Free evolution of self-gravitating, spherically symmetric waves

Mirta S. Iriondo and Oscar A. Reula
FaMAF, Medina Allende y Haya de la Torre, Ciudad Universitaria, 5000 Córdoba, Argentina
(Received 30 January 2001; published 28 January 2002)

We perform a numerical free evolution of a self-gravitating, spherically symmetric scalar field satisfying the wave equation. The evolution equations can be written in a very simple form and are symmetric hyperbolic in the Eddington-Finkelstein coordinates. The simplicity of the system allows us to display and deal with the typical gauge instability present in these coordinates. The numerical evolution is performed with a standard method of lines, fourth order in space and time. The evolution is performed using the standard Runge-Kutta method while the space discrete derivative is symmetric (nondissipative). The constraints are preserved, within numerical errors, by the evolution and we are able to reproduce several known results.

DOI: 10.1103/PhysRevD.65.044024 PACS number(s): 04.30.−w

I. INTRODUCTION

The aim of this work is to study *instabilities* which are recurrent in many numerical simulations of black hole collapse when those evolutions are completely free, that is, when only evolution equations are solved for, instead of a mixture of evolution and constraint equations. These instabilities are relevant for they preclude the computation of the final phases of the collapse of orbiting black hole systems.

The nature of these instabilities is not well understood; they do not seem to be of a numerical nature; they also lack the well posedness of the initial-boundary value problem for the system of evolution equations. Rather they are suspected to be a consequence of the exponential growth of certain gauge quantities whose propagation is poorly understood. Since no system of equations with a manifestly well posed initial-boundary value formulation has been used so far, and in many cases *ad hoc* discretization schemes had been used, the understanding of the nature of the instabilities has not been possible.[1]

Thus, to make a reliable study of these instabilities it is necessary to find a version of Einstein's equations which admits a well posed initial-boundary value formulation and to use a standard, well understood discretization scheme. To achieve this we decided to study the simple problem of the collapse of a spherically symmetric self-gravitating scalar field. Of course this is an oversimplified problem and therefore many of the findings and remedies for it might not admit generalizations to the full 3D case. Nevertheless, there are instabilities present in this simple problem and, furthermore, there are strong indications [1,2,3] that the instabilities are predominantly due to a longitudinal or Newtonian mode, namely the only one present in the spherically symmetric case.

In the past decade, due to the results of several numerical experiments, there has been a considerable understanding of the phenomena of the collapse of a spherically symmetric, self-gravitating scalar field. In particular, universal scaling properties of the final mass of the black hole were discovered and then explained. For a review and references see [4].

There are few cases in which free evolution has been successful. Some of these approaches [5–8] use light cone coordinates and in fact some of the equations being solved in these cases could be considered as constraints.[2] Others use conformal space-time methods [11,12] where no instability seems to be present. Finally, in [9,10] a symmetric hyperbolic formulation is used, and certain freedom, still available in that setting, is used to suppress the main instability found there, reaching in that way a stable propagation. Unfortunately the discretization scheme has first-order numerical dissipation and the boundary-value problem for the system is not known to be well posed. Our work can be considered as a continuation to this approach where boundary conditions are now well posed and the equation system is much simpler.

We use Eddington-Finkelstein coordinates in order to have smooth field components at the horizon. These coordinates, or suitable generalizations of them, are widely used for studying numerical collapse. In these coordinates we are able to choose equations and variables so that the final system of equations is remarkably simple and manifestly symmetric hyperbolic, in fact diagonal. Such simplicity is useful to study the nature of the instability and ways to avoid it.

To study instabilities in a reasonable way it is necessary to work with a well posed evolution system, that is, a system which is hyperbolic and for which one can impose correct, i.e., stable in the analytic sense, boundary conditions. As it is well known, there is a considerable freedom in choosing the set of evolution Einstein's equations, for it is always possible to add to any given set of equations new terms proportional to the constraints. In Sec. II we introduce the equations and discuss which evolution systems, among all possible equivalent ones, are suitable for free evolution. It turns out that, by adding terms proportional to the constraints, two out of the five characteristics of the system can be specified at will. These characteristics can even become complex, signaling the failure of the evolution equations to be hyperbolic, and so to have a well posed initial value formulation.

In the free evolution proposed one has to make sure that

[1]Note that the conformal systems are well posed (no boundary is present and the systems are symmetric hyperbolic) and there no instability seems to be present.

[2]Note that in [8] the gauge breaks at about 50 M.

the remaining Einstein's equations, namely the constraint equations, continue to hold along evolution if they hold initially. To ensure this one has to study the evolution system that the constraints satisfy whenever the free evolution equations are satisfied. It is well known that in evolutions which cover only the domain of dependence of an initial Cauchy surface it suffices to impose the constraint equations on that initial surface. Uniqueness of the constraints evolution equations then guarantees the constraints remain satisfied. In cases where the integration domain exceeds that domain of dependence, and boundary conditions are therefore imposed, constraints preservation is a much more delicate business, for its preservation depends on the boundary conditions imposed.

The change on the characteristics of the evolution system alluded to above implies an identical change on the characteristics of the evolution equations satisfied by the constraints equations. Thus, it is possible to understand that the change in boundary conditions induced by the change on the direction of the characteristics is just the one needed to ensure the correct preservation of the constraints equations under evolution. Among all possible evolution systems we take a specific one which is both the simplest for dealing with constraints preservation and for discretization purposes. We also consider other interesting systems where the point above is well illustrated.

In Sec. II A we look at the residual gauge freedom of the Eddington-Finkelstein coordinates. That freedom is the only dynamics which is present in the vacuum case, and is the mode that can generate instabilities.

After that, in Sec. II B we discuss the initial-boundary value problem for the specific evolution system chosen and determine which fields must be given in order to have a unique evolution; in particular, we discuss a case in which the characteristics change at the inner boundary according to whether that boundary is inside or outside the horizon.

We then discuss the instabilities of the system. Due to the simplicity of the evolution system chosen this task is trivial and we can display explicitly the main instability. We can decide in which cases there will be problems and of what sort. In particular near the stationary regime, that is, when most of the scalar field has dissipated away or fallen into the black hole, the system can be made stable using simple initial-boundary values for the gauge fields.

Having completed the analytic study of the equations in Sec. III we discuss the numerical methods employed. We use standard discretization methods since their numerical stability and convergence are well understood, and because we believe that to simulate a well posed, and stable system, which is well understood analytically, no fancier numerical methods are needed. We use a method of lines with a fourth-order symmetric discretization for space derivatives and a fourth-order Runge-Kutta for time integration.

