Understanding critical collapse of a scalar field

Carsten Gundlach*

LAEFF-INTA (Laboratorio de Astrofísica Espacial y Física Fundamental – Instituto Nacional de Tecnología Aeroespacial),
P.O. Box 50727, 28080 Madrid, Spain

I construct a spherically symmetric solution for a massless real scalar field minimally coupled to general relativity which is discretely self-similar (DSS) and regular. This solution coincides with the intermediate attractor found by Choptuik in critical gravitational collapse. The echoing period is $\Delta = 3.4453 \pm 0.0005$. The solution is continued to the future self-similarity horizon, which is also the future light cone of a naked singularity. The scalar field and metric are C^1 but not C^2 at this Cauchy horizon. The curvature is finite nevertheless, and the horizon carries regular null data. These are very nearly flat. The solution has exactly one growing perturbation mode, thus confirming the standard explanation for universality. The growth of this mode corresponds to a critical exponent of $\gamma = 0.374 \pm 0.001$, in agreement with the best experimental value. I predict that in critical collapse dominated by a DSS critical solution, the scaling of the black hole mass shows a periodic wiggle, which like γ is universal. My results carry over to the free complex scalar field. Connections with previous investigations of self-similar scalar field solutions are discussed, as well as an interpretation of Δ and γ as anomalous dimensions. [S0556-2821(97)03402-4]

PACS number(s): 04.25.Dm, 04.20.Dw, 04.40.Nr, 04.70.Bw

I. INTRODUCTION

A. Critical phenomena in gravitational collapse

In an astrophysical context, gravitational collapse normally begins with a star. This means that the initial data are almost stationary, and that they have a characteristic scale which is provided by the matter. Therefore astrophysical black holes have a minimum mass, namely, the Chandrasekhar mass. Abandoning the restriction to almost stationary initial data, or alternatively to realistic matter, one should be able to make arbitrarily small black holes. One may then ask what happens if one tries to make a black hole of infinitesimal mass by fine-tuning the initial data.

The investigation is simplified by choosing a matter model that does not admit stable stationary solutions. Then, for *any* initial data, there are only two possible outcomes: formation of a black hole, or dispersion leaving behind flat spacetime. The first systematic numerical examination of the limit between the two (the ''critical surface'' in phase space) was carried out by Choptuik [1] for a massless minimally coupled real scalar field in spherical symmetry. He evolved members of various one-parameter families of initial data each of which comprised both collapsing and dispersing data, and searched for the critical parameter value by bisection. For all families he investigated he was able to make arbitrarily small black holes by tuning the parameter p of the data: there was no evidence for a ''mass gap.'' Instead he found two unexpected new phenomena.

For marginal data, both supercritical and subcritical, the time evolution approaches a certain universal solution which is the same for all one-parameter families of data. This solution is an ''intermediate attractor'' in the sense that the time evolution first converges onto, but eventually diverges from it, to either form a black hole or to disperse. This universal solution (also called the ''critical solution'') has a curious symmetry: it is periodic in the logarithm of spacetime scale, with a period of $\Delta \simeq 3.44$. This is also referred to as ''echoing,'' or discrete self-similarity (DSS).

Moreover, for marginally *super*critical data, the final black hole mass scales as $M \sim (p - p_*)^\gamma$, where p is the parameter of the family of initial data, and p_* its critical value. The ''critical exponent'' γ has the value $\simeq 0.37$ for the scalar field, and like the critical solution it is universal in the sense that it is the same for all one-parameter families of data.

Both phenomena were then also found in the axisymmetric collapse of pure gravity [2], indicating that they are an artifact of neither the choice of matter nor of spherical symmetry. There, Δ was found to be $\simeq 0.6$, and the critical exponent γ to be $\simeq 0.36$. For a perfect fluid with equation of state $p = \rho/3$ in spherical symmetry [3], the universal attractor has a different symmetry: it is not discretely, but continuously self-similar (CSS). γ is found to be $\simeq 0.36$ once more.

Choptuik's results for the scalar field have been duplicated, to varying precision, in [4–6].

Subsequently, the matter models were generalized. For a fluid with $p = k\rho$ in spherical symmetry, now with arbitrary constant k, γ was found to be strongly k-dependent [7,8]. The real scalar field model was generalized to a one-parameter family of two-component nonlinear σ models [9]. This family includes the cases of a free complex scalar field [10,11], a real scalar field coupled to Brans-Dicke gravity [12] and, as a special case of the latter, a string-inspired axion-dilaton model [13]. For the axion-dilaton model, $\gamma \simeq 0.264$ is found in collapse simulations [14]. From these new examples it is clear that γ is not universal with respect to different kinds of matter, but only with respect to the initial data for any one matter model.

*Current address: Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Schlaatzweg 1, D-14473 Potsdam, Germany. Electronic address: gundlach@aei-potsdam.mpg.de

TABLE I. A comparison between the Kerr-Newman and critical collapse solutions.

Kerr-Newman solutions	Critical collapse solutions
Length scale M	Scale-invariant
Stationary	Self-similar
Horizon	Naked singularity
$\Rightarrow$ vacuum	$\Rightarrow$ matter dependent
Quasinormal modes $e^{i\omega_i t/M}$	Perturbations t^{λ_i}
Attractor	Attractor of codimension one
$\Rightarrow$ ''no hair''	$\Rightarrow$ universality

B. The emerging picture

Let us now examine some general features of the ''critical solutions'' which appear as intermediate attractors in collapse, and which seem to describe the limiting case of the formation of a zero mass black hole. First of all, they must be scale invariant in some way, and in fact will show homotheticity, or ''self-similarity of the first kind'' [15]. Because of self-similarity, they must have a curvature singularity, but they should not have an event horizon. The absence of a horizon means that the collapsing matter always remains visible in the solution.

The unique end point of gravitational collapse is given by the Kerr-Newman family of solutions because, roughly speaking, they only admit damped perturbations: they are attractors in phase space. If the critical collapse solutions were also attractors, we would see many naked singularities in nature, and this is clearly not the case. In fact they are attractors of codimension one, and we shall see that this gives rise to an analogue of the ''no hair'' theorem: universality. The Kerr-Newman and critical collapse solutions are briefly contrasted in Table I.

In a schematic picture of phase space [16], the critical solution remains within the critical surface. The observed universality suggests that it is an attractor within the critical surface, and in consequence an attractor of codimension one in phase space. This attractor could either be a fixed point, corresponding to CSS, or a limit cycle, corresponding to DSS. Figure 1 illustrates the DSS case.

Nearly critical Cauchy data are situated close to the critical surface, but may be far from the critical point. Under time evolution they are attracted towards the critical point. While they approach it, their ''distance'' from the critical surface increases exponentially but remains small until they are close to the critical point because, by the assumption of near-criticality, it is initially very small. Near the critical point, the exponential increase takes over, and the phase space trajectories are repelled from the critical surface, all into the same direction. This constitutes the mechanism of universality with respect to initial data. The formula for the black hole mass follows essentially by dimensional analysis, and the critical exponent γ can be read off from the linear perturbations of the critical solution [16]. As we shall see in Sec. IVB, the periodicity of the critical solution in the DSS case gives rise to a periodic wiggle in the scaling law.

What confuses the naive phase space picture is gauge invariance: the same spacetime corresponds to very different trajectories in superspace, depending on how it is sliced. The
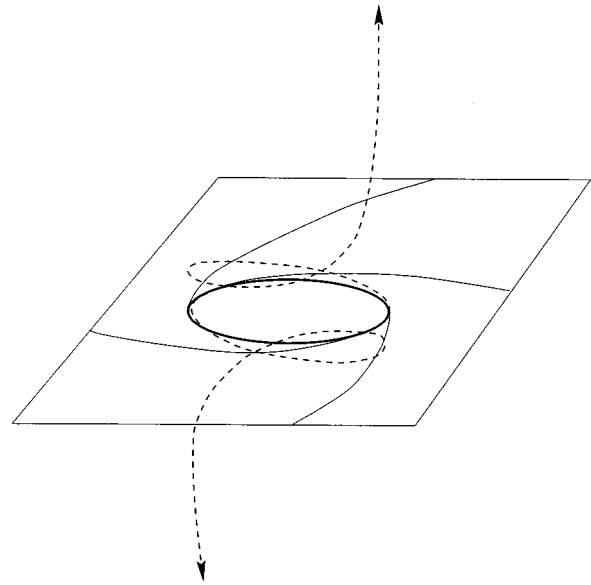

FIG. 1. The phase space picture for discrete self-similarity. The plane represents the critical surface. (In reality this is a hypersurface of codimension one in an infinite-dimensional space.) The circle (fat unbroken line) is the limit cycle representing the critical solution. The thin unbroken curves are spacetimes attracted to it. The dashed curves are spacetimes repelled from it. There are two families of such curves, labeled by one periodic parameter τ, one forming a black hole, the other dispersing to infinity. Only one member of each family is shown.

naive picture must therefore be used with care pending its formulation in geometric terms.

CSS solutions and their linear perturbations have already been calculated for various matter models. For the $p = \rho/3$ model [3,16] and for the axion-dilaton model [13,14] they agree with the critical solution found in collapse simulations, and give the correct critical exponent. For the $p = k\rho$ model with $k \neq 1/3$, γ has been calculated on the basis of the perturbations of a CSS solution [7,8]. No CSS solution exists for $k \gtrsim 0.888$ [7], but the CSS solution appears to be the critical solution right up to that limit [8].

For the complex scalar field a CSS solution has been found [10], but it is only an attractor of codimension three [11]. The CSS solution and its perturbations have been calculated also for the family of nonlinear σ models [9]. The number of unstable modes of the CSS solution changes from one to three at some (known) value of the parameter. This probably indicates the changeover from CSS to DSS in the critical solution. Why some critical solutions are CSS and others DSS is not yet understood, however.

The present paper gives the first calculation of a DSS critical solution, together with its maximal extension and its linear perturbations. This is technically much more difficult than CSS, but DSS is the most generic case of self-similarity, and the mathematical and numerical methods developed here will be useful in other applications. The critical mass scaling generalizes to contain a universal periodic wiggle.

The plan of the paper is as follows. In Sec. II, I define DSS and construct the DSS solution of the real scalar field model in the past light cone of the singularity as a nonlinear eigenvalue problem. It agrees with the critical solution found

by Choptuik [1]. This section, together with Appendixes A, C, D, and F, is an expansion of [17]. Sections III and IV both build on Sec. II, but are independent of each other.

In Sec. III, I extend the critical solution up to the future light cone of the naked singularity. I find that this Cauchy horizon is regular, and that it is in fact nearly flat. In Sec. IV, I calculate the linear perturbations, and generalize the calculation of the critical exponent to the DSS case. My value for γ agrees with the experimental one, but I also predict the existence of a (small) universal wiggle overlaid on the power-law scaling. I show that the Choptuik solution is the critical solution not only for the real, but also the complex scalar field. In Sec. V, I put my results into the context of results for other collapsing systems on one hand, and of the study of (continuously) self-similar scalar field models on the other. I then discuss the next steps to be taken, and the possible connection with critical phenomena in statistical mechanics and quantum field theory. Various details are given in the Appendixes.

II. THE CRITICAL SOLUTION

A. Field equations

In this section I construct an isolated solution of general relativity minimally coupled to a massless real scalar field with the following properties: (1) spherical symmetry, (2) discrete self-similarity (to be defined below), (3) analytic at the center of spherical symmetry, (4) analytic at the past self-similarity horizon, (5) the scalar field is bounded. [It is likely that there is no solution obeying (1)–(4) that does not also obey (5), but I have not shown this.]

The Einstein equations we consider here are

$$G_{ab} = 8\pi G\left(\phi_{,a}\phi_{,b} - \frac{1}{2}g_{ab}\phi_{,c}\phi^{,c}\right), \qquad (1)$$

in spherical symmetry. The matter equation $\phi_{,c}{}^{;c} = 0$ follows from the Einstein equations as the contracted Bianchi identity. The spacetime metric is

$$ds^2 \equiv -\alpha(r,t)^2 dt^2 + a(r,t)^2 dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \qquad (2)$$

This form of the metric is invariant under transformations $t \to \tilde{t}(t)$, $\alpha \to \tilde{\alpha}$, such that $\alpha dt = \tilde{\alpha} d\tilde{t}$. In order to write the matter equations in first-order form, we introduce the auxiliary matter fields

$$X(r,t) \equiv \sqrt{2\pi G}\frac{r}{a}\phi_{,r}, \quad Y(r,t) \equiv \sqrt{2\pi G}\frac{r}{\alpha}\phi_{,t}. \qquad (3)$$

The combinations $X_\pm = X \pm Y$ of these fields propagate along characteristics. The radial null geodesics, which are also the matter characteristics, are characterized by the quantity $g \equiv a/\alpha$.[1] The scalar wave equation in $X_\pm$ is then

---

[1] I have made one change of notation. In [17] I defined $g \equiv \exp(\xi_0)a/\alpha$, while $g \equiv a/\alpha$ here. This is for greater convenience in the remainder of the paper.

$$r(X_{\pm,r} \mp gX_{\pm,t}) = \left[\pm r\frac{a_{,t}}{\alpha} - r\frac{\alpha_{,r}}{\alpha}\right]X_\pm - X_\mp. \qquad (4)$$

In the following we use X_+, X_-, g, and a as our basic variables. A complete set of Einstein equations in these variables is

$$rg_{,r} = (1-a^2)g, \qquad (5)$$

$$ra_{,r} = \frac{1}{2}a[(1-a^2) + a^2(X_+^2 + X_-^2)] \equiv C_1, \qquad (6)$$

$$gra_{,t} = \frac{1}{2}a^3(X_+^2 - X_-^2) \equiv C_2, \qquad (7)$$

and the matter equations become

$$r(X_{\pm,r} \mp gX_{\pm,t}) = \left[\frac{1}{2}(1-a^2) - a^2X_\mp^2\right]X_\pm - X_\mp \equiv C_\pm \qquad (8)$$

when we eliminate the metric derivatives with the help of the Einstein equations. The five first-order equations (5)–(8) are our field equations. I have defined the expressions $C_\pm$, C_1, and C_2 for later use. The absence of $g_{,t}$ in the equations reflects the fact that α, and hence g, contains a gauge degree of freedom not determined by the Cauchy data.

