

Gravitational properties of monopole spacetimes near the black hole threshold

Arthur Lue* and Erick J. Weinberg†

Department of Physics, Columbia University, New York, New York 10027

(Received 24 January 2000; published 3 May 2000)

Although nonsingular spacetimes and those containing black holes are qualitatively quite different, there are continuous families of configurations that connect the two. In this paper we use self-gravitating monopole solutions as tools for investigating the transition between these two types of spacetimes. We show how causally distinct regions emerge as the black hole limit is achieved, even though the measurements made by an external observer vary continuously. We find that near-critical solutions have a naturally defined entropy, despite the absence of a true horizon, and that this has a clear connection with the Hawking-Bekenstein entropy. We find that certain classes of near-critical solutions display naked black hole behavior, although they are not truly black holes at all. Finally, we present a numerical simulation illustrating how an incident pulse of matter can induce the dynamical collapse of a monopole into an extremal black hole. We discuss the implications of this process for the third law of black hole thermodynamics.

PACS number(s): 04.70.Dy, 04.70.Bw

I. INTRODUCTION

Nonsingular spacetimes and those containing black holes are qualitatively quite different. Nevertheless, it is possible to find sequences of spacetimes that, while remaining nonsingular, come arbitrarily close to having horizons [1–4]. In a previous paper [5] we studied a class of such solutions that are associated with self-gravitating monopoles in a spontaneously broken Yang-Mills theory. The emphasis there was on the detailed behavior of the fields as one approaches the critical solution in which a horizon first appears. In this paper, we concentrate instead on the geometrical aspects of the spacetimes associated with these objects near criticality and on using these to gain insights into the properties of true black holes.

As in Ref. [5], we restrict ourselves to spherically symmetric spacetimes and write the metric in the form

$$ds^2 = B dt^2 - A dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2). \quad (1.1)$$

A horizon corresponds to a zero of $1/A$; the horizon is extremal if $d(1/A)/dr$ also vanishes. We work in the context of an $SU(2)$ gauge theory with gauge coupling e and a triplet Higgs field ϕ whose vacuum expectation value v breaks the symmetry down to $U(1)$. In flat spacetime this theory possesses a finite energy monopole solution with magnetic charge $Q_M = 4\pi/e$ and mass $M \sim v/e$. It has a core region, of radius $\sim 1/ev$, with nontrivial Higgs and massive vector boson fields. Beyond this core is a Coulomb region in which all massive fields approach their vacuum values exponentially rapidly, leaving only the Coulomb magnetic field. The effects of adding gravitational interactions depend on the value of v . For v much less than the Planck mass M_{Pl} , one finds that $1/A$ is equal to unity at the origin, decreases to a minimum at a radius of order $1/ev$, and then increases again with $A(\infty) = 1$. As v is increased, this minimum becomes deeper, until an extremal horizon develops at a critical value v_{cr} of

the order of M_{Pl} . As we describe in more detail in Sec. II, two distinct types of critical behavior are possible, depending on the ratio of the Higgs and gauge boson masses. For lower values of this ratio, one finds ‘‘Coulomb-type’’ critical solutions, in which the horizon occurs in the Coulomb region of the monopole at $r_0 = \sqrt{4\pi G/e^2}$. Outside the horizon, the metric is that of an extremal Reissner-Nordström black hole, with

$$B(r) = \frac{1}{A} = 1 - \frac{2MG}{r} + \frac{Q^2 G}{4\pi r^2}, \quad (1.2)$$

while the massive fields take on their vacuum values. As the Higgs self-coupling increases, there is a transition to ‘‘core-type’’ critical solutions that have a horizon inside the monopole core and nontrivial matter fields (or hair) outside the horizon.

In both types of critical monopole solutions the fields remain nonsingular at $r=0$. However, it is also possible to have solutions with singularities at $r=0$ that can be viewed as self-gravitating monopoles with Schwarzschild black holes at their center. As long as the mass of the central black hole is not too great, the variation of these solutions with v is quite similar to that of the nonsingular monopoles, and one finds the same two types of critical behavior [6].

After reviewing these solutions, we discuss in Sec. III how near-critical monopoles might appear to an outside observer. One would expect the measurements made by such an observer to vary continuously with the parameters of the monopole and to show no discontinuity at the critical solution. An external observer could probe the monopole with either particles or waves. In the case of the particle, we find that the time needed for the particle to emerge from the interior (as measured by a static external observer using Schwarzschild time) diverges as the critical solution is approached. When a wave is sent in, there is a reflected wave due to the gravitational field just outside the horizon and a transmitted wave that passes through the interior and then emerges with a time delay. As before, the time delay diverges as $v \rightarrow v_{\text{cr}}$, while the reflected wave becomes indis-

*Email address: lue@phys.columbia.edu

†Email address: ejw@phys.columbia.edu

tinguishable from that due to a black hole. Using either type of probe, an observer whose lifetime is finite cannot distinguish between a true black hole and a nonsingular, subcritical solution that is sufficiently close to being critical. We discuss the implications for our understanding of black hole entropy.

We also find that the near-critical Coulomb-type solutions display what Horowitz and Ross [7,8] have termed naked-black-hole behavior, even though there is no black hole at all. This is characterized by the fact that a freely falling observer passing through the minimum of $1/A$ (we shall refer to the location of this minimum as the quasi-horizon) feels a tidal force that diverges as the critical solution is approached. For core-type solutions, on the other hand, no such behavior is observed.

In Sec. IV, we consider the effect of having additional matter fall into a near-critical solution, addressing in particular the question of whether this process could produce an extremal black hole. Extremal black holes are especially interesting from the standpoint of black hole thermodynamics because they have vanishing Hawking temperature. The analogies between black hole dynamics and thermodynamics thus suggest that they should be rather difficult, if not impossible, to create. Indeed, one of the formulations [9] of the third law of black hole dynamics asserts the impossibility (under certain technical assumptions) of making a nonextremal black hole extremal. One could also envision producing an extremal black hole starting from a nonsingular spacetime. Boulware [10] showed that this can be done by the collapse of a charged shell of matter. However, this mechanism relies critically on the shell being infinitely thin; shells of finite thickness and density do not collapse to an extremal configuration.

It is easy to understand the difficulty of making an extremal black hole if one recalls that the extremal Reissner-Nordström black hole is characterized by having a mass and a charge that (in appropriately rescaled Planck units) are equal. Forming such an object by the collapse of a shell with equal charge and mass densities involves a delicate balance between electromagnetic and gravitational forces. One could instead try to achieve extremality by adding matter to a pre-existing nonextremal Reissner-Nordström black hole (i.e. one with greater mass than charge). However, because the added matter would have to have more charge than mass, the Coulomb repulsion between the black hole and the infalling matter would tend to overcome their gravitational attraction.