Three types of initial-boundary data are considered. The first type is normal initial data with homogeneous boundary data, the type of data normally used for collapse studies. The second type of data is pure gauge boundary data, the space-time is just Schwarzschild in nonstandard gauge. Finally the third type of data is pure incoming scalar field boundary data.

We perform the standard convergence tests for all types of data. They confirm the order of the discretization: as long as the bulk of the field amplitudes are away from the boundaries the method is fourth order. When amplitudes are important at the boundaries the order is lower since the space derivatives is not fourth order there, but just third order.

We check constraint preservation under evolution for several types of initial-boundary data sets. The constraint quantities are very small at the initial time, but then grow a bit and stay bounded along the whole evolution. As expected, the growth of the constraint quantities diminishes as the step sizes diminish.

Finally, in Sec. IV we comment on several results obtained by the simulations; they do not pretend to be novel nor interesting, but are done just to confirm the good properties of the evolution system chosen and to confirm our expectations mentioned above about the success of a standard numerical method. These results include ringing and tail studies for the scalar field, the relation mass vs flux for boundary data, and the relation total mass vs the black hole mass for different types of collapse.

II. THE EQUATIONS

The most general spherically symmetric metric is

$$ds^2=(-\alpha^2+a^2\beta^2)dt^2+2a^2\beta dtdr+a^2dr^2+r^2b^2d\Omega^2,$$

with a, b, α, and β being functions of r and t, and $d\Omega^2$ the metric of the unit sphere. We partially fix the gauge freedom by requiring these coordinates to be of the type Eddington-Finkelstein [13]. Namely, we require (a) the areal locking condition ($\dot{b}=0$), and (b) the condition that the vector $\partial_t-\partial_r$ be null. Both conditions are achieved by choosing the lapse and the shift as

$$\beta=\frac{raK^{\theta}{}_{\theta}}{1+raK^{\theta}{}_{\theta}}, \qquad \alpha=\frac{a}{1+raK^{\theta}{}_{\theta}},$$

where $K^{\theta}{}_{\theta}$ is the corresponding coordinate component of the extrinsic curvature of the constant time hypersurfaces. Moreover, rescaling the coordinate r at some initial surface we can set $b=1$. Thus, the metric becomes

$$ds^2=a^2[(2\beta-1)dt+dr](dt+dr)+r^2d\Omega^2.$$

In this spherically symmetric space-time we consider a massless scalar field ϕ, satisfying the wave equation. In the $3+1$ decomposition this equation becomes the first-order (in time) system

$$\partial_t\Phi=\partial_r[\beta\Phi+(1-\beta)\Pi],$$

$$\partial_t\Pi=\frac{1}{r^2}\partial_r\{r^2[\beta\Pi+(1-\beta)\Phi]\}, \tag{1}$$

where $\Pi=(a/\alpha)(\partial_t\phi-\beta\partial_r\phi)$ and $\Phi=\partial_r\phi$.

The full set of evolutions and constraint equations for the geometric variables $(a, K^{\theta}{}_{\theta}, K^{r}{}_{r})$ are[3]

$$\partial_t a = \partial_r(a\beta) - a^2(1-\beta)K^r{}_r,$$

$$\partial_t K^{\theta}{}_{\theta} = \beta\partial_r K^{\theta}{}_{\theta} + \frac{\partial_r\beta}{ar} + a(1-\beta)K^{\theta}{}_{\theta}(K^r{}_r + 2K^{\theta}{}_{\theta}) + \frac{1-\beta}{r^2}\left(a - \frac{1}{a}\right),$$

$$\partial_t K^r{}_r = \beta\partial_r K^r{}_r + \frac{\partial_r^2\beta}{a} + \frac{\beta-1}{a}\left[\frac{\partial_r^2 a}{a} - \left(\frac{\partial_r a}{a}\right)^2 - \frac{2\partial_r a}{ra}\right] + \frac{\partial_r\beta\partial_r a}{a^2} + a(1-\beta)K^r{}_r(K^r{}_r + 2K^{\theta}{}_{\theta}) + \frac{8}{a}\Phi^2(\beta-1), \tag{2}$$

and

$$C_a = \partial_r a + \frac{1}{2r}(a^3 - a) + \frac{a^3 r}{2}K^{\theta}{}_{\theta}(2K^r{}_r + K^{\theta}{}_{\theta}) - 2ra(\Phi^2 + \Pi^2) = 0,$$

$$C_{K^{\theta}{}_{\theta}} = \partial_r K^{\theta}{}_{\theta} + \frac{K^{\theta}{}_{\theta} - K^r{}_r}{r} - \frac{4}{a}\Phi\Pi = 0, \tag{3}$$

respectively.

In order to obtain first-order evolution equations, we introduce new variables $(f, g, h, \chi_+, \chi_-)$, which are related to the former ones as follows:

$$f = K^{\theta}{}_{\theta}\, ar + 1, \quad g = \frac{a^2}{f}, \quad h = \frac{1}{af}\left(\frac{a}{2}\partial_r f + \partial_r a - a^2 K^r{}_r\right),$$

$$\chi_+ = \frac{(\Phi + \Pi)r}{2}, \quad \chi_- = \frac{(\Phi - \Pi)r}{2}.$$

With this change the shift becomes $\beta = (f-1)/f$ and

$$ds^2 = g[(f-2)dt + fdr](dr + dt) + r^2 d\Omega^2.$$

Then the full set of evolution and constraint equations for the metric variables (f, g, h) are

$$\partial_t f = \frac{\partial f}{\partial r} + \frac{g - 2 + f}{r},$$

$$\partial_t g = -\frac{g}{f}\frac{\partial f}{\partial r} + \frac{f-2}{f}\frac{\partial g}{\partial r} + 2gh - \frac{g(g-2+f)}{rf}, \tag{4}$$

$$\partial_t h = \frac{\partial h}{\partial r} - \frac{1}{2fr}\frac{\partial f}{\partial r} + \frac{g-2}{2fhr}\frac{\partial g}{\partial r} - \frac{(g-f)h}{fr} - \frac{(g-2+f)}{2fr^2} + 2\left(\frac{\chi_+ + \chi_-}{fr}\right)^2,$$

and

$$C_f = \frac{\partial f}{\partial r} - 2(f-2)h + \frac{f+g-2}{r} - 4\frac{\chi_+^2}{fr} = 0,$$

$$C_g = \frac{\partial g}{\partial r} - \frac{2g(rfh - \chi_+^2 + \chi_-^2)}{fr} = 0, \tag{5}$$

while the wave equation for the massless scalar field (χ_+, χ_-) becomes

$$\partial_t\chi_+ = \frac{\partial\chi_+}{\partial r} - \frac{(f-2)\chi_-}{fr},$$

$$\partial_t\chi_- = \frac{f-2}{f}\frac{\partial\chi_-}{\partial r} + \frac{2\chi_-}{f^2 r}[2r(f-2)h + 2 - g - f] + \frac{8(\chi_-)\chi_+^2}{f^3 r} - \frac{\chi_+}{r}, \tag{6}$$

where we have already used one of the constraint equations to substitute a spatial derivative of f by terms without derivatives.