The two scalars made from the Ricci curvature, using the Einstein equations, are

$$R = 4r^{-2}X_+X_-, \quad R_{ab}R^{ab} = R^2. \qquad (9)$$

The Riemann tensor will be considered in Appendix G.

B. Discrete self-similarity

The concept of (continuous) self-similarity (CSS) (or homotheticity) has been defined in a relativistic context [18,15] as the presence of a vector field χ such that

$$\mathcal{L}_\chi g_{ab} = 2g_{ab}, \qquad (10)$$

where $\mathcal{L}_\chi$ denotes the Lie derivative. I now introduce the concept of discrete self-similarity (DSS). In this symmetry there exist a diffeomorphism ϕ and a real constant Δ such that, for any integer n,

$$(\phi_*)^n g_{ab} = e^{2n\Delta}g_{ab}, \qquad (11)$$

where ϕ_* is the pullback of ϕ.

To see what DSS looks like in coordinate terms, we introduce coordinates (σ, x^α), such that if a point p has coordinates (σ, x^α), its image $\phi(p)$ has coordinates $(\sigma + \Delta, x^\alpha)$. One can verify that DSS in these coordinates is equivalent to

$$g_{\mu\nu}(\sigma, x^\alpha) = e^{2\sigma}\tilde{g}_{\mu\nu}(\sigma, x^\alpha),$$

where

$$\tilde{g}_{\mu\nu}(\sigma, x^\alpha) = \tilde{g}_{\mu\nu}(\sigma + \Delta, x^\alpha). \qquad (12)$$

In other words, the DSS acts as a discrete isomorphism on the rescaled metric $\tilde{g}_{\mu\nu}$. σ is intuitively speaking the logarithm of spacetime scale.

One can formally construct such a coordinate system in the following way: Fix a hypersurface Σ such that its image Σ' under ϕ does not intersect Σ. Introduce coordinates x^α in Σ, and copy them to Σ' with ϕ. Introduce coordinates (σ, x^α) in the region between Σ and Σ' such that σ has range $[0, \Delta]$, their restriction to Σ is $(0, x^\alpha)$ and their restriction to Σ' is (Δ, x^α). Finally, copy these coordinates to the entire spacetime, such that if p has coordinates (σ, x^α), its nth image $\phi^n(p)$ is assigned coordinates $(\sigma + n\Delta, x^\alpha)$. Clearly there is enormous freedom in defining such coordinates.

In order to clarify the connection between CSS and DSS, one may define a vector field $\chi \equiv \partial/\partial\sigma$, although there is no unique χ associated with a given ϕ. The discrete diffeomorphism ϕ is then realized as the Lie dragging along χ by a distance Δ. Clearly, CSS corresponds to DSS for infinitesimally small Δ, and hence for all Δ, and is in this sense a degenerate case of DSS. In this limit, χ becomes unique. In the following I speak of DSS only in the absence of CSS.

In order to see what form DSS takes in spherical symmetry in the particular coordinates defined in Eq. (2), we make a coordinate transformation $t \equiv e^\sigma T(\sigma, z)$ and $r \equiv e^\sigma R(\sigma, z)$, where T and R are periodic in σ with period Δ. Here σ is the same as in the general construction, and $x^\alpha = (z, \theta, \varphi)$. In the new coordinates the metric (2) takes the form

$$ds^2 = e^{2\sigma}\{-\alpha^2[(T+T_{,\sigma})d\sigma + T_{,z}dz]^2$$
$$+ a^2[(R+R_{,\sigma})d\sigma + R_{,z}dz]^2 + R^2(d\theta^2 + \sin^2\theta d\varphi^2)\}. \tag{13}$$

This is of the form (12) if and only if $\alpha(z, \sigma)$ and $a(z, \sigma)$ are also periodic in σ with period Δ. In terms of t and r this periodicity corresponds to

$$a(r, t) = a(e^{n\Delta}r, e^{n\Delta}t), \quad \alpha(r, t) = \alpha(e^{n\Delta}r, e^{n\Delta}t). \tag{14}$$

This is not yet the most general way to impose DSS in Eq. (2). We obtain that by also admitting the transformations $t \to \bar{t}(t)$, with $\alpha \to \bar{\alpha} = dt/d\bar{t}\,\alpha$. $\bar{\alpha}$ need no longer be periodic, but it must be related in this way to some α that is.

C. Formulation as an eigenvalue problem

We now introduce specific coordinates of the kind just discussed. The following choice will be sufficiently general for our purpose:

$$\tau \equiv \ln\left(\frac{t}{r_0}\right), \quad \zeta \equiv \ln\left(\frac{r}{t}\right) - \xi_0(\tau). \tag{15}$$

Here r_0 is an arbitrary fixed scale, and $\xi_0(\tau)$ is a periodic function with period Δ.[2] Both are to be determined later. In the new coordinates, the matter and Einstein equations are

$$X_{\pm,\zeta} = \frac{C_\pm \pm e^{\zeta + \xi_0}g X_{\pm,\tau}}{1 \pm (1 + \xi_0')e^{\zeta + \xi_0}g}, \tag{16}$$

$$g_{,\zeta} = (1 - a^2)g, \tag{17}$$

$$a_{,\zeta} = C_1, \tag{18}$$

$$a_{,\tau} = e^{-(\zeta + \xi_0)}g^{-1}C_2 + (1 + \xi_0')C_1. \tag{19}$$

Here $\xi_0' = d\xi_0(\tau)/d\tau$, and in this paper a prime always denotes the derivative of a function of one variable with respect to its formal argument. The equations are invariant under a translation in τ, corresponding to a change in the arbitrary scale r_0.

In order to impose discrete homotheticity, we demand the periodicity (14). We also impose the regularity condition $a = 1$ at $r = 0$ and the gauge condition $\alpha = 1$ at $r = 0$. (Both are compatible with the periodicity.) In our choice of dependent and independent variables we therefore impose boundary conditions

$$a(\zeta, \tau + n\Delta) = a(\zeta, \tau), \quad g(\zeta, \tau + n\Delta) = \alpha(\zeta, \tau), \tag{20}$$

and

$$a(\zeta = -\infty, \tau) = 1, \quad g(\zeta = -\infty, \tau) = 1. \tag{21}$$

Note that we continue to describe the metric with the variables a and α, but that they are not the coefficients of the metric associated with coordinates (τ, ζ).

From the Einstein equations it follows that the periodicity condition must hold also for X_+ and X_-. From the equations defining X and Y, we obtain

$$\phi_{,\tau} = (2\pi G)^{-1/2}a[(1 + \xi_0')X + g^{-1}e^{-(\zeta + \xi_0)}Y],$$

$$\phi_{,\zeta} = (2\pi G)^{-1/2}aX. \tag{22}$$

Because the right-hand sides of both equations are periodic in τ, the scalar field ϕ itself is of the form

$$\phi(\zeta, \tau + \Delta) = \text{(periodic in } \tau\text{)} + \kappa\tau, \tag{23}$$

where κ is a constant. κ is not an independent parameter, but is determined through the first of Eqs. (22). ϕ is bounded and periodic if and only if $\kappa = 0$.

As we have a $(1+1)$-dimensional hyperbolic problem, we can interchange space and time. In this view, near $r = 0$, Eqs. (16), (17), and (18) form a first-order system of evolution equations for a, g, X_+, and X_- with "time" coordinate ζ and (periodic) "space" coordinate τ. The data are subject to one constraint (19), which is propagated by the evolution equations. As $\zeta \to -\infty$, we impose boundary conditions corresponding to a regular center $r = 0$.

The equations become singular where the denominator of Eq. (16) vanishes. The treatment of this singular surface, both analytical and numerical, is simplified if we use the coordinate freedom incorporated in the choice of $\xi_0(\tau)$ to make this happen "for all τ at once," namely on the line $\zeta = 0$. We therefore impose the coordinate condition

$$[1 - (1 + \xi_0')e^{\xi_0}g]_{\zeta=0} = 0. \tag{24}$$

---

[2] As in [17] I assume $t > 0$. In a collapse context, where the spacetime region we are about to calculate is to the past of the singularity, t then decreases to the future, but this is purely a matter of convention. The convention can be reversed in the results simply by changing the sign of Y.

A necessary condition for regularity is now that the numerator also vanishes at $\zeta=0$, or

$$[C_- - e^{\xi_0}gX_{-,\tau}]_{\zeta=0}=0. \tag{25}$$

We may look at this from a different angle: The coordinate condition means that $\zeta=0$ is null, and hence a characteristic of the equations. Such a characteristic which is mapped onto itself by the self-similarity map ϕ is called a "self-similarity horizon" (SSH), or a "sonic line" [19]. For our equations, with periodic boundary conditions in the "space" coordinate τ, it constitutes a Cauchy horizon. On any other $\zeta=$const surface, X_+ and X_- would be free Cauchy data, but on $\zeta=0$ they are overdetermined, as one requires only characteristic data. This gives again rise to the condition (25).

We still have to see what kind of regularity this condition enforces, and if it is sufficient as well as necessary. To investigate the possible behavior of the solution at the past SSH $\zeta=0$, we assume for the moment that a, g, and X_+ are at least C^0 there, and expand the equation for X_- to leading order in ζ. The resulting approximate equation, using the coordinate condition (24) and Eq. (17), is

$$X_{-,\zeta}=\frac{A(\tau)X_- - X_{-,\tau}+C(\tau)}{\zeta B(\tau)}, \tag{26}$$

where the coefficients

$$A=(1+\xi_0')\left[\frac{1}{2}(1-a_0^2)-a_0^2X_{+0}^2\right], \quad B=-(1+\xi_0')(1-a_0^2),$$

$$C=-(1+\xi_0')X_{+0} \tag{27}$$

are evaluated at $\zeta=0$. This approximate equation admits an exact general solution: namely,

$$X_-=X_-^{\mathrm{inhom}}(\tau)+X_-^{\mathrm{hom}}(\zeta,\tau). \tag{28}$$

The particular inhomogeneous solution X_-^{inhom} is defined as the unique solution of the ordinary differential equation (ODE)

$$AX_-^{\mathrm{inhom}}-X_{-,\tau}^{\mathrm{inhom}}+C=0 \tag{29}$$

with periodic boundary conditions. This solution exists and is unique, unless the average value of A vanishes. The general homogeneous solution X_-^{hom} is of the form

$$X_-^{\mathrm{hom}}=\zeta^{\hat{A}_0/\hat{B}_0}e^{I_0A-(\hat{A}_0/\hat{B}_0)I_0B}F\left[\tau+\frac{I_0B-\ln|\zeta|}{\hat{B}_0}\right]. \tag{30}$$

The notation here is that of Appendix C: $\hat{A}_0$ is the constant part, or average value, of the periodic function A. I_0A denotes the principal function of $A-\hat{A}_0$, that is of A minus its average, with the integration constant defined so that I_0A itself has zero average. F is a periodic function of one variable with period Δ. It is the free parameter of the solution. We do not need to determine it here. What is important is that the solution is analytic if and only if F vanishes identically. If F does not vanish, there are two possibilities. If $\hat{A}_0/\hat{B}_0<0$, the solution blows up at $\zeta=0$. If $\hat{A}_0/\hat{B}_0>0$, the solution is C_0 but not C_1 there. It is easy to see that $\hat{A}_0<0$

and $\hat{B}_0>0$ in our case, so that the solution either blows up or, for $F\equiv0$, is analytic. So with the one condition (25) we automatically enforce analyticity.

We shall impose one extra symmetry on our ansatz. The results of numerical collapse simulations [1] indicate that ϕ itself is periodic. Moreover, even if one adds a potential to the scalar field action, the attractor is found to be the same as for the massless field [20]. This requires ϕ to remain bounded in the critical solution, because only then can a polynomial of ϕ be neglected with respect to the space and time derivatives of ϕ, which are unbounded because of the echoing on an exponentially decreasing spacetime scale [20]. For this reason we look for a solution with $\kappa=0$. κ is the average value (averaged over one period in τ, for any ζ) of $\phi_{,\tau}$, as defined in the first of Eqs. (22). This average may vanish by a cancellation on the right-hand side of that equation, but this appears unlikely. If the even frequencies of $X_\pm$ and the odd frequencies of a and g vanish for all ζ, however, κ vanishes automatically, without any unexplained cancellation. In our ansatz we impose this as an assumption: a and g, as well as ξ_0, are composed only of even frequency terms in τ, and $X_\pm$ only of odd frequency terms. This reflection-type symmetry is compatible with the field equations. Imposing it is justified *a posteriori*: a solution with the symmetry exists (and is the same as the one found by Choptuik).