The situation is rather different in our case, because the nonsingular monopole solutions are *overcharged*; i.e., their long range fields are those of a Reissner-Nordström solution with greater charge than mass.¹ Allowing uncharged matter to fall into these objects increases their mass and should

bring them closer to criticality. If the amount of matter entering is just sufficient to create a zero of $1/A$, one would expect an extremal solution to result. We will present numerical arguments that support this expectation. Finally, we make some concluding remarks in Sec. V.

II. REVIEW OF PREVIOUS RESULTS

As in Ref. [5], we consider an $SU(2)$ gauge theory that is spontaneously broken to $U(1)$ by the vacuum expectation value v of a triplet Higgs field ϕ . This theory has magnetic monopole solutions that can be described by the spherically symmetric ansatz

$$\phi^a = v \hat{r}^a h(r) \quad (2.1)$$

$$A_{ia} = \epsilon_{iak} \hat{r}^k \frac{1-u(r)}{er}. \quad (2.2)$$

Finiteness of the energy requires that $u(\infty)=0$ and $h(\infty)=1$. If the solutions are also required to be nonsingular at $r=0$, then $u(0)=1$ and $h(0)=0$.

In a spacetime with a metric of the form of Eq. (1.1), the static field equations for these matter fields can be derived from a $(1+1)$ -dimensional action of the following form [11]:

$$S_{\text{matter}} = -4\pi \int dt dr r^2 \sqrt{AB} \left[\frac{K(u,h)}{A} + U(u,h) \right] \quad (2.3)$$

where $U(u,h)$ involves the fields but not their derivatives, while

$$K = \frac{1}{e^2 r^2} \left(\frac{du}{dr} \right)^2 + \frac{v^2}{2} \left(\frac{dh}{dr} \right)^2. \quad (2.4)$$

The Euler-Lagrange equations for the matter fields that follow from Eq. (2.3) must be supplemented by the gravitational field equations. For static, spherically symmetric field configurations these reduce to

$$G_{\hat{t}\hat{t}} = \frac{1}{2r^2} \frac{d}{dr} \left[r \left(\frac{1}{A} - 1 \right) \right] = -4\pi G \left(\frac{K}{A} + U \right) \quad (2.5)$$

and

$$G_{\hat{r}\hat{r}} + G_{\hat{\theta}\hat{\theta}} = - \left(\frac{2}{rA} \right) \frac{1}{\sqrt{AB}} \frac{d\sqrt{AB}}{dr} = - \frac{16\pi GK}{A}. \quad (2.6)$$

Here carets indicate orthonormal components.

Equation (2.6) can be immediately integrated to obtain

$$B(r) = \frac{1}{A(r)} \exp \left[-16\pi G \int_r^\infty dr' r' K \right]. \quad (2.7)$$

¹In the pure Reissner-Nordström case, this leads to a naked singularity. The singularity is avoided here by the same mechanism that makes the mass of the flat-space monopole finite: the orientation of the massive gauge fields in the monopole core is such that their magnetic dipole energy just cancels the singular Coulomb energy at the origin.

Using this to eliminate B , one is left with two second order and one first order equation for the functions u , h , and A . These equations must be solved numerically. Up to a rescaling of distances, the solutions of these equations depend only on the two dimensionless parameters $a = 8\pi G v^2$ and $b = (m_H/2m_W)^2$.

For small values of b (roughly $b \lesssim 25$ for a quartic Higgs field) [5,6] one finds Coulomb-type solutions in which the minimum of $1/A$ is located outside the monopole core. This minimum decreases² with increasing a , until the critical solution is reached at $a_{\text{cr}} = 8\pi G v_{\text{cr}}^2$. In the critical solution, the matter fields u and h reach their asymptotic values $u=0$ and $h=1$ at the horizon and are then constant for all $r > r_0$; because both fields fall as fractional powers of $r_0 - r$, the derivatives du/dr and dh/dr both diverge as r approaches r_0 from below. (This nonanalytic behavior is possible because an extremal horizon is a singular point of the matter field equations.)

The metric of the critical solution is identical to the extremal Reissner-Nordström metric outside the horizon, but differs from it for $r < r_0$. The metric function $1/A$ varies relatively smoothly, falling monotonically from unity at the origin to a zero at the horizon. Just inside the horizon $1/A \sim k(r_0 - r)^2$, with k being larger than for the corresponding Reissner-Nordström solution. The behavior of B contrasts sharply with this. Equation (2.7) shows that the product AB (which is identically equal to unity in both the Schwarzschild and Reissner-Nordström solutions) is given by an integral of the functional $K(u, h)$. The singularities in the derivatives of u and h at the horizon are strong enough to cause this integral to diverge, so that the ratio

$$c \equiv \frac{\sqrt{AB}|_{\text{outside } r_0}}{\sqrt{AB}|_{\text{inside } r_0}} \quad (2.8)$$

is infinite in the critical limit. If we adopt the conventional normalization $B(\infty) = 1$, then B vanishes identically inside the horizon. If we instead set $B(0) = 1$, then B is finite and varying inside the horizon and infinite for $r > r_0$; depending on the value of b , the minimum of B may be at $r=0$ or at some finite radius, but in neither case does B have a zero. For near-critical solutions where the minimum value $(1/A)_{\text{min}} \equiv \epsilon$ is small but nonzero, we find that the ratio c varies as ϵ^{-q} , where q ranges from about 0.7 to unity.

A rather different type of critical solution is found for larger b . For these core-type solutions the horizon occurs at a radius $r_* < r_0$, with the values u_* and h_* of the matter fields at this point being different than their asymptotic values. Although the solutions are still nonanalytic at the horizon, this nonanalyticity occurs only in subdominant terms. Thus, $1/A$ again vanishes as $(r - r_0)^2$ as one approaches the horizon, but the coefficient is now the same inside and outside

the horizon.³ The radial derivatives of both matter fields are finite at the horizon, so K remains finite and there is no sharp change in AB at the horizon. Because AB remains finite and nonzero, B has a zero at the horizon that coincides with the zero of $1/A$.

These solutions can be generalized to include a black hole in the center of the monopole. Instead of requiring that the fields be nonsingular at $r=0$, one instead requires that there be a zero of $1/A$ at a nonzero radius r_H . At this zero, the equations for the matter fields become constraint equations relating the fields and their first derivative; solving these constraints yields enough boundary conditions to determine a solution.

If r_H is not too large, the effect of increasing a is similar to what it is in the absence of a central black hole.⁴ There is an outer minimum of $1/A$ that moves downward, finally reaching zero and becoming an extremal horizon at some critical value $a_{\text{cr}}(r_H)$. For small values of b the solutions are Coulomb type, while for large b one finds core-type critical solutions.