Clearly the evolution equations for the scalar field are symmetric hyperbolic as a subsystem and their characteristics are fixed, they are physical and do not change when the evolution equations are changed, and they correspond to the two null directions. On the other hand, the set of evolution equations for the metric coefficients is not unique, since we can add to them terms proportional to the constraint equations and change some of their characteristics at will. To study these possibilities we add to the evolution equations for f, g, and h, respectively, the following terms $(K_{ff} - 1)C_f + K_{fg}C_g$, $(K_{gf} + g/f)C_f + \{K_{gg} - [(f-2)/f]\}C_g$, and $\{K_{hf} - [(g-2)/2frh]\}C_f + [K_{hg} + (1/2fr)]C_g$. We get a system with the following principal part matrix:[4]

$$\begin{pmatrix} K_{ff} & K_{fg} & 0 \\ K_{gf} & K_{gg} & 0 \\ K_{hf} & K_{hg} & 1 \end{pmatrix}.$$

[3]Notice some corrections with respect to the equations in [13]; also notice the difference in the choice of gravitational constant.

[4]Recall that the principal part matrix is that one formed by the coefficients of the derivative terms appearing in the equations.

The characteristics of this system are given by the one forms (1, 1), $(1,\lambda_+)$, and $(1,\lambda_-)$, where

$$\lambda_\pm=[K_{ff}+K_{gg}\pm\sqrt{(K_{ff}-K_{gg})^2+4K_{fg}K_{gf}}]/2.$$

The first one corresponds to the propagation of h and it does so along the null incoming direction. It is the only fixed characteristic. The other two can take any values. For instance, if we take $K_{fg}=K_{fg}=0$ then the characteristics are $(1,K_{ff})$, and $(1,K_{gg})$. Thus, choosing these two values we can prescribe any propagation direction we please. They can be incoming, outgoing, timelike, null or even spacelike. Even more, choosing $K_{ff}=K_{gg}=0$, $K_{fg}=-K_{gf}\neq 0$ the characteristics become imaginary and so the system is not even hyperbolic. So we see that the different ways of writing Einstein's equations are only equivalent in the sense that solutions of one are solutions of the others, but some of the resulting evolution equations do not even have a well posed initial value problem. Just for amusement we have tried to evolve numerically a system as the last one (namely, with $K_{fg}=-K_{gf}=1$) and as expected it is completely unstable, giving an explosive enlargement of higher wave number modes. To avoid the fact that no good boundary conditions are known for nonsymmetric hyperbolic systems this system was tested using periodic boundary conditions. The initial data imposed did not satisfy the constraint equations. In any case the *explosive type* instability found for ill-posed systems does not seems to be sensitive to the failure of the constraint equations to be fulfilled nor to the particular boundary values chosen, both global properties, since this instability appears locally, at every point of the evolution hypersurface.

We must, therefore, choose coefficients such that the system is hyperbolic. Since the coefficients K_{hf}, K_{hg} do not alter the characteristics we shall take them to vanish. This simplifies a bit the numerical computations. The sign of the characteristics at the boundary points determines whether or not values for the different fields must be given or not as boundary conditions. However, not all these values are free, since they must be consistent with constraint propagation. In order for the constraints to remain enforced along evolution we must make sure that the resulting boundary conditions for the evolution equations obeyed by the constraints have unique solutions, for given initial and boundary data, and it is the trivial one when the appropriate initial-boundary data vanishes. It easily follows from the form of the constraints that the characteristic matrix for the constraints evolution equations is

$$\begin{pmatrix} K_{ff} & K_{fg} \\ K_{gf} & K_{gg} \end{pmatrix}.$$

So the propagation directions for them coincide with the propagation directions of the pair (f,g). This fact seems to imply that the freedom in choosing part of the pair (f,g) at the boundaries is just the one needed for getting maximally dissipative boundary conditions [14] for the constraint system, and so ensuring the correct constraint propagation. Unfortunately, even for this simple case it seems one needs to impose boundary conditions involving some combination of space derivatives of f and g, but there is no rigorous proof (not even a sound argument) that space derivatives can be given as boundary data and still retain well posedness.[5] In most cases, the resulting conditions are complicated to implement numerically.[6] Thus, it seems that the best one can do retaining simplicity is to choose coefficients so that the characteristics for the pair (f,g) are incoming or tangent in both boundaries. In this way no boundary condition is needed for the pair (f,g) and therefore no boundary condition must be satisfied by the constraint, since the characteristics for the constraints evolution equations would also be outgoing. Thus, there is no danger that the constraints would cease to be satisfied during evolution. The simplest case of all is when both characteristics are tangent to the boundary and this is the case, for instance, if we make all K's vanish. With this choice the evolution system reduces to

$$\begin{aligned}
\partial_t f &= 2(f-2)h+\frac{4\chi_+^2}{rf},\\
\partial_t g &= 2gh-\frac{2g\chi_+^2}{fr}+\frac{2g(f-2)\chi_-^2}{rf^2},\\
\partial_t h &= \frac{\partial h}{\partial r}-\frac{(g-2)\chi_+^2}{f^2r^2}+\frac{4\chi_+\chi_-}{f^2r^2}+\frac{g\chi_-^2}{f^2r^2}.
\end{aligned} \tag{7}$$

Similarly the evolution equations for the constraints are just ordinary differential equations,

$$\partial_t C_f=\frac{2}{rf^2}(rf^2h-2\chi_+^2)C_f, \tag{8}$$

$$\begin{aligned}
\partial_t C_g=&\frac{2g}{rf^3}[f\chi_+^2-(f-4)\chi_-^2]C_f+\frac{2}{rf^2}[rf^2h-f\chi_+^2\\
&+(f-2)\chi_-^2]C_g.
\end{aligned} \tag{9}$$

Since the constraints are homogeneous ordinary differential equations, they will remain satisfied along evolution if they are initially satisfied. No extra condition is need at (nor can be imposed on) the boundaries. This is the system we shall study in detail from now on. Note that this system not only is a symmetric hyperbolic, but it is also diagonal. In fact, we reach such a simple system when looking for variables which would give diagonal systems. In Fig. 1 we show all characteristics of the chosen system.

