass M of the hole and the thermodynamic energy E of the hole embedded in a finite cavity. The difference is a self-energy term which is negligible if $r\gg M$. I deduce the scaling laws of black-hole thermodynamics, which play an important role in the theory of black-hole phase transitions, a subject to be dealt with in detail in another work though some aspects are treated here. (These transitions are first order in a thermodynamic description.) Naturally, the heat capacity of the locally stable hole is positive and, consequently, its energy fluctuations and density of states are real and nonpathological. The latter properties were shown recently to hold for black holes in anti—de Sitter spaces ($\Lambda<0$) under certain conditions.[8] However, I shall not introduced a cosmological constant Λ here.

(5) For fixed r and T, I introduce a generalized real free energy function $\bar{F}$ for black holes of any mass M. When $rT\geq$ (a critical value), $\bar{F}$ has extrema corresponding to the two equilibrium values of M, whose real Euclidean sections are topologically regular with Euler characteristic $\chi=2$. For other values of M, $\bar{F}$ is well behaved but the corresponding geometries are topologically defective. From E, $\bar{F}$, and S, one can construct all the other thermodynamic potentials for the black-hole phase. (One can include in these the modifications caused by equilibrium back-reaction effects, based on the results of Ref. 9.) This makes it possible to treat the formation (and disappearance) of black holes from the relativistic Jeans instability[4] or the nucleation instability[4] under a wide variety of ambient or boundary conditions. For instance, one can dispense altogether with "fixed walls" by using the Gibbs functions. This will be treated in detail elsewhere.

I shall limit attention to cases of explicit spherical symmetry. Thus, any black hole that is present will be assumed to be at the center of the (by definition) spherical cavity. This seems reasonable because, with the wall of the cavity at a fixed uniform temperature, one expects the locally stable equilibrium states to be manifestly spherically symmetric. Extension of the results by a "dilute-gas" approximation seems feasible, but I shall not attempt that here.

Throughout, I shall use absolute units ($G=\hbar=c=k_B=1$) except where restoration of conventional units is helpful.

II. CANONICAL BOUNDARY CONDITIONS AND DOUBLE VALUEDNESS OF THE SCHWARZSCHILD MASS

Two key problems that arise in describing self-gravitating systems thermodynamically are the following. (1) Such systems under certain conditions may collapse completely to form a black hole. However, Hawking's discovery ascribes a temperature to the hole and equilibrium thermodynamics can therefore be employed to describe the exterior of the collapsed object. (2) In equilibrium, a self-gravitating object does not have a spatially constant temperature in its vicinity; one has for a local observer at rest a local temperature that depends upon the observer's position. This is a well-known consequence of the principle of equivalence that implies that temperature is red- or blue-shifted in the same manner as the frequency of a quantum of energy.

Consider, then, an observer perched at the radius r of a spherical cavity enclosing (say) a black hole. If a bit of energy δE is added to the system so as to maintain equilibrium, it should have the local temperature $T(r)\equiv T$ prevailing at the boundary radius r of the cavity. It follows that $(\delta E)T^{-1}$ is independent of position if r and $T(r)$ are fixed and the formulation of black-hole thermodynamics in a system of finite size must reflect this fact. If there is a black hole of mass M at the center of the cavity, then

$$T(r)\equiv T=(8\pi M)^{-1}\left[1-\frac{2M}{r}\right]^{-1/2} \tag{1}$$

according to Hawking's result.

A canonical ensemble is defined by temperature and a variable measuring the size of the elements of the ensemble; this variable is ordinarily the volume. The above discussion motivates the following definition of the canonical ensemble, under conditions of spherical symmetry, when a black hole may be present at the center of the cavity. The *size* of the system is defined by the geometrically well-defined invariant area $A=4\pi r^2$ of the cavity wall, where r is either the standard flat-space radius or the standard Schwarzschild radial coordinate. It would seem to be pointless to use spatial volume as a measure of size when a black hole is present because the volume of a black hole is not defined at constant Schwarzschild time. The interior of a hole is not static in the Lorentzian description and the hole has no interior on its real Euclidean section. Correspondingly, the *temperature* is defined as the uniform temperature T of the *wall* (boundary) of the spherical cavity. This T is also uniform throughout the interior of the cavity if there is no hole inside; otherwise it is given by (1). In the latter case the usually purely "intensive" variable T becomes scale dependent and is therefore no longer purely intensive.