Rather than increasing a with r_H fixed, one can instead increase r_H with a held fixed; this is much more analogous to the process of actually dropping matter into a near-critical solution that we will consider in Sec. IV. For initial values of a that are sufficiently close to $a_{\text{cr}}(r_H=0)$, this gives a family of solutions with a critical limit. In Figs. 1 and 2 we illustrate this with a Coulomb-type solution with $b=1.0$ and a core-type solution with $b=10^6$.

III. PROBING THE QUASI-BLACK HOLE

For any $a < a_{\text{cr}}$, the self-gravitating monopole solution is a nonsingular spacetime with a Penrose diagram of the same form as that of Minkowski spacetime (Fig. 3a). The critical solution, on the other hand, can be extended beyond the original coordinate patch to yield a spacetime with the Penrose diagram shown in Fig. 3b. This diagram is quite similar to that of an extremal Reissner-Nordström black hole, but differs from it by not having a singularity⁵ at $r=0$. The difference between the two diagrams is striking and seems to indicate a discontinuity at $a = a_{\text{cr}}$, in contradiction with the usual expectation that physical quantities should vary continuously with the parameters of a theory.

However, this discontinuity can be seen as an artifact of the conformal transformation that produces the Penrose diagram from an infinite spacetime. This can be illustrated by

³For intermediate values of b , subcritical monopoles exhibit both core-type and Coulomb-type quasi-horizons. However, as one approaches criticality for a given value of b , only one quasi-horizon actually becomes a horizon. The other quasi-horizon, though interesting, is essentially irrelevant for our purposes.

⁴For larger values of r_H , see the discussion in [3].

⁵Although there is no singularity at the origin for a critical monopole black hole, there are singularities at the extremal horizon resulting from nonanalytic behavior of the monopole fields. These singularities are relatively mild in the core-type case, but are more dramatic in the Coulomb-type case.

²This behavior is modified slightly for very small b . For a detailed description, see [2,3,12,13].

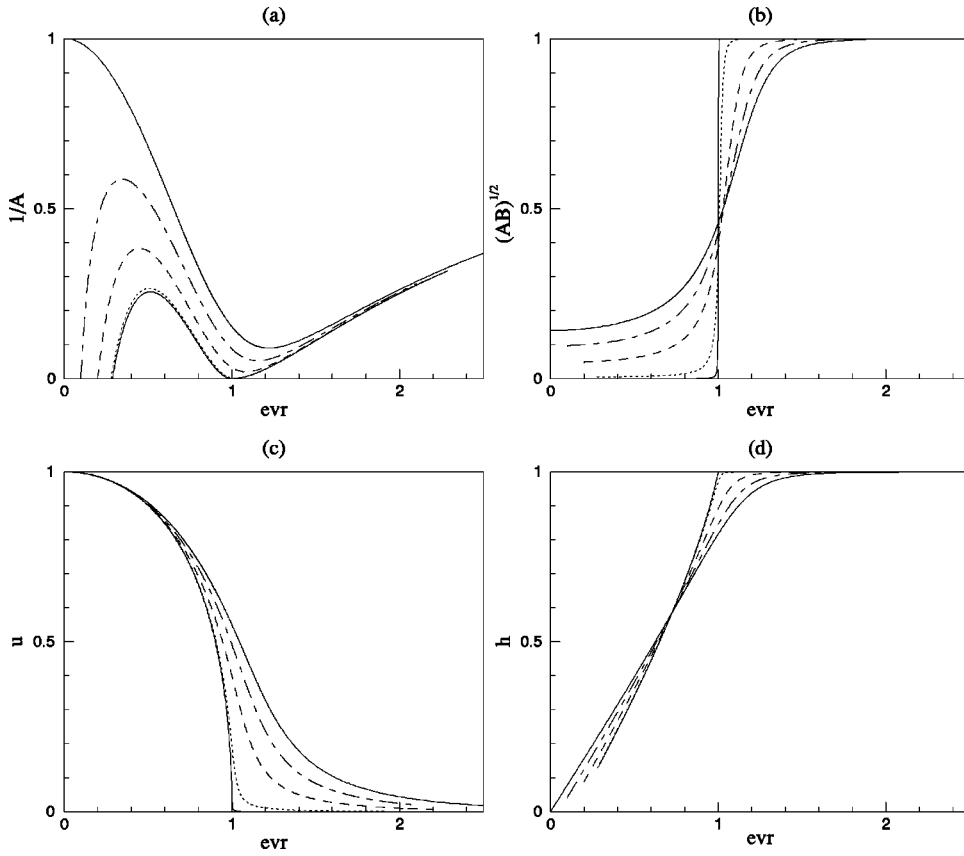


FIG. 1. Monopole solutions for $a=2.0$, $b=1.0$ and various values of central black hole radius, r_H . The progression from solid line, dot-dashed line, to dashed line, to dotted line and to solid line corresponds to $evr_H = 0.0, 0.1, 0.2, 0.28$ and 0.288 . The panels depict the functions (a) $1/A(r)$, (b) $(AB)^{1/2}(r)$, (c) $u(r)$ and (d) $h(r)$.

