Universality and Scaling at the Onset of Quantum Black Hole Formation

Andrew Strominger*
Department of Physics, University of California at Santa Barbara, Santa Barbara, California 93106-9530

Lárus Thorlacius†
Institute for Theoretical Physics, University of California at Santa Barbara, Santa Barbara, California 93106-4030
(Received 9 December 1993)

In certain two-dimensional models, collapsing matter forms a black hole if and only if the incoming energy flux exceeds the Hawking radiation rate. Near the critical threshold, the black hole mass is given by a universal formula in terms of the distance from criticality, and there exists a scaling solution describing the formation and evaporation of an arbitrarily small black hole.

PACS numbers: 04.70.Dy, 04.60.Kz

Recently remarkable critical behavior has been discovered at the onset of classical black hole formation. Choptuik [1] considered a two-dimensional theory obtained from the S-wave sector of four-dimensional general relativity coupled to a massless scalar field. A sufficiently weak incoming S-wave pulse is simply reflected through the origin to an outgoing pulse. This behavior changes qualitatively as the initial amplitude is increased. Above a certain threshold, the pulse crosses its own Schwarzschild radius before it reaches the origin, and a black hole is formed. Choptuik [1] numerically investigated the mass $M_{\rm BH}$ of the resulting black hole as a function of the distance δ in the initial data space from the threshold. For small δ he finds

$$\ln M_{\rm BH} = \gamma \ln\delta + O(\delta^0)\,, \qquad (1)$$

where the critical exponent γ is numerically found to be near 0.37. This result appears quite universal and is insensitive to the precise definition of δ or scalar field couplings. In addition the near-threshold scaling solution was found to have fascinating self-similar oscillations. Related work can be found in [2]. At present there appears to be little analytic or conceptual understanding of these interesting phenomena.

Two-dimensional theories obtained by S-wave reduction from four dimensions have also been of recent interest as simplified arenas for the study of quantum black hole evaporation [3]. In this context null matter which reduces to a free conformal field in two dimensions is usually considered because it is much simpler than a scalar field (which reduces to a field with complicated gravitational couplings). At the classical level, every incoming S-wave pulse of such null matter, no matter how weak, forms a black hole. A threshold appears, however, when quantum effects are incorporated: Energy must be thrown at the origin at a rate faster than a newly formed black hole wants to Hawking evaporate.

In this paper we study the onset of quantum black hole formation using the semiclassically soluble two-dimensional Russo-Susskind-Thorlacius (RST) model [4]. We find analytically that

$$\ln\frac{M_{\rm BH}}{\lambda} = \frac{1}{2}\ln\delta - \frac{1}{2}\ln\alpha + O(\delta^0)\,. \qquad (2)$$

$M_{\rm BH}$ is defined here as the incoming energy of the null matter swallowed by the black hole during its lifetime and λ is the dimensionful parameter of the model. The $O(\delta^0)$ term depends universally on α, the second derivative of the energy density at the point where the critical threshold is exceeded. [It would be interesting to determine if the $O(\delta^0)$ term of (1) has a similar universal dependence.] The $O(\delta^{1/2})$ corrections depend nonuniversally on the shape of the incoming pulse. We also find a scaling solution near criticality which corresponds to the formation and evaporation of an arbitrarily small black hole. We have not determined whether the scaling (2) is universal with respect to small changes in the coupling constants of the theory. Presumably numerical work is required to answer this interesting question.

There are obvious similarities between our results and those of Ref. [1], but there are also apparent differences. First of all, our critical exponent is a rational number whereas Choptuik's at least appears to be irrational. Second, there is no analog of the self-similar oscillations in our work. This could be a special feature arising from the linear nature of the RST equations, while more general two-dimensional models (which are not analytically soluble) might exhibit such oscillations.