Suppose, then, that we are given the boundary radius r, or boundary area $A=4\pi r^2$, and the temperature T at this locus. We ask what topologically regular spherically symmetric Euclidean solutions of the Einstein equation fit these boundary conditions. One is flat space with boundary $S^1\times S^2$, where S^2 denotes a sphere of fixed area $4\pi r^2$ and S^1 is the circle corresponding to periodically identified Euclidean time. The *proper* length around this circle has a fixed value $\beta=T^{-1}$. Another solution is a Euclidean Schwarzschild geometry of mass M with boundary $S^1\times S^2$ geometrically identical to that described above. This geometry must be topologically regular (no conical singularity at its "axis" or horizon), a geometrical condition equivalent to the physical requirement of thermal equilibrium. The mass M is found by solving (1) for M as a function of r and T. An important finding is that the solution of (1) is in general *double valued,* as I have pointed out elsewhere.[9]

Analysis of (1) proceeds by squaring it, obtaining the roots of the cubic equation thereby obtained, and inserting these roots back into (1). I assume here and throughout that $r\geq 0$ and $T\geq 0$. One finds that if

$$rT\geq\frac{\sqrt{27}}{8\pi}\left[\frac{\hbar c}{k_B}\right], \tag{2}$$

then the only solutions of (1) are real and positive. They are given by

$$\frac{GM_1}{c^2}=\frac{1}{6}r\left[1-2\cos\left[\frac{\alpha}{3}+\frac{\pi}{3}\right]\right], \tag{3}$$

$$\frac{GM_2}{c^2}=\frac{1}{6}r\left[1+2\cos\frac{\alpha}{3}\right], \tag{4}$$

$$\cos\alpha=1-54\left[\frac{\hbar c}{8\pi rTk_B}\right]^2, \tag{5}$$

$$0\leq\alpha\leq\pi\ . \tag{6}$$

If the *inequality* in (2) holds, then $M_1<M_2$. If the *equality* holds, then $M_1=M_2$ and $rM_1{}^{-1}=rM_2{}^{-1}=3$; that is, the cavity wall coincides with the circular photon orbit of the real Lorentzian black-hole geometry.

On the other hand, if

$$rT<\frac{\sqrt{27}}{8\pi}\left[\frac{\hbar c}{k_B}\right], \tag{7}$$

then there are no real solutions M of (1). [There is, however, a conjugate pair of *complex* mass solutions of (1) with the given r and T real and positive. I do not know if these have any physical significance. In any case, I shall not regard them as "black holes" and shall ignore them in this paper, though future work might show them to be important if complex metrics are included in the functional integral.]

We see that a region of the T-r phase plane exists from which black holes are excluded *ab initio.* It is significant that the critical length scale implied by (7) varies as T^{-1} and consequently is *independent of the gravitational constant G.* This stands in sharp contrast with the critical length scale defining the relativistic Jeans length[4] for massless quanta, which varies as T^{-2} and depends explicitly on G. In fact (7) implies an absolute criterion: No black hole whatsoever can exist in thermal equilibrium, whether stable or unstable, inside a spherical region bounded by area $4\pi r^2=A$ with temperature T at r if (in absolute units)

$$AT^2 < \frac{27}{16\pi} \simeq 0.537\,148 \ . \tag{8}$$

If $M(r)$ is plotted against r for fixed $T>0$, one sees that the curve has two branches that join smoothly at $r=\sqrt{27}(8\pi T)^{-1}$ and $M=\sqrt{3}(8\pi T)^{-1}$. In the limit of large r, only the light-mass or lower (M_1) branch approaches $(8\pi T)^{-1}$, while on the upper branch, M_2 approaches $(0.5)r$ from below. Previously, people have in effect only been treating the large-r limit on the lower branch. This leads to the misleading conclusion that T and M are essentially synonymous. It will be seen below that the hole M_1 is thermodynamically unstable and M_2 is locally stable. Because (2) implies $(4\pi rT) \gtrsim (0.5)(27)^{1/2}$, useful approximations for M_1 and M_2 are

$$M_1 \simeq (8\pi T)^{-1}[1+(8\pi rT)^{-1}] \ , \tag{9}$$

$$M_2 \simeq \tfrac{1}{2} r[1-(4\pi rT)^{-2}] \ . \tag{10}$$

III. ENTROPY, ENERGY, AND ACTION

It is desirable to construct the thermodynamics of black holes in a manner that incorporates the role of quantum gravity, as stressed by Hawking.[3] For this purpose, I shall follow Gibbons and Hawking,[2] and Hawking,[3] in assuming that the partition function Z contains the first-order classical Euclidean Einstein action of a hole as its leading term. I extend their calculation to spherical cavities of finite size. I shall consider here only the contribution of *one* Schwarzschild geometry because, of the two permissible cases, it turns out that only one of them (M_2) is in fact thermodynamically locally stable. Only for the locally stable solution does it make sense to derive thermodynamic properties from the essentially classical contribution $\exp(-I)$ to Z. The unstable solution is, however, important as the mediator of phase transitions from hot flat space to locally stable black holes. The critical mass nucleated is M_2, not M_1.