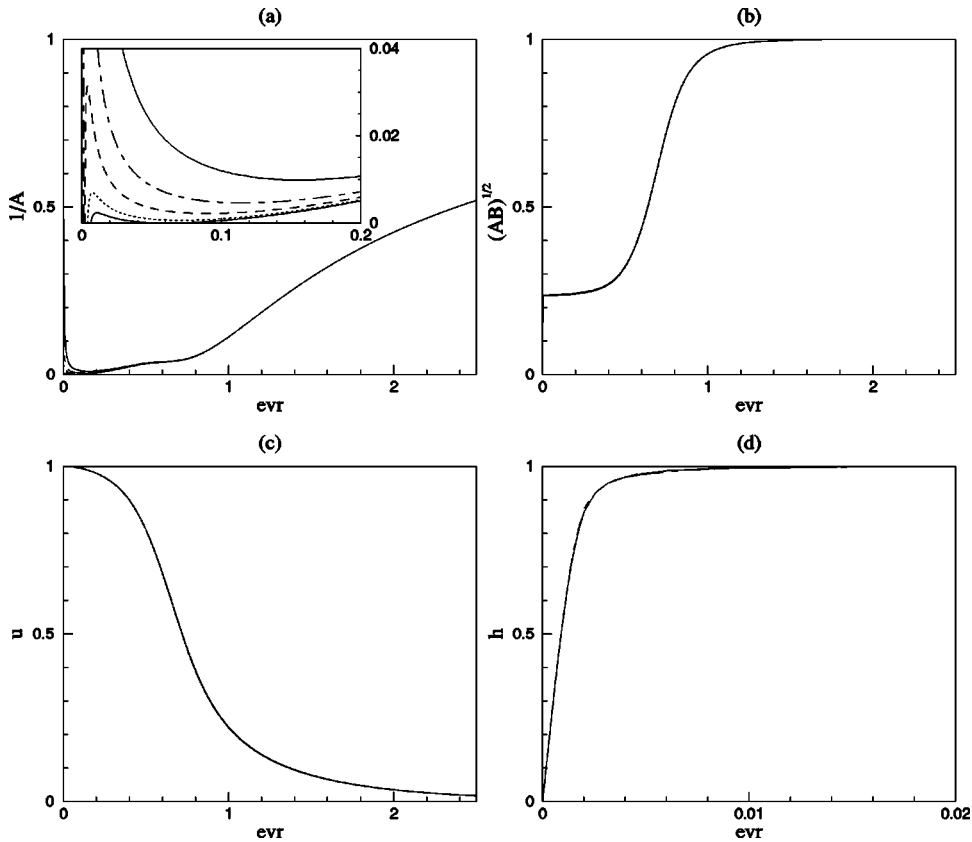


FIG. 2. Monopole solutions for $a=1.002$ and $b=10^6$ and various values of central black hole radius, r_H . Here $a_{cr}(r_H=0) = 1.011654$ and the minimum a using this scenario is 1.001. The progression from solid line, dot-dashed line, to dashed line, to dotted line and to solid line corresponds to $evr_H = 0, 0.001, 0.002, 0.004$ and 0.00628 . The panels depict the functions (a) $1/A(r)$, (b) $(AB)^{1/2}(r)$, (c) $u(r)$ and (d) $h(r)$. Note that the monopole fields are virtually unchanged as the internal black hole size is varied. Note that the radial scale for $h(r)$ is exaggerated to show detail.

considering the points A and B that lie on the curves of constant r that are shown in Figs. 3a and 3b. These have been chosen so that it is possible for an object to start at A, move in to $r=0$ at C, and then travel out again to B. The total proper time along this world line (or the total affine parameter, if the world line is lightlike) is finite. This should be compared with the proper time along the world line of constant r . This is finite for the subcritical case, whereas in the critical case the proper times along the segments AD and DB are both infinite, corresponding to the fact that an observer in the exterior region containing A cannot detect objects behind the horizon PD. In order to obtain consistency with our physical expectations of continuity, we should require that the proper time along the worldline of constant r should diverge as $a \rightarrow a_{cr}$. More generally, the time required for an external observer to receive information from a probe of the interior region should diverge in the critical limit.

A. Particle and wave probes

To see how this works out, we consider probing the interior region (i.e., the region $r < r_*$, where r_* is the quasi-horizon) of a near-critical solution by sending in either a particle or a wave. In both cases, we assume that the probe interacts only gravitationally, and has no direct interaction with the monopole fields.

To begin, imagine releasing a massive particle from an initial radius $r_1 \gg r_*$ that is large enough that we may approximate $B(r_1) \approx 1$. The rotational and time-translation symmetries of the metric allow us to take the motion to lie in the $\theta = \pi/2$ plane and guarantee the conservation of the angular momentum per unit mass

$$J = r^2 \frac{d\phi}{d\tau} \quad (3.1)$$

and the energy per unit mass

$$E = B(r) \frac{dt}{d\tau}. \quad (3.2)$$

These, together with Eq. (1.1), imply that

$$\frac{dr}{d\tau} = \frac{1}{\sqrt{AB}} \left[E^2 - B \left(\frac{J^2}{r^2} + 1 \right) \right]^{1/2}. \quad (3.3)$$

If $J=0$, the particle falls radially in, passes through the origin, and emerges on the other side of the monopole. If instead $J \neq 0$, the particle turns around after reaching a minimum radius $r_{\min}(J)$ and returns to r_1 with its trajectory advanced by an angle

$$\begin{aligned} \Delta\phi &= 2 \int_{r_{\min}}^{r_1} dr \frac{d\phi/d\tau}{dr/d\tau} \\ &= 2 \int_{r_{\min}}^{r_1} dr \frac{J}{r^2 \sqrt{AB}} \left[1 - \frac{B}{E^2} \left(\frac{J^2}{r^2} + 1 \right) \right]^{-1/2}. \end{aligned} \quad (3.4)$$

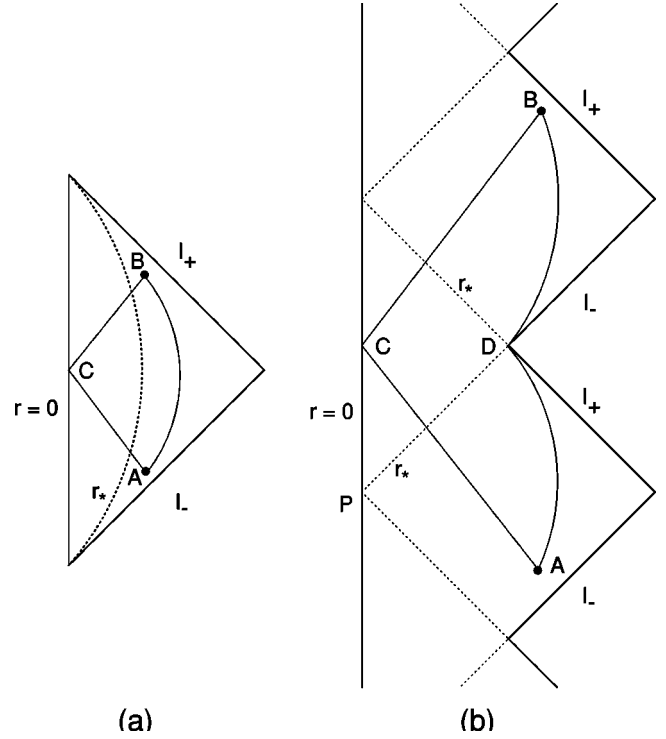


FIG. 3. Penrose diagrams for (a) subcritical monopole and (b) critical monopole black hole. In the former case r_* represents the quasi-horizon, whereas in the latter case that radius represents a true horizon.