We have now completed the formulation of a hyperbolic boundary value problem. Its Cauchy data are values for the four fields a, g, X_+, and X_-, up to one constraint, and up to a translation invariance in τ, plus the unknown value of Δ and the unknown function $\xi_0(\tau)$. The count is therefore $4N-N-1+1+N=4N$. (Here N stands for the countably infinite number of degrees of freedom of a periodic function of one variable.) These free data are balanced by two boundary conditions at $\zeta=-\infty$ and two at $\zeta=0$, or $4N$ degrees of freedom again. One would therefore expect this boundary value problem to have at most a discrete set of solutions.

Numerically I have constructed one such solution. Locally it is unique. To solve the boundary value problem numerically, I have expanded all periodic fields in Fourier components of τ, truncating the expansion at some relatively small number N of Fourier components. This takes advantage of the fact that the solution is smooth, so that the Fourier series converges rapidly. The field a is not evolved in ζ, but reconstructed at each step from the constraint (19). Details are given in the appendices. I find that $\Delta=3.4453\pm0.0005$. In a previous publication [17], I had given a value of $\Delta=3.4439\pm0.0004$. The difference, corresponding to 3.5 standard errors, is due to a change in numerical details of the algorithm. These changes correct what I now regard as an inconsistency in my original pseudo-Fourier method, and are therefore "systematic error." The quoted error is in both cases estimated discretization error, which is discussed in Appendix F. The present Fourier methods are outlined in Appendix C.

I have shown in [17] that the DSS solution I have constructed agrees with the intermediate attractor observed by Choptuik [1] to the numerical precision of the latter. The error in the DSS solution I have corrected here is small enough not to affect this agreement.

III. MAXIMAL EXTENSION OF THE CRITICAL SPACETIME

A. From the past self-similarity horizon to $t=0$

The coordinates (ζ, τ) become singular at $t=0$ ($\zeta=\infty$). Clearly it will be necessary to replace $\ln t$ with $\ln r$ as the periodic coordinate σ to regularize the equations there. Before we do this, it is useful to examine the asymptotic behavior of the solution in the old variables as $\zeta \to \infty$, or $t \to 0$. Neglecting terms of order $\exp(-\zeta)$, the field equations in this limit reduce to

$$(1+\xi_0')X_{\pm,\zeta}=X_{\pm,\tau}, \tag{31}$$

$$(1+\xi_0')a_{,\zeta}=a_{,\tau}, \tag{32}$$

$$a_{,\zeta}=C_1, \tag{33}$$

$$g_{,\zeta}=(1-a^2)g, \tag{34}$$

with the general solution

$$X_\pm=X_{\pm 0}(\rho), \tag{35}$$

$$a=a_0(\rho), \tag{36}$$

$$g=g_0(\tau)e^{\sigma(\rho)+\nu\zeta}, \tag{37}$$

where

$$\rho\equiv\zeta+\tau+\xi_0(\tau)\equiv\ln\left(\frac{r}{r_0}\right), \tag{38}$$

and the periodic functions $X_{\pm 0}$ and g_0 are free parameters of the general solution. The periodic function a_0 is uniquely determined by $X_{\pm 0}$ through the ODE

$$a_0'=\frac{1}{2}a_0[(1-a_0^2)+a_0^2(X_{+0}^2+X_{-0}^2)], \tag{39}$$

and the periodic function σ and constant ν are determined from a_0 by integration, as

$$\sigma=I_0(1-a_0^2), \quad \nu=\widehat{(1-a_0^2)}_0, \tag{40}$$

where the notation is that of Appendix C.

From this asymptotic expansion we see that a and $X_\pm$ are regular at $t=0$, while g is not and will have to be replaced by another dependent variable.

We replace ζ and τ by ρ as above and w given by

$$w\equiv\exp[(1+n)(\tau-\rho)+f(\tau)+h(\rho)], \tag{41}$$

where n is a constant and $h(\rho)$ and $f(\tau)$ are periodic functions of one variable, all yet to be determined. Among the dependent variables, we replace g by

$$\bar{g}\equiv[1+n+f'(\tau)]e^{n(\tau-\rho)+f(\tau)+h(\rho)}g. \tag{42}$$

We now have DSS if the dependent variables a, $\bar{g}$, X_+, and X_- are periodic in ρ at constant w, with the same period Δ we have determined previously. Figure 2 summarizes the different coordinate systems we use in this paper: (ζ, τ),
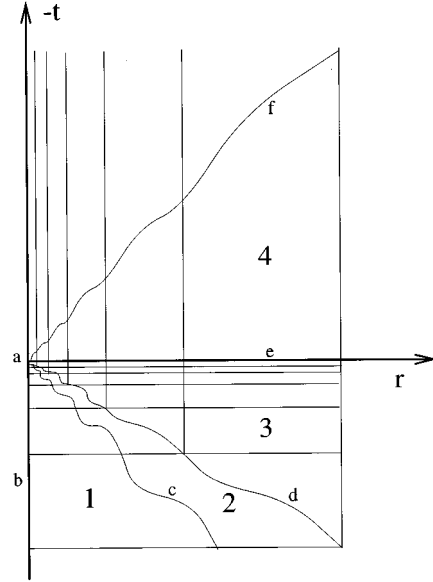

FIG. 2. Schematic diagram of the coordinate systems I use. The horizontal and vertical axes are r and $-t$. Horizontal accumulating lines are $\tau=0$, $\tau=\Delta$, $\tau=2\Delta$, etc., from bottom up. Vertical accumulating lines are $\rho=0$, $\rho=\Delta$, $\rho=2\Delta$, etc., from right to left. (a) The curvature singularity ($r=0, t=0$). (b) The regular center ($r=0, t>0$). (c) The past self-similarity horizon $\zeta=0$. (d) The matching line $\zeta=\zeta_0 \Leftrightarrow w=1$. (e) The matching line $w=0 \Leftrightarrow t=0$. (f) The future self-similarity horizon $w=-1$. I use four separate coordinate patches: (1) The nonlinear eigenvalue problem for Δ. (2) Continuation to ζ_0. (3) Continuation to $t=0$. (4) Continuation to the future SSH.

(w, ρ), and a second version of (w, ρ), with different functions h and f, to be introduced later.

The field equations in the new dependent and independent variables are

$$X_{\pm,w}=\frac{C_\pm-X_{\pm,\rho}}{-\Gamma w \mp \bar{g}}, \tag{43}$$

$$\bar{g}_{,w}=\frac{\bar{g}_{,\rho}+(a^2-2+\Gamma)\bar{g}}{\Gamma w}, \tag{44}$$

$$a_{,w}=\bar{g}^{-1}C_2, \tag{45}$$

$$a_{,\rho}=C_1+\Gamma w\bar{g}^{-1}C_2, \tag{46}$$

where

$$\Gamma(\rho)\equiv1+n-h'(\rho). \tag{47}$$

We are dealing again with four evolution equations (now in w instead of ζ) (43)–(45) and one constraint (46). Singular points arise where the denominator of Eq. (43) or of Eq. (44) vanishes.

We first examine the singular point $w=0$, corresponding to $t=0$. We assume that a is at least C^0 there, as suggested

by the asymptotics, and takes the value $a_0(\rho)$. If we set $a = a_0(\rho)$ in Eq. (44), this approximate equation admits an exact general solution, which we can write as

$$\bar{g}(w,\rho) = w^{(n-\nu)/(1+n)} \exp\left\{ -\frac{n-\nu}{1+n} h(\rho) - I_0(a^2 - 2 + \gamma)(\rho) \right.$$
$$\left. + F\left[\rho + \frac{\ln|w| - h(\rho)}{1+n} \right] \right\}. \tag{48}$$

Here ν is defined by Eq. (40), and F is a periodic function of one variable with period Δ, which serves as the free parameter of the solution. We see that $\bar{g}$ will either blow up or vanish at $w=0$, unless we impose $n=\nu$. Furthermore $\bar{g}$ will not be C^0 unless $F \equiv 0$. But imposing these two conditions is equivalent to the one condition

$$[\bar{g}_{,\rho} + (a^2 - 2 + \Gamma)\bar{g}]_{w=0} = 0. \tag{49}$$

(This is so because the expression $a^2 - 2 + \Gamma$ can only be the derivative of the periodic function $\ln\bar{g}$ if its average value vanishes. This corresponds precisely to $n=\nu$, although ν does of course not appear explicitly in the equations.) This is just the regularity condition one would have expected from inspecting the equations, but now we have shown that imposing it actually corresponds to imposing analyticity. (Note that this is not a condition on g, but a partial fixing of the free parameters in the new coordinate system.)

Before we discuss the other singular points, we use $h(\rho)$ to identify $\zeta = \zeta_0$ with $w=1$, thus simplifying the matching between the two coordinate systems. (Then we have to interpolate only in one, not two dimensions, in order to match the grids. $0 < \zeta_0 < \infty$ is arbitrary.) We define the periodic function $\widetilde{\xi}_0$ implicitly in terms of ξ_0 and ζ_0 by the equation

$$\widetilde{\xi}_0[\tau + \zeta_0 + \xi_0(\tau)] \equiv \xi_0(\tau). \tag{50}$$

(Numerically, this is solved by iteration.) Next we define

$$\widetilde{f}(\rho) \equiv f[\rho - \zeta_0 - \widetilde{\xi}_0(\rho)]. \tag{51}$$

This definition implies that $\widetilde{f}(\rho)$ and $f(\tau)$ coincide when restricted to the line $\zeta = \zeta_0$. Now we fix h as

$$h(\rho) \equiv (1+n)[\zeta_0 + \widetilde{\xi}_0(\rho)] - \widetilde{f}(\rho), \tag{52}$$

and it can be verified that with this definition $w=1$ for $\zeta = \zeta_0$. Now let $a_0(\tau) \equiv a(\zeta_0, \tau)$, and define $\widetilde{a}_0$ from a_0 in the same way as $\widetilde{f}$ from f. Proceed similarly for g, X_+, and X_-. $\widetilde{a}_0$ and $\widetilde{X}_{\pm 0}$ now constitute initial data for a and $X_\pm$ on $w=1$. The initial data for $\bar{g}$ on $w=1$ can be expressed in terms of $\widetilde{g}_0$ as

$$\bar{g}(w=1,\rho) = \frac{\Gamma}{1 - \widetilde{\xi}_0'} e^{\zeta_0 + \widetilde{\xi}_0(\rho)} \widetilde{g}_0(\rho). \tag{53}$$

Now we come back to the other singular points of the equations, namely where $-\Gamma w \mp \bar{g} = 0$. Straightforward calculation shows that this happens where $dw = 0$ is null, just as the denominator of Eq. (16) vanishes where $d\zeta = 0$ is null. Either occurrence indicates a SSH. If we choose ζ_0 large

enough, we do not encounter the past SSH. Why do we not simply choose $\zeta_0 = 0$, however? As one expects, the denominator of the matter equation then vanishes on the entire line $w=1$, but it vanishes also at two more lines (per period in ρ) crossing $w=1$. This behavior would give us the numerical problems we avoided in the old coordinate system by making the SSH a line of constant ζ. [The denominator of the matter equation in the (ζ, τ) coordinate system is everywhere increasing with ζ, which means that it vanishes only at $\zeta = 0$ and nowhere else.]

We have now formulated the analytic continuation at $t=0$ as a boundary value problem mathematically quite similar to the one we have solved above. We consider a as determined by $\bar{g}$ and $X_\pm$ via the constraint (46). The unknowns are the fields $\bar{g}$ and $X_\pm$ between $w=1$ and $w=0$, and the function $\Gamma(\rho)$. The three boundary conditions on the left are the matching of $\bar{g}$ and $X_\pm$ to the data at $\zeta = \zeta_0$, and the one boundary condition on the right is Eq. (49). Note that the unknown functions and constant appear in the field equations only in the one combination Γ, and that $\widetilde{g}(w=1)$ is determined from Γ, $\widetilde{\xi}_0$, and $\widetilde{g}_0$. This breaks up our numerical procedure naturally into two steps. In a first step, the functions $\widetilde{\xi}_0$, $\widetilde{a}_0$, $\widetilde{g}_0$, and $\widetilde{X}_{\pm 0}$ are once and for all determined from the data at $\zeta = \zeta_0$. Then we vary Γ in a relaxation algorithm, until we have solved the system (43)–(46) with boundary condition (49).