As mentioned above, I shall take $Z=\exp(-I) = \exp(-\beta F)$, where I is the first-order Euclidean Einstein action including a subtraction term. The present calculation is at finite radius.

The action I is defined by[2]

$$I = I_1 - I_{\text{subtract}} \ , \tag{11}$$

where[10]

$$I_1 = -\frac{1}{16\pi}\int R\sqrt{g}\, d^4x + \frac{1}{8\pi}\oint \sqrt{\gamma} K\, d^3x \ . \tag{12}$$

Here, K is the trace of the extrinsic curvature tensor K_{ij} of the boundary $S^1 \times S^2 (r=\text{const})$ and γ_{ij} is its induced three-metric. The four-space metric is

$$ds^2 = \left[1-\frac{2M}{r}\right]dt^2 + \left[1-\frac{2M}{r}\right]^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2 \ . \tag{13}$$

The volume term in (12) is zero. The Euclidean time is integrated over the period $\beta_* = T_\infty^{-1} = 8\pi M$ and the proper length of the S^1 of the boundary is therefore

$$\beta = T^{-1} = \int_0^{\beta_*} g_{tt}^{1/2} dt = 8\pi M \left[1-\frac{2M}{r}\right]^{1/2} , \tag{14}$$

which depends on the boundary radius r. We find

$$\sqrt{\gamma} = r^2 \sin\theta \left[1-\frac{2M}{r}\right]^{1/2} , \tag{15}$$

$$K = -\frac{2}{r}\left[1-\frac{2M}{r}\right]^{1/2} - \frac{M}{r^2}\left[1-\frac{2M}{r}\right]^{-1/2} , \tag{16}$$

$$I_1 = 12\pi M^2 - 8\pi M r \ . \tag{17}$$

The action subtracted in (11) is simply I_1 evaluated for a flat four-metric with boundary $S^1 \times S^2$ identical to that of the Schwarzschild geometry; that is, it has proper area $4\pi r^2$ for the S^2 and proper circumference β for the S^1. A suitable flat four-metric is (note $\tau \neq t$)

$$ds^2 = d\tau^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2 \ , \tag{18}$$

where τ has period β. We have $\sqrt{\gamma} = r^2 \sin\theta$ and $K = -2r^{-1}$, yielding[11]

$$I_{\text{subtract}} = -\beta r = -8\pi M r \left[1-\frac{2M}{r}\right]^{1/2} , \tag{19}$$

$$I = 12\pi M^2 - 8\pi M r + \beta r \ . \tag{20}$$

The important point about (19) is that it is linear in β (β and r, or T and r, are the independent variables). As we shall see below, this means that the subtraction term has *no effect* on the entropy but "normalizes" the thermal energy E to zero for the Schwarzschild geometry with $M=0$. Hence, the subtraction is analogous to removing the zero-point energy for a collection of harmonic oscillators at a fixed temperature, which does not influence the entropy. These two features may well give guidance in constructing subtraction terms in cases where the bounding three-geometry cannot be embedded at a finite distance in a given background "reference" geometry.

From (20), we see that $I \to 4\pi M^2$ as $rM^{-1} \to \infty$ and $I \to -4\pi M^2$ for $rM^{-1} \to 2$. At $rM^{-1} = 2.25$, $I=0$. This is important, for it means that

$$2 \le rM^{-1} < 2.25 \Longleftrightarrow F < 0 \ . \tag{21}$$

The free energy of a black hole can thus be negative. This fact allows a thermodynamically consistent description of black-hole nucleation.