We are interested in the proper time measured by an observer who remains at $r=r_1$. Assuming that the observer is not moving at relativistic speeds, this is approximately the same as the Schwarzschild coordinate time t , which over the course of the entire trajectory increases by an amount

$$\Delta t = 2 \int_{r_{\min}}^{r_1} dr \frac{dt/d\tau}{dr/d\tau} = 2 \int_{r_{\min}}^{r_1} dr \frac{A}{\sqrt{AB}} \left[1 - \frac{B}{E^2} \left(\frac{J^2}{r^2} + 1 \right) \right]^{-1/2}. \quad (3.5)$$

There are two potential sources of divergences in this integral as $\epsilon = (1/A)_{\min} \rightarrow 0$ and the critical solution is approached. In both types of critical solutions there is a contribution from $r \approx r_*$ associated with the growth of $A(r_*)$. In the Coulomb-type solutions the near-vanishing of \sqrt{AB} [see Eq. (2.7)] gives a second contribution from the entire region $r < r_*$. Let us examine these in more detail.

For core-type solutions, in the region $r \approx r_*$ we can write

$$A \approx k_1 \left[\left(\frac{r-r_*}{r_*} \right)^2 + \epsilon \right]^{-1} \quad (3.6)$$

with k_1 of order unity, while \sqrt{AB} is roughly constant and independent of ϵ . Because $B(r_*)$ is small, the J -dependent term in Δt can be neglected for any J such that the particle could have reached r_* . By a similar argument, we see that any particle that reaches r_* goes through the peak of A before turning around. Hence,

$$\Delta t \approx \frac{2k_1 \pi r_*}{\sqrt{AB}|_{r=r_*}} \epsilon^{-1/2} + \dots \quad (3.7)$$

where the ellipsis represents subdominant terms.

For Coulomb solutions the dominant effect is due to the fact that $\sqrt{AB} \sim \epsilon^q$ is almost vanishing throughout the interior region. Because our numerical solutions show that q ranges between 0.7 and unity, the divergence due to this effect is greater than that from the region near r_* . Furthermore, the near-vanishing of B in the interior implies that any particle that enters the interior almost reaches the origin, so that $r_{\min} \approx 0$. Thus,

$$\Delta t \approx k_2 r_* \epsilon^{-q} + \dots \quad (3.8)$$

where k_2 is of order unity.

Rather than sending in a particle, one can also probe the quasi black hole by sending in a wave packet. As an example, let us consider a free massive scalar field ϕ , whose field equation in curved spacetime takes the form

$$0 = \frac{1}{\sqrt{g}} \partial_\mu [\sqrt{g} g^{\mu\nu} \partial_\nu \phi] + m^2 \phi. \quad (3.9)$$

This can be put into a more tractable form by writing

$$\psi = r \phi \quad (3.10)$$

and defining a new coordinate $y(r)$ satisfying

$$\frac{dr}{dy} = \frac{\sqrt{AB}}{A}. \quad (3.11)$$

Equation (3.9) then takes the form of a one-dimensional wave equation

$$0 = \frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial y^2} + [U(r) + m^2 B] \psi \quad (3.12)$$

with a scattering potential

$$U(r) = \frac{1}{2r} \frac{d}{dr} \left[\frac{AB}{A^2} \right] + \frac{J(J+1)B}{r^2}. \quad (3.13)$$

When a wave packet incident from large r reaches the region near the quasi-horizon, a portion is reflected by the scattering potential, while a portion is transmitted and emerge with some time delay. If a near-critical solution is to appear essentially indistinguishable from a black hole to an outside observer, two conditions must hold. First, the reflection coefficient as a function of wave number must approach that of the black hole as $\epsilon \rightarrow 0$. Second, the time delay in the emergence of the transmitted wave should diverge in the critical limit.

To see how the first of these conditions comes about, let us use Eq. (2.6) to rewrite the scattering potential as

$$U(r) = \frac{AB}{rA} \left[\frac{8\pi GK}{A} - \frac{d}{dr} \left(\frac{1}{A} \right) + \frac{J(J+1)}{r} \right]. \quad (3.14)$$

For both core- and Coulomb-type solutions the quantity K/A remains finite in the critical limit, while the second term in the brackets is zero at the quasi-horizon. Since $AB \leq 1$, it is clear that $U(r_*)$ vanishes at least as fast as ϵ as the critical limit is approached. Hence, in the limit the scattering potential splits into two parts, one inside and one outside the horizon. The outer potential is equal either to that of an extremal Reissner-Nordström black hole (in the Coulomb case) or that of a black hole with hair (in the core case). Because the variation of the outer potential with ϵ is smooth in both cases, our conditions on the reflection coefficients are satisfied if we can ignore reflection from the inner part of the potential.

This can be understood by noting that the natural distance variable in which to discuss the motion of the waves is y . By integrating Eq. (3.11) inward from some reference point $r_1 \gg r_*$, we obtain

$$y(r) = y(r_1) - \int_r^{r_1} dr \frac{A}{\sqrt{AB}}. \quad (3.15)$$

The behavior of this integral as the critical limit is approached is very similar to that of the integral in the expression for Δt , Eq. (3.5). For either type of solution, the region near the quasi-horizon gives a contribution that diverges at least as fast as $\epsilon^{-1/2}$. There is a corresponding growth in both the effective distance from the inner portion of the potential to any external point and in the time delay of the corresponding reflected wave. As ϵ is increased, an external observer at fixed r first sees the reflections from the inner and outer parts of the potential split into two distinct reflected waves, and then finds that the time delay of the second reflected wave (from the inner potential) grows without bound.

The portion of the wave that is transmitted through the region near the quasi-horizon either continues through the origin and then outward or reflects off a central centripetal barrier, according to whether or not J vanishes. In either case, the time delay accumulated by this wave before it returns to the quasi-horizon grows in essentially the same manner as the travel time for a massive particle traversing the same path: as $\epsilon^{-1/2}$ for a core-type solution and as ϵ^{-q} for a Coulomb-type solution.

B. Information and entropy

Thus, regardless of the type of probe used, an external observer at fixed r_{obs} must wait for at least a time $\Delta t \geq O(\epsilon^{-1/2})$ before the probe emerges from the region inside the quasi-horizon. To leading order, this time delay is independent of the energy or angular momentum of the probe, and is instead determined solely by the spacetime geometry. Hence, to an observer with a finite lifetime T , the interior region of any near-critical configuration with $\epsilon \lesssim T^{-2}$ is inaccessible.⁶ He would most naturally describe any larger

⁶Note that once ϵ is less than T^{-2} , the boundary of the inaccessible region depends only very weakly on T and r_{obs} , and is essentially indistinguishable from the quasi-horizon.

system containing this configuration by a density matrix ρ obtained by tracing over the degrees of freedom inside the quasi-horizon. From this density matrix one can derive an entropy $S_{\text{interior}} = -\text{Tr} \rho \ln \rho$ that can be associated with the interior of the quasi black hole.