We have also considered the following equivalent system:

$$\partial_t f=2(f-2)h+\frac{4\chi_+^2}{rf},$$

[5]In simple cases one can trade, using the evolution equations, those space derivatives by time derivatives. Still, even in those cases, difficulties arise when the K's depend on the metric components.

[6]See, nevertheless the second case considered below.

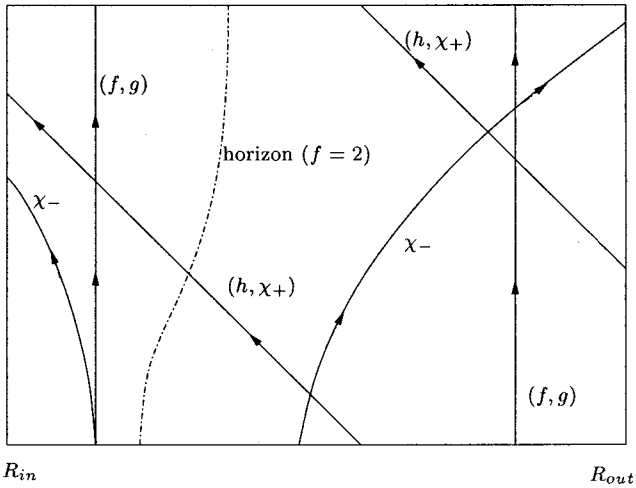

FIG. 1. Integration region with characteristics lines.

$$\partial_t g = \frac{\partial g}{\partial r} - \frac{4 g \chi_-^2}{f^2 r}, \tag{10}$$

$$\partial_t h = \frac{\partial h}{\partial r} - \frac{(g-2)\chi_+^2}{f^2 r^2} + \frac{4\chi_+\chi_-}{f^2 r^2} + \frac{g\chi_-^2}{f^2 r^2}.$$

In this case the characteristic diagram is similar to the one in Fig. 1 except that now the characteristic for g coincides with the ones of h, and χ_-.

The constraint evolution equations are now

$$\partial_t C_f = \left(2h - \frac{4}{r f^2}\chi_+^2\right) C_f + \frac{1}{r} C_g,$$

$$\partial_t C_g = \frac{\partial C_g}{\partial r} + \frac{2g}{r f^3}[f\chi_+^2 - (f-4)\chi_-^2] C_f - \frac{4}{f^2 r}\chi_-^2 C_g.$$

So the characteristic for C_f is still along $r=\text{const}$ but now the one for C_g is along $t+r=\text{const}$. The latter is an incoming characteristic at the right boundary and so in order for the constraint to remain enforced we must find boundary conditions which set $C_g=0$ at the R_{out} boundary. Substituting the normal derivative appearing in C_g, $\partial g/\partial r$, from the evolution equation for g we find that the correct boundary condition to impose is

$$\partial_t g = 2hg - \frac{2g}{f^2 r}[f\chi_+^2 - (f-2)\chi_-^2] \quad \text{at } R_{\text{out}}=0. \tag{11}$$

Numerical simulations show that with this boundary condition the constraints propagate correctly and so they vanish (within numerical errors) along the whole evolution. Notice that since g cannot be prescribed arbitrarily the boundary value freedom for both systems is the same.

A. The gauge freedom

The chosen lapse and shift functions depend on the dynamical variables and therefore do not fix completely the gauge. To display the remaining freedom we introduce the null coordinate $u=(t+r)$.[7] In (u,r) coordinates the metric becomes

$$ds^2 = g[(f-2)du + 2dr]du + r^2 d\Omega^2.$$

For Schwarzschild the choice $g=1$, $f-2=-(1-2M/r)$ gives the standard form of the metric. If we choose some other coordinate $\tilde{u}=\tilde{u}(u)$ the functional form remains the same if $\tilde{g}=g/(d\tilde{u}/du)$ and $(\tilde{f}-2)/g=(f-2)/g$.

The invariant ratio $(f-2)/g$ is related to the mass function, indeed,

$$m(r,t) = \frac{r}{2}\left(1 + \frac{f-2}{g}\right)$$

is the usual form of the local mass for spherical symmetric space-times. Another invariant is the surface where f takes the value 2, this can be seen to be an apparent horizon.

From the first constraint equation (5), it follows that under this gauge transformation, in the vacuum case,

$$\tilde{h} = h + \frac{1}{2}\frac{d\bar{u}/du}{d^2\bar{u}/du^2}.$$

Thus to fix completely the gauge we must uniquely define the coordinate u. This can be done by choosing their values at the initial time slice and at the outer boundary of the integration region. We shall do this by choosing the values for h at those points.

B. Initial-boundary value problem

We are interested in performing a time evolution between two fixed (areal) radiae, R_{in}, R_{out}, starting at a given initial surface, $t=t_0$. Thus we have to see what initial-boundary values can be prescribed.

With the standard choice of Eqs. (6) and (7) it is clear that we need to give initial values for the fields (f,g,h,χ_+,χ_-). Since they must satisfy the constraints not all initial data can be freely specified. We have taken a set of initial data sets for which the constraints are easily solved. We give an arbitrary initial value for χ_+ and solve

$$\frac{\partial m}{\partial r} = {\chi_+}^2, \quad m|_{r=R_{in}} = m_0.$$

Then, the following choice is a solution of both constraint equations:

$$\chi_- := 0, \tag{12}$$

$$g := 1,$$

$$f := 1 + 2m/r, \tag{13}$$

$$h := \frac{{\chi_+}^2}{rf}.$$

[7] We thank M. Tiglio for pointing out to us this way to proceed.

In our numerical simulations we have taken χ_+ to be a Gaussian or a power of r times an exponential decay, so that the mass integration could be done exactly. The integration constant for m, the mass at R_{in}, can take any value. In general we have taken either[8] $m|_{R_{in}}=1$ or $m|_{R_{in}}=0$.

We now discuss the boundary conditions (see Fig. 1). If the inner boundary is inside the apparent horizon (which is located at $f=2$) there is no incoming characteristic at it and so no boundary condition is needed (nor can be prescribed) there. Note that in this case this boundary is actually spacelike. In fact, inside the horizon the evolution equations for f and g can be considered as constraint equations on the spacelike hypersurfaces of constant t. If the inner boundary is outside the horizon there is one characteristic incoming into the integration region, namely the one for χ_-, so we have to prescribe some value for it. To allow for an apparent horizon marching across the inner boundary, the numerical code checks for the value of f at that boundary and, whenever its value is smaller than 2, a boundary condition for χ_- is enforced. We have normally used a null incoming radiation condition, $(\chi_-|_{R_{in}}=0)$ but it is clear that in that case there is no natural boundary condition which can mimic the physical collapse one is trying to simulate. So in general we have concentrated in simulations where the inner boundary is inside the apparent horizon and so it is spacelike.