We have seen that in the boundary value problem we are only dealing with one free function $\Gamma(\rho)$ to be determined (which plays the same role as ξ_0 before), while f, h, and n do not appear explicitly or separately. I have introduced them because they are useful in deriving an initial guess for Γ for use in the numerical algorithm. Substituting the asymptotic form (37) of g into the definition of $\bar{g}$ we obtain

$$\bar{g} \simeq g_0(\tau)[1 + n + f'(\tau)] \exp[f(\tau) - \nu\xi_0(\tau)$$
$$+ (n - \nu)(\tau - \rho) + \sigma(\rho) + h(\rho)]. \tag{54}$$

(Note that, exceptionally, I have mixed coordinate systems, in using ρ and τ as independent variables.) This is regular at $\tau = \infty$ if $n = \nu$, and if

$$g_0(\tau)[1 + n + f'(\tau)] e^{f(\tau) - n\xi_0(\tau)} = 1. \tag{55}$$

This latter condition can be considered as a linear ODE for $\exp[f(\tau)]$, given g_0 and $n=\nu$. By evolving in coordinates (ζ, τ) to large ζ, and comparing with the asymptotic form (37) we can estimate ν and g_0. From there we can calculate an estimate for f via Eq. (55) and then for h and hence Γ via Eqs. (50)–(52).

Now that the problem has been solved in the new variables, we can calculate $\alpha = a/g$ from Eq. (42). It contains a singular t-dependent factor that we can absorb into the (singular) redefinition $t \to \bar{t} = t^{1+n} e^{f[\ln(t/r_0)]}$. The resulting regular $\bar{\alpha}$ is

$$\bar{\alpha} = e^{-n\rho + h(\rho)} \frac{a}{g}. \tag{56}$$

n and h are given in terms of the now determined function Γ by $n = \widetilde{\Gamma}_0 - 1$ and $h = -I_0\Gamma$. Note that $\bar{\alpha}$ is no longer periodic in ρ, although the spacetime is DSS. We have had to

use the most general form of the metric compatible with Eq. (2), abandoning the gauge condition $\alpha = 1$ at $r = 0$. In fact $\bar{\alpha}$ is singular at $r = 0$, but then we only use it for a patch around spacelike infinity.

We obtain a clearer picture of the behavior at spacelike infinity by ignoring the ''wiggles'' in a and α. Then they simplify to the expressions

$$a \sim \sqrt{1-n}, \quad \bar{\alpha} \sim r^{-n}, \tag{57}$$

which are valid for $r \gg |t|$. The Hawking mass is proportional to the radius, and the geometry at spacelike infinity is conical. Our constant n coincides with ''$-1+1/n$'' in the notation of [3]. Numerically I find $n \simeq -0.16$.

B. From $t = 0$ to the future self-similarity horizon

It remains to extend the spacetime all the way to the future SSH, the future light cone of the singularity at $(t=0, r=0)$, beyond which the continuation is no longer unique. *A priori*, the future SSH might itself be singular, because we have no free parameters left to make it regular.

At $t = 0$, the periodic data $X_{\pm}(\rho)$ are fixed, and they determine $a(\rho)$ through the constraint (46). $\bar{g}$, on the other hand, is pure gauge at $t = 0$, depending on our choice of Γ through Eq. (49). We now continue to evolve in w with the same equations as before, but with a new choice of Γ. This means that we introduce a third coordinate system, although one of the same class (38) and (41) as before. Figure 2 shows where this third coordinate system is used. This time we choose Γ so as to make the line $w = -1$ the future SSH, which is also the future light cone of the singularity. Our reason for this is twofold: on the one hand, we are better placed to control the vanishing of the denominator in the matter equation (now for X_+ instead of X_-) when it takes places on a coordinate line, on the other hand, we want the edge of the domain of dependence of our data to coincide with a coordinate line, so that we can evolve right up to it. Our boundary conditions are now the data for $X_{\pm}$ and a at $w = 0$, the constraint (49) which determines $\bar{g}$ from Γ at $w = 0$ (up to a constant factor), and the coordinate condition $\bar{g} = \Gamma$ (vanishing of the denominator of $X_{+,w}$) at $w = 1$. *A priori* this new boundary value problem is not well posed. We have no freedom left to impose the vanishing of the numerator of $X_{+,w}$ at $w = 1$ as an additional boundary condition, so that the solution should be genuinely singular at this point.

To investigate what happens we once more make an analytic approximation. Let us assume that X_- is at least C^0 at $w = -1$, and takes value $X_{-0}(\rho)$. By definition, as our new coordinate condition, $\bar{g}$ takes the value $\Gamma(\rho)$. In consequence X_+ drops out of the constraint (45), and independently of the value of X_+, a takes the value $a_0(\rho)$, which is the solution of the ODE:

$$a_0' = \frac{1}{2} a_0 (1 - a_0^2) + a_0^3 X_{-0}^2 \tag{58}$$

with periodic boundary conditions. The solution exists and is unique. We can now calculate $\bar{g}_{,w}$ at $w = -1$, and hence the linear approximation to the denominator $-\Gamma w - \bar{g}$ near $w = -1$. With these expressions in hand we write out the

equation for X_+, in the leading terms in both the numerator and the denominator. We obtain an approximate equation of the same form as Eq. (26): namely,

$$X_{+,w} = \frac{A(\rho)X_+ - X_{+,\rho} + C(\rho)}{(w+1)B(\rho)}, \tag{59}$$

where the coefficients are

$$A(\rho) = \frac{1}{2}(1 - a_0^2) - a_0^2 X_{-0}^2, \quad B(\rho) = a_0^2 - 2 + (\ln\Gamma)',$$

$$C(\rho) = -X_{-0}. \tag{60}$$

The exact general solution to the approximate equation is, once more,

$$X_+ = X_+^{\text{inhom}}(\rho) + X_+^{\text{hom}}(w, \rho), \tag{61}$$

where

$$A X_+^{\text{inhom}} - X_{+,\rho}^{\text{inhom}} + C = 0 \tag{62}$$

and

$$X_+^{\text{hom}} = (w+1)^{\hat{A}_0/\hat{B}_0} e^{I_0 A - (\hat{A}_0/\hat{B}_0)I_0 B} F\left[\rho + \frac{I_0 B - \ln|w+1|}{B_0}\right]. \tag{63}$$

As we have no freedom to impose any boundary conditions, F does not vanish and the solution is not analytic at $w = -1$. But we see that $\hat{B}_0 < 0$, while $\hat{A}_0 < 0$ unless $(a_0 - 1)$ and X_{-0} vanish identically. The infinitely oscillating term therefore vanishes at $w = -1$ as $(w+1)^\epsilon$, where ϵ is positive and small. X_+ is therefore C^0, although not C^1. This is a remarkable result: The presence of even a small amount of radiation crossing the future SSH (the component X_-) regularizes the radiation running along the horizon (the component X_+), by damping its oscillations. A similar result was found by Horne in the axion-dilaton [13] and free complex scalar [10] (see note added in print) CSS solutions, using a similar analytic approximation. I have no explanation of this phenomenon in physical terms.

From the field equations it follows that a, $\bar{g}$, and X_- are C^1 but not C^2 (with respect to w, differentiation with respect to ρ is not a problem). As not all second derivatives of the metric exist, one must ask if the spacetime curvature is finite. In Appendix G, I show that all components of the Riemann tensor are in fact C^0.

Although the numerical problem is ill defined because of the presence of an infinite number of oscillations in X_+, I have run a naive relaxation algorithm on it. The algorithm does in fact converge, and I even see convergence in the values of Γ, a_0, and X_{-0} with decreasing step size. I find $\Gamma \simeq 1 + \nu$, and $a_0 \simeq 1$ and $X_{-0} \simeq 0$. This means that spacetime is approximately flat on the future SSH, and very little scalar field radiation is crossing it. X_+, however, oscillates more and more rapidly. These oscillations appear to be at constant amplitude, and I do not see their eventual decay numerically. This is consistent with the fact that $\epsilon \ll 1$, so that the decay is very slow.

In order to make the problem numerically well defined, in spite of the solution being singular, one would have to subtract the singular part, solving for F in another boundary value problem. (This caveat may apply also to Horne's numerical results [13].) I have not attempted this nontrivial step, as I do not see an immediate need for quantitative results. Although the problem is not numerically well defined as it stands, I am confident in the qualitative result that the null data at the future SSH are regular, and nearly, but not quite flat.

While all this explains in hindsight why the future SSH is, in some sense, regular, it does not explain why it is nearly flat, that is, why X_{-0} is so small. I do not see a mechanism in the field equations that would drive arbitrary data at $w=0$ to very nearly flat space values. And in fact the SSH tends to be less flat when I slightly perturb X_+ or X_- at $w=0$ (and adjust a accordingly). This indicates that the near-flatness at the future SSH is a property of the special DSS solution which is regular at the past SSH. (Incidentally, it is also an argument against near-flatness being a numerical artifact.)

IV. LINEARIZED PERTURBATIONS AND CRITICAL EXPONENT

A. The linear eigenvalue problem

Now we turn to the study of the linearized perturbations of the critical solution, specifically those which leave the perturbed solution regular at $r=0$ and $\zeta=0$. If such a regular perturbation existed that was also periodic in τ, the critical

solution would not be isolated. The perturbations must therefore break the discrete homotheticity, that is, the periodicity in τ. The coefficients of the linearized equations, however, are periodic, and therefore the general linear perturbation is a superposition of terms of the form

$$\delta Z(\zeta,\tau)=\sum_{i=1}^{\infty} C_i e^{\lambda_i \tau}\delta_i Z(\zeta,\tau), \tag{64}$$

where each $\delta_i Z$ is periodic in τ with period Δ (although δZ is not). Z is shorthand for (a,g,X_+,X_-). The exponents λ_i and hence the $\delta_i Z$ must be allowed to be complex even for real δZ. As the equations are real, they form complex conjugate pairs, corresponding to sine and cosine oscillations in τ with frequency $\text{Im}\lambda_i$. Because the ansatz already contains all frequencies that are integer multiples of $2\pi/\Delta$, and because values of λ_i come in complex conjugate pairs, we need consider only $0 \leqslant \text{Im}\lambda_i < \pi/\Delta$.

If the equation for δZ is of the form

$$\delta Z_{,\zeta}=A\,\delta Z+B\,\delta Z_{,\tau}, \tag{65}$$

the equation for $\delta_i Z$ is of the form

$$\delta_i Z_{,\zeta}=(A+\lambda_i B)\delta_i Z+B\,\delta_i Z_{,\tau}. \tag{66}$$

This indicates how we obtain the equations for $\delta_i Z$ from the linearized equations for δZ. In the following I denote the components of $\delta_i Z$ by $\delta_i a$, $\delta_i g$, and $\delta_i X_\pm$. The equations for the periodic quantities $\delta_i Z$ are then

$$\delta_i X_{\pm,\zeta}=\left\{\left[\frac{1}{2}(1-a^2)-a^2 X_\mp^2\right]\delta_i X_\pm-(2a^2 X_+ X_-+1)\delta_i X_\mp-aX_\pm(1+2X_\mp^2)\delta_i a\pm e^{\zeta+\xi_0}[X_{\pm,\tau}\delta_i g+g(\delta_i X_{\pm,\tau}+\lambda_i \delta_i X_\pm)\right.$$

$$\left.-(1+\xi_0')X_{\pm,\zeta}\delta_i g]\right\}[1\pm(1+\xi_0')e^{\zeta+\xi_0}g]^{-1}, \tag{67}$$

$$\delta_i g_{,\zeta}=(1-a^2)\delta_i g-2ag\,\delta_i a, \tag{68}$$

$$\delta_i a_{,\zeta}=\left[\frac{1}{2}+\frac{3}{2}a^2(X_+^2+X_-^2-1)\right]\delta_i a+a^3(X_+\delta_i X_++X_-\delta_i X_-), \tag{69}$$

$$\delta_i a_{,\tau}=-\lambda_i \delta_i a+e^{-(\zeta+\xi_0)}g^{-1}\left[\frac{3}{2}a^2(X_+^2-X_-^2)\delta_i a+a^3(X_+\delta_i X_+-X_-\delta_i X_-)-\frac{1}{2}g^{-1}a^3(X_+^2-X_-^2)\delta_i g\right]$$

$$+(1+\xi_0')\left\{\left[\frac{1}{2}+\frac{3}{2}a^2(X_+^2+X_-^2-1)\right]\delta_i a+a^3(X_+\delta_i X_++X_-\delta_i X_-)\right\}. \tag{70}$$

We now consider the boundary conditions for these equations. At $\zeta\to-\infty$ we have one free function $\delta_i Y_1(\tau)$ (compare Appendices A and B). At the boundary $\zeta=0$, the denominator of the $\delta_i X_-$ equation vanishes, because it is the same as that of the X_- equation. We therefore have to impose the vanishing of the numerator. Imposing the linear constraint as well, we can freely specify $\delta_i g$ and $\delta_i X_+$ at $\zeta=0$, and calculate from them $\delta_i a$ and $\delta_i X_-$. Details are given in Appendix B.

We now have three free functions $\delta_i Y_1$, $\delta_i g_0$, and $\delta_i X_{-0}$ at the boundaries. At some intermediate point the three functions $\delta_i g$, $\delta_i X_+$, and $\delta_i X_-$ will have to match. The fourth, $\delta_i a$, will match automatically by virtue of the linearized constraint (70). We also have the free constant λ_i to solve for. Its presence is balanced by the fact that because of linearity an overall factor in the perturbations is arbitrary and has to be fixed in some way. The λ_i are the eigenvalues in a new hyperbolic boundary value problem, this time a linear

one. One would therefore expect them to form a discrete set.

Details of the numerical method are given in Appendix E, and an error estimate in Appendix F. Because of the even-odd symmetry of the background, perturbations which have the same symmetry decouple from perturbations with the opposite symmetry, and we can consider them separately.

In the left half plane of λ_i, corresponding to modes that grow towards the singularity $\tau = -\infty$, I have found one perturbation of the first type. In the following I refer to this mode by $i=1$. It is real, with $\lambda_1 = -2.674 \pm 0.009$, corresponding to a critical exponent $\gamma = 0.374 \pm 0.001$. I find no perturbations of the second type in the left half plane. There is one perturbation, of the first type, at $\lambda_i = 0$. It is the gauge mode $\delta Z \propto Z_{*,\tau}$, corresponding to a translation of the background in τ. On the positive real axis I find perturbations of both types, and presumably there are more elsewhere in the right half-plane.

B. Critical scaling of the black hole mass

The derivation of the scaling of the black hole mass for near-critical initial data which I present here was suggested in [3], and first made explicit in [16] for the CSS case. The DSS case requires a subtle generalization, and I find a new phenomenon: a "wiggle" in the mass scaling law.

In preparation, let us consider a family of Cauchy data at constant t: namely,

$$Z_\tau(r) = Z_* \left(\ln\frac{r}{r_0}, \tau \right) + \epsilon \delta_1 Z \left(\ln\frac{r}{r_0}, \tau \right), \qquad (71)$$

where τ is the parameter of the family. Here $Z_*(\zeta, \tau)$ is the critical solution that we have just constructed, and $e^{\lambda_1 \tau} \delta_1 Z(\zeta, \tau)$ is the one linear perturbation mode that is growing on small scales, that is, which has negative λ. r_0 is an arbitrary length scale and ϵ is a fixed small constant, small enough so that the linear approximation is good initially. $\ln r/r_0$ is the value the formal argument ζ takes. Clearly this family is periodic in τ with period Δ.

What happens at late times when these data are evolved in t? For one sign of ϵ, say $\epsilon < 0$, we must find dispersion, for the other a black hole. The key observation [11] is that the data depend on r only through the dimensionless combination r/r_0, while the evolution equations themselves have no scale. The entire solution evolved from these data scales with r_0 as

$$Z_\tau(r,t) = f \left(\frac{r}{r_0}, \frac{t}{r_0}, \tau \right). \qquad (72)$$

Therefore we know, without having to rely on the linear approximation, that the black hole mass, which has dimension length, must be proportional to r_0 (with at most a periodic dependence on τ). More precisely, we have

$$M = r_0 e^{\mu(\tau)}, \qquad (73)$$

where $\mu(\tau)$ is an unknown periodic function of period Δ. It can be calculated numerically by evolving members of the family Z_τ which span one period in τ.

Now we consider a generic solution. If the initial data are sufficiently close to criticality, there is a spacetime region in

their evolution where the solution "echos," that is, where it is close to the critical solution. There the solution is well approximated by the critical solution plus linear perturbations, as

$$Z(\zeta,\tau) \simeq Z_*(\zeta,\tau) + \sum_{i=1}^{\infty} C_i(p) e^{\lambda_i \tau} \delta_i Z(\zeta,\tau). \qquad (74)$$

The amplitudes C_i of the perturbations depend in a complicated way on the initial data in general and hence on the parameter p of a given one-parameter family of initial data. As $t \to 0$ but while the perturbations are still small, we can neglect all perturbations but the growing one, and we have

$$Z(\zeta,\tau) \simeq Z_*(\zeta,\tau) + C_1(p) e^{\lambda_1 \tau} \delta_1 Z(\zeta,\tau). \qquad (75)$$

By definition we obtain the precisely critical solution for $p = p_*$, and so we must have $C_1(p_*) = 0$. Linearizing $C_1(p)$ around p_*, we obtain

$$Z(\zeta,\tau) \simeq Z_*(\zeta,\tau) + (p - p_*) \frac{\partial C_1(p_*)}{\partial p} e^{\lambda_1 \tau} \delta_1 Z(\zeta,\tau). \qquad (76)$$

We define

$$\gamma \equiv -\frac{1}{\lambda_1} \qquad (77)$$

(as the notation suggests, this will turn out to be the critical exponent), and

$$\tau_*(p) \equiv \gamma \ln \left[(p - p_*) \frac{1}{\epsilon} \frac{\partial C_1}{\partial p}(p_*) \right],$$

$$r(p) \equiv r_0 e^{\tau_*(p) + \xi_0[\tau_*(p)]}, \qquad (78)$$

where ϵ is the same arbitrary small constant as in definition (71). If we now fix $\tau = \tau_*(p)$ in the approximate solution (76), we obtain a p-dependent family of Cauchy data, namely

$$Z_p(r) = Z_* \left(\ln\frac{r}{r(p)}, \tau_*(p) \right) + \epsilon \delta_1 Z \left(\ln\frac{r}{r(p)}, \tau_*(p) \right). \qquad (79)$$

But this is of the form for which we know the black hole mass. Therefore we have, for the mass of the black hole as a function of p,

$$M(p) = r(p) e^{\mu[\tau_*(p)]}$$

$$= r_0 (p - p_*)^\gamma e^{\tau_f + \mu[\gamma \ln(p - p_*) + \tau_f] + \xi_0[\gamma \ln(p - p_*) + \tau_f]}$$

$$\equiv r_0 (p - p_*)^\gamma e^{\tau_f + \bar{\mu}[\gamma \ln(p - p_*) + \tau_f]}, \qquad (80)$$

where

$$\tau_f \equiv \gamma \ln \left[\frac{1}{\epsilon} \frac{\partial C_1}{\partial p}(p_*) \right] \qquad (81)$$

is a family-dependent constant, while μ and ξ_0, and hence $\bar{\mu} \equiv \mu + \xi_0$ are not family dependent. Let us first consider the CSS case. Then μ and ξ_0 degenerate into constants. We

recover the well-known exact scaling of the black hole mass. The unknown, family-dependent, constant τ_f corresponds to an unknown overall factor in the mass.

In the DSS case we find a ''wiggle'' overlaid on the scaling law, unless the function $(\mu + \xi_0)$ vanishes identically. The period of both μ and ξ_0 is nominally Δ, but ξ_0 has only even frequencies, and so does μ, being based only on the metric coefficients a and g. The real period in τ of $(\mu + \xi_0)$ is therefore $\Delta/2$, and hence the period of the wiggle in the directly measured parameter $\ln(p - p_*)$ is $\Delta/(2\gamma) \simeq 4.61$.

The offset of the wiggle is the same constant τ_f that already appears in the overall factor. Given the function $\mu(\tau)$ [or equivalently $\bar{\mu}(\tau)$], the black hole mass is therefore as much determined in the DSS case, namely up to one family-dependent constant, as in the CSS case, and $\mu(\tau)$ has the same universal significance as γ. It would therefore be interesting to determine $\mu(\tau)$ from evolving Eq. (71) and to test the expression (80) against collapse simulations.

C. The free complex scalar field model

As mentioned already, there is a regular spherically symmetric CSS solution for the free complex scalar field, that is a complex scalar field with neither a mass term nor a coupling to an electromagnetic field [10]. Later it was discovered that this solution has not one but three unstable modes [11], and that the critical solution for the free complex scalar field is in fact the real DSS solution we have discussed here (up to a global complex phase) [21].

The action is, in loose notation, $R + |\partial\Phi|^2$. Writing the complex scalar Φ as $\phi + i\psi$, the action becomes $R + (\partial\phi)^2 + (\partial\psi)^2$, or that of two real scalar fields. Any solution of the real scalar field model is therefore also a solution of the free complex scalar field model.

We now consider the real scalar DSS solution as a solution of the complex scalar model, by setting $\phi = \phi_*$ and $\psi = 0$. Its linear perturbations decouple into two kinds. We have already obtained the purely real perturbations (with $\delta\psi = 0$), and the accompanying metric perturbations. Purely imaginary scalar field perturbations, with $\delta\phi = 0$, must be considered separately. They do not give rise to metric perturbations (to linear order) because the stress tensor is the sum of a term quadratic in ϕ and one quadratic in ψ. The first-order perturbation of this second term vanishes if the background value of ψ is zero. The perturbations $\delta\psi$ therefore obey the real free wave equation on the background spacetime Z_*. This also means that they decouple from the real perturbations $\delta\phi$.

Using the same numerical methods as for the real perturbations, only with different field equations, I have checked that all imaginary modes $\delta\psi$ are damped. With a little extra work I have thus confirmed perturbatively that the real DSS solution is an attractor of codimension one even in the free complex model. Details will be given elsewhere [22].

V. CONCLUSIONS

A. Critical solutions and matter models

As I have argued in the Introduction, critical solutions of the kind discussed here play a role in critical collapse similar to the role the Kerr-Newman solution plays in generic, noncritical, collapse. The crucial difference is the presence of matter: A critical solution does not have a horizon, and therefore depends explicitly on the choice of matter. This reduces the importance of any one such solution. Two other differences are mainly of technical importance. While the Kerr-Newman solutions are known in closed form, all critical solutions discovered so far are only known numerically. Furthermore, CSS solutions, like stationary ones, effectively depend on one dimension fewer, but this simplification does not hold for the DSS ansatz: it only makes all fields periodic in one coordinate. The labor involved in constructing the real scalar field critical solution, which is DSS, is therefore much greater than for the other self-similar solutions found so far, all of which are CSS. There are three reasons for investing it.

First, the real scalar field model is the first in which critical phenomena were observed, and has raised much interest, probably as much because of the ''echos'' as because of the critical mass scaling. Here we have explained the echos as a natural generalization of (continuous) self-similarity.

Second, this investigation has predicted a new phenomenon, the ''wiggle'' in the mass scaling law, which had not been seen in Choptuik's pioneering work.

Note added: After submission of this paper I learned of an independent prediction, together with a numerical confirmation in collapse simulations, of the ''wiggle'' in the mass scaling [23].

Finally, the real scalar field is a test bed of methods for dealing with DSS which can now be applied to other DSS critical solutions. The most interesting of these is probably that of pure gravity in axisymmetry [2].

This raises the fundamental problem of angular momentum and electric charge in the initial data for gravitational collapse. On the one hand, the resulting black hole can have angular momentum and charge. On the other, both must be smaller than the mass. So what happens to the black hole charge and angular momentum if one tries to fine-tune the black hole mass to zero? Clearly the case of rotation is the more interesting one, but it cannot be treated in spherical symmetry. Only in one case so far has critical collapse been considered in axisymmetry [2], but in vacuum and therefore without angular momentum.

Another restriction of the matter models which have so far been considered in the study of critical phenomena is that most of them, in marked difference from any realistic macroscopic matter, do not introduce a preferred length scale. (In units where $G = c = 1$, this is equivalent to the absence of dimensionful parameters in the action.) This guarantees the existence of an exactly self-similar solution. Even in the presence of a preferred scale, however, a self-similar solution may still be a good approximation towards the singularity, as $s^2 \equiv r^2 + t^2 \to 0$. An example for this is adding a mass term $m^2\phi^2$ is added to the scalar field action. Then the DSS solution we have constructed here is a good approximation for $s \ll m^{-1}$, simply because ϕ is bounded while $\partial\phi$ is not. In collapse simulations, it is found that the Choptuik solution is in fact the attractor for the massive scalar field [20]. It remains to be investigated how other matter models introducing a scale react to the attempt to make small black holes. A very recent example is the Einstein-Yang-Mills system, which has one intrinsic scale, and which shows both a mass

gap, and critical mass scaling, for different classes of initial data [24]. The critical solution for this system has now been constructed [36]. A rather different, and interesting, one is the attempt to consider semiclassical Einstein equations, with the new scale being the Planck length [25].

B. Other self-similar scalar field solutions

In this paper, I have considered a scalar field ϕ which is (a) bounded on the entire spacetime ($\kappa=0$), (b) real, and (c) discretely self-similar (DSS). These conditions are suggested by the universal attractor that Choptuik [1] found in critical collapse of a real scalar field, and they are justified by the fact that the unique solution of the resulting eigenvalue problem coincides with Choptuik's attractor to numerical precision [17]. I now review some related work on self-similar scalar field solutions departing from one or another of these assumptions.

CSS real scalar field solutions have been studied in [26]. The solution for $\kappa=0$ can be derived in closed form and was apparently first published in [27], then rediscovered in the context of critical phenomena in [28,29]. CSS arises as a degenerate case of our formalism when a, g, and $X_\pm$ are not only periodic in τ, but do not depend on τ at all. In our notation, continuous self-similarity implies that $Y_1(\tau)$ in Eq. (A2) is a constant, namely $Y_1=\kappa$. (This is incompatible with the even-odd symmetry I have assumed.)