One could, of course, proceed in this manner to define an entropy for any arbitrary region in space, just as one can choose to make the information in any subsystem inaccessible by putting the subsystem behind a locked door. The crucial difference here is that the inaccessibility is due to the intrinsic properties of the spacetime, and that the boundary of the inaccessible region is defined by the system itself rather than by some arbitrary external choice. It is thus reasonable to define S_{interior} as *the* entropy of the quasi black hole.

A precise calculation of this entropy is clearly infeasible. Among other problems, such a calculation would require a correct implementation of an ultraviolet cutoff, which presumably would require a detailed understanding of how to perform the calculation in the context of a consistent theory of quantum gravity. As an initial effort, one can take the ultraviolet cutoff as the Planck mass M_{Pl} and ask for an order of magnitude calculation. Such a calculation was done by Srednicki [14], who showed that tracing over the degrees of freedom of a scalar field inside a region of flat spacetime with surface area A led to an entropy $S = \kappa M^2 A$ where M is the ultraviolet cutoff and κ is a numerical constant. Furthermore, although the precise calculations depend on the details of the theory, Srednicki gave general arguments suggesting that such an entropy should always be proportional to the surface area. This leads us to expect that $S_{\text{interior}} \sim M_{\text{Pl}}^2 A$.

This result is, of course, consistent with the possibility that in the critical limit S_{interior} goes precisely to the standard black hole result $S_{\text{BH}} = M_{\text{Pl}}^2 A/4$. However, in contrast with the usual black hole case, our spacetime configurations are topologically trivial. The ‘‘interior’’ region enclosed by the quasi-horizon is nonsingular and static. Furthermore, this region can be unambiguously defined, so that it is conceptually clear what is meant by tracing over the interior degrees of freedom, even though it may not yet be possible to implement this calculation in complete detail. We find it quite striking that by this approach one can arrive so nearly at the standard entropy result.

C. Curvature and naked-black-hole behavior

In our discussion above of the trajectory of a particle probe, we focused on the coordinate time that elapses over the course of the particle’s passage through the monopole. However, it also of interest to consider the elapsed proper time, which can be found by integrating $d\tau/dr$ [see Eq. (3.3)]. For core-type solutions this gives a finite nonzero result with no unusual behavior as the critical limit is approached. The situation with Coulomb-type solutions is, on the other hand, quite striking. The sharp decrease in AB at

the quasi-horizon leads to a corresponding decrease in $d\tau/dr$, so that the proper time elapsed while the probe is within the quasi-horizon is⁷

$$\Delta\tau \approx \frac{2r_*}{E} \sqrt{AB} \Big|_{r=0} \sim \epsilon^q. \quad (3.16)$$

In the critical limit AB vanishes identically for $r < r_*$, and $\Delta\tau = 0$.

This vanishing of $\Delta\tau$ is related to another interesting property of these solutions. It is well known that the Riemann tensor is nonsingular at a black hole horizon. It therefore does not seem surprising that in the most familiar black hole solutions, the Schwarzschild and Reissner-Nordström, a particle suffers no unusual effects as it crosses the horizon. However, Horowitz and Ross [7] showed that this is not always the case. Because of the acceleration of a particle as it approaches the horizon, the components of the Riemann tensor in a coordinate frame that is freely falling with the particle can be quite different from the components measured in a static frame. With a metric of the form of Eq. (1.1), the components $R_{t'kt'k}$ (where k denotes a transverse spatial direction and t' the time in the boosted frame) are given by

$$R_{t'kt'k} = -\frac{1}{2r} \frac{d}{dr} \left[\frac{E^2}{AB} - \frac{1}{A} \right], \quad (3.17)$$

where E is the energy per unit mass of the infalling particle.

The fact that this curvature component is never large (with E of order unity) for the Schwarzschild and Reissner-Nordström black holes is a consequence of the fact that AB is constant in both cases. This is not true in general. Horowitz and Ross found several examples of dilaton black holes for which $R_{t'kt'k}$, and thus the tidal forces felt by an infalling particle, could be made arbitrarily large near the horizon by taking the solution to be sufficiently close to extremality. They introduced the term ‘‘naked black hole’’ to indicate the fact that this (almost) singular behavior occurs outside the horizon. Subsequently [8], they showed that in these examples $R_{t'kt'k}$ was inversely proportional to the square of the proper time remaining before the particle reached the singularity at $r=0$.

Applying their results to our solutions, we find that near-critical Coulomb-type solutions display naked black hole behavior, even though they are not black holes at all. This can be seen by noting that Eq. (2.6) implies that

⁷The drop in AB also has consequences for the shape of the trajectory through the interior. By combining Eqs. (3.1) and (3.3), one finds that in the critical limit all probes follow a straight line passing through the origin, regardless of the incident angle with which they hit the horizon.

$$\frac{d}{dr} \left(\frac{1}{AB} \right) = \frac{16\pi G r K}{AB}. \quad (3.18)$$

Inside the quasi-horizon $AB \sim \epsilon^{2q}$, while the radial derivatives of u and h , and hence K , are of order unity. Inserting this result into Eq. (3.17), gives

$$R_{t'kt'} \sim \epsilon^{-2q}. \quad (3.19)$$

Note that this is proportional to $(\Delta\tau)^{-2}$, giving a relationship between tidal forces and proper time reminiscent of the examples described in Ref. [8].

IV. COLLAPSE TO AN EXTREMAL BLACK HOLE

To gain an insight into the third law of thermodynamics, as noted in Sec. I, it is of interest to determine whether systems with either initially nonsingular spacetimes or initially non-extremal black holes can evolve into systems with an extremal horizon. It appears that under reasonable conditions of finiteness and causality this cannot be done by adding charge to an undercharged object [9,10]. The discussion in the previous two sections, however, suggests an alternative. Recall that our subcritical monopoles are overcharged; i.e., they have a charge larger than their mass. In the normal Reissner-Nordström case, such a system would exhibit a naked singularity. However, in the monopole the Coulomb core is screened by the massive particles in such a way that no gravitational singularity exists. This suggests a scenario by which an extremal black hole is dynamically formed from a monopole by dropping in uncharged matter.

Let us add to our theory a chargeless matter field that is coupled to the monopole fields only through gravity. We then allow a spherical shell (of small, but finite thickness) of this matter fall into the monopole. If the mass of the shell is sufficiently small, we do not expect a horizon to be formed. On the other hand, if the shell contains enough matter, the system should collapse to form a black hole. It seems plausible that threshold case between these two regimes should produce an extremal horizon.