At the outer boundary, and provided the apparent horizon is inside the integration region, $(f|_{R_{out}}<2)$, there are two incoming characteristics, one corresponding to χ_+ the other to h. We can prescribe arbitrary values for them there, and we have performed several runs prescribing either incoming scalar field radiation, $(h=0)$, or pure gauge modes $(\chi_+=0)$.

With these initial-boundary value conditions the problem is well posed and so, given smooth data, we obtain a smooth solution valid for a finite time interval.

For the alternative system, Eq. (10), the only difference on the initial-boundary value problem is that one can now prescribe the value of g at the outer boundary. But that value cannot be given arbitrarily, otherwise the constraints would not propagate correctly. At that boundary one has to prescribe[9] g as given in Eq. (11).

C. Analytic instabilities

The evolution system is well posed, so there are no instabilities of the explosive type, that is, those whose growth rate increases unboundly with the frequency. Rather, the expected instabilities must grow at most exponentially in time with an exponent whose positive real part does not grow with the frequency. These instabilities essentially affect the pair (f,g) which controls the characteristics of the system. If they are strong, or act for a long enough time they can ruin the hyperbolicity of the system—by causing some of the propagation speeds to diverge—and so its well posedness. But much before that happens one of the characteristic speeds grows so much that the system becomes numerically unstable due to violations of the Courant-Friedrichs-Lewy stability condition.

For simplicity we analyze the instability for the vacuum case. In that case the equations are

$$\partial_t f = 2(f-2)h,$$
$$\partial_t g = 2gh, \tag{14}$$
$$\partial_t h = \frac{\partial h}{\partial r}.$$

Given the initial data for (f,g,h) and the boundary data at the outer boundary for h, we can first solve for h, $h(t,r)=h_b(t+r-R_{out})$ and then for (f,g). We have

$$f(t,r)-2=[f_0(r)-2]e^{2\int_0^t h_b(\tilde{t}+r-R_{out})d\tilde{t}},$$
$$g(t,r)=g_0(r)e^{2\int_0^t h_b(\tilde{t}+r-R_{out})d\tilde{t}}.$$

Thus, we see that g remains always positive, and the horizon, the radius at which $f(t,r)=2$, remains fixed. But the equal time surfaces tilt, and so if h takes big positive values, f and g grow exponentially and we have an analytical instability, which in numerical simulations would result in a numerical one if no appropriate methods are used. If a scalar field is present, no matter how small, the tilt of the equal time surfaces would increase the propagation speed of the outgoing modes, $v=(f-2)/f$. At points inside the horizon $f_0-2>0$ and therefore at those points f grows exponentially, and the propagation speed v goes to 1. At points outside the horizon $f_0-2<0$ and therefore f will eventually become negative. Notice that when f vanishes the equal time hypersurface becomes null and so the time evolution of all outgoing modes (in this case only the one from the scalar field) is ill posed in this gauge.

On the other hand, if h becomes large but negative, $f\to 2$ and now the hypersurface of constant r approaches null surfaces. In that case the evolution still can be done, but the loss of accuracy is very important, even for moderate values of h.

We need therefore to keep h moderately small during evolution and propagating inwards without any important residual part left over due to the inaccuracies of the numerical method. This can be achieved in many cases by just prescribing the initial-boundary conditions for h so that it starts small, in that case, and for the choice of variables we made, the residual part stays small.[10] There are, nevertheless, situations for which h becomes big and the simulation runs into

[8]Because of the scaling properties of the system there is no loss of generality in choosing only these two values.

[9]Note that we give the time derivative of g as a boundary condition; this is very convenient, and all that is needed, at the level of the numerical code.

[10]This is not automatic, and indeed for some other choices of variables it happens that the residual part starts to grow in time, unleashing the instability. Since typically the residual part is of high frequency, numerical dissipation can also be used to keep it small.

the instabilities. This is so, for instance, if one takes boundary data for the incoming component of the scalar field with a very big amplitude (final masses of order 10). In this case the h generated along the evolution of the scalar field is important and the systems become unstable. On the other hand, for situations describing the final stages of collapse (after most of the scalar field has fallen into the black hole and the geometry has settled in a near stationary regime) the instabilities do not show up. We have tested that with evolutions lasting for several thousand masses.

III. NUMERICAL METHODS

To perform the simulations we use a standard fourth-order method of lines, see for instance [14]. The space derivatives discretization is done with standard fourth-order (we have also used second-order discretizations with similar results) discretized derivatives which are one sided at the boundaries. We have used symmetric, and also nonsymmetric, discretizations. The symmetric ones [14,15] are guaranteed to be stable, but are only second- or third-order accurate at the last points of the grid near the boundary. This accounts, for instance, for spikes at the error on the constraints propagation check, some of these errors sometimes propagate inwards, but stay bounded and diminish as the grid spacing is decreased. The nonsymmetric discretizations used were full fourth order (but not uniformly, with bigger error coefficients near the boundary) and they give smaller errors on constraints evolution near the boundaries. But there is no guarantee that the scheme is stable. Most of the calculations, including all the results and tests given in this paper, were done with the symmetric ones. Their symmetry implies they are norm preserving and so they do not have any dissipation. Fourth-order centered differences approximations introduce high frequency modes with backwards propagations speeds (group velocities) which are 5/3 bigger than the exact speeds. These modes, although very small in amplitude, are important when analyzing the tail behavior of the scalar field, since it also become very small. The higher than light propagation velocity of them implies that boundary effects take place before they should. These high frequency modes become smaller as the grid size is decreased, thus tail decay studies are still possible if enough grid points are included, but this should be very costly in higher dimensions. For these studies it seems better to use centered difference schemes of second order, since they have a maximum group velocity of one.

Time discretization is done using the standard fourth-order Runge-Kutta method.

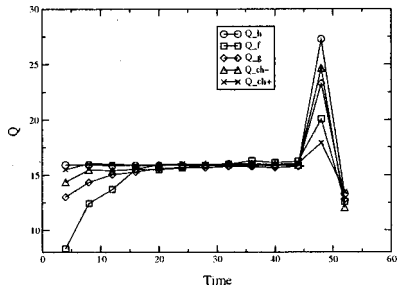

FIG. 2. Q factor for L^2 norms of type-I data.

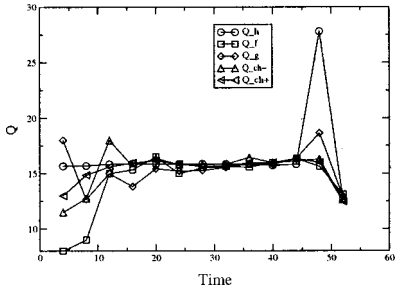

FIG. 3. Q factor for L^∞ norms of type-I data.

We have done three types of simulations: The first type of simulation has no-incoming boundary conditions ($\chi_+ = h = 0$ at the outer boundary) and an initial data satisfying Eq. (12). One starts with a solution representing an inner black hole (usually of unit mass) and a certain amount of in-falling scalar field. The evolution should result in a bigger black hole and some amount of the scalar field going outwards. Thus, a ringing and later a power law tail should be observed in the outgoing component of the scalar field, χ_-.

The second type of data are vacuum solutions representing pure gauge dynamics. They have Schwarzschild or Minkowski initial data and the dynamics is created by allowing the incoming gauge $[h|_{R_{out}} = h^b(t)]$ to enter the outer boundary. The evolution should have some dynamics and later on, if the incoming gauge has a finite duration, should relax into a static black hole in a different gauge. Throughout the simulation the mass should stay constant, as should the horizon position.

The third type of data starts with initial data corresponding to a Schwarzschild black hole of unit mass and later on the evolution an incoming scalar field mode is injected into the system from the outer boundary during a finite amount of time $[\chi_+|R_{out} = \chi_+^b(t)]$. The evolution should create a bigger black hole and some outgoing radiation should be seen. Also for this case a ringing and later a power law tail should be observed in the outgoing component of the scalar field, χ_-.

We shall discuss the results for these three types of simulations in the next section.

A. Convergence test

We have performed convergence tests for the three types of simulations described above. In general it is difficult to analyze the convergence order in detail due to the fact that the discrete space derivative we use is not of the same order near the boundary (third or second order rather than fourth). This gives rise to Q values[11] different than the corresponding to fourth order ($Q = 16$). Typically after boundary effects become important (that is, when the fields reach the bound-

[11]Recall that Q is defined as $Q(\phi) = [\|\phi(h) - \phi(h/2)\|]/[\|\phi(h/2) - \phi(h/4)\|]$, where $\phi(h)$ is a solution to the discretized system using a step size of h. We have used for the norm both the L^2 and the L^∞ ones. Error propagation theory indicates that $Q = 2^n$, where n is the order of the approximation scheme.

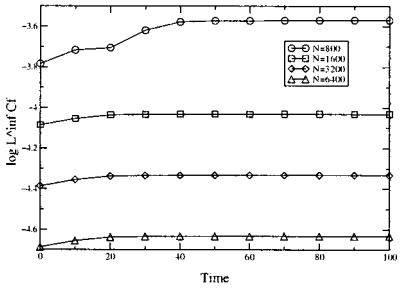

FIG. 4. Logarithm of the L^∞ norm of C_f of type-I data.

ary) we get values over 12, but sometimes they drop to about 4 when the bulk of the solution is near the boundary. In Figs. 2 and 3 we give values for the Q factors obtained using the L^2 and L^∞ norms, respectively, for a run with homogeneous boundary conditions and a Gaussian peak of the in-going component of the scalar field midway in the integration region. This run has an initial black hole mass $M_{Initial}=1$, a total mass, $M_{Total}=8.90$ and a final black hole mass of $M_{Final}=8.67$ so it is well into the nonlinear regime. The peak observed at the end of the integration period is due to boundary effects, after that the values of the different Q's go to values above 12, but never return to 16. The peak is probably due to a larger coefficient in the expansion over the grid size of the numerical solution. To resolve this we should go to smaller grid spacings, but then our method of collecting data (in single precision) produces mostly truncation errors on these tests.[12] The subsequent drop on Q to values of the order of 12 is due to the fact that our discrete derivative operators are not fourth order near the boundary, but just third order. These Q values were obtained using step sizes corresponding to 800, 1600, and 3200 grid points. For smaller step sizes fourth-order methods give solution differences which are mostly rounding errors and so the Q does not make sense.

IV. RESULTS

A. Constraints evolution

The constraint equations are not solved for in any step of our simulations, and although they are satisfied along evolution for the exact equations here we have to check this happens at the discrete level. Figures 4–9 show the L^∞ and L^2 norm of the constraint expressions along evolution of the different types of data set used. For type-I data we had $M_{Initial}=1$, $M_{Total}=1.15$, $M_{Final}=1.12$. For type-II data we introduced at the outer boundary 10 cycles of gauge mode of frequency 2π, starting at $t=10$. The black hole mass was of course constant and had the value 1. For the type-III data we introduced at the outer boundary 10 cycles of the χ_+ mode with frequency 2π, starting at $t=10$. We had $M_{Initial}=1$, $M_{Total}=2.6$, $M_{Final}=2.4$.

[12]Calculations are done with double precision, but output data are stored in single precision (ten significant figures), so some of the convergence tests must be done with limited grid point numbers in order that the subtractions involved remain significant.

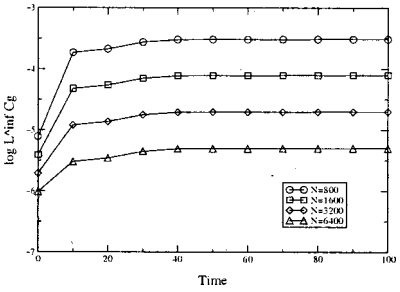

FIG. 5. Logarithm of the L^∞ norm of C_g of type-I data.

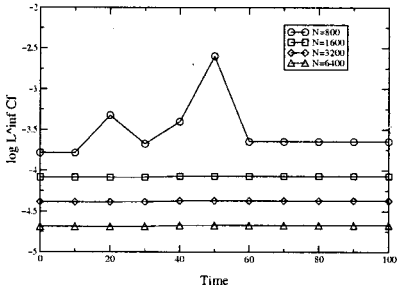

FIG. 6. Logarithm of the L^∞ norm of C_f of type-II data.

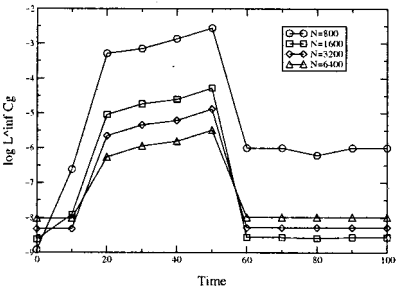

FIG. 7. Logarithm of the L^∞ norm of C_g of type-II data.

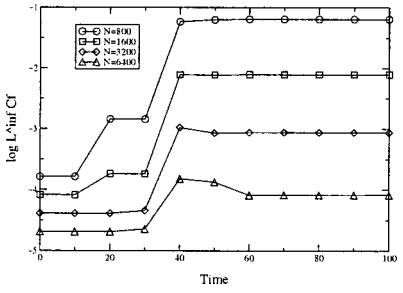

FIG. 8. Logarithm of the L^∞ norm of C_f of type-III data.

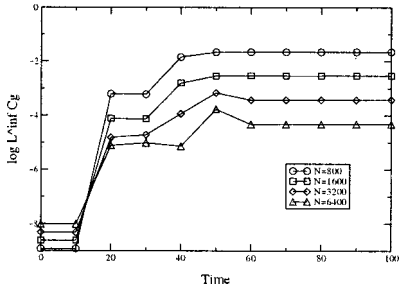

FIG. 9. Logarithm of the L^∞ norm of C_g of type-III data.

It is clear that the constraint quantities remain bounded along evolution and that they go to zero as the grid size diminishes. Most of the contribution to the L^∞ norm comes from the boundary, where the derivative operator used is only second or third order. These peaks near the boundary can be reduced to the same level as the rest if one uses a discrete derivative operators which are fourth order at all grid points.

If one looks at the ratios $C_f/(df/dr)$, $C_g/(dg/dr)$, which in some sense measures better the failure of the constraint to be satisfied, one sees that these ratios do not change very much along evolution. For C_f the change is of only about 7%, and is of the order of 10^{-4}. For C_g, since g is chosen initially to be zero, the change at the beginning is rather big, but after a transient it gets to a plateau of the order of 10^{-4}.

We have performed long-time simulations of the order of 1000 crossing times for type-I initial data and the constraints remain constant (after the initial rise seen on the above plots) for the whole evolution. In these long runs, with grid spacing of the order of one-tenth of the mass scale, the constraints stabilize to a value (in the L^∞ norm) of 6.0×10^{-3}.

B. Decay of the scalar field, ringing, and power law decay

We reproduce standard results on ringing and tail (power law decay) for the first type of simulation. In most of the runs $\Delta t/\Delta x \approx 2$, so we just quote the value of Δx, or rather $\Delta x/M$, where M is the initial mass. This value of $\Delta t/\Delta x$ does not violate the Courant-Friedrichs-Lewy condition because we use Runge-Kutta for time evolution.[13] In Fig. 10 the logarithm of the absolute value of χ_- at $r=25.5$ is plotted with respect to time. It is a run where the outer boundary is at $R_{\rm outer}=6000$, the space grid was $\delta x/M=0.0234$. The initial scalar field is given by a Gaussian χ_+ pulse centered at $r=10$, giving the Initial black hole mass $M_{\rm Initial}=1$, the total mass $M_{\rm Total}=2.56$, final black hole mass $M_{\rm Final}=2.45$. At the beginning the characteristic ringing is seen and then a power law decay of the type t^α (this can be seen best plotting the local decay power $t[d\ln(\chi_-)/dt]$ as defined in [7,18], Fig. 11). The power law decay computed is $t^{-4.01}$ (at $t=7200$), which agrees quite well with the one expected for linear perturbation theory, namely t^{-4}, see [16,17]. In the plot we also show a lower resolution ($\Delta x/M=0.0468$) computation where it is clear that an instability starts to appear and become important at about half the time. This instability diminishes when the grid resolution is incremented, but still seems to be present at longer times. At the resolution used, a feature departing from the power law decay is seen after $t=7200$. This corresponds to a higher frequency mode propagating backwards at 1.66 the speed of light. These high frequency modes are common on all centered fourth-order methods. Their amplitude decreases with increased precision. The time of arrival of that feature implies it originates near the inner boundary probably due to imperfect in-going boundary conditions. After that feature, we again see (not shown in the figure) a even bigger feature. This one propagates at the correct speed of the problem, but its presence at that time implies that it was generated by the one traveling at higher speed when it enters the region near the outer boundary.

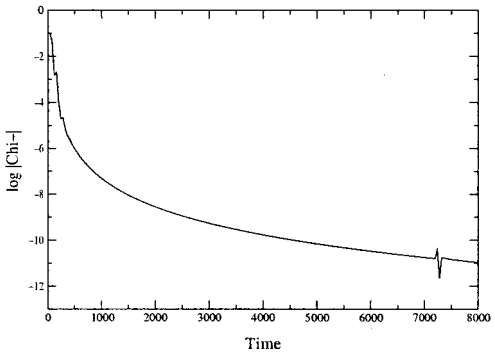

FIG. 10. Decay of χ_- at $r=25.5$.

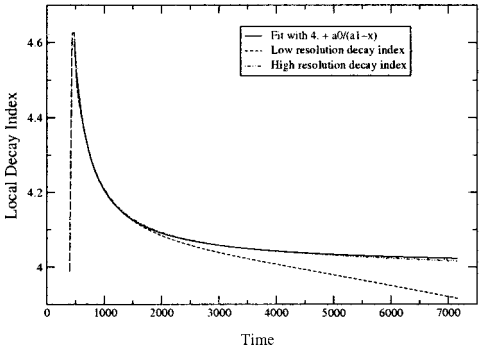

FIG. 11. Power law decay of χ_- at $r=25.5$.

The possibility of giving boundary conditions which would automatically satisfy the constraint equations allow us to perform several interesting runs. In particular we made a couple of runs of type III with

$$\chi_+|R_{out}=A_+[1-\cos(\omega_+ t)] \quad t\in[10,20]$$

with $A_+=1/2\pi$, and in one case $\omega=2\pi$, in the other $\omega=2\pi/10$. For these two boundary data we have

$$\int_{10}^{20}\chi_+^2\,dt=15A_+^2.$$

[13]For Runge-Kutta the Courant-Friedrichs-Lewy stability condition is $\Delta t/\Delta x<2.06$ [14].

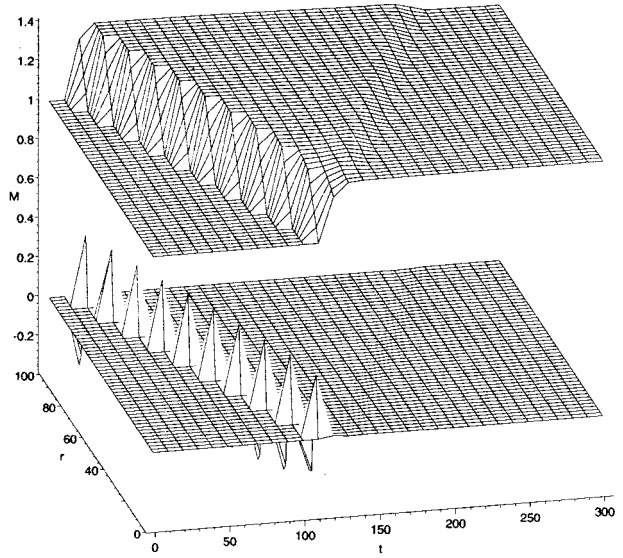

FIG. 12. Mass of type-III run and mass difference between runs augmented five times.

Since the energy flux is given by

$$\frac{dm}{dt}=\frac{1}{g}\left(\chi_+^2-\frac{(f-2)^2}{f^2}\chi_-^2\right).$$

and $dg/dt\approx(2g/fr)\chi_+^2$, $df/dt\approx(4/fr)\chi_+^2$, $d\chi_-/dt\approx(-1/r)\chi_+$. We see that to a 1% accuracy the flux of both solutions should coincide and so the black hole masses should be similar. Even more, the ringing of both solutions should be very similar, for it mostly depends on the geometry. In Fig. 12 we see the mass of one of the simulations and the mass difference between both simulations multiplied by 5. We see that most of the difference is bounded to the region where the wave of scalar field is moving, leaving behind the same geometry. In Fig. 13 we see one of the ringings (the value of χ_- at $r=50$) and the difference between the ringing of both solutions augmented by a factor 10. Again we see that most of the difference is at the moment where the two different wave fronts pass the point and then on the precise location of the maxima of the ringing.

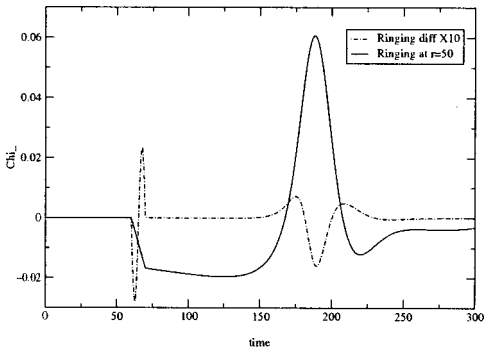

FIG. 13. Values of χ_- at $r=50$ for a type-III run, bellow the difference between runs augmented 10 times.

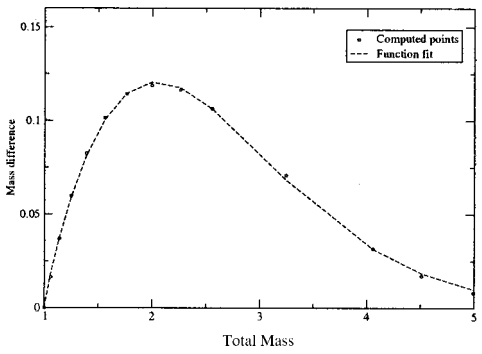

FIG. 14. ΔM vs M for type-I data, $\Delta M=0.21*(M-1)^{0.86}e^{-0.55(M-1)^{1.47}}$.

C. Relation between the initial mass and the black hole mass

We have tested the relation between the mass gap given by the total initial mass minus the final black hole mass as a function of the total initial mass for type-I and type-III initial-boundary data sets. In Figs. 14 and 15 we display the results. In both cases we start with a black hole of mass unit, and put our inner boundary inside its horizon.

For type-III data, namely data given at the outer boundary, we cannot construct very big black holes, the problem being that the scalar field injected the produce h field and it leads to instabilities.

V. CONCLUSIONS

It is difficult to see whether some of the ideas presented in this work can be extended to the full three-dimensional case, where many more variables are present and where there is no preferred center in which to anchor a gauge as the one used here. But the model is so simple that perhaps the observations made in solving it can shed some light into the more difficult problem. Among the observations we have the following: (a) If the gauge prescription does not fix the gauge completely it is expected there would be gauge modes propagating. In this case there was just one gauge freedom left (the value of h at the initial surface and at the outer boundary). Nevertheless there were three equations which acquired a nontrivial propagation. (b) Gauge modes, being nonphysical, can have singular behavior, although the four-dimensional geometry of the solution can be regular. It is clear in the model studied that the different components of

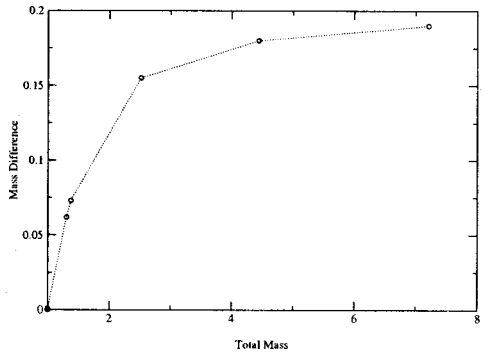

FIG. 15. ΔM vs M for type-III data.

the metric grow exponentially while the solution is still Schwarzschild. Thus, it seems to be necessary to isolate and keep under control the potentially unstable gauge modes. In our case this was done by choosing variables for which it was clear where the instability was, and so, choosing conveniently the initial-boundary data we could keep it small for many cases, although not always. (c) The choice of gauge modes can also complicate the boundary value problem. Ideally some of the fields should acquire values at the boundary so that the constraints propagate correctly. In the model it can be seen that if part of the data for the pair (fg) must be given at the boundary [according to the values chosen for the coefficients $(K_{ff},K_{fg},K_{gg},K_{gf})$] that part must satisfy some evolution equation along the boundary, so actually there is no freedom left. But other gauge quantities, like h, can take any value. Thus, if the variables are not chosen appropriately it could become very cumbersome to find the correct boundary conditions which would, at the same time, keep the constraint equations satisfied and the gauge instabilities under control. Note that the procedure we used to impose boundary conditions, consisting of obtaining equations which are intrinsic to the boundary for some of the variables, is similar to the one used in [19], the only case where well posedness of the initial-boundary value problem in full general relativity has been asserted. (d) Note also that the problem of giving the correct boundary condition for the gauge quantities cannot be resolved by just moving the boundary far away. Indeed, if by doing so we make the unstable gauge of the order $1/R$, since it is going to propagate a time R until it disappears into the inner boundary, the net effect would not depend on R, but on how far away we fix the boundary.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge several important discussions with S. Frittelli, H. O. Kreiss, P. Laguna, and M. Tiglio. This work was supported by Antorchas, CONICOR, CONICET and Se.CyT, UNC. M.S.I was partially supported by AIT, Córdoba, Argentina. O.A.R. is associated with CONICET.

[1] M. Shibata and T. Nakamura, Phys. Rev. D **52**, 5428 (1995).
[2] T. W. Baumgarte and S. L. Shapiro, Phys. Rev. D **59**, 024007 (1999).
[3] M. Alcubierre, G. Allen, B. Brügmann, E. Seidel, and Wai-Mo Suen, Phys. Rev. D **62**, 124011 (2000).
[4] Carsten Gundlach, Living Rev. Relativ. **2**, 4 (1999).
[5] Roberto Gomez and J. