An action for black hole membranes

Maulik K. Parikh

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544

Frank Wilczek

School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540

The membrane paradigm is the remarkable view that, to an external observer, a black hole appears to behave exactly like a dynamical fluid membrane, obeying such pre-relativistic equations as Ohm's law and the Navier-Stokes equation. It has traditionally been derived by manipulating the equations of motion. Here we provide an action formulation of this picture, clarifying what underlies the paradigm and simplifying the derivations. Within this framework, we derive previous membrane results, and extend them to dyonic black hole solutions. We discuss how it is that an action can produce dissipative equations. Using a Euclidean path integral, we show that familiar semi-classical thermodynamic properties of black holes also emerge from the membrane action. Finally, in a Hamiltonian description, we establish the validity of a minimum entropy production principle for black holes. [S0556-2821(98)00616-X]

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

I. INTRODUCTION

The event horizon of a black hole is a peculiar object: it is a mathematically defined, locally undetectable boundary, a surface-of-no-return inside which light cones tip over and ''time'' becomes spatial [1]. Otherwise natural descriptions of physics often have trouble accommodating the horizon; as the most primitive example, the familiar Schwarzschild metric has a coordinate singularity there. Theories of fields that extend to the horizon face the additional challenge of having to define boundary conditions on a surface that is infinitely red-shifted, has a singular Jacobian, and possesses a normal vector which is also tangential. These considerations might induce one to believe that black hole horizons are fundamentally different from other physical entities.

On the other hand, further work has established a great variety of analogies between the horizon and more familiar, pre-relativistic bodies. In addition to the famous four laws of black hole thermodynamics [2–5], which are global statements, there is also a precise local mechanical and electrodynamic correspondence. In effect, it has been shown [6–10] that an observer who remains outside a black hole perceives the horizon to behave according to equations that describe a fluid bubble with electrical conductivity as well as shear and bulk viscosities. Moreover, it is possible to define a set of local surface densities, such as charge or energy-momentum, which inhabit the bubble surface and which obey conservation laws. Quite remarkably, a general-relativistically *exact* calculation then leads, for arbitrary non-equilibrium black holes, to equations for the horizon which can be precisely identified with Ohm's law, the Joule heating law, and the Navier-Stokes equation.

These relations were originally derived for the mathematical, or true, event horizon. For astrophysical applications it became more convenient to consider instead a ''stretched horizon,'' a (2+1)-dimensional time-like surface located slightly outside the true horizon. Because it has a non-singular induced metric, the stretched horizon provides a more tractable boundary on which to anchor external fields; outside a complicated boundary layer, the equations governing the stretched horizon are to an excellent approximation [11,12] the same as those for the true horizon. This view of a black hole as a dynamical time-like surface, or membrane, has been called the membrane paradigm [13].

Most of the mentioned results have been derived through general-relativistic calculations based on various intuitive physical arguments. In this paper, we show that the gravitational and electromagnetic descriptions of the membrane can be derived systematically, directly, and more simply from the Einstein-Hilbert or Maxwell actions. Aside from the appeal inherent in a least action principle, an action formulation is a unifying framework which is easily generalizable and has the advantage of providing a bridge to thermodynamics and quantum mechanics (see [14] for related work). In a follow-up paper, we exploit these advantages to evaluate some effects of the back-reaction of Hawking radiation on the spacetime geometry [15].

The key idea in what follows is that, since (classically) nothing can emerge from a black hole, an observer who remains outside a black hole cannot be affected by the dynamics inside the hole. Hence the equations of motion ought to follow from varying an action restricted to the external universe. However, the boundary term in the derivation of the Euler-Lagrange equations does not in general vanish on the stretched horizon as it does at the boundary of spacetime. In order to obtain the correct equations of motion, we must add to the external action a surface term that cancels this residual boundary term. The membrane picture emerges in interpreting the added surface term as electromagnetic and gravitational sources residing on the stretched horizon.

In the rest of this paper, we examine individually the boundary terms for the electromagnetic, gravitational, and axidilaton cases. We also discuss dissipation and provide thermodynamic and Hamiltonian descriptions. We use lowercase indices for four-dimensional tensor indices and uppercase indices for the two-dimensional tensors that occupy

space-like sections of the horizon. We use geometrized units ($G \equiv c \equiv 1$), and a spacetime metric with positive signature ($-+++$). Our sign conventions are those of Misner, Thorne, and Wheeler (MTW) [16], with the exception of the extrinsic curvature which we define to have a positive trace for a convex surface.

II. HORIZON PRELIMINARIES

In this section, we fix our conventions, first in words, then in equations. Through every point on the true horizon there exists a unique null generator l^a which we may parameterize by some regular time coordinate whose normalization we fix to equal that of time-at-infinity. Next, we choose a time-like surface just outside the true horizon. This is the stretched horizon, $\mathcal{H}$, whose location we parameterize by $\alpha \ll 1$ so that $\alpha \to 0$ is the limit in which the stretched horizon coincides with the true horizon. We will always take this limit at the end of any computation. Since many of the useful intermediate quantities will diverge as inverse powers of α, we renormalize them by the appropriate power of α. In that sense, α plays the role of a regulator.

For our purposes, the principal reason for preferring the stretched horizon over the true horizon is that the metric on a time-like—rather than null—surface is nondegenerate, permitting one to write down a conventional action. Generically (in the absence of horizon caustics), a one-to-one correspondence between points on the true and stretched horizons is always possible via, for example, ingoing null rays that pierce both surfaces (see [12] for details).

We can take the stretched horizon to be the world-tube of a family of time-like observers who hover just outside the true horizon. These nearly light-like "fiducial" observers are pathological in that they suffer an enormous proper acceleration and measure quantities that diverge as $\alpha \to 0$. However, although we take the mathematical limit in which the true and stretched horizons conflate, for physical purposes the proper distance of the stretched horizon from the true horizon need only be smaller than the length scale involved in a given measurement. In that respect, the stretched horizon, although a surrogate for the true horizon, is actually more fundamental than the true horizon, since measurements at the stretched horizon constitute real measurements that an external observer could make and report, whereas accessing any quantity measured at the true horizon would entail the observer's inability to report back his or her results.

We take our fiducial observers to have world lines U^a, parameterized by their proper time, τ. The stretched horizon also possesses a space-like unit normal n^a which for consistency we shall always take to be outward-pointing. Moreover, we choose the normal vector congruence on the stretched horizon to emanate outwards along geodesics. We define α by requiring that $\alpha U^a \to l^a$ and $\alpha n^a \to l^a$; hence αU^a and αn^a are equal in the true horizon limit. This is nothing more than the statement that the null generator l^a is both normal and tangential to the true horizon, which is the defining property of null surfaces. Ultimately, though, it will be this property that will be responsible for the dissipative behavior of the horizons. The 3-metric, h_{ab}, on $\mathcal{H}$ can be

written as a 4-dimensional tensor in terms of the spacetime metric and the normal vector, so that h^a_b projects from the spacetime tangent space to the 3-tangent space. Similarly, we can define the 2-metric, γ_{AB}, of the space-like section of $\mathcal{H}$ to which U^a is normal, in terms of the stretched horizon 3-metric and U^a, thus making a 2+1+1 split of spacetime. We denote the 4-covariant derivative by ∇_a, the 3-covariant derivative by $|a$, and the 2-covariant derivative by $\|A$. For a vector in the stretched horizon, the covariant derivatives are related by $h^c_d \nabla_c w^a = w^a_{|d} - K^c_d w_c n^a$ where $K^a_b \equiv h^c_b \nabla_c n^a$ is the stretched horizon's extrinsic curvature, or second fundamental form. In summary,

$$l^2 = 0 \tag{2.1}$$

$$U^a = \left(\frac{d}{d\tau}\right)^a, \quad U^2 = -1, \quad \lim_{\alpha \to \infty} \alpha U^a = l^a \tag{2.2}$$

$$n^2 = +1, \quad a^c = n^a \nabla_a n^c = 0, \quad \lim_{\alpha \to \infty} \alpha n^a = l^a \tag{2.3}$$

$$h^a_b = g^a_b - n^a n_b, \quad \gamma^a_b = h^a_b + U^a U_b = g^a_b - n^a n_b + U^a U_b \tag{2.4}$$

$$K^a_b \equiv h^c_b \nabla_c n^a, \quad K_{ab} = K_{ba}, \quad K_{ab} n^b = 0 \tag{2.5}$$

$$w^c \epsilon \mathcal{H} \Rightarrow h^c_d \nabla_c w^a = w^a_{|d} - K^c_d w_c n^a \Rightarrow \nabla_c w^c = w^c_{|c} + w^c a_c = w^c_{|c}. \tag{2.6}$$