In the case of a Coulomb-type solution, one might run into difficulties because of the naked-black-hole behavior it exhibits. This concern, however, should not be an issue for the case of a core-type solution. Here the fields of the critical solution are much better behaved; what nonanalyticities exist at the horizon are mild, and become increasingly so as one increases b . Nothing unusual happens to the near-critical core solutions as one parametrically approaches criticality. Moreover, adding a small Schwarzschild black hole at the center of the self-gravitating monopole should not significantly change the scenario. The infall of a spherical shell of appropriate mass should still turn the quasi-horizon of a near-critical solution into an extremal horizon. The black hole would essentially play a spectator role, as its presence is largely irrelevant to the dynamics of the system.

We have carried out numerical simulations to test these ideas. We begin with a core-type monopole solution that, for numerical convenience, has a small Schwarzschild black hole with horizon radius r_H at its center. The parameters are chosen to be such that the solution is near criticality, so that only a small amount of additional matter is needed for the quasi-horizon at $r=r_*$ to collapse to a true horizon. We add to the theory a massive scalar field $\chi(r,t)$ that is coupled only to gravity. We then send a spherically symmetric Gaussian pulse of χ field in toward the monopole and watch the system evolve. To simplify the computation, we freeze the matter field variables u and h at their initial values; because the fields for near-critical core-type monopoles are not very sensitive to the metric (cf. Fig. 2), this approximation should cause little error.

When the pulse amplitude is larger than some threshold value, the pulse falls into the monopole until the metric function $1/A$ develops a simple zero near the location of the quasi-horizon of the initial monopole configuration. This newly formed horizon is non-extremal. More interesting is the situation where the pulse amplitude is at, or just below, this threshold value. Figure 4 shows a sequence of snapshots illustrating this scenario at four different points during the pulse's motion inward.⁸ (The distortion of the pulse is due to the backreaction of the monopole metric.)

As required by Birkhoff's theorem, $1/A$ is undisturbed ahead of the pulse, but undergoes a shift as the pulse passes. Thus, in the region ahead of the pulse $1/A$ is the same as it was in the original configuration, while behind the pulse it has the form corresponding to a static monopole with a central black hole whose horizon radius exceeds r_H by an amount determined by the energy-momentum of the infalling pulse. These two are joined by a kink at the pulse position. As the pulse passes through the quasi-horizon, $1/A$ reaches its minimum.

Similarly, the plot of \sqrt{AB} appears as nearly a step function centered at the pulse position, with the inner and outer regions corresponding to the initial and final configurations. The jump in \sqrt{AB} at the pulse position varies with time. It reaches its maximum when the pulse is passing through the quasi-horizon, and then decreases. (The resulting variation in the value of \sqrt{AB} at large r may appear to violate causality, but is actually a just a consequence of the gauge choice implicit in our choice of coordinates.)

As the pulse continues past the quasi-horizon into the monopole core, the metric function $1/A$ remains static outside the pulse. However, $\sqrt{AB}_{\text{outside}}$ decreases from its maximum value as the pulse continues inward. Eventually, the

⁸Note that we have chosen a normalization of time such that $AB \rightarrow 1$ as $r \rightarrow r_H$, the horizon of the internal black hole; this corresponds to using a time variable appropriate to an observer in the interior of the monopole. If we had used the more conventional normalization with AB equal to unity at spatial infinity, the time coordinate would be that appropriate to an external observer and would grow rapidly as the pulse approached the quasi-horizon.