We now present a derivation of the scaling relation (2). The semiclassical effective action for the RST model is

$$S = \frac{1}{2\pi}\int d^2x\sqrt{-g}\left[\left(e^{-2\phi} - \frac{N}{24}\phi\right)R + 4e^{-2\phi}[(\nabla\phi)^2+\lambda^2] - \frac{1}{2}\sum_{i=1}^{N}(\nabla f_i)^2\right]$$
$$- \frac{N}{96\pi}\int d^2x\sqrt{-g(x)}\int d^2x'\sqrt{-g(x')}R(x)G(x;x')R(x')\,, \qquad (3)$$

where $g_{\mu\nu}$ is the two-dimensional metric, ϕ is a scalar field called the dilaton, f_i are N minimally coupled scalar matter

0031-9007/94/72(11)/1584(4)$06.00

fields and G is a Green function for the operator ∇^2. (For more details on this model consult [4,5]. We use the conventions of [5].) The effective action includes the one-loop Liouville term due to the matter fields and if N is large this term provides the dominant quantum backreaction on the geometry. This model differs from the original Callan-Giddings-Harvey-Strominger model [6] by a finite local counterterm, which restores a global symmetry of the classical theory and enables writing down exact semiclassical solutions in a rather simple form. Numerical analyses of the original model [7] indicate, however, that the two models are similar, i.e., that the qualitative behavior of semiclassical solutions is not sensitive to the existence of the global symmetry.

It is convenient to work in conformal gauge, $g_{+-} = -\frac{1}{2}e^{2\rho}$, $g_{++}=g_{--}=0$, and use the global symmetry to further fix the coordinates to "Kruskal gauge," where $\rho=\phi+\frac{1}{2}\ln(N/12)$. This eliminates the conformal factor from the discussion. If we define a new dilaton field,

$$\Omega=\frac{12}{N}e^{-2\phi}+\frac{1}{2}\phi+\frac{1}{4}\ln\frac{N}{48}, \tag{4}$$

the semiclassical equations reduce to

$$\partial_+\partial_- f_i=0\,,\quad \partial_+\partial_-\Omega=-\lambda^2\,,\quad -\partial_\pm^2\Omega=T^f_{\pm\pm}+t_\pm\,, \tag{5}$$

where

$$T^f_{\pm\pm}=(6/N)\sum_{i=1}^{N}(\partial_\pm f_i)^2\,.$$

The function $t_+(x^+)$ takes the value $t_+=-1/4(x^+)^2$ in Kruskal coordinates for any incoming matter energy flux which vanishes sufficiently rapidly at asymptotic early and late times. The field redefinition (4) is degenerate at $\Omega=\frac{1}{4}$ and $\Omega<\frac{1}{4}$ does not correspond to a real value of ϕ. The curve $\Omega=\frac{1}{4}$ is the analog of the origin of radial coordinates in higher dimensional gravity and solutions should not be continued beyond it. (This analogy can be made precise when the model is interpreted as an effective theory for radial modes of near extreme magnetic dilaton black holes in four-dimensional gravity [3].) Instead, RST impose the following boundary conditions at this curve, wherever it is timelike [4]:

$$\partial_\pm\Omega|_{\Omega=1/4}=0\,. \tag{6}$$

Incoming energy flux is then reflected off the boundary and the total outgoing flux (including the anomalous part t_-) can be determined by using (6). The boundary curve undergoes dynamical motion in response to the incoming matter and this gives rise to some nontrivial behavior in this model, in spite of the extreme simplicity of the field equations (5).

The boundary conditions (6) ensure semiclassical energy conservation and also that the physical curvature remains finite at the boundary curve as long as it is timelike. It should be noted, however, that these boundary conditions are not the most general ones allowed, and our results may depend qualitatively on this choice. Indeed (6) is incompatible with Dirichlet or Neumann boundary conditions on the f_i, and may not be realizable as the semiclassical limit of any fully quantum mechanical boundary conditions [8]. Alternate possibilities are currently being explored.

The solution corresponding to incoming matter energy flux, which tapers off at early and late times, but is otherwise quite general, is given by

$$\begin{aligned}\Omega(x^+,x^-)=&-x^+[\lambda^2x^-+P_+(x^+)]+(1/\lambda)M(x^+)\\&-(1/\lambda)M(x_B^+(x^-))\\&-\tfrac{1}{4}\ln[x^+/x_B^+(x^-)]\,,\end{aligned} \tag{7}$$

where

$$\begin{aligned}M(x^+)&=\lambda\int_0^{x^+}du\,uT^f_{++}(u)\,,\\P_+(x^+)&=\int_0^{x^+}du\,T^f_{++}(u)\,,\end{aligned} \tag{8}$$