Brady [26] has shown that for all κ in the CSS case, boundedness at the past SSH is automatic, in contrast to the DSS case. As in the DSS case, continuation beyond the past SSH is not unique. Brady has considered the one-parameter family of possible continuations (for each value of κ). Unfortunately, it is not clear which of his continuations is the analytic one. In the DSS case, I have investigated only the analytic continuation at the past SSH, because the past SSH is in the Cauchy development of regular data when we think of it as arising as an attractor in collapse simulations, and therefore I believe it should be analytic. Of Brady's results I review here only the special case $\kappa=0$, because it appears to be most closely related to our case. It is qualitatively different from $\kappa\neq0$, in that no CSS solution with a regular center exists, except for flat empty space. There is a one-parameter family of solutions with a singular center $r=0$. One branch of $r=0$ is always timelike and has negative (Hawking) mass, the other is either timelike with negative mass (called subcritical solutions in [28]), or spacelike with positive mass, and in the latter case it is preceded by an apparent horizon (supercritical solutions). In either case both the past and the future SSH carry flat-space null data, and one can therefore replace the negative mass part of the solution with flat space in both the past and future light cone of the singularity. Subcritical solutions thus pieced together are qualitatively like the maximally extended Choptuik solution in that $r=0$ is a regular center, except at the single point ($r=0,t=0$), which is a naked singularity. Both the past and the future SSH are regular and flat, however. (In the Choptuik solution the past SSH is not flat, and the future SSH is only nearly flat.)

What would happen if we allowed $\kappa\neq0$ in DSS, while still imposing regularity at the center and on the past SSH? Because of mode coupling, this would mean giving up the even-odd symmetry, and doubling the (infinite) number of

degrees of freedom in the boundary value problem. Therefore, one might obtain an infinity of new solutions, some, or no new solutions at all. If any new solutions exist, one would wonder next why it is only the one with $\kappa=0$ that serves as a universal attractor. If there is a family of such solutions continuous with the Choptuik solution, that question would be even more acute. I leave these questions to future work.

A CSS solution (with $\kappa=0$ assumed implicitly) has also been constructed for the free complex scalar field [10]. In this solution, only the metric and the scalar field modulus are CSS, but the complex phase of the scalar field is $i\omega\tau$ for some constant ω, so that the complex scalar field might be considered DSS with (in our notation) $\Delta=2\pi/\omega$. Not surprisingly, our solution shares more features with this one than with the CSS real scalar field models: There is a unique solution with a regular center and regular past SSH, and the null data on the future SSH are for nearly flat space. The same qualitative picture was found for the CSS solution with axion-dilaton matter [13]. It is remarkable that now three different self-similar solutions are known which are nearly, but not quite, flat at the future SSH. The Roberts solution is clearly the limiting case, with a flat SSH (but consequently a scalar field that is not C^0), and may be of help in finding an explanation.

C. Perturbations, universality, and ''renormalization''

I have found exactly one unstable mode of the critical solution. The picture of an attractor of codimension one is thus confirmed perturbatively. The Lyapunov exponent of $\lambda\simeq-2.674\pm0.009$ gives rise to the value $\gamma\simeq0.374\pm0.001$ of the critical exponent. This value agrees with the most precise value obtained from collapse calculations [6], which is given as 0.374. Allowing for complex scalar field perturbations around the real Choptuik solution does not add any unstable modes. The Choptuik solution is therefore an attractor of codimension one even for the complex scalar field.

It would be interesting to run collapse calculations for a one-parameter family of matter models, such as the $p=k\rho$ family [7] or the nonlinear σ model [9], where the Lyapunov exponents of the perturbations change continuously with the parameter. One might thus be able to find parameter regions where two equally strong attractors coexist, and new interesting phenomena would arise. It would also be extremely interesting if one could find a way of calculating or estimating the number of unstable modes of a given self-similar solution other than by constructing the solution and all its perturbations explicitly.

Another point should be addressed in future theoretical work. As I mentioned in the Introduction, the phase space picture of universality, although apparently correct in some aspects, is strictly speaking wrong because the same spacetime corresponds to many different trajectories in superspace, according to the way it is sliced. At the very least one should be able to derive a universal geometrical prescription fixing the lapse and shift from Cauchy data in superspace, such that the intuitive phase space flow is realized as a Hamiltonian flow. Furthermore, this Hamiltonian flow should admit a geometric interpretation as a scaling, or ''renormalization group,'' flow. The evolution in the time variable τ (at constant ζ) does in fact go from one set of

Cauchy data to the same, only changed in overall scale.

Clearly the idea of approximate scale invariance is at the core of renormalization group ideas and methods. Furthermore, the calculation of γ as given in Sec. IV B is identical with the calculation of the critical exponent governing the divergence of the correlation length given in any textbook on critical phenomena in statistical mechanics, for example [30]. As far as I can see, however, the ''critical phenomena'' analogy is not with critical phenomena in a system which is described by a partition function (such as a system in thermal equilibrium or a quantum field). There is neither a nonvanishing Hamiltonian, nor an inverse temperature β, or quantum of action $\hbar$, such that one could construct a weight on phase space.

The analogy seems to be rather with the application of renormalization group methods to deterministic partial differential equations (PDE's) [31]. This takes up ideas of Barenblatt's [32] of generalizing a self-similar ansatz such as $f(r^2/t)$ for a PDE such as a generalized diffusion equation to $t^\alpha f(r^2/t)$. Here, α is a noninteger ''anomalous dimension'' that cannot be determined by dimensional analysis. But the factor t^α gives rise to terms proportional to α in the equation for f, and this equation will admit regular solutions only for certain values of α: one has a nonlinear eigenvalue problem.

In the critical collapse case we see this twice. In our ansatz for the perturbations we clearly have an explicit factor t^λ. But Δ has the same function in the background solution. There we expand all fields, which are periodic in τ, in terms of $e^{in\omega\tau}Z_n(\zeta)$, where $\omega \equiv 2\pi/\Delta$. In consequence we obtain terms proportional to ω in the equations for the expansion coefficients $Z_n(\zeta)$. The CSS case corresponds to these terms being absent. We may therefore consider the DSS ansatz as a generalization of the CSS ansatz parallel to the anomalous dimension in Barenblatt's examples, and consider $i\omega$ a complex anomalous dimension. These parallels certainly put the investigation of critical gravitational collapse into a wider context, and may yet give rise to new predictions or a simplified theory.

ACKNOWLEDGMENTS

APPENDIX A: BACKGROUND BOUNDARY CONDITIONS

$\zeta = -\infty$ and $\zeta = 0$ are singular points of the equations. The boundary conditions are therefore implemented by expanding the field equations in powers of e^ζ around $\zeta = -\infty$ and of ζ around $\zeta = 0$. The resulting equations are given here. Using these expansions we can calculate Cauchy data for a very large and a very small negative value of ζ, thus avoiding the vanishing of $X_\pm$ at $\zeta = -\infty$ and the breakdown of the Cauchy scheme at $\zeta = 0$.

The regularity conditions at $\zeta \to -\infty$ can be solved in closed form. We expand in powers of e^ζ, as

$$a(\zeta,\tau) = a_0(\tau) + a_1(\tau)e^\zeta + a_2(\tau)e^{2\zeta} + \cdots. \quad \text{(A1)}$$

It turns out to be more natural to expand X and Y instead of X_+ and X_-. As discussed in Sec. II, we impose $a_0 = 1$ and $g_0 = 1$. Expanding the field equations, we find that the two (periodic) functions $Y_1(\eta)$ and $\xi_0(\eta)$ can be chosen freely. Their significance here is the following. ξ_0 parametrizes the class of spacetime coordinates we use, while a combination of the two contains free boundary data for the scalar field ϕ at $r = 0$:[3]

$$e^{-\xi_0(\tau)}Y_1(\tau) = \sqrt{2\pi G}\frac{\partial\phi}{\partial(\ln t)}\bigg|_{r=0}. \quad \text{(A2)}$$

The other nonvanishing expansion coefficients up to order $e^{3\zeta}$ in X, Y, a, and g are

$$a_2 = -g_2 = \frac{1}{3}Y_1^2, \quad \text{(A3)}$$

$$X_2 = \frac{1}{3}e^{\xi_0}[Y_1' - (1+\xi_0')Y_1], \quad \text{(A4)}$$

$$Y_3 = -\frac{2}{3}Y_1^3 + e^{\xi_0}\left[\frac{1}{2}X_2' - (1+\xi_0')X_2\right], \quad \text{(A5)}$$

while the other coefficients vanish. These coefficients were calculated from $a_0 = g_0 = 1$ and the evolution equations alone, but they also obey the constraint order by order.

Regularity conditions at the boundary $\zeta = 0$ give rise to an ODE system. We again make a power-law ansatz of the form

$$a(\zeta,\tau) = a_0(\tau) + a_1(\tau)\zeta + a_2(\tau)\zeta^2 + \cdots. \quad \text{(A6)}$$

Here we find that convenient fields to expand in are a, X_+, X_- and, instead of g, the quantity

$$D = (1+\xi_0')e^{\zeta+\xi_0}g. \quad \text{(A7)}$$

Its evolution equation is easily seen to be

$$D_{,\zeta} = D(2-a^2). \quad \text{(A8)}$$

The expansion coefficients can be calculated recursively, by the solution of first-order ODE's, if we use $X_{+0}(\tau)$ and $\xi_0(\tau)$ as the free parameters. Our coordinate condition (24) is simply

$$D_0 = 1. \quad \text{(A9)}$$

We note that X_- drops out of the constraint (19) to leading order, which is

$$a_0' = (1+\xi_0')a_0\left[\frac{1}{2} + a_0^2\left(X_{+0}^2 - \frac{1}{2}\right)\right]. \quad \text{(A10)}$$

Given X_{+0}, this is a nonlinear, inhomogeneous, first-order ODE for a_0. The one integration constant is fixed uniquely by the requirement that the solution be periodic. The regularity condition (25) gives us another ODE, this time linear, for X_{-0}:

---

[3]This quantity was called $Y_0(\tau)$ in [17].

$$X'_{-0} + (1+\xi'_0)\left[a_0^2\left(\frac{1}{2}+X_{+0}^2\right)-\frac{1}{2}\right]X_{-0} + (1+\xi'_0)X_{+0} = 0. \quad \text{(A11)}$$

We can now calculate algebraically,

$$D_1 = 2 - a_0^2, \quad \text{(A12)}$$

$$a_1 = \frac{1}{2}a_0[(1-a_0^2)+a_0^2(X_{+0}^2+X_{-0}^2)], \quad \text{(A13)}$$

$$X_{+1} = \frac{1}{2}\left[-a_0^2 X_{+0}\left(\frac{1}{2}+X_{-0}^2\right)+\frac{1}{2}X_{+0}-X_{-0}\right.$$
$$\left. + (1+\xi'_0)^{-1}X'_{+0}\right]. \quad \text{(A14)}$$

For X_{-1} we obtain again a linear ODE,

$$X'_{-1} + (1+\xi'_0)\left[-\frac{5}{2}+a_0^2\left(\frac{3}{2}+X_{+0}^2\right)\right]X_{-1} + (1+\xi'_0)\left[2a_0 a_1 X_{-0}\left(\frac{1}{2}+X_{+0}^2\right)+2a_0^2 X_{-0}X_{+0}X_{+1}+X_{+1}\right] + (2-a_0^2)X'_{-0} = 0. \quad \text{(A15)}$$

To quadratic order, we have three algebraic expressions and one more linear ODE, which we do not give here. [We need the previous equations to determine X_{-1}, which will be needed in Eq. (B6) below.] We have used explicitly only the zeroth order of the constraint (19), but the first and second orders are satisfied identically as expected.

APPENDIX B: BOUNDARY CONDITIONS FOR THE PERTURBATIONS

The perturbed boundary conditions at $\zeta \to -\infty$ in terms of the free perturbation $\delta_i Y_1(\tau)$ are

$$\delta_i a_2 = -\delta_i g_2 = \frac{2}{3}Y_1 \delta_i Y_1, \quad \text{(B1)}$$

$$\delta_i X_2 = \frac{1}{3}e^{\xi_0}[\delta_i Y'_1 + (\lambda_i - 1 - \xi'_0)\delta_i Y_1], \quad \text{(B2)}$$

$$\delta_i Y_3 = -2Y_1^2 \delta_i Y_1 + e^{\xi_0}\left[\frac{1}{2}\delta_i X'_2 + \left(\frac{1}{2}\lambda_i - 1 - \xi'_0\right)\delta_i X_2\right]. \quad \text{(B3)}$$

At $\zeta = 0$ the perturbed constraint (70) simplifies and no longer contains $\delta_i X_-$:

$$\delta_i a_{0,\tau} + \lambda_i \delta_i a_0 = (1+\xi'_0)\left\{\left[\frac{1}{2}+\frac{3}{2}a_0^2(2X_{+0}^2-1)\right]\delta_i a_0 \right.$$
$$+ 2a_0^3 X_{+0}\delta_i X_{+0} - \frac{1}{2}g_0^{-1}a_0^3$$
$$\left. \times (X_{+0}^2 - X_{-0}^2)\delta_i g_0\right\}. \quad \text{(B4)}$$

We can specify $\delta_i X_{+0}$ and $\delta_i g_0$ freely and solve this equation for $\delta_i a_{+0}$.

One detail requires consideration. The average value of the coefficient of $\delta_i a_0$ in Eq. (B4) is $\lambda_i + 1$. When this vanishes, the equation has no solution (with periodic boundary conditions), as discussed in Appendix C. This means that for $\lambda_i = -1$, there are no perturbations that are regular on

$\zeta = 0$. In the λ plane this gives rise to a pole at $\lambda = -1$, see Appendix E. [To calculate the average value of the coefficient, we note that the background constraint (19), evaluated at $\zeta = 0$ with the boundary condition (24), reduces to

$$(\ln a_0)_{,\tau} = (1+\xi'_0)\left[\frac{1}{2}(1-a_0^2)+a_0^2 X_{+0}^2\right]. \quad \text{(B5)}$$

As the left-hand side is the derivative of a periodic function, the right-hand side has vanishing average.]

The vanishing of the numerator of the $\delta_i X_-$ equation is an ODE that can be solved for $\delta_i X_{-0}$; namely,

$$\delta_i X'_{-0} + \left[\lambda_i - (1+\xi'_0)\left(\frac{1}{2}(1-a_0^2)-a_0^2 X_{+0}^2\right)\right]\delta_i X_{-0}$$
$$+ (1+\xi'_0)\{a_0 X_{-0}(1+2X_{+0}^2)\delta_i a_0$$
$$+ [(1+\xi'_0)^{-1}X'_{-0}-X_{-1}]g_0^{-1}\delta_i g_0$$
$$+ (2a_0^2 X_{+0}X_{-0}+1)\delta_i X_{+0}\}$$
$$= 0. \quad \text{(B6)}$$

This equation in turn has no solution if the coefficient of $\delta_i X_{-0}$ has vanishing average. This is the case if λ_i equals the average of $(1+\xi'_0)(1-a_0^2)$, which numerically has value $\simeq -0.385$. This gives rise to another pole.

We do not need to expand the perturbations away from $\zeta = 0$, because in the discretization (D1) of the linearized equations we do not need to evaluate the ζ derivatives of the perturbations at $\zeta = 0$.

APPENDIX C: PSEUDO-FOURIER METHOD

The discrete Fourier transform of the N complex numbers f_n is defined by

$$\hat{f}_k = \frac{1}{N}\sum_{n=0}^{N-1}f_n e^{2\pi ikn/N}, \quad \text{(C1)}$$

and its inverse is

$$f_n = \sum_{k=0}^{N-1} \hat{f}_k e^{-2\pi ikn/N}. \tag{C2}$$

The motivation of this definition is of course that the f_n represent the values of a smooth periodic (complex) function $f(x)$ at N equidistant points over one period. The essential idea of pseudospectral (here pseudo-Fourier) methods is to carry out algebraic operations pointwise on the f_n, and integration and differentiation on the $\hat{f}_k$, switching from one to the other with a fast Fourier transform algorithm. A detailed description of pseudospectral methods can be found in [33]. I give here only the technical information necessary to specify my numerical method.

We shall need to define various operations on the Fourier components $\hat{f}_k$ which represent operations on $f(x)$. To do this in a consistent way, we have to start from a *definition* of $f(x)$ in terms of the $\hat{f}_k$. We choose

$$f(x) = \hat{f}_0 + \sum_{k=1}^{N/2-1} (\hat{f}_k e^{-(2\pi ik/\Delta)x} + \hat{f}_{N-k} e^{(2\pi ik/\Delta)x})$$
$$+ f_{N/2}\cos\left(\frac{N\pi}{\Delta}x\right), \tag{C3}$$

where Δ is the period of $f(x)$. This expression obeys $f(x_n) = f_n$ for $x_n = n\Delta/N$, but this requirement alone does not uniquely define it. We can now *derive* how any operation on the fictitious function $f(x)$ is represented as an operation on the $\hat{f}_k$, for example, interpolation of $f(x)$ to arbitrary values of x, differentiation or integration. Let us begin with differentiation. The corresponding operator D is given by

$$(D\hat{f})_k = -\frac{2\pi k}{\Delta}\hat{f}_k, \quad 0 \le k \le N/2-1,$$

$$(D\hat{f})_{N/2} = 0,$$

$$(D\hat{f})_k = -\frac{2\pi(k-N)}{\Delta}\hat{f}_k, \quad N/2+1 \le k \le N-1. \tag{C4}$$

We see that on differentiation the high-frequency cosine, which takes values of ± 1 on the x_n, is annihilated, because the high-frequency sine takes the value 0 on the x_n. For the purpose of integration we define a more general operator, namely I_λ given by

$$(I_\lambda\hat{f})_0 = \frac{\hat{f}_0}{\lambda}, \quad \lambda \ne 0,$$

$$(I_\lambda\hat{f})_0 = 0, \quad \lambda = 0,$$

$$(I_\lambda\hat{f})_k = \frac{1}{\lambda - 2\pi k/\Delta}\hat{f}_k, \quad 1 \le k \le N/2-1,$$

$$(I_\lambda\hat{f})_{N/2} = \frac{\lambda}{\lambda^2 - (\pi N/\Delta)^2}\hat{f}_{N/2},$$

$$(I_\lambda\hat{f})_k = \frac{1}{\lambda - 2\pi(k-N)/\Delta}\hat{f}_k, \quad N/2+1 \le k \le N-1. \tag{C5}$$

Clearly, for $\lambda = 0$ and $\hat{f}_0 = 0$, this is simply integration, with the integration constant fixed so that the integral has vanishing zero-frequency component. If $f(x)$ has a nonvanishing average (zero-frequency part) $\hat{f}_0$, then its principal function has a part $\hat{f}_0 x$ and is no longer periodic. In this case I_λ enforces periodicity by setting $\hat{f}_0$ equal to zero. When $\lambda \ne 0$, however, even for $\hat{f}_0 \ne 0$, $I_\lambda f$ solves the ODE

$$(I_\lambda f)' + \lambda I_\lambda f = f, \tag{C6}$$

with the integration constant chosen such that $I_\lambda f$ is itself periodic. With the help of this definition we can write the unique periodic solution f of

$$f' + gf + h = 0 \tag{C7}$$

in closed form, namely as

$$f = -e^{-I_0 g} I_{\hat{g}_0}(e^{I_0 g} h). \tag{C8}$$

This is, of course, only the standard use of an integration factor, written in Fourier components, and so that it obeys periodic boundary conditions. The expression diverges as $1/\hat{g}_0$ as $\hat{g}_0 \to 0$, and indeed Eq. (C7) has no solution with *periodic* boundary conditions for $\hat{g}_0 = 0$. All the ODE's we have to solve are of the form (C7) except the constraint (19), which is nonlinear, of the form

$$-2a' + ga + ha^3 = 0. \tag{C9}$$

Fortunately, it can be reduced to the linear form (C7) by the substitution $f = a^{-2}$.

We also need to define an operation that doubles N while representing the same function $f(x)$, and its inverse. These operations will be required for de-aliasing—doubling the $\hat{f}_k$ before going to the f_n, carrying out the necessary algebraic operations on the doubled f_n and going back to the $\hat{f}_k$, then halving the $\hat{f}_k$ and thus throwing away high-frequency noise. With our definition of $\hat{f}(x)$, doubling means splitting the high-frequency cosine. The doubling operation B from N to $2N$ components is

$$(B\hat{f})_k = \hat{f}_k, \quad 0 \le k \le N/2-1,$$

$$(B\hat{f})_{N/2} = (1/2)\hat{f}_{N/2},$$

$$(B\hat{f})_k = 0, \quad N/2+1 \le k \le 3N/2-1,$$

$$(B\hat{f})_{3N/2} = (1/2)\hat{f}_{N/2},$$

$$(B\hat{f})_k = \hat{f}_{k-N}, \quad 3N/2+1 \le k \le 2N-1. \tag{C10}$$

The halving operation S from N to $N/2$ components is

$$(S\hat{f})_k = \hat{f}_k, \quad 0 \le k \le N/4-1,$$

$$(S\hat{f})_{N/4} = \hat{f}_{N/4} + \hat{f}_{3N/4},$$

$$(S\hat{f})_k = \hat{f}_{k+N/2}, \quad N/4+1 \le k \le N/2-1. \tag{C11}$$

I stress once more that the definitions of D, I_λ, B, and S follow uniquely from the definition (C3). With our choice,

D and I_0 are not the inverse of one another because of the treatment of the high-frequency cosine. $S \circ B$ is the identity operator, while $B \circ S$ is a smoothing operator. Our choice of definition is motivated by the requirement that none of the operations mix the real and imaginary part, that is Df, $I_\lambda f$, Bf, and Sf are all real if f is real. I require this because I typically encode two independent real functions as the real and imaginary parts of one complex function.

In a previous code [17], I did not use one consistent definition of $f(x)$ in the Fourier discretization. More seriously, the discretization mixed real and imaginary parts in some places even where the real and imaginary parts represent two unrelated real functions. I have therefore replaced it by the one described above. This results in a small change in the numerical results for the same differential equations discretized on the same grid, in particular a slightly changed value of Δ.

APPENDIX D: NUMERICAL METHODS FOR THE CRITICAL SOLUTION

I use Fourier transformation to transform $1+1$ hyperbolic PDE's into an ODE system, and ODE boundary conditions into algebraic boundary conditions. I use the assumption that $X_\pm$ are odd while a, g, and ξ_0 are even in τ to encode the variables $X_\pm$, g, and ξ_0 as a single complex one, as $\xi_0 + i g + X_+ + i X_-$. The discrete Fourier transform of this function constitutes my independent variables. Let us assume that the information contained in them is to correspond to a sampling of $X_\pm$, g, and ξ_0 at N points each. Then the total number of complex variables is also N, or $2N$ real variables. Typically I use $2N = 128$. Out of these $2N$ variables, I set the one combination corresponding to the fundamental sine mode of ξ_0 equal to zero, storing Δ in its place. As mentioned in the main text, this fixes the translation invariance in τ of the problem and balances the degrees of freedom against the constraints. In the analytic continuation problem the complex variable function is $\Gamma + i G + X_+ + i X_-$. For the linearized perturbations it is $i \delta_i g + \delta_i X_+ + i \delta_i X_-$, plus the constant λ_i, corresponding to $3N/2 + 1$ real variables. The linear perturbations are normalized by an additional condition at the right boundary, namely that the root-mean square of X_{+0} be 1.

ξ_0 and Δ are formally part of the dependent variables of the background problem, although with vanishing ζ derivative. In the perturbation problem they are supplied as external fixed parameters, together with the background solution. In each case, a or $\delta_i a$ is reconstructed at each step from the other three fields by the solution of the constraint. This solution can be given in closed form in Fourier components, as explained in Appendix C. Previously, we have doubled the number of Fourier components, transformed back to real space, and separated the real and imaginary and the even and odd parts to obtain our four fields sampled at $2N$ points each. At the same time we have calculated the necessary τ derivatives. Now we calculate the necessary algebraic expressions pointwise in the $2N$ sampling points, put the result back together in complex form, transform back, and halve the number of Fourier components.

The ODE system $df/d\zeta = F(f,\zeta)$ is discretized in a standard way as

$$f_{n+1} - f_n = (\zeta_{n+1} - \zeta_n) F\left[\frac{f_{n+1} + f_n}{2}, \frac{\zeta_{n+1} + \zeta_n}{2}\right]. \quad \text{(D1)}$$

To solve this system of algebraic equations to machine precision together with algebraic boundary conditions on two sides, I use a standard relaxation algorithm [34]. If I have to integrate these equations forward (that is, with boundary conditions (''initial data'') only on one side), I use a second-order Runge-Kutta step as a first guess for f_{n+1}, and then solve the discretized equations to machine precision using Newton's method. On the one hand, this is necessary because an explicit algorithm is unstable for these equations even for very small time steps. On the other, I want the equations to be discretized in the same way as in the relaxation algorithm when I continue the solution from $\zeta = 0$ to large ζ.

Schematically, $F(f,\zeta) = N(f,\zeta)/D(f,\zeta)$. As discussed in the main text, the SSH announces itself by $D = 0$. As a boundary condition we impose that $D = 0$ at $\zeta = 0$ for all τ. We then impose $N = 0$ as a second boundary condition at $\zeta = 0$. $\zeta = 0$ is itself the last grid point. We see from Eq. (D1) that F is never evaluated at $D = 0$. There is therefore no need to treat the singular point $\zeta = 0$ in a special way. If we want to shoot away from $\zeta = 0$, however, for example in order to prime the relaxation algorithm, we need to expand the field equations in powers of ζ. This is done in Appendix A.

The other set of boundary conditions are imposed at $\zeta = -\infty$. We do this by expanding in powers of $\exp \zeta$ (see Appendix A).

APPENDIX E: NUMERICAL TREATMENT OF THE PERTURBATIONS

I have studied the perturbations with two partially independent codes. One code assumes that λ_i is real and that the perturbations have the same symmetry ($X_\pm$ odd and a and g even) as the background. A second code allows for complex λ_i and perturbations, and the perturbations either having the same symmetry as the background, or the opposite symmetry. The two kinds of perturbations decouple because of the background symmetry, and can be considered separately.

I used the following discretization of the linearized version $d \delta f/d\zeta = \partial F(f,\zeta)/\partial f \delta f$ of the ODE system $df/d\zeta = F(f,\zeta)$:

$$\delta f_{n+1} = \left(1 - \frac{h}{2} \frac{\partial F}{\partial f}\right)^{-1} \left(1 + \frac{h}{2} \frac{\partial F}{\partial f}\right) \delta f_n. \quad \text{(E1)}$$

The matrix $\partial F(f,\zeta)/\partial f$ is extracted in Fourier components, by varying one Fourier component of δf at a time in the linearized derivatives $\partial F(f,\zeta)/\partial f \delta f$. Its coefficients are evaluated on the background taken at the midpoint between f_n and f_{n+1}, and $h \equiv \zeta_{n+1} - \zeta_n$. This scheme is implicit (and in fact stable) and second-order accurate.