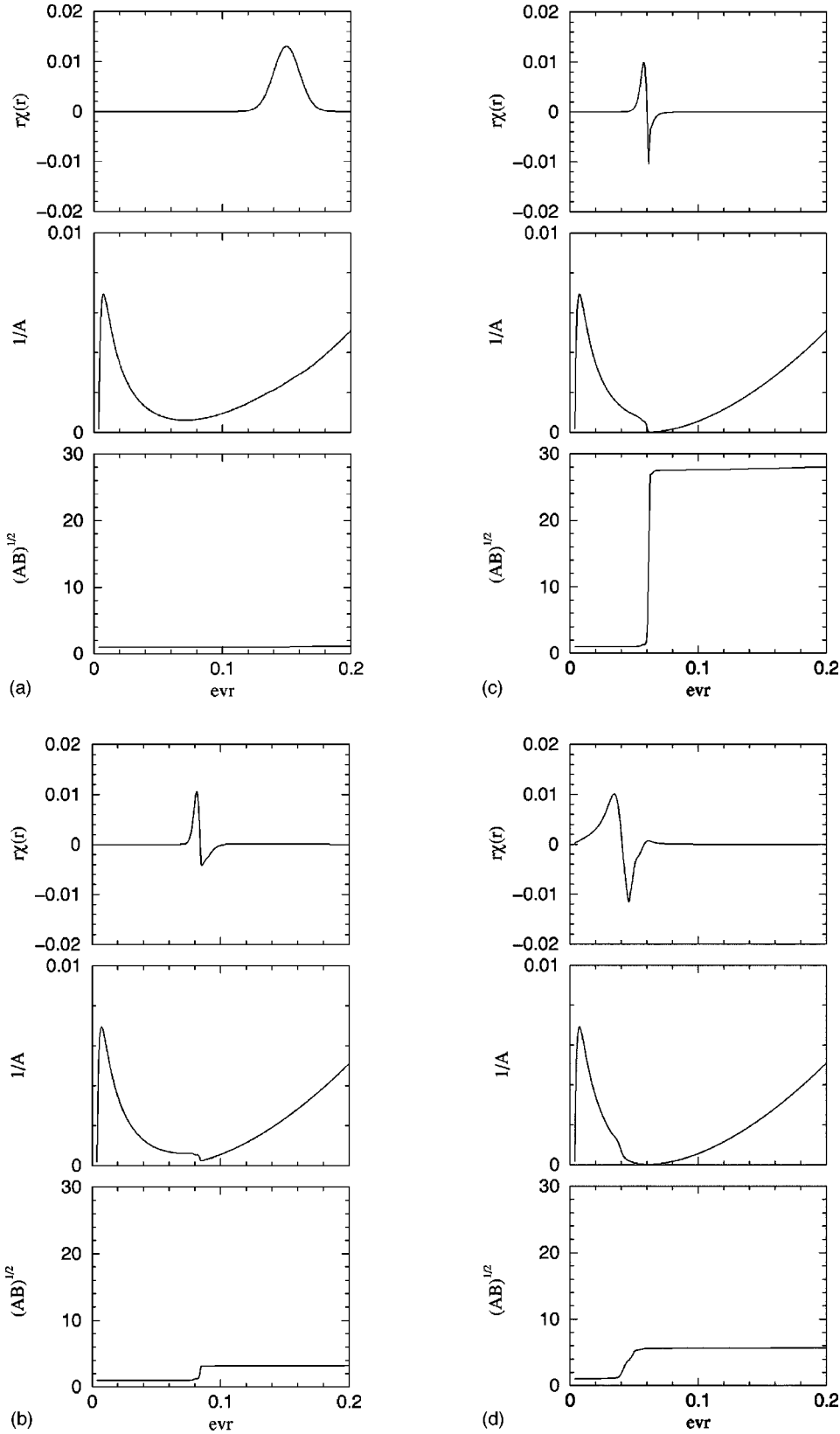


FIG. 4. Evolution of $r\chi$ and the metric functions $1/A$ and $(AB)^{1/2}$ with time. (a) At $t=0$, a radial Gaussian pulse in χ is sent into a monopole with $a=1.002$, $b=10^6$, and a small Schwarzschild black hole of radius $r_H=0.004(ev)^{-1}$ at its center. The scalar field is coupled only to gravity and has a mass $m=1.0ev$. (b) The configuration at $t=56.07(ev)^{-1}$. (c) At $t=90.72(ev)^{-1}$, $(1/A)_{\min}$ first attains the smallest value it exhibits in this process, with $(1/A)_{\min} \approx 2.3 \times 10^{-5}$. At the same time, $(AB)^{1/2}$ behind the pulse achieves its maximum value. (d) At $t=113.87(ev)^{-1}$, the pulse first hits the internal black hole.

pulse bounces off the central black hole (with a small amount of its energy being absorbed) and the process reverses itself. The pulse passes by the quasi-horizon and retreats to infinity. As it does so, one sees the metric variables

restore themselves to almost the original values they had before the insertion of the pulse.

As the initial pulse amplitude is increased towards its threshold value, the minimum value achieved by $1/A$ ap-

proaches zero, while the maximum value of \sqrt{AB} appears to grow without bound. The time t_{qh} at which the pulse passes through the quasi-horizon does not vary appreciably.⁹

A numerical simulation will not, of course, be able to produce a precisely extremal horizon. In our simulations, we have been able to adjust the pulse amplitude to make the minimum $1/A$ be as small as 2.3×10^{-5} . In analyzing the behavior of the solutions as the pulse amplitude is varied, we see no indication of any singularity as the threshold is approached. We therefore expect that a pulse precisely at threshold would produce a nonsingular extremal horizon. In this case of critical collapse, the subsequent evolution of the system would be quite similar to the subcritical case, with the pulse continuing inward, bouncing off the central black hole, and then retreating outward. However, in this case the spacetime into which it moves is causally distinct from the one where the pulse had originated; i.e., it is a new sector of the Penrose diagram.

As a final comment, in all this analysis, dropping in pressureless dust should give analogous results. One can invoke Birkhoff's law so that the metric behind the (radially thick) dust shell must be represented by a static metric. But one should expect the same sorts of naked singularity behavior since this results from the interaction of metric variables.

V. CONCLUDING REMARKS

In this paper we have used near-critical self-gravitating monopoles as tools for studying the transition from a nonsingular spacetime to one with a horizon. By analyzing the properties of trajectories that pass through the quasi-horizon and then emerge again, we have seen that the observations made by an external observer vary continuously and show no evidence of discontinuity when the critical limit is reached. This analysis also shows how the many causally distinct regions of the extremal black hole spacetime naturally emerge from the simple Penrose diagram of the subcritical monopole.

⁹This would not be the case if we had fixed the normalization of AB at spatial infinity; t_{qh} would then diverge as the threshold amplitude was approached.

A somewhat unexpected result from this analysis is that for Coulomb-type solutions the proper time required to traverse the interior region vanishes in the critical limit. This is closely associated with the fact that near-critical Coulomb solutions display naked black hole behavior; these are the first examples of configurations without horizons that do so. However, the absence of this behavior in core-type solutions shows that this is not a universal property of near-critical solutions.

Our analysis also sheds light on some aspects of black holes themselves. We have seen that the region bounded by the quasi-horizon becomes effectively inaccessible to outside observers when the solution is sufficiently close to criticality. The interior degrees of freedom thus become unmeasurable. Tracing over them then leads a naturally defined entropy that can be attributed to this configuration. An order of magnitude estimate of this entropy agrees with the Hawking-Bekenstein formula for the entropy of a black hole; it seems plausible that in the critical limit the agreement would be precise. These nonsingular near-critical solutions thus provide a concrete and unambiguous framework for implementing the old idea that black hole entropy might be understood in terms of the degrees of freedom hidden behind the horizon.

Finally, we have argued that an extremal black hole can be produced by allowing additional matter to fall into a near-critical monopole. We have illustrated this by numerical simulations. Starting with an initially nonsingular monopole, this leads to a zero-temperature black hole where there had previously been no horizon at all. Alternatively, one can start with a small black hole at the center of the monopole. In this latter case, a configuration with a nonzero Hawking temperature evolves into one with $T=0$. The existence of these possibilities gives additional clues for, and constraints on, a more precise formulation of the third law of black hole thermodynamics.

ACKNOWLEDGMENTS

We wish to thank Gary Horowitz and Krishna Rajagopal for helpful conversations. This work was supported in part by the U.S. Department of Energy.

-
- [1] K. Lee, V. P. Nair, and E. J. Weinberg, *Phys. Rev. D* **45**, 2751 (1992).
 - [2] M. E. Ortiz, *Phys. Rev. D* **45**, R2586 (1992).
 - [3] P. Breitenlohner, P. Forgács, and D. Maison, *Nucl. Phys.* **B383**, 357 (1992).
 - [4] P. Breitenlohner, P. Forgács, and D. Maison, *Nucl. Phys.* **B442**, 126 (1995).
 - [5] A. Lue and E. J. Weinberg, *Phys. Rev. D* **60**, 084025 (1999).
 - [6] Y. Brihaye, B. Hartmann, and J. Kunz, hep-th/9911148; Y. Brihaye, F. Grard, and S. Hoorelbeke, hep-th/9912023.
 - [7] G. T. Horowitz and S. F. Ross, *Phys. Rev. D* **56**, 2180 (1997).
 - [8] G. T. Horowitz and S. F. Ross, *Phys. Rev. D* **57**, 1098 (1998).
 - [9] W. Israel, *Phys. Rev. Lett.* **57**, 397 (1986).
 - [10] D. G. Boulware, *Phys. Rev. D* **8**, 2363 (1973).
 - [11] P. van Nieuwenhuizen, D. Wilkinson, and M. J. Perry, *Phys. Rev. D* **13**, 778 (1976).
 - [12] H. Hollmann, *Phys. Lett. B* **338**, 181 (1994).
 - [13] T. Tachizawa, K. Maeda, and T. Torii, *Phys. Rev. D* **51**, 4054 (1995).
 - [14] M. Srednicki, *Phys. Rev. Lett.* **71**, 666 (1993).