and $x_B^+(x^-)$ is the x^+ value of the point on the boundary curve from which the reflected signal propagates to (x^+,x^-). Wherever the boundary curve is timelike its shape is a simple function of the incoming energy flux,

$$\lambda^2x_B^-=-P_+(x_B^+)-1/4x_B^+\,. \tag{9}$$

If, however, the incoming energy flux becomes larger than the outgoing Hawking flux of a two-dimensional black hole, $T^f_{++}(x^+)>1/4(x^+)^2$, for some value of x^+ then the boundary curve becomes spacelike and it is less trivial to determine its shape. [The black hole temperature is independent of mass in this model and an energy flux of $T^f_{++}(x^+)=1/4(x^+)^2$ in Kruskal coordinates corresponds to a uniform incoming flux, $T^f_{++}(\sigma^+)=\lambda^2/4$, in the coordinate, $\lambda\sigma^+=\ln\lambda x^+$, appropriate to asymptotic inertial observers.]

Spacelike segments of the boundary curve are curvature singularities. They form inside regions of future trapped points which are bounded by an apparent horizon, located where $\partial_+\Omega=0$ [9]. Applying ∂_+ to (7) one finds that the apparent horizon curve (x_H^+,x_H^-) satisfies (9) for all values of x_H^+. In other words, the apparent horizon coincides with the boundary curve where the latter is timelike, but where the boundary becomes spacelike the two curves separate and the apparent horizon cloaks the singularity, as shown in Fig. 1. Once the incoming energy flux falls below the threshold value the black hole evaporates and the apparent horizon approaches the singularity. The curves join again at the end point of the evaporation, which is denoted by E in Fig. 1. The null line segment $x^-=x_E^-$, $x_B^+(x_E^-)<x^+<x_E^+$, is the global event horizon of the geometry. We define the black hole mass to be the total energy (as measured at $\mathcal{I}^-$) which enters the black hole

$$M_{\rm BH}=M(x_E^+)-M(x_B^+(x_E^-))\,. \tag{10}$$

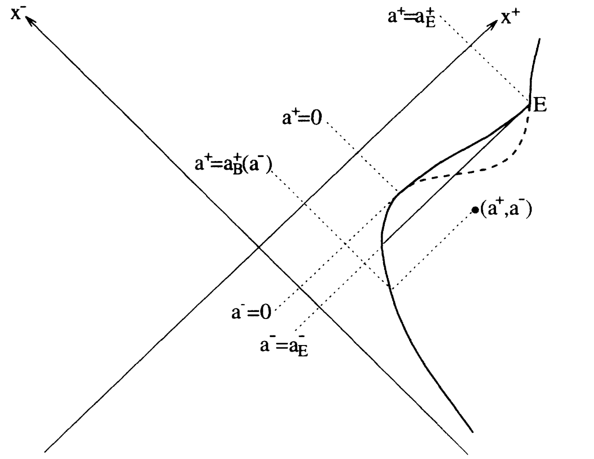

FIG. 1 Kruskal diagram for black hole formation and evaporation in the scaling regime. A near-critical flux of matter energy is incident from $x^- = -\infty$. The solid curve is the $\Omega = \frac{1}{4}$ boundary and the dashed curve is the apparent horizon. The spacelike portion of the solid curve is the black hole singularity.

Since both ends of the global horizon are on the boundary curve and also on the apparent horizon it immediately follows from (7) and (9) that

$$M_{\rm BH} = \frac{\lambda}{4} \ln[x_E^+ / x_B^+(x_E^-)] \,. \tag{11}$$

Now consider black hole formation just above threshold. For concreteness assume that the incoming energy flux has a maximum at $\lambda\sigma^+ = \ln\lambda x^+ = 0$, where its value is $T^f_{++} = \lambda^2(\frac{1}{4} + \delta)$, and that the flux is below threshold everywhere except near $\sigma^+ = 0$ so that only a single small black hole is formed. In the scaling limit, $\delta \to 0$, a generic incoming flux of this type may be parametrized as follows near $\sigma^+ = 0$:

$$T^f_{++}(\sigma^+) = \lambda^2(\tfrac{1}{4} + \delta)(1 - \alpha\lambda^2\sigma^{+2} + \cdots) \,. \tag{12}$$

It turns out that the higher terms in the Taylor expansion of T^f_{++} do not contribute in the scaling limit.