The thermodynamic relations for a black hole can be deduced straightforwardly from the usual formulas employing either Z or F or I; here the action I will be used. One has only to be careful in computing the derivatives because I has to be regarded as a function of r and T rather than of r and M that appear in (20). The thermodynamic energy will be denoted by E. We obtain ($A = 4\pi r^2$)

$$E = \left[\frac{\partial I}{\partial \beta}\right]_A = r - r\left[1-\frac{2M}{r}\right]^{1/2} , \tag{22}$$

$$S=\beta\left[\frac{\partial I}{\partial \beta}\right]_A - I = 4\pi M^2 \ . \tag{23}$$

First, we note that the entropy equals one-fourth of the area of the event horizon for all choices of r and T satisfying (2). This is in accord with the idea that the entropy of a black hole can be regarded as a property of the horizon.[6]

The formula for the thermodynamic energy E is worth comment because $E>M$ except in the limit $rM^{-1}\to\infty$, when $E=M$. First, we see that if r is fixed, then

$$(\delta E)_r=\delta M\left[1-\frac{2M}{r}\right]^{-1/2} \tag{24}$$

as anticipated in Sec. II, that is, $(\delta E)_r T^{-1}$ is independent of r. Second, we solve (22) for M and obtain

$$M=E-\frac{1}{2}\frac{E^2}{r} \ , \tag{25}$$

which states that the Arnowitt-Deser-Misner energy is the thermal energy plus the gravitational self-energy associated with the thermal energy. We find $M=E$ in the limits $rM^{-1}\to\infty$ or $rE^{-1}\to\infty$. That it is necessary in general to distinguish E and M will be apparent in Sec. VIII, which contains an estimate of the density of states $\nu(E)$ for a locally stable black hole.

Note that the subtraction term in I gives rise to the term r in (22). Without this term, the thermal energy E would have been $(-r)$ for $M=0$. However, $M=0$ in (22) yields $E=0$, as it should. Alternatively, we note that the subtraction term makes the *classical* action $I=0$ for flat space with arbitrary temperature and radius. Because it is linear in β, the subtraction term did not affect the entropy calculated in (23).

IV. STABILITY AND THE THERMODYNAMIC IDENTITY

Black-hole thermodynamics as deduced above will only make sense if a black hole can be in locally stable equilibrium in the canonical ensemble. *Thermal* stability is determined by the heat capacity. In keeping with the definition of the canonical ensemble, the heat capacity is defined at constant area $A=4\pi r^2$ of the cavity boundary. We find

$$C_A\equiv T\left[\frac{\partial S}{\partial T}\right]_A=\left[\frac{\partial E}{\partial T}\right]_A$$
$$=8\pi M^2\left[1-\frac{2M}{r}\right]\left[\frac{3M}{r}-1\right]^{-1} . \tag{26}$$

Hence, if $2M<r<3M$, the heat capacity is positive and one has thermal stability. The root-mean-square energy fluctuations ΔE defined by the square root of

$$\langle(\Delta E)^2\rangle=C_A T^2=\frac{1}{8\pi}\left[\frac{3M}{r}-1\right]^{-1} \tag{27}$$

are real in this case. We see from (23) that ΔE is finite even when $r\to 2M$, and therefore when $T\to\infty$. We find that, regardless of the value of M,

$$(\Delta E)_{\text{horizon}}=\frac{1}{\sqrt{4\pi}} \ . \tag{28}$$

A finite result is obtained as $r\to 2M$ because $C_A\to 0$ in this limit, that is, because no energy can be added to a hole in equilibrium without changing its radius.

If $r>3M$, $C_A<0$ and approaches $-8\pi M^2$ as $rM^{-1}\to\infty$. At $rM^{-1}=3$, C_A suffers an infinite discontinuity and changes signs. This behavior in itself does not indicate a phase transition in the canonical ensemble. It determines the relative scale (rM^{-1}) at which a hole can be in (meta)stable equilibrium when placed in thermal contact with a heat reservoir that holds T fixed at the given r. Referring back to the double valuedness of M, we see that the light hole M_1 is unstable while the heavy hole M_2 is thermally stable or at least metastable.

To complete the thermodynamic description of a black hole we need to construct the thermodynamic identity. Clearly $dE\neq T\,dS$; there is another term arising from the stresses or "pressure" associated with the self-energy. However, because volume is not defined, I replace ordinary pressure by a suitable "surface pressure" conjugate to the area $A=4\pi r^2$. This is defined in analogy in ordinary pressure by

$$\sigma=-\left[\frac{\partial E}{\partial A}\right]_S$$
$$=\frac{1}{8\pi r}\left[\left[1-\frac{M}{r}\right]\left[1-\frac{2M}{r}\right]^{-1/2}-1\right] . \tag{29}$$

We see that σ is non-negative and that

$$dE=T\,dS-\sigma\,dA \tag{30}$$

is an exact differential. Both E and dE are regular for all $r\geq 2M$; the divergences in T and σ cancel at $r=2M$. In the limit $rM^{-1}\to\infty$, we recover $dM=T_\infty dS$. Alternatively, one verifies that $dS=\beta\,dE+\beta\sigma\,dA$ is equivalent to $dS=T_\infty{}^{-1}dM$.