For finding the eigenvalues λ_i, both codes generalize ideas already used in similar calculations [11], using the linearity of the equations. All possible linear perturbations compatible with the boundary conditions at either $\zeta = 0$ or $\zeta = -\infty$ are evolved together to a midpoint, say $\zeta = -1$. One obtains a complex square matrix ∂(mismatch at the midpoint)$/\partial$(free parameters at the end points). As a test, I have

varied the real and imaginary part of the free parameters independently. The Cauchy-Riemann equations (with respect to the free parameters) are obeyed without any numerical error.

This matrix, say A, depends on λ_i. If for a given λ_i a linear perturbation exists which obeys all boundary conditions at both end points, $\det A = 0$, and this perturbation is the eigenvector with eigenvalue zero. We therefore search for zeros of $\det A(\lambda_i)$ in the complex plane. Because λ_i^* does not appear in the equations, $\det A(\lambda_i)$ is holomorphic. Complex conjugation of A corresponds to a certain interchange of both rows and columns (the interchange of positive and negative frequencies in τ), and therefore $\det A(\lambda_i^*) = \det A(\lambda_i)^*$. I have checked both symmetries numerically. I find a relative numerical error in the Schwarz reflection principle of 10^{-10} and in the Cauchy-Riemann equations (with respect to λ_i) of 10^{-9}.

It is sufficient to consider the strip $0 \leqslant \mathrm{Im}\lambda_i < \Delta/\pi$. Because $\det A(\lambda_i)$ is holomorphic, we can use the well-known contour integral $\int A'/A\, d\lambda_i = 2\pi i (N_z - N_p)$ to count the number of zeroes minus the number of poles within the contour. Apart from the perturbations discussed in Sec. IV, I find a simple pole at $\lambda_i = -1$, for modes with the same symmetry as the background, and a simple pole at $\lambda \simeq -0.385$ for modes with the opposite symmetry. Their origin is discussed in Appendix B.

Once I have found a zero of $\det A(\lambda_i)$, I determine the zero eigenvector, and hence the corresponding perturbation field. (Numerically I find the eigenvector by singular value decomposition.) To refine this result, and to provide an independent check, I use it as the starting point of a relaxation algorithm. In this algorithm, λ_i is one of the dependent variables, and its presence is balanced by a boundary condition fixing the norm of the linear perturbation. I have implemented this algorithm only for real λ_i, as I am mainly interested in the unstable mode, which is real.

APPENDIX F: NUMERICAL ERROR ESTIMATES

For an ODE discretization to second order such as Eq. (D1) one would expect the solution to converge globally to second order in the step size. I had incorrectly claimed to observe this in [17]. The real situation is more complicated: the root-mean-squared or maximum measure of convergence show convergence that is quadratic *on the average* over a wide range of step sizes. By quadratic on the average I mean that the log-log plot of step size versus the difference of two numerical solutions with different step sizes is a somewhat wiggly line, but with an average slope of two, as illustrated in Fig. 3. This result is puzzling, because one would expect a much tighter result to hold, namely that the numerical solution of the equations at a given step size h is equal to the continuum solution plus a discretization error which itself is a power series in even powers of h only (the Richardson form). This clearly does not hold here. I believe that the irregularity is due either to the presence of boundary conditions, or in particular to the singular nature of the boundaries.

The fact that the numerical solution really obeys Eq. (D1) to machine precision is easily established, independently of the (complicated) relaxation algorithm that found the solution in the first place. Furthermore the convergence behavior
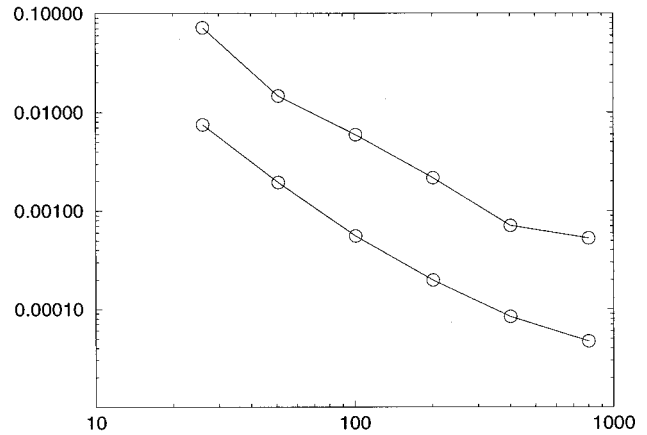

FIG. 3. Convergence of the solution from $r = 0$ to the past SSH. The horizontal axis gives the number of grid points (26, 51, …, 801). The vertical axis gives the difference between the numerical solution and the reference solution with 1601 points. The upper line shows the maximum error over all components and grid points, while the lower line shows the root-mean-square error. The upper line coincides with the error in Δ, as this is the component which has the largest error. For comparison, note that the diagonal of the box has slope 2, the expected convergence behavior.

should be independent of what is inside the (complicated) routines that supply the algebraic boundary conditions and the derivative $F(f,\zeta)$. Therefore it appears unlikely that the cause of the irregular convergence behavior is a programming error.

When a numerical error has the Richardson form, one obtains a sharp error estimate based on a theory of the numerical solution process. Here I cannot use this theory. Nevertheless, there clearly is something like quadratic convergence with decreasing step size, and so I feel justified to use the difference between the numerical solution with N grid-points and the numerical solution with $2N$ gridpoints as a measure of the numerical error of the solution with N grid-points. Adding a safety factor of 4, I take this value as the error of the $2N$ solution, and the $2N$ solution itself as the best value.

Figure 3 plots maximum and root-mean-square measures of convergence of the entire solution. As it happens, the maximum measure is dominated by the error in Δ, and so doubles as an error plot for Δ. As described above, I obtain the best value and error bar $\Delta = 3.4453 \pm 0.0005$.

Table II gives the values of the parameters Δ, λ_1, and λ_2 for various N. λ_2 is the gauge mode corresponding to a translation of the background solution, and we should find $\lambda_2 = 0$. Its absolute numerical value gives a bound on the error of λ of ± 0.009, while from the convergence of λ_1 I would have estimated a much smaller error of ± 0.0002. I opt for the larger error estimate, and obtain $\lambda_1 = -2.674 \pm 0.009$, corresponding to a critical exponent $\gamma = -1/\lambda_1 = 0.374 \pm 0.001$.

APPENDIX G: CURVATURE AT THE FUTURE SSH

The Riemann tensor in spherical symmetry in general coordinates has been calculated in [35]. Substituting the form

TABLE II. Convergence of Δ and λ with step size in ζ. Δ is the echoing period. λ_1 is the Lyapunov exponent of the one growing mode. Its negative inverse is the critical exponent γ. λ_2 is the exponent of the translation gauge mode. It must be zero and serves as a check on the numerical error. Note that the numerical grid (and number of steps) is the same for the background as for the perturbations in each case.

Number of steps	Δ	λ_1	λ_2
51	3.4321725669119	-2.6858281957399	-3.8648655101422D-02
101	3.4513015765429	-2.6700545097384	6.3598135051540D-02
201	3.4431664827331	-2.6741157762787	-2.4122047436405D-02
401	3.4446384162424	-2.6748914760819	-2.3339465009179D-02
801	3.4458770431665	-2.6738878803912	2.9009984141595D-02
1601	3.4453484479734	-2.6740958070987	-9.4268323369794D-03

(2) of the metric, it is straightforward to verify that the only appearance of any second derivatives of a or α with respect to r and t in any component of the Riemann tensor is in the combination

$$\frac{a}{\alpha} r \left(r \frac{a_{,t}}{\alpha} \right)_{,t} - r \left(r \frac{\alpha_{,r}}{a} \right)_{,r}. \tag{G1}$$

We now use the Einstein equations

$$r \frac{a_{,t}}{\alpha} = \frac{1}{2} a^2 (X_+^2 - X_-^2),$$

$$r \frac{\alpha_{,r}}{\alpha} = \frac{1}{2}(a^2 - 1) + \frac{1}{2} a^2 (X_+^2 - X_-^2), \tag{G2}$$

to transform the terms in round brackets into algebraic expressions, and we transform the outer derivatives from coordinates (r,t) to coordinates (w,ρ) via

$$g r \frac{\partial}{\partial t} = G \frac{\partial}{\partial w}, \quad r \frac{\partial}{\partial r} = \frac{\partial}{\partial \rho} - \Gamma w \frac{\partial}{\partial w}. \tag{G3}$$

The only derivative in the resulting expression that potentially does not exist is $X_{+,w}$. But it arises as

$$(G + \Gamma w)(a^2 X_+^2)_{,w}, \tag{G4}$$

and while $X_{+,w}$ blows up as $(w+1)^{\epsilon-1}$ as $w \to -1$, the coefficient vanishes as $(w+1)$, because of the coordinate condition $\Gamma = G$ at $w = -1$. We conclude that the curvature remains finite at $w = -1$ although not all second derivatives of the metric are finite there.

[1] M. W. Choptuik, Phys. Rev. Lett. **70**, 9 (1993).
[2] A. M. Abrahams and C. R. Evans, Phys. Rev. Lett. **70**, 2980 (1993).
[3] C. R. Evans and J. S. Coleman, Phys. Rev. Lett. **72**, 1782 (1994).
[4] C. Gundlach, J. Pullin, and R. Price, Phys. Rev. D **49**, 890 (1994).
[5] D. Garfinkle, Phys. Rev. D **51**, 5558 (1995).
[6] R. S. Hamadé and J. M. Stewart, Class. Quantum Grav. **13**, 497 (1996).
[7] D. Maison, Phys. Lett. B **366**, 82 (1996).
[8] T. Hara, T. Koike, and S. Adachi, ''Renormalization group and critical behavior in gravitational collapse,'' Report No. gr-qc/9607010 (unpublished).
[9] E. W. Hirschmann and D. M. Eardley, ''Criticality and bifurcation in the gravitational collapse of a self-coupled scalar field,'' Report No. gr-qc/9511052 (unpublished).
[10] E. W. Hirschmann and D. M. Eardley, Phys. Rev. D **51**, 4198 (1995).
[11] E. W. Hirschmann and D. M. Eardley, Phys. Rev. D **52**, 5850 (1995).
[12] S. L. Liebling and M. W. Choptuik, Phys. Rev. Lett. **77**, 1424 (1996).
[13] D. M. Eardley, E. W. Hirschmann, and J. H. Horne, Phys. Rev. D **52**, 5397 (1995).
[14] R. S. Hamade, J. H. Horne, and J. M. Stewart, Class. Quantum Grav. **13**, 2241 (1996).
[15] B. Carter and R. N. Henriksen, Ann. Phys. (France) Suppl. **6**, 47 (1989).
[16] T. Koike, T. Hara, and S. Adachi, Phys. Rev. Lett. **74**, 5170 (1995).
[17] C. Gundlach, Phys. Rev. Lett. **75**, 3214 (1995).
[18] M. E. Cahill and A. H. Taub, Commun. Math. Phys. **21**, 1 (1969).
[19] A. Ori and T. Piran, Phys. Rev. D **42**, 1068 (1990); T. Foglizzo and R. N. Henriksen, *ibid.* **48**, 4645 (1993).
[20] M. W. Choptuik, *Deterministic Chaos in General Relativity*, edited by D. Hobill, A. Burd, and A. Coley (Plenum, New York, 1994), pp. 155–175.
[21] M. W. Choptuik (private communication). See also [9].
[22] C. Gundlach and J. M. Martín García, Phys. Rev. D **54,** 7353 (1996).
[23] S. Hod, M.Sc. thesis, The Hebrew University in Jerusalem, December 1995; S. Hod and T. Piran, ''Fine-structure of Choptuik's mass scaling relation,'' Report No. gr-qc/9606087 (unpublished).
[24] M. W. Choptuik, T. Chmaj, and P. Bizon, Phys. Rev. Lett. **77**, 424 (1996).
[25] T. Chiba and M. Siino, ''Disappearance of critical behavior in semiclassical general relativity,'' Kyoto University Report No. KUNS 1384, 1996 (unpublished).
[26] P. Brady, Phys. Rev. D **51**, 4168 (1995).
[27] M. D. Roberts, Gen. Relativ. Gravit. **21**, 907 (1989).

[28] P. Brady, Class. Quantum Grav. **11**, 1255 (1994).

[29] Y. Oshiro, K. Nakamura, and A. Tomimatsu, Prog. Theor. Phys. **91**, 1265 (1994).

[30] J. M. Yeomans, *Statistical Mechanics of Phase Transitions* (Oxford University Press, New York, 1992).

[31] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Addison-Wesley, Reading, MA, 1992).

[32] G. I. Barenblatt, *Similarity, Self-Similarity and Intermediate Asymptotics* (Consultants Bureau, New York, 1979).

[33] C. Canuto *et al.*, *Spectral Methods in Fluid Dynamics* (Springer-Verlag, Berlin, 1988).

[34] W. H. Press *et al.*, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, England, 1992).

[35] E. Poisson and W. Israel, Phys. Rev. D **41**, 1796 (1990).

[36] C. Gundlach, ''Echoing and scaling in Einstein-Yang-Mills critical collapse,'' Report No. gr-qc/9610069 (unpublished).