To obtain the shape of the apparent horizon curve, we transform to Kruskal coordinates, compute $P_+(x^+)$ for this flux distribution, and insert the result into (9). We are interested in the $\delta \ll 1$ limit and for that purpose it is convenient to shift and rescale the Kruskal coordinates as follows:

$$\begin{aligned} \lambda x^+ &= 1 + \sqrt{\delta/\alpha}\,(a^+ - 2) \,, \\ \lambda x^- &= -(1/\lambda)P_+(1/\lambda) - \tfrac{1}{4} + \sqrt{\delta^3/\alpha}\,(a^- + \tfrac{4}{3}) \,. \end{aligned} \tag{13}$$

The origin of the (a^+, a^-) coordinate system has been chosen where, to leading order in δ, the boundary curve turns spacelike, as shown in Fig. 1. In the scaling region where higher order terms in δ can be dropped, the apparent horizon takes a simple and universal form,

$$a_H^-(a_H^+) = -\tfrac{1}{2} a_H^{+2} + \tfrac{1}{12} a_H^{+3} \,. \tag{14}$$

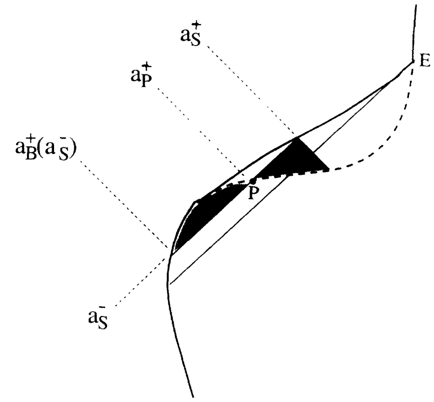

FIG. 2. The two shaded regions have equal area for any point (a_S^+, a_S^-) on the singularity curve.

The only reference made to the parameter α is in the definition of the scaling variables (13). Higher orders in the Taylor expansion of $T^f_{++}(\sigma^+)$ in (12) only contribute terms carrying positive powers of δ, which can be ignored in the scaling region.

It is straightforward to show that the singularity curve is also universal in this limit. Expressing (7) in terms of (a^+, a^-) coordinates and applying d/da^+ to both sides yields the following differential equation for the $\Omega = \frac{1}{4}$ curve:

$$[a_S^+ - a_B^+(a_S^-)] \frac{da_S^-}{da_S^+} = -a_S^- + a_H^-(a_S^+) \,. \tag{15}$$

This is a nonlinear differential equation and we do not have an analytic solution, but there is no explicit dependence on the flux parameters so the shape of the curve must be universal. However, we do not need the whole singularity curve in order to determine the black hole mass (11). It is sufficient to locate the end point of evaporation and this can be achieved by the following geometric argument, which is illustrated in Fig. 2.

Take any point on the singularity curve and consider the past-directed null line $a^- = a_S^-$ which extends from

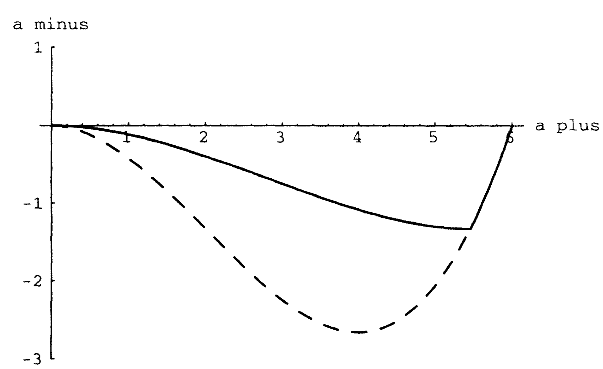

FIG. 3. Singularity curve obtained from numerical calculation. The dashed curve is the apparent horizon (14).