Integrating (30) shows that the Euler relation for black holes is $E=2(TS-\sigma A)$; E is homogeneous of degree one-half in S and A and of degree one in M and r. Hence the scaling laws for black-hole thermodynamics are defined by $M\to\lambda M\,(S\to\lambda^2 S)$, $r\to\lambda r\,(A\to\lambda^2 A)$, $T\to\lambda^{-1}T$, $\sigma\to\lambda^{-1}\sigma$, and $E\to\lambda E$. The normally purely "intensive" variable T therefore has to be scaled in the theory of hot black holes. This has interesting consequences in the description of phase transitions (Sec. VII).

Having defined the basic "mechanical" variable σ, we can investigate the *mechanical* stability of a black hole in a hot cavity. Ordinarily, for this purpose one would compute the isothermal compressibility $\kappa_T(V)$ (using pressure and volume),[12] but I will replace that by an appropriate analogous quantity. Stability requires that when the area of the wall is increased at constant temperature, then σ should also increase. Therefore, I define and calculate

$$\kappa_T(A)\equiv\frac{1}{A}\left[\frac{\partial A}{\partial\sigma}\right]_T=16\pi r\left[\frac{r}{M}\right]^3\left[\frac{3M}{r}-1\right]\left[1-\frac{2M}{r}\right]^{3/2}$$
$$\times\left\{1+\left[\frac{r}{M}\right]^3\left[\frac{3M}{r}-1\right]\left[\left[1-\frac{2M}{r}\right]^{3/2}-\left[1-\frac{3M}{r}+\frac{3M^2}{r^2}\right]\right]\right\}^{-1} . \quad (31)$$

This quantity is non-negative for $2M\leq r\leq 3M$, indicating stability, and is negative for $r>3M$, indicating instability.

It is perhaps useful to compare the above description of the mechanical stability of a black hole with that of pure massless radiation (hot flat space) using σ and A rather than p and V. From $-p\,dV=-\sigma\,dA$, or definition (29) applied to the usual radiation formulas, one has

$$\sigma=-\left[\frac{\partial E}{\partial A}\right]_S=\tfrac{1}{2}rp=\tfrac{1}{2}r(\tfrac{1}{3}aT^4) . \quad (32)$$

Because p is independent of V, the usual (volume) isothermal compressibility is not defined. However, applying the definition (31) of $\kappa_T(A)$ to the radiation gives

$$\kappa_T(A)=\frac{1}{A}\left[\frac{\partial A}{\partial\sigma}\right]_T=\frac{4}{rp}=\frac{2}{\sigma} , \quad (33)$$

which is positive as expected.

I conclude that it is reasonable to deduce the thermodynamic relations for the larger mass hole from the hypothesis of Gibbons and Hawking concerning the role of the Euclidean action in the partition function. Once obtained, the differential relations can also be applied to the smaller mass hole or to any spherical hole. The thermodynamic identity (30) is generally applicable. However, if [radius of boundary] > [3×(mass of black hole)], the hole will be unstable in the canonical ensemble. Its fate under the influence of small fluctuations will be either to decay to hot flat space or to grow until it reaches the stable mass value. With high probability, the final state will be the configuration of lower free energy.

V. STABILITY OF THE ACTION

The behavior of the thermodynamic variables C_A and $\kappa_T(A)$ suggests that the action of a black hole should be a strict minimum if $2M\leq r<3M$. This question has been studied recently by Allen[11] in an extension of the results of Gross, Perry, and Yaffe to the case of a boundary at finite radius. Allen found that when $r\lesssim 2.89M$, the saddle point in the action is eliminated. However, we must note that in the gauge used by Allen to describe "isothermal perturbations," the *area* of the surface S^2 on which T was fixed to have the value (1) did not remain equal to $4\pi r^2$. This is because his formulation of the perturbations required that $\delta g_{\phi\phi}=(\sin^2\theta)\delta g_{\theta\theta}(r)$ not be zero on the boundary. Hence, his value ($\cong 2.89$) and mine ($=3$) are in no obvious disagreement.

However, there is a more basic problem that arises from a nonzero variation of the geometry of the boundary $S^1\times S^2$. Because this boundary has nonvanishing extrinsic curvature, the surface term

$$(\delta I)_{\text{surface}}=\text{const}\times\oint_{S^1\times S^2}\sqrt{\gamma}(K^{ij}-K\gamma^{ij})\delta\gamma_{ij}d^3x \quad (34)$$

in the first variation of (11) will not be zero if $\delta\gamma_{ij}$ is not zero on the boundary. Hence, even though the Schwarzschild geometry satisfies the Einstein equation, the action does not have an extremum at boundary radius r (of the background geometry) in Allen's analysis. Hence the eigenvalue spectrum relevant to behavior of the second variation was not calculated at an extremum of the action. This does not matter if $rM^{-1}\to\infty$. However, this defect can of course be repaired by a more careful formulation of the problem.