The last expression relates the covariant divergence associated with g_{ab} to the covariant divergence associated with h_{ab}.

For example, the Reissner-Nordström solution has

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)dt^2$$
$$+ \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 d\Omega^2, \tag{2.7}$$

so that a stretched horizon at constant r would have

$$\alpha = \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{1/2}, \tag{2.8}$$

$$U_a = -\alpha(dt)_a, \tag{2.9}$$

and

$$n_a = +\alpha^{-1}(dr)_a. \tag{2.10}$$

III. ACTION FORMULATION

To find the complete equations of motion by extremizing an action, it is not sufficient to set the bulk variation of the action to zero: one also needs to use the boundary conditions. Here we take our Dirichlet boundary condition to be $\delta\phi = 0$ at the boundary of spacetime, where ϕ stands for any field.

Now since the fields inside a black hole cannot have any classical relevance for an external observer, the physics must follow from varying the part of the action restricted to the spacetime outside the black hole. However, this external action is not stationary on its own, because boundary conditions are fixed only at the singularity and at infinity, but not at the stretched horizon. Consequently, we rewrite the total action as

$$S_{\text{world}} = (S_{\text{out}} + S_{\text{surf}}) + (S_{\text{in}} - S_{\text{surf}}), \qquad (3.1)$$

where now $\delta S_{\text{out}} + \delta S_{\text{surf}} \equiv 0$, which implies also that $\delta S_{\text{in}} - \delta S_{\text{surf}} = 0$. The total action has been broken down into two parts, both of which are stationary on their own, and which do not require any new boundary conditions.

The surface term, S_{surf}, corresponds to sources, such as surface electric charges and currents for the Maxwell action, or surface stress tensors for the Einstein-Hilbert action. The sources are fictitious: an observer who falls through the stretched horizon will not find any surface sources and, in fact, will not find any stretched horizon. Furthermore, the field configurations inside the black hole will be measured by this observer to be entirely different from those posited by the membrane paradigm. On the other hand, for an external fiducial observer the source terms are a very useful artifice; their presence is consistent with all external fields. This situation is directly analogous to the method of image charges in electrostatics, in which a fictitious charge distribution is added to the system to implement, say, conducting boundary conditions. By virtue of the uniqueness of solutions to Poisson's equation with conducting boundary conditions, the electric potential on one—and only one—side of the boundary is guaranteed to be the correct potential. An observer who remains on that side of the boundary has no way of telling through the fields alone whether they arise through the fictitious image charges or through actual surface charges. The illusion is exposed only to the observer who crosses the boundary to find that not only are there no charges, but the potential on the other side of the boundary is quite different from what it would have been had the image charges been real.

In the rest of this section, we shall implement Eq. (3.1) concretely in important special cases.

A. The electromagnetic membrane

The external Maxwell action is

$$S_{\text{out}}[A_a] = \int d^4 x \sqrt{-g} \left(-\frac{1}{16\pi} F^2 + J \cdot A \right), \qquad (3.2)$$

where F is the electromagnetic field strength. Under variation, we obtain the inhomogeneous Maxwell equations

$$\nabla_b F^{ab} = 4\pi J^a, \qquad (3.3)$$

as well as the boundary term

$$\frac{1}{4\pi} \int d^3 x \sqrt{-h} F^{ab} n_a \delta A_b, \qquad (3.4)$$

where h is the determinant of the induced metric, and n^a is the outward-pointing space-like unit normal to the stretched horizon. We need to cancel this term. Adding the surface term

$$S_{\text{surf}}[A_a] = +\int d^3 x \sqrt{-h} j_s \cdot A, \qquad (3.5)$$

we see that we must have

$$j_s^a = +\frac{1}{4\pi} F^{ab} n_b. \qquad (3.6)$$

The surface 4-current, j_s^a, has a simple physical interpretation. We see that its time-component is a surface charge, σ, that terminates the normal component of the electric field just outside the membrane, while the spatial components, $\vec{j}_s$, form a surface current that terminates the tangential component of the external magnetic field:

$$E_\perp = -U_a F^{ab} n_b = 4\pi\sigma \qquad (3.7)$$

$$\vec{B}_\parallel^A = \epsilon_B^A \gamma_a^B F^{ab} n_b = 4\pi (\vec{j}_s \times \hat{n})^A. \qquad (3.8)$$

It is characteristic of the membrane paradigm that σ and $\vec{j}_s$ are *local* densities, so that the total charge on the black hole is the surface integral of σ over the membrane, taken at some constant universal time. This is in contrast to the total charge of a Reissner-Nordström black hole, which is a global characteristic that can be defined by an integral at spatial infinity.

From Maxwell's equations and Eq. (3.6), we obtain a continuity equation for the membrane 4-current which, for a stationary hole, takes the form

$$\frac{\partial \sigma}{\partial \tau} + \vec{\nabla}_2 \cdot \vec{j}_s = -J^n, \qquad (3.9)$$

where $\vec{\nabla}_2 \cdot \vec{j}_s \equiv (\gamma_a^A j_s^a)_{\parallel A}$ is the two-dimensional divergence of the membrane surface current, and $-J^n = -J^a n_a$ is the amount of charge that falls into the hole per unit area per unit proper time, τ. Physically, this equation expresses local charge conservation in that any charge that falls into the black hole can be regarded as remaining on the membrane: the membrane is impermeable to charge.

The equations we have so far are sufficient to determine the fields outside the horizon, given initial conditions outside the horizon. A plausible requirement for initial conditions at the horizon is that the fields measured by freely falling observers (FFO's) at the stretched horizon be finite. There being no curvature singularity at the horizon, inertial observers who fall through the horizon should detect nothing out of the ordinary. In contrast, the fiducial observers (FIDO's) who make measurements at the membrane are infinitely accelerated. Their measurements, subject to infinite Lorentz boosts, are singular. For the electromagnetic fields we have, with γ the Lorentz boost and using orthonormal coordinates,

$$E_\theta^{\text{FIDO}} \approx \gamma (E_\theta^{\text{FFO}} - B_\phi^{\text{FFO}}), \quad B_\phi^{\text{FIDO}} \approx \gamma (B_\phi^{\text{FFO}} - E_\theta^{\text{FFO}}),$$
$$(3.10)$$

$$B_\theta^{\text{FIDO}} \approx \gamma(B_\theta^{\text{FFO}} - E_\phi^{\text{FFO}}), \quad E_\phi^{\text{FIDO}} \approx \gamma(E_\phi^{\text{FFO}} - B_\theta^{\text{FFO}}),$$