$a^+ = a_S^+$ to the timelike boundary curve at $a^+ = a_B^+ \times (a_S^-)$. Denote by P the point where this null line intersects the apparent horizon. The value of $\Omega(a_P^+, a_P^-)$ can be obtained by integrating $\partial_+ \Omega$ along this null line starting either from the spacelike singularity at a_S^+ or from the timelike boundary curve at $a_B^+(a_S^-)$. The value of $\partial_+ \Omega$ on $a^- = a_S^-$ can in turn be obtained by integrating $\partial_- \partial_+ \Omega$ along null lines of constant a^+ starting from the apparent horizon curve (where $\partial_+ \Omega$ vanishes by definition). By using the equation of motion, $\partial_- \partial_+ \Omega = -\lambda^2$, inside the double integral one thus finds that $\Omega(a_P^+, a_P^-) - \frac{1}{4}$ equals the area of each of the shaded regions in Fig. 2 and therefore these areas must be equal. Interestingly, this "equal area rule" is an exact relationship which also holds outside of the scaling limit.

The equal area rule, when applied at the global event horizon, $a^- = a_E^-$, gives $a_E^+ - 2 = 2 - a_B^+(a_E^-)$. Both end points of the global event horizon are also points on the apparent horizon curve (14) so we have $-\frac{1}{2} a_E^{+2} + \frac{1}{12} a_E^{+3} = -\frac{1}{2} a_B^+(a_E^-)^2 + \frac{1}{12} a_B^+(a_E^-)^3$. From these two relations it follows that $a_E^+ = 2 + \sqrt{12}$, $a_B^+(a_E^-) = 2 - \sqrt{12}$, and the black hole mass (11) is to leading order seen to be

$$M_{\mathrm{BH}}/\lambda = \sqrt{3\delta/\alpha}\,, \tag{16}$$

which is the scaling relation (2).

Finally, it is straightforward to integrate (15) numerically to obtain the entire singularity curve. The result is plotted in Fig. 3.

This work was supported in part by DOE Grant No. DOE-91ER40618 and NSF Grant No. PHY-89-04035.

*Electronic address: andy@denali.physics.ucsb.edu
†Electronic address: larus@nsfitp.itp.ucsb.edu

[1] M. W. Choptuik, Phys. Rev. Lett. **70**, 9 (1993).
[2] D. Christodolou, Commun. Math. Phys. **105**, 337 (1986); **106**, 587 (1986); **109**, 613 (1987); A. M. Abrahams and C. R. Evans, Phys. Rev. Lett. **70**, 2980 (1993).
[3] For reviews see J. A. Harvey and A. Strominger, in *Proceedings of the TASI Summer School, 3–28 June 1992, Boulder, Colorado* (World Scientific, Singapore, 1993); S. B. Giddings, in *Proceedings of the International Workshop on Theoretical Physics, 6th Session, 21–28 June, 1992, Erice, Italy* (World Scientific, Singapore, 1993).
[4] J. G. Russo, L. Susskind, and L. Thorlacius, Phys. Rev. D **46**, 3444 (1992); D **47**, 533 (1993).
[5] A. Strominger, Phys. Rev. D **48**, 5769 (1993).
[6] C. G. Callan, S. B. Giddings, J. A. Harvey, and A. Strominger, Phys. Rev. D **45**, R1005 (1992).
[7] D. Lowe, Phys. Rev. D **47**, 2446 (1993); T. Piran and A. Strominger, Phys. Rev. D **48**, 4729 (1993).
[8] A. Strominger and S. Trivedi (unpublished).
[9] J. G. Russo, L. Susskind, and L. Thorlacius, Phys. Lett. B **292**, 13 (1992).

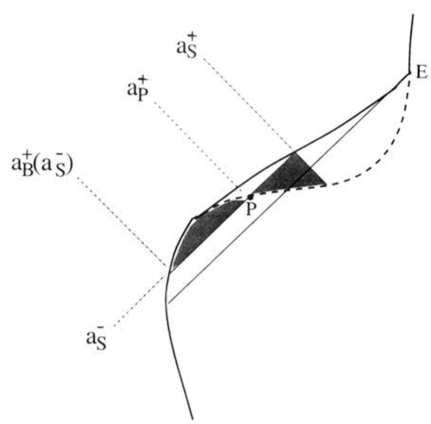

FIG. 2. The two shaded regions have equal area for any point (a_S^+, a_S^-) on the singularity curve.