Furthermore, an argument was made to the effect that from the absence of a negative mode when $rM^{-1}<$ (some number), one can conclude that hot flat space is stable against nucleation of a black hole when $AT^2=(4\pi r^2)T^2<$ (another number). This inference is clearly impossible on dimensional grounds. It is correct that a criterion involving AT^2 will decide the issue and one has been given in (8). But (8) follows from the nonexistence of *any* black hole in this case; when AT^2 satisfies (8), rM^{-1} is not defined. Below I shall give a result ruling out nucleation that is stronger than (8) and applies even when rM^{-1} is defined.

VI. FREE ENERGY FUNCTION FOR BLACK HOLES

The Helmholtz free energy function F_{BH} deduced from the action $I=\beta F_{\text{BH}}$ applies to the two equilibrium values of mass M_1 and M_2. It is desirable to obtain a "generalized" free energy $\bar{F}$ for any value of M. We regard r and T as fixed and employ the dimensionless parameter $x=Mr^{-1}$ in the standard definition $F=E-TS$ to obtain

$$\bar{F}(x,r,T)=r-r(1-2x)^{1/2}-4\pi r^2Tx^2 . \quad (35)$$

The extreme values of $\bar{F}$ are obtained from

$$\left[\frac{\partial\bar{F}}{\partial x}\right]_{r,T}=0 . \quad (36)$$

This yields an equation for x equivalent to the relation governing T, M, and r given by (1). Thus, extrema occur when and only when $rT\geq(27)^{1/2}(8\pi)^{-1}$. Then the values x_1 and x_2 of the extrema are given by (3) and (4). One finds $\bar{F}(x_1,r,T)=F_{\text{BH}}(M_1)$ and $\bar{F}(x_2,r,T)=F_{\text{BH}}(M_2)$.

The extremum at x_1 is a maximum ($C_A<0$) and that at x_2 a minimum ($C_A>0$) except when $x_1=x_2$

$[rT=(27)^{1/2}(8\pi)^{-1}]$, which is an inflection point. Actually, of course, the extremum at x_1 corresponds to a saddle point, not a maximum, of the *action*. Here we are able to use such a simple $\bar{F}$ and to consider only static spherical perturbations of finite mass because, as shown by Gross, Perry, and Yaffe, only such perturbations are relevant to the question of unstable equilibrium configurations. (However, a black hole in empty space—not in equilibrium—is quantum-mechanically unstable with respect to time-dependent nonspherical fluctuations of quadrupole or higher angular order. This provides an alternative view of the origin of the Hawking effect.[13])

The free energy $\bar{F}$ is also defined at values of x for which it does not take an extreme value. The corresponding Schwarzschild geometries, however, do not have the "correct" temperature. Geometrically, this means they are topologically defective, possessing conical-type singularities on the real Euclidean section at their "axes" at $2M$. Formation of a stable black hole from hot flat space, by way of the nucleation instability, will involve the passing of x from zero through x_1 to x_2. The Euler characteristics χ of the corresponding sequence of geometries take the values $\chi=0$ at $x=0$, $\chi=2$ at x_1 and x_2, and are not defined at other values of x. Thus one sees the familiar fact that in general there is no insuperable effective potential barrier to prevent a change of topology in quantum gravity at finite temperature. (The free energy plays the role of an effective potential in the canonical ensemble.) It should be noted, however, that the barrier is infinitely high (in effect) and cannot be surmounted beginning at $x=0$ if $rT<(27)^{1/2}(8\pi)^{-1}$. In this part of the T-r phase plane, black holes are excluded. One may imagine that a hole can begin to form as a small "bubble" in this region but that it must disappear quickly.

One recognizes that when the cavity contains a black hole, it must also contain some residual radiation, characterized by a nonzero renormalized stress-energy tensor, in order actually to be in equilibrium. Thus, back-reaction effects should be included. I have obtained these from the results of Ref. 9. For the stable hole, the corrections are typically small. For the unstable hole, the corrections can be significant. However, if $r \gtrsim 1$ and $T \lesssim 1$, the essential features survive. The heat capacity still changes signs at the (corrected) circular photon orbit $rM^{-1}=3(1+\delta)$, $\delta>0$. The increase of the action of (light hole plus residual radiation) above the action of hot flat space is still rather well estimated by $I(M_1)=\beta F(M_1)$. From (9) and (20) we obtain therefore

$$B \approx I(M_1) \approx \frac{1}{16\pi T^2}\left[1+\frac{1}{8\pi r T}\right] . \quad (37)$$

A thermodynamically necessary condition for nucleation of the locally stable hole from hot flat space clearly is that $F(M_2)<0$. From (21) and (4) we see that this means $rT>27(32\pi)^{-1}$. Hence, if

$$rT \le 27(32\pi)^{-1} \Longleftrightarrow AT^2 \le \frac{1}{\pi}(\tfrac{27}{16})^2 \simeq 0.906\,437 , \quad (38)$$

then nucleation of a hole from hot flat space will *never* occur in a thermodynamic sense.

VII. GROUND STATE OF THE CANONICAL ENSEMBLE

The ground state will be the one of least free energy. When $F<0$ for the black hole, one has that $(4\pi rT)^2>(27)^2 8^{-2} \simeq 11.4$. Combining (20) and (10) we obtain an estimate for $I(M_2)$ valid through order $(4\pi rT)^{-2}$:

$$I(M_2) \simeq \beta r - \pi r^2 - \frac{\beta^2}{8\pi} \simeq \beta F(M_2) . \quad (39)$$

The action (39) is to be compared to that of hot flat space:

$$I_{\rm HFS} = -\tfrac{1}{3} N a_0 \beta^{-3} V , \quad (40)$$

where $a_0=\pi^2(30)^{-1}$ and N is the effective number of massless states of zero helicity. The actions (39) and (40) can be compared directly, for any fixed value of N, if we employ the scaling rules for hot gravity introduced in Sec. IV. Thus, take $\lambda=N^{1/2}$ and define scaled variables $\hat{r}=N^{-1/2}r$, $\hat{T}=N^{1/2}T$, $\hat{M}=N^{-1/2}M$, and so on. Note that $rT=\hat{r}\hat{T}$ and $rM^{-1}=\hat{r}\hat{M}^{-1}$. We find that $\hat{I}(\hat{M}_2)=N^{-1}I(M_2)$ and $\hat{I}_{\rm HFS}=N^{-1}I_{\rm HFS}$ can be compared directly, with

$$\hat{I}_{\rm HFS} = -\tfrac{1}{3} a_0 \hat{\beta}^{-3} \hat{V} . \quad (41)$$

Inspection of $\hat{I}(\hat{M}_2)$ and $\hat{I}_{\rm HFS}$ reveals that at sufficiently large $\hat{T}$ and *any* value of $\hat{r}$, hot flat space must always be the dominant phase. Ignoring possible "finite-size" corrections to the value of a_0, we find that HFS dominates thermodynamically when

$$\frac{2\pi^3}{135}\hat{T}^2 \gtrsim -(\hat{r}\hat{T})^{-2}+\pi(\hat{r}\hat{T})^{-1}+\frac{1}{8\pi}(\hat{r}\hat{T})^{-3} . \quad (42)$$

The right-hand side has a maximum with respect to $(\hat{r}\hat{T})$ at $(\hat{r}\hat{T})=\pi^{-1}[1+0.25(10)^{1/2}]$, which implies that when

$$\hat{T} \gtrsim \hat{T}_{\max} \cong 2.401 , \quad (43)$$

then HFS is the dominant phase, for all $\hat{r}$. Exact calculations yield $(\hat{r}\hat{T})=(343)(192\pi)^{-1}$ and $\hat{T}_{\max} \simeq 2.402$.

Similarly, HFS will be dominant for all $\hat{T}$ if $\hat{r}$ is sufficiently small. This occurs when

$$\frac{2\pi^3}{135}\hat{r}^{-2} \gtrsim (\hat{r}\hat{T})^{-4}+\pi(\hat{r}\hat{T})^{-3}+\frac{1}{8\pi}(\hat{r}\hat{T})^{-5} . \quad (44)$$

The right-hand side has a minimum with respect to $(\hat{r}\hat{T})$ when $(\hat{r}\hat{T})=(6\pi)^{-1}[4+(8.5)^{1/2}]$, which implies that when

$$\hat{r} \lesssim \hat{r}_{\min} \simeq 0.179 , \quad (45)$$

then HFS is dominant for all $\hat{T}$. Exact calculations yield $(\hat{r}\hat{T})=(2197)(1920\pi)^{-1}$ and $\hat{r}_{\min} \simeq 0.178$.

We see that in the semiclassical approximation, when $\hat{r}$ and $\hat{T}$ are near the Planck scale, hot flat space will be the dominant phase of spacetime structure. Because N could be very large, the physical values could have $r_{\min}$ significantly larger than the Planck length and $T_{\max}$ significantly smaller than the Planck temperature. Therefore, the dominance of hot flat space would seem to emerge as a nonperturbative phenomenon. Obviously, new physics could intervene to alter these conclusions because the critical values of $\hat{r}$ and $\hat{T}$ are close to where the semiclassical approximation including back reaction breaks down. However, an important point is that where new physics might enter should be determined by the scaling laws for hot gravity, that is, in terms of $\hat{r}$ and $\hat{T}$ rather than by the physical values r and T.

It will be interesting to apply these results in the early Universe. This will be more realistic if the implicit "immovable heat-conducting walls" are avoided. For this purpose one uses the appropriate Gibbs function $G = E - TS + \sigma A$ and the role of "boundary conditions" is replaced by the ambient values of T and σ. This will be treated elsewhere. It is worth mentioning, however, that the analogs of $\hat{T}_{\max}$ and $\hat{r}_{\min}$ still exist. One finds the same $\hat{T}_{\max}$ and a certain $\hat{\sigma}_{\max}$ above which hot flat space is again the dominant phase in the semiclassical approximation.

VIII. DENSITY OF STATES

Let $\nu(E)dE$ be the number of states of the gravitational field with energies between E and $E+dE$ in a spherical cavity of radius r. From a well-defined partition function $Z(\beta,r)$, such as the one corresponding to the locally stable black-hole geometry, one can obtain the density of states by a suitably defined "inverse Laplace transform:"

$$\nu(E) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} Z(\beta,r)\exp(\beta E)\,d\beta \ . \tag{46}$$

In this integration, E and r are held fixed. To simplify the calculation, and obtain an estimate, we can employ the approximation (39), yielding

$$Z(\beta,r) \simeq \exp\left[-\left[\beta r - \pi r^2 - \frac{\beta^2}{8\pi}\right]\right] . \tag{47}$$

The integration is straightforward and gives

$$\nu(E) \cong \sqrt{2}\exp[\pi r^2 - 2\pi(E-r)^2] \ . \tag{48}$$

However, in the same order of approximation that gave (39) and (47), one has, using (10),

$$\pi r^2 - 2\pi(E-r)^2 = \frac{1}{4}\left[4\pi r^2 - \frac{\beta^2}{2\pi}\right] = 4\pi M^2 \ , \tag{49}$$

that is, one-fourth of the area of the event horizon. Thus,

$$\nu(E) \cong \sqrt{2}\exp(4\pi M^2) \ . \tag{50}$$

A steepest-descent estimate of (46) gives a similar conclusion. These calculations can be contrasted with similar ones employing in Z the action of the unstable hole. There one obtains a divergent result for $\nu(E)$ if one uses (46).[3]

ACKNOWLEDGMENTS

I thank Paul Frampton, Gary Gibbons, Don Page, Bernard Whiting, and especially Bernard Kay for helpful remarks. Financial support was received from the National Science Foundation.

---

[1]S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).

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

[3]S. W. Hawking, in *General Relativity*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979).

[4]D. J. Gross, M. J. Perry, and L. G. Yaffe, Phys. Rev. D **25**, 330 (1982).

[5]J. B. Hartle and S. W. Hawking, Phys. Rev. D **13**, 2188 (1976); W. Israel, Phys. Lett. **57A**, 107 (1976); G. W. Gibbons and M. J. Perry, Phys. Rev. Lett. **36**, 985 (1976).

[6]J. D. Bekenstein, Phys. Rev. D **12**, 3077 (1975).

[7]An argument suggesting the negative mode was given by D. N. Page (unpublished).

[8]S. W. Hawking and D. N. Page, Commun. Math. Phys. **87**, 577 (1983).

[9]J. W. York, Jr., Phys. Rev. D **31**, 775 (1985). This work made use of results given by D. N. Page, *ibid.* **25**, 1499 (1982).

[10]The Lorentzian-spacetime version of I_1 was given by J. W. York, Jr., Phys. Rev. Lett. **28**, 1082 (1972).

[11]A result equivalent to (20) was obtained previously by B. Allen, Phys. Rev. D **30**, 1153 (1984).

[12]See, for example, L. E. Reichel, *A Modern Course in Statistical Mechanics* (University of Texas Press, Austin, 1980).

[13]J. W. York, Jr., Phys. Rev. D **28**, 2929 (1983).