$$(3.11)$$

or, more compactly,

$$\vec{E}_\parallel^{\text{FIDO}} = \hat{n} \times \vec{B}_\parallel^{\text{FIDO}}.$$

$$(3.12)$$

That is, the regularity condition states that all radiation in the normal direction is ingoing; a black hole acts as a perfect absorber. Combining the regularity condition with Eq. (3.8) and dropping the FIDO label, we arrive at

$$\vec{E}_\parallel = 4\pi \vec{j}_s.$$

$$(3.13)$$

That is, black holes obey Ohm's law with a surface resistivity of $\rho = 4\pi \approx 377$ Ω. Furthermore, the Poynting flux is

$$\vec{S} = \frac{1}{4\pi}(\vec{E} \times \vec{B}) = -j_s^2 \rho \hat{n}.$$

$$(3.14)$$

We can integrate this over the black hole horizon at some fixed time. However, for a generic stretched horizon, we cannot time-slice using fiducial time as different fiducial observers have clocks that do not necessarily remain synchronized. Consequently we must use some other time for slicing purposes, such as the time at infinity, and then include in the integrand a (potentially position-dependent) factor to convert the locally measured energy flux to one at infinity. With a clever choice of the stretched horizon, however, it is possible to arrange that all fiducial observers have synchronized clocks. In this case, two powers of α, which is now the lapse, are included in the integrand. Then, for some given universal time, t, the power radiated into the black hole, which is also the rate of increase of the black hole's irreducible mass, is given by

$$\frac{dM_{\text{irr}}}{dt} = -\int \alpha^2 \vec{S} \cdot d\vec{A} = +\int \alpha^2 j_s^2 \rho \; dA.$$

$$(3.15)$$

That is, black holes obey the Joule heating law, the same law that also describes the dissipation of an ordinary Ohmic resistor.

B. The gravitational membrane

We turn now to gravity. The external Einstein-Hilbert action is

$$S_{\text{out}}[g^{ab}] = \frac{1}{16\pi}\int d^4x \sqrt{-g}R + \frac{1}{8\pi}\oint d^3x\sqrt{\pm h} \; K + S_{\text{matter}},$$

$$(3.16)$$

where R is the Ricci scalar and K is the trace of the extrinsic curvature, and where for convenience we have chosen the field variable to be the inverse metric g^{ab}. The surface integral of K is only over the outer boundary of spacetime, and not over the stretched horizon. It is required in order to obtain the Einstein equations because the Ricci scalar contains second order derivatives of g_{ab}. When this action is varied, the bulk terms give the Einstein equations

$$R_{ab} - \frac{1}{2}g_{ab}R = 8\pi T_{ab}.$$

$$(3.17)$$

We are interested, however, in the interior boundary term. This comes from the variation of the Ricci tensor. We note that

$$g^{ab}\delta R_{ab} = \nabla^a[\nabla^b(\delta g_{ab}) - g^{cd}\nabla_a(\delta g_{cd})],$$

$$(3.18)$$

where $\delta g_{ab} = -g_{ac}g_{bd}\delta g^{cd}$. Gauss' theorem now gives

$$\int d^4x\sqrt{-g}(g^{ab}\delta R_{ab})$$

$$= -\int d^3x\sqrt{-h}n^a h^{bc}[\nabla_c(\delta g_{ab}) - \nabla_a(\delta g_{bc})],$$

$$(3.19)$$

where the minus sign arises from choosing n^a to be outward-pointing. Applying the Leibniz rule, we can rewrite this as

$$\int d^4x\sqrt{-g}(g^{ab}\delta R_{ab})$$

$$= \int d^3x\sqrt{-h}h^{bc}[\nabla_a(n^a\delta g_{bc}) - \delta g_{bc}\nabla_a(n^a)$$

$$- \nabla_c(n^a\delta g_{ab}) + \delta g_{ab}\nabla_c(n^a)].$$

$$(3.20)$$

Now, in the limit that the stretched horizon approaches the null horizon, the first and third terms on the right-hand side vanish:

$$\int d^3x\sqrt{-h}h^{bc}[\nabla_a(n^a\delta g_{bc}) - \nabla_c(n^a\delta g_{ab})] = 0.$$

$$(3.21)$$

A proof of this identity is given in the Appendix. With $K^{ba} = h^{bc}\nabla_c n^a$, the variation of the external action is

$$\delta S_{\text{out}}[g^{ab}] = \frac{1}{16\pi}\int d^3x\sqrt{-h}(Kh_{ab} - K_{ab})\delta g^{ab}.$$

$$(3.22)$$

Since the expression in parentheses contains only stretched horizon tensors, the normal vectors in the variation $\delta g^{ab} = \delta h^{ab} + \delta n^a n^b + n^a \delta n^b$ contribute nothing. As in the electromagnetic case, we add a surface source term to the action to cancel this residual boundary term. The variation of the required term can therefore be written as

$$\delta S_{\text{surf}}[h^{ab}] = -\frac{1}{2}\int d^3x\sqrt{-h}t_{sab}\delta h^{ab}.$$

$$(3.23)$$

We shall see later that this variation is integrable; i.e., an action with this variation exists. Comparison with Eq. (3.22) yields the membrane stress tensor

$$t_s^{ab} = +\frac{1}{8\pi}(Kh^{ab} - K^{ab}).$$

$$(3.24)$$

Now just as a surface charge produces a discontinuity in the normal component of the electric field, a surface stress term creates a discontinuity in the extrinsic curvature. The relation between the discontinuity and the source term is given by the Israel junction condition [16],

$$t_s^{ab} = \frac{1}{8\pi}([K]h^{ab} - [K]^{ab}),\qquad (3.25)$$

where $[K] = K_+ - K_-$ is the difference in the extrinsic curvature of the stretched horizon between its embedding in the external universe and its embedding in the spacetime internal to the black hole. Comparing this with our result for the membrane stress tensor, Eq. (3.24), we see that

$$K_-^{ab} = 0,\qquad (3.26)$$

so that the interior of the stretched horizon molds itself into flat space. The Einstein equations, Eq. (3.17), can be rewritten via the contracted Gauss-Codazzi equations [16] as

$$t_{s|b}^{ab} = -h_c^a T^{cd} n_d.\qquad (3.27)$$

Equations (3.24) and (3.27) taken together imply that the stretched horizon can be thought of as a fluid membrane, obeying the Navier-Stokes equation. To see this, recall that as we send α to zero, both αU^a and αn^a approach l^a, the null generator at the corresponding point on the true horizon. Hence, in this limit we can equate αU^a and αn^a, permitting us to write the relevant components of K_b^a, in terms of the surface gravity, g, and the extrinsic curvature, k_B^A, of a space-like 2-section of the stretched horizon:

$$U^c \nabla_c n^a \rightarrow \alpha^{-2} l^c \nabla_c l^a \equiv \alpha^{-2} g_H l^a$$
$$\Rightarrow K_U^U = -g, \quad K_U^A = \gamma_a^A K_b^a U^b = 0, \quad (3.28)$$

where $g_H \equiv \alpha g$ is the renormalized surface gravity at the horizon, and

$$\gamma_A^c \nabla_c n^b \rightarrow \alpha^{-1} \gamma_A^c \nabla_c l^b \Rightarrow K_A^B = \gamma_A^a K_a^b \gamma_b^B = \alpha^{-1} k_A^B,\qquad (3.29)$$

where k_{AB} is the extrinsic curvature of a space-like 2-section of the true horizon,

$$k_{AB} \equiv \gamma_A^d l_{B\|d} = \frac{1}{2}\,\pounds_{l^a} \gamma_{AB},\qquad (3.30)$$

where $\pounds_{l^a}$ is the Lie derivative in the direction of l^a. We can decompose k_{AB} into a traceless part and a trace, $k_{AB} = \sigma_{AB} + \frac{1}{2}\gamma_{AB}\theta$, where σ_{AB} is the shear and θ the expansion of the world lines of nearby horizon surface elements. Then

$$t_s^{AB} = \frac{1}{8\pi}\left[-\sigma^{AB} + \gamma^{AB}\left(\frac{1}{2}\theta + g\right)\right].\qquad (3.31)$$

But this is just the equation for the stress of a two-dimensional viscous Newtonian fluid [17] with pressure $p = g/8\pi$, shear viscosity $\eta = 1/16\pi$, and bulk viscosity $\zeta = -1/16\pi$. Hence we may identify the horizon with a two-

dimensional dynamical fluid, or membrane. Note that, unlike ordinary fluids, the membrane has negative bulk viscosity. This would ordinarily indicate an instability against generic perturbations triggering expansion or contraction. It can be regarded as reflecting a null hypersurface's natural tendency to expand or contract [10]. Below we shall show how for the horizon this particular instability is replaced with a different kind of instability.

Inserting the A-momentum density $t_{sa}^b \gamma_A^a U_b = t_{s\,A}^U \equiv \pi_A$ into the Einstein equations, Eq. (3.27), we arrive at the Navier-Stokes equation

$$\pounds_\tau \pi_A = -\nabla_A p + \zeta \nabla_A \theta + 2\eta \sigma_{A\,\|B}^B - T_A^n,\qquad (3.32)$$

where $\pounds_\tau \pi_A = \partial \pi_A / \partial \tau$ is the Lie derivative (which is the general-relativistic equivalent of the convective derivative) with respect to proper time, and $-T_A^n = -\gamma_A^a T_a^c n_c$ is the flux of A-momentum into the black hole.

Inserting the U-momentum (energy) density $t_{s\,b}^a U_a U^b \equiv \Sigma = -\theta/8\pi$ gives

$$\pounds_\tau \Sigma + \theta\Sigma = -p\theta + \zeta\theta^2 + 2\eta\sigma_{AB}\sigma^{AB} + T_b^a n_a U^b,\qquad (3.33)$$

which is the focusing equation for a null geodesic congruence [18]. We might now suspect that if the analogy with fluids extends to thermodynamics, then Eq. (3.33), as the equation of energy conservation, must be the heat transfer equation [17] for a two-dimensional fluid. Writing the expansion of the fluid in terms of the area, ΔA, of a patch,

$$\theta = \frac{1}{\Delta A}\frac{d\Delta A}{d\tau},\qquad (3.34)$$

we see that we can indeed rewrite Eq. (3.33) as the heat transfer equation (albeit with an additional relativistic term on the left)

$$T\left(\frac{d\Delta S}{d\tau} - \frac{1}{g}\frac{d^2\Delta S}{d\tau^2}\right) = (\zeta\theta^2 + 2\eta\sigma_{AB}\sigma^{AB} + T_b^a n_a U^b)\Delta A,\qquad (3.35)$$

with T the temperature and S the entropy, provided that the entropy is given by

$$S = \eta\frac{k_B}{\hbar}A,\qquad (3.36)$$

and the temperature by

$$T = \frac{\hbar}{8\pi k_B \eta}g,\qquad (3.37)$$

where η is some proportionality constant.

Thus, the identification of the horizon with a fluid membrane can be extended to the thermodynamic domain. Nonetheless, the membrane is an unusual fluid. The focusing equation itself, Eq. (3.33), is identical in form to the equation of energy conservation for a fluid. However, because the energy density, Σ, is proportional to the expansion, $\Sigma =$

$-\theta/8\pi$, one obtains a nonlinear first-order differential equation for θ which has no counterpart for ordinary fluids. The crucial point is that, owing to the black hole's gravitational self-attraction, the energy density is negative, and the solution to the differential equation represents a horizon that grows with time. For example, the source-free solution with a time-slicing for which the horizon has constant surface gravity is

$$\theta(t) = \frac{2g}{1 + \left(\dfrac{2g}{\theta(t_0)} - 1\right)e^{g(t_0 - t)}}. \tag{3.38}$$

Because of the sign of the exponent, this would represent an ever-expanding horizon if $\theta(t_0)$ were an initial condition; the area of the horizon, which is related to θ by $\theta = (d/d\tau)\ln\sqrt{\gamma}$, expands exponentially with time. To avoid this runaway, one must impose "teleological boundary conditions" (that is, final conditions) rather than initial conditions. Hence, the horizon's growth is actually acausal; the membrane expands to intercept infalling matter that is yet to fall in [13]. This is because the membrane inherits the global character of the true horizon: the stretched horizon covers the true horizon whose location can only be determined by tracking null rays into the infinite future. In fact, the left-hand side of the heat transfer equation, Eq. (3.35), is of the same form as that of an electron subject to radiation reaction; the acausality of the horizon is therefore analogous to the preacceleration of the electron.

At this classical level, using only the equations of motion, the parameter η in Eq. (3.36) is undetermined. However, because we have an action we hope to do better, since the normalization in the path integral is now fixed. By means of a Euclidean path integral, we should actually be able to derive the Bekenstein-Hawking entropy, including the coefficient η, from the membrane action. We do this in a later section.

C. The axidilaton membrane

Another advantage of the action formulation is that it is easily generalized to arbitrary fields. For example, we can extend the membrane paradigm to include the basic fields of quantum gravity. Here we use the tree-level effective action obtained from string theory after compactification to four macroscopic dimensions. This action is a generalization of the classical Einstein-Hilbert-Maxwell action to which it reduces when the axidilaton, λ, is set to $i/16\pi$. The action is

$$S[\lambda, \bar{\lambda}, A_a, g_{ab}]$$
$$= \int d^4x \sqrt{-g}\left(\frac{R}{16\pi} - \frac{|\partial\lambda|^2}{2\lambda_2^2} + \frac{i}{4}(\lambda F_+^2 - \bar{\lambda}F_-^2)\right), \tag{3.39}$$

where R is the four-dimensional Ricci curvature scalar, $F_\pm \equiv F \pm i\tilde{F}$ are the self- and anti-self-dual electromagnetic field strengths, and $\lambda \equiv \lambda_1 + i\lambda_2 = a + ie^{-2\phi}$ is the axidilaton, with

a the axion and ϕ the dilaton. Solutions to the equations of motion arising from this action include electrically (Reissner-Nordström) and magnetically charged black holes [19,20], as well as their duality-rotated cousins, dyonic black holes [21], which carry both electric and magnetic charge.

The equations of motion are

$$\nabla_a\left(\frac{\partial^a\lambda}{\lambda_2^2}\right) + i\frac{|\partial\lambda|^2}{\lambda_2^3} - \frac{i}{2}F_-^2 = 0 \tag{3.40}$$

and

$$\nabla_a(\lambda F_+^{ab} - \bar{\lambda}F_-^{ab}) = 0, \tag{3.41}$$

besides the Einstein equations.

As before, we require the external action to vanish on its own. Integration by parts on the axidilaton kinetic term leads to a variation at the boundary,

$$\int d^3x \sqrt{-h}\left[\delta\lambda\left(\frac{n_a\partial^a\bar{\lambda}}{2\lambda_2^2}\right) + \delta\bar{\lambda}\left(\frac{n_a\partial^a\lambda}{2\lambda_2^2}\right)\right], \tag{3.42}$$

where n^a is again chosen to be outward-pointing. To cancel this, we add the surface term

$$S_{\text{surf}} = \int d^3x \sqrt{-h}(\lambda\bar{q} + \bar{\lambda}q), \tag{3.43}$$

so that

$$q = -\frac{n_a\partial^a\lambda}{\lambda_2^2}. \tag{3.44}$$

To interpret this, we note that the kinetic term in λ is invariant under global $SL(2,\mathbb{R})$ transformations of the form

$$\lambda \rightarrow \frac{a\lambda + b}{c\lambda + d}, \quad ad - bc = 1, \tag{3.45}$$

which are generated by Peccei-Quinn shifts, $\lambda_1 \rightarrow \lambda_1 + b$, and duality transformations, $\lambda \rightarrow -1/\lambda$. The Peccei-Quinn shift of the axion can be promoted to a classical local symmetry to yield a Nöther current:

$$J_{P-Q}^a = -\frac{1}{2\lambda_2^2}(\partial^a\lambda + \partial^a\bar{\lambda}). \tag{3.46}$$

Therefore, under a Peccei-Quinn shift,

$$\delta S_{\text{surf}} = \int d^3x \sqrt{-h}\,\delta\lambda(q + \bar{q}) = \int d^3x \sqrt{-h}\,\delta\lambda(n_a J_{P-Q}^a). \tag{3.47}$$

The sum of the q and $\bar{q}$ terms induced at the membrane, Eq. (3.44), is the normal component of the Peccei-Quinn current. Hence, at the membrane,

$$(h_b^a J_{P-Q}^b)|_a = -F\tilde{F} - \nabla_a[(q + \bar{q})n^a]. \tag{3.48}$$

That is, the membrane term $\nabla_a[(q+\bar{q})n^a]$ augments the dyonic $F\tilde{F}$ term as a source for the three-dimensional Peccei-Quinn current, $h_b^a J_{P-Q}^b$, at the membrane.

The membrane is again dissipative with the Peccei-Quinn charge accounting for the dissipation in the usual $\alpha \rightarrow 0$ limit. The local rate of dissipation is given by the bulk stress tensor at the membrane:

$$T_{ab}U^a n^b = \frac{1}{16\pi} \frac{\partial_a\lambda \partial_b\bar{\lambda} + \partial_a\bar{\lambda}\partial_b\lambda}{2\lambda_2^2} U^a n^b \rightarrow \frac{\lambda_2^2|q|^2}{16\pi}. \tag{3.49}$$

In addition, the presence of the axidilaton affects the electromagnetic membrane. (The gravitational membrane is unaffected since the surface terms come from the Ricci scalar which has no axidilaton factor.) The electromagnetic current is now

$$j_s^a = -2i(\lambda F_+^{ab} - \bar{\lambda}F_-^{ab})n_b. \tag{3.50}$$

The surface charge is therefore

$$\sigma = 4(\lambda_2 E_\perp + \lambda_1 B_\perp), \tag{3.51}$$

and the surface current is

$$\vec{j}_s = 4(\lambda_2 \hat{n} \times \vec{B}_\parallel - \lambda_1 \hat{n} \times \vec{E}_\parallel), \tag{3.52}$$

which, by the regularity of the electromagnetic field, Eq. (3.12), satisfies

$$\begin{pmatrix} j_s^\theta \\ j_s^\phi \end{pmatrix} = 4 \begin{pmatrix} \lambda_2 & \lambda_1 \\ -\lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} E^\theta \\ E^\phi \end{pmatrix}. \tag{3.53}$$

The conductivity is now a tensor. When the axion is absent, the resistivity is

$$\rho = \frac{1}{4\lambda_2}. \tag{3.54}$$

The inverse dependence on λ_2 is to be expected on dimensional grounds. The pure dilaton action can be derived from Kaluza-Klein compactification of pure gravity in five dimensions, where the fifth dimension is curled into a circle of radius $e^{-2\phi} = \lambda_2$. In five dimensions, with $c \equiv 1$, resistance (and hence resistivity for a two-dimensional resistor such as the membrane) has dimensions of inverse length. Using the regularity condition, Eq. (3.12), the rate of dissipation, for a stretched horizon defined to have uniform lapse α with respect to time at infinity, t, is

$$\frac{dM_{\text{irr}}}{dt} = -\int \alpha^2 \vec{S} \cdot d\vec{A} = \int 4\alpha^2 \lambda_2 E_\parallel^2 dA = \int \alpha^2 \frac{\lambda_2}{4|\lambda|^4} \vec{j}_s^2 dA, \tag{3.55}$$

which is the Joule heating law in the presence of an axidilaton.

IV. DISSIPATION

Given that the bulk equations of motion are manifestly symmetric under time-reversal, the appearance of dissipation, as in Joule heating and fluid viscosity, might seem mysterious, all the more so since it has been derived from an action.

The procedure, described here, of restricting the action to some region and adding surface terms on the boundary of the region cannot be applied with impunity to any arbitrary region: a black hole is special. This is because the region outside the black hole contains its own causal past; an observer who remains outside the black hole is justified in neglecting (indeed, is unaware of) events inside. However, even ''past sufficiency'' does not adequately capture the requirements for our membrane approach. For instance, the past light cone of a spacetime point obviously contains its own past, but an observer in this light cone must eventually leave it. Rather, we define the notion of a future dynamically closed set:

A set S in a time-orientable globally hyperbolic spacetime (M, g_{ab}) is *future dynamically closed* if $J^-(S) = S$, and if, for some foliation of Cauchy surfaces Σ_t parameterized by the values of some global time function, we have that $\forall t_0 \forall p \in (S \cap \Sigma_{t_0}) \forall (t > t_0) \exists q \in (I^+(p) \cap S \cap \Sigma_t)$.

That is, S is future dynamically closed if it contains its own causal past and if from every point in S it is possible for an observer to remain in S. Classically, the region outside the true horizon of a black hole is dynamically closed. So too is the region on one side of a null plane in flat space; this is just the infinite-mass limit of a black hole. The region outside the stretched horizon is strictly speaking *not* dynamically closed since a signal originating in the thin region between the stretched horizon and the true horizon can propagate out beyond the stretched horizon. However, in the limit that the stretched horizon goes to the true horizon, $\alpha \rightarrow 0$, this region becomes vanishingly thin so that in this limit, which is in any case assumed throughout, we are justified in restricting the action.

The breaking of time-reversal symmetry comes from the definition of the stretched horizon; the region exterior to the black hole does not remain future dynamically closed under time-reversal. In other words, we have divided spacetime into two regions whose dynamics are derived from two different simultaneously vanishing actions, $\delta(S_{\text{out}} + S_{\text{surf}}) = \delta(S_{\text{in}} - S_{\text{surf}}) = 0$. Given data on some suitable achronal subset we can, for the exterior region, predict the future but not the entire past, while, inside the black hole, we can ''postdict'' the past but cannot determine the entire future. Thus, our choice of the horizon as a boundary implicitly contains the irreducible logical requirement for dissipation, that is, asymmetry between past and future.

Besides the global properties that logically permit one to write down a time-reversal asymmetric action, there is also a local property of the horizon which is the proximate cause for dissipation, namely that the normal to the horizon is also tangential to the horizon. Without this crucial property—which manifests itself as the regularity condition, or the identification of the stretched horizon extrinsic curvature with intrinsic properties of the true horizon—there would

still be surface terms induced at the stretched horizon, but no dissipation.

The regularity condition imposed at the boundary is not an operator identity, but a statement about physical states: all tangential electromagnetic fields as measured by a fiducial observer must be ingoing. Such a statement is not rigorously true. For any given value of $\alpha = (1 - 2M/r)^{1/2}$, there is a maximum wavelength, $\lambda_{\max}$, for outgoing modes that are invisible to the observer:

$$\lambda_{\max} = \frac{r - 2M}{\left(1 - \dfrac{2M}{r}\right)^{1/2}} \to 2M\alpha. \tag{4.1}$$

Dissipation occurs in the membrane paradigm because the finite but very high-frequency modes that are invisible to the fiducial observer are tacitly assumed not to exist. The regularity condition amounts to a coarse-graining over these modes. It is conceivable that for a theory with benign ultraviolet behavior, the amount of information lost is finite. Einstein gravity is not such a theory, but one may ask abstractly whether an effective horizon theory could exist at a quantum level [22,23]. Quantum effects cause the black hole to emit radiation. In order to preserve time-evolution unitarity, we might require the emitted radiation to be correlated with the interior state of the black hole. In this case, the membrane viewpoint remains valid only as a classical description, since quantum-mechanically the external universe receives information from the black hole in the form of deviations of the radiation from thermality; the crucial premise that the outside universe is emancipated from the internal state of the black hole is violated. It is important to emphasize, however, that correlations between the radiation and the horizon itself do not preclude the membrane paradigm. Indeed, the fact that the Bekenstein-Hawking entropy is proportional to the surface area of the black hole suggests that, even at the quantum level, an effective horizon theory may not be unfeasible.

V. THERMODYNAMICS

To make contact with thermodynamics, we perform an analytic continuation to imaginary time, $\tau = it$, so that the path integral of the Euclideanized action becomes a partition function. For a stationary hole, regularity (or the removal of a conical singularity) dictates a period $\beta = \int d\tau = 2\pi/g_H$ in imaginary time [24], where g_H is the surface gravity; for a Schwarzschild hole, $\beta = 8\pi M$. This is the inverse Hawking temperature in units where $\hbar = c = G = k_B = 1$. The partition function is then the path integral over all Euclidean metrics which are periodic with period $2\pi/g_H$ in imaginary time. Since the dominant contribution to the path integral comes from the classical solution, we can evaluate the partition function in a stationary phase approximation:

$$Z = \int Dg_E^{ab} \exp\left(-\frac{1}{\hbar}(S_{\text{out}}^E[g_E^{ab}] + S_{\text{surf}}^E[h_E^{ab}]) \right)$$

$$\approx \exp\left(-\frac{1}{\hbar}(S_{\text{out}}^E[g_{E\,\text{cl}}^{ab}] + S_{\text{surf}}^E[h_{E\,\text{cl}}^{ab}]) \right). \tag{5.1}$$

The external action itself can be written as $S_{\text{out}} = S_{\text{bulk}} + S_\infty$, where S_{bulk} is zero for a black hole alone in the universe. The boundary term S_∞ is the integral of the extrinsic curvature of the boundary of spacetime. In fact, a term proportional to the surface area at infinity can be included in S_∞ without affecting the Einstein equations since the metric is held fixed at infinity during variation. In particular, the proportionality constant can be chosen so that the action for all of spacetime is zero for Minkowski space:

$$S_\infty = \frac{1}{8\pi} \int d^3x \sqrt{-h}[K], \tag{5.2}$$

where $[K]$ is the difference in the trace of the extrinsic curvature at the spacetime boundary for the metric g_{ab} and the flat-space metric η_{ab}. With this choice, the path integral has a properly normalized probabilistic interpretation. The Euclideanized value of S_∞ for the Schwarzschild solution is then [24]

$$S_\infty^E = \lim_{r \to \infty} \frac{1}{8\pi}(-32\pi^2 M)\left[(2r - 3M) - 2r\left(1 - \frac{2M}{r}\right)^{1/2} \right]$$

$$= +4\pi M^2. \tag{5.3}$$

To obtain an explicit action for the membrane, we must integrate its variation, Eq. (3.23):

$$\delta S_{\text{surf}}[h^{ab}] = -\frac{1}{16\pi} \int d^3x \sqrt{-h}(Kh_{ab} - K_{ab})\delta h^{ab}. \tag{5.4}$$

We see that

$$S_{\text{surf}}[h^{ab}] = \int d^3x \sqrt{-h}(B_{ab}h^{ab} - b) \tag{5.5}$$

is a solution, provided that the (undifferentiated) source terms are $B_{ab} = (+1/16\pi)K_{ab}$ and $b = (-1/16\pi)K$. This action has the form of surface matter plus a negative cosmological constant in three dimensions. The value of the membrane action for a solution to the classical field equations is then

$$S_{\text{surf}}[h_{\text{cl}}^{ab}] = +\frac{1}{8\pi} \int d^3x \sqrt{-h_{\text{cl}}}K_{\text{cl}}. \tag{5.6}$$

To evaluate this, we can take our fiducial world-lines U^a to be normal to the isometric time-slices of constant Schwarzschild time. The stretched horizon is then a surface of constant Schwarzschild r. Hence $\alpha = (1 - 2M/r)^{1/2}$, $\theta = 0$, and $K = g + \theta = g$, the unrenormalized surface gravity of the stretched horizon. Inserting these into Eq. (5.6), we find that the Euclidean action is

$$S_{\text{surf}}^E = \lim_{r \to r_H} \frac{1}{8\pi}\left(\int -d\tau \right)\alpha 4\pi r^2 g = -\pi r_H^2 = -4\pi M^2, \tag{5.7}$$

where $r_H = 2M$ is the black hole's radius, and $g_H = \alpha g = 1/4M$ is its renormalized surface gravity.

The Euclidean membrane action exactly cancels the external action, Eq. (5.3). Hence the entropy is zero! That, however, is precisely what makes the membrane paradigm attractive: to an external observer, there is no black hole—only a membrane—and so neither a generalized entropy nor a strictly obeyed second law of thermodynamics. The entropy of the outside is simply the logarithm of the number of quantum states of the matter outside the membrane. This number decreases as matter leaves the external system to fall through and be dissipated by the membrane. When all matter has fallen into the membrane, the outside is in a single state—vacuum—and has zero entropy, as above.

To recover the Bekenstein-Hawking entropy, we must then use not the combination of external and membrane actions, which gave the entropy of the external system, but the combination of the *internal* and membrane actions,

$$Z_{B\text{-}H} = \int Dg_E^{ab} \exp\left(-\frac{1}{\hbar}(S_{\text{in}}^E[g_E^{ab}] - S_{\text{surf}}^E[h_E^{ab}])\right),$$
(5.8)

where now S_{surf} is subtracted [see Eq. (3.1)]. With $S_{\text{in}} = \int d^4x \sqrt{-g}R = 0$, the partition function for a Schwarzschild hole in the stationary phase approximation is

$$Z_{B\text{-}H} \approx \exp\left(-\frac{1}{\hbar}(+4\pi M^2)\right),$$
(5.9)

from which the Bekenstein-Hawking entropy, $S_{B\text{-}H}$, immediately follows:

$$S_{B\text{-}H} = \beta\left(M + \frac{\ln Z_{B\text{-}H}}{\beta}\right) = 8\pi M\left(M - \frac{1}{8\pi M}4\pi M^2\right) = \frac{1}{4}A,$$
(5.10)

which is the celebrated result.

For more general stationary (Kerr-Newman) holes, the Helmholtz free energy contains additional "chemical potential" terms corresponding to the other conserved quantities, Q and J,

$$F = M - TS - \Phi Q - \Omega J,$$
(5.11)

where $\Phi = Q/r_+$ and $\Omega = J/M$, where r_+ is the Boyer-Lindquist radial coordinate at the horizon. For a charged hole, the action also contains electromagnetic terms. The surface electromagnetic term, Eq. (3.5), has the value $(1/4\pi)\int d^3x \sqrt{-h} F^{ab} A_a n_b$. However, in order to have a regular vector potential, we must gauge transform it to $A_a = (Q/r - \Phi)\nabla_a t$ which vanishes on the surface. Hence, the surface action is again given by the gravitational term, which has the Euclideanized value $S_{\text{surf}}^E = -\pi r_+^2$. It is easy to verify using Eq. (5.11) that this again leads to a black hole entropy equal to one-fourth of the horizon surface area and an external entropy of zero.

For nonstationary black holes, the extrinsic curvature also includes a term for the expansion of the horizon, $K = g + \theta$. Inserting this into the surface action enables us to calculate the instantaneous entropy as matter falls into the membrane in a nonequilibrium process. Of course, like the horizon itself, the entropy grows acausally.

VI. HAMILTONIAN FORMULATION

The equations of motion can equally well be derived within a Hamiltonian formulation. This involves first singling out a global time coordinate, t, for the external universe, which is then sliced into space-like surfaces, Σ_t, of constant t. We can write, in the usual way,

$$t^a \equiv \left(\frac{d}{dt}\right)^a = \alpha U^a - v^a,$$
(6.1)

where U^a is the unit normal to Σ_t, $U^2 = -1$, and α and $-v^a$ are Wheeler's lapse and shift, respectively, with $v^a = dx^a/dt$ the ordinary 3-velocity of a particle with world-line U^a. For convenience we choose the stretched horizon to be a surface of constant lapse so that α, which goes to zero at the true horizon, serves as the stretched horizon regulator. The external Hamiltonian for electrodynamics, obtained from the Lagrangian via a Legendre transform and written in ordinary three-dimensional vector notation, is

$$H_{\text{out}}[\phi, \vec{A}, \vec{\pi}] = \frac{1}{4\pi} \int_{\Sigma_t} d^3x \sqrt{^3g}\left(\frac{1}{2}\alpha(\vec{E}\cdot\vec{E} + \vec{B}\cdot\vec{B})\right.$$
$$\left. + \vec{v}\cdot(\vec{E}\times\vec{B}) - \phi(\vec{\nabla}\cdot\vec{E})\right),$$
(6.2)

where $^3g_{ab}$ is the 3-metric on Σ_t, $\phi \equiv -A_a t^a$ is the scalar potential, $\vec{A}_a \equiv {}^3g_a{}^b A_b$ is the three-dimensional vector potential, and $\vec{\pi}^a \equiv -\sqrt{^3g}\vec{E}^a$ its canonical momentum conjugate. Note that $E^a = F^{ab}U_b$ is the co-moving electric field; $\vec{E}$ and $\vec{B}$ above refer to the fields measured by a fiducial observer with world-line U^a. Finally, the scalar potential is nondynamical; its presence in the Hamiltonian serves to enforce Gauss' law as a constraint. The equations of motion are now determined by Hamilton's equations and the constraint:

$$\frac{\delta H}{\delta \vec{\pi}} = \dot{\vec{A}}, \quad \frac{\delta H}{\delta \vec{A}} = -\dot{\vec{\pi}}, \quad \frac{\delta H}{\delta \phi} = 0.$$
(6.3)

In the bulk these equations are simply Maxwell's equations but, because of the inner boundary, the usually discarded surface terms that arise during integration by parts now need to be canceled. It is easy to show then that the above equations hold only if additional surface terms are added to the Hamiltonian:

$$H = H_{\text{out}} - \int d^2x \sqrt{\gamma}\, j_s \cdot A.$$
(6.4)

For Maxwell's equations to be satisfied in the bulk, the surface terms are once again the surface charges and currents necessary to terminate the normal electric and tangential magnetic fields at the stretched horizon. Thus, the membrane paradigm is recovered.

However, it is perhaps more interesting to proceed in a slightly different fashion. Instead of adding new terms, we can use the external Hamiltonian to prove the validity of a principle of minimum heat production. Such a principle, which holds under rather general circumstances for stationary dissipative systems, holds for black holes also in slightly nonstationary situations.

Now the time derivative of the external Hamiltonian is not zero, again because of the inner boundary. We can use Hamilton's equations to substitute expressions for the time derivative of the field and its momentum conjugate. Hamilton's equations are

$$\dot{\vec{A}} = -\alpha\vec{E} + \vec{v}\times\vec{B} - \vec{\nabla}\phi \qquad (6.5)$$

$$\dot{\vec{E}} = \vec{\nabla}\times(\alpha\vec{B} + \vec{v}\times\vec{E}), \qquad (6.6)$$

so that, making repeated use of the vector identity

$$\vec{\nabla}\cdot(\vec{V}\times\vec{W}) = \vec{W}\cdot(\vec{\nabla}\times\vec{V}) - \vec{V}\cdot(\vec{\nabla}\times\vec{W}), \qquad (6.7)$$

we find that the energy loss is

$$-\dot{H} = -\frac{1}{4\pi}\int d^2x\sqrt{\gamma}[\hat{n}\cdot(\alpha\vec{E}_\parallel\times\alpha\vec{B}_\parallel)$$

$$+ \vec{v}\cdot(E_\perp\alpha\vec{E}_\parallel + B_\perp\alpha\vec{B}_\parallel)]. \qquad (6.8)$$

So far, we have used only Hamilton's equations. It remains, however, to implement the constraint. Hence we may regard $-\dot{H}$ as a functional of the Lagrange multiplier, ϕ. We therefore have

$$-\frac{\delta\dot{H}}{\delta\phi} = -\frac{d}{dt}\frac{\delta H}{\delta\phi} = 0. \qquad (6.9)$$

That is, the equations of motion follow from minimizing the rate of mass increase of the black hole with respect to the scalar potential. This is an exact statement; we now show that this reduces to a minimum heat production principle in a quasi-stationary limit. Now we note that the first law of black hole thermodynamics allows us to decompose the mass change into irreducible and rotational parts:

$$\frac{dM}{dt} = \frac{dQ}{dt} + \Omega_H\frac{dJ}{dt}, \qquad (6.10)$$

where Ω_H is the angular velocity at the horizon, and J is the hole's angular momentum. Since $|\vec{v}|\to\Omega_H$ at the horizon, we see that the second term on the right in Eq. (6.8) corresponds to the torquing of the black hole. When this is small, we may approximate the mass increase as coming from the first, irreducible term. Hence, in the quasi-stationary limit, for a slowly rotating black hole, the black hole's rate of mass increase is given by the dissipation of external energy. Invoking the regularity condition, Eq. (3.12), then gives

$$D[\phi] = +\frac{1}{4\pi}\int d^2x\sqrt{\gamma}\,(\alpha\vec{E}_\parallel)^2, \quad \frac{\delta D}{\delta\phi} = 0, \quad (6.11)$$

where $\alpha\vec{E}_\parallel$ is given by Eq. (6.5). This is the principle of minimum heat production: minimizing the dissipation functional leads to the membrane equation of motion.

We observe that we could have anticipated this answer. The numerical value of the Hamiltonian is the total energy of the system as measured at spatial infinity (assuming an asymptotically flat spacetime). The time derivative is then simply the rate, as measured by the universal time of distant observers, that energy changes. The rate of decrease of energy is the integral of the Poynting flux as measured by local observers, multiplied by two powers of α, one power to convert local energy to energy-at-infinity and one power to convert the rate measured by local clocks to the rate measured at infinity. Thus we can immediately define a dissipation functional:

$$D[\phi] \equiv -\frac{1}{4\pi}\int d^2x\sqrt{\gamma}\,\hat{n}\cdot(\vec{E}_H\times\vec{B}_H), \qquad (6.12)$$

where the subscript H denotes that a power of α has been absorbed to renormalize an otherwise divergent fiducial quantity.

In this manner, we can easily write down the dissipation functional for gravity for which time-differentiating the Hamiltonian is a much more laborious exercise. The local rate of energy transfer is given by the right-hand side of the heat transfer equation, Eq. (3.35). The Hamiltonian for gravity satisfies two constraint equations with the lapse and shift vector serving as Lagrange multipliers. Since the membrane picture continues to have a gauge freedom associated with time-slicing, the constraint equation associated with the lapse is not implemented. This implies that the dissipation is a functional only of the shift. Hence we have

$$D[v^A] = \int d^2x\sqrt{\gamma}(\zeta\theta_H^2 + 2\eta\sigma_H^2 + \alpha^2 T_b^a n_a U^b), \qquad (6.13)$$

where again the two powers of α have been absorbed to render finite the quantities with the subscript H. Extremizing D with respect to v^A leads to the membrane equations of motion, enforcing the gauge constraint or, equivalently, obeying the principle of minimum heat production.

VII. CONCLUSION

We have derived the equations for the membrane paradigm of black holes from an action principle directly by demanding that both terms in Eq. (3.1) be extremized separately. This brings advantages of conceptual unity and ease of generalization over the traditional approach of manipulating the equations of motion. Specifically, the derivation makes it clear why a membrane picture, including dissipative behavior, is possible. A fundamental advantage of having an action principle is the guidance it offers for quantization—a property we used to fix the constant in the Bekenstein-Hawking formula.

APPENDIX

In this appendix, we shall prove that Eq. (3.21) is zero in the limit that the stretched horizon approaches the true horizon. In that limit, $\alpha n^a \to l^a$. We shall make liberal use of Gauss' theorem, the Leibniz rule, and the fact that $h^{ab} n_b$

$= K^{ab} n_b = 0$. In order to use Gauss' theorem, we note that since the "acceleration" $a^c \equiv n^d \nabla_d n^c$ of the normal vector (not to be confused with the fiducial acceleration $U^d \nabla_d U^c$) is zero, the 4-covariant divergence and the 3-covariant divergence of a vector in the stretched horizon are equal, Eq. (2.6).

Now, variations in the metric that are in fact merely gauge transformations can be set to zero. Using a vector v^a where v^a vanishes on the stretched horizon, we can gauge away the variations in the normal direction so that $\delta g_{ab} \to \delta h_{ab}$. Then the left-hand side of Eq. (3.21) becomes

$$\int d^3 x \sqrt{-h}\, h^{bc}[\nabla_a(n^a \delta h_{bc}) - \nabla_c(n^a \delta h_{ab})]$$

$$= \int d^3 x \sqrt{-h}\,[\nabla_a(h^{bc} n^a \delta h_{bc}) - (\nabla_a h^{bc}) n^a \delta h_{bc} - \nabla_c(h^{bc} n^a \delta h_{ab}) + (\nabla_c h^{bc}) n^a \delta h_{ab}]$$

$$= \int d^3 x \sqrt{-h}\,[\nabla_a(h^{bc} n^a \delta h_{bc}) + (n^c a^b + n^b a^c)\delta h_{bc} - (h^{bc} n^a \delta h_{ab})_{|c} - h^{bc} n^a \delta h_{ab} a_c - K n^b n^a \delta h_{ab} - a^b n^a \delta h_{ab}]$$

$$\text{(using } h^{bc} = g^{bc} - n^b n^c,\ K_{ab} = +h_a^c \nabla_c n_b,\ \text{and } \nabla_c w^c = w^c_{|c} + w^c a_c \text{ for } w^c \epsilon \mathcal{H})$$

$$= \int d^3 x \sqrt{-h}\,[\nabla_a(h^{bc} n^a \delta h_{bc}) - K n^b n^a \delta h_{ab}] \quad \text{(using Gauss' theorem, and } a^c = 0)$$

$$= \int d^3 x \sqrt{-h}\left[\nabla_a\left(h^{bc}\frac{\alpha}{\alpha} n^a \delta h_{bc}\right) - K[\delta(n^b n^a h_{ab}) - n^a h_{ab}\delta n^b - n^b h_{ab}\delta n^a]\right]$$

$$\to \int d^3 x \sqrt{-h}\,\nabla_a\left(h^{bc}\frac{1}{\alpha} l^a \delta h_{bc}\right) \quad \text{(using } h_{ab} n^b = 0, \text{ and } \alpha n^a \to l^a)$$

$$= \int d^3 x \sqrt{-h}\left(h^{bc}\frac{1}{\alpha} l^a \delta h_{bc}\right)_{|a}$$

$$= 0. \tag{A1}$$

[1] For a review see, e.g., S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-time* (Cambridge University Press, Cambridge, England, 1973).

[2] D. Christodoulou, Phys. Rev. Lett. **25**, 1596 (1970).

[3] D. Christodoulou and R. Ruffini, Phys. Rev. D **4**, 3552 (1971).

[4] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).

[5] J. M. Bardeen, B. Carter, and S. W. Hawking, Commun. Math. Phys. **31**, 161 (1973).

[6] R. S. Hanni and R. Ruffini, Phys. Rev. D **8**, 3259 (1973).

[7] T. Damour, Phys. Rev. D **18**, 3598 (1978).

[8] R. L. Znajek, Mon. Not. R. Astron. Soc. **185**, 833 (1978). Znajek's approach differs from ours in that he considers a membrane with a nonzero thickness and volume conductivities.

[9] T. Damour, Ph.D. thesis, University of Paris VI, 1979.

[10] T. Damour, in *Proceedings of the Second Marcel Grossman Meeting on General Relativity*, edited by R. Ruffini (North-Holland, Amsterdam, 1982), p. 587.

[11] D. A. Macdonald and K. S. Thorne, Mon. Not. R. Astron. Soc. **198**, 345 (1982).

[12] R. H. Price and K. S. Thorne, Phys. Rev. D **33**, 915 (1986).

[13] *Black Holes: The Membrane Paradigm*, edited by K. S. Thorne, R. H. Price, and D. A. Macdonald (Yale University Press, London, 1986). It should be noted that these authors imply by "the membrane paradigm" not only the existence of a stretched horizon, but also a 3+1 split of spacetime. Their motivation for severing space from time is pragmatic: astrophysical processes are usually not modeled in the full tensor formalism. In this paper, we require such a split only to write down Ohm's law, the Joule law, and the Navier-Stokes equation, in their familiar Galilean form.

[14] S. Carlip, Phys. Rev. D **51**, 632 (1995).

[15] M. K. Parikh and F. Wilczek, Princeton University Report No. PUPT-1775.

[16] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, New York, 1973).

[17] See, e.g., L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1987).

[18] S. W. Hawking and J. B. Hartle, Commun. Math. Phys. **27**, 283 (1972).

[19] G. Gibbons and K. Maeda, Nucl. Phys. **B298**, 741 (1988).

[20] D. Garfinkle, G. Horowitz, and A. Strominger, Phys. Rev. D **43**, 3140 (1991).

[21] A. Shapere, S. Trivedi, and F. Wilczek, Mod. Phys. Lett. A **6**, 2677 (1991).

[22] G. 't Hooft, Nucl. Phys. **B335**, 138 (1990).

[23] L. Susskind, L. Thorlacius, and J. Uglum, Phys. Rev. D **48**, 3743 (1993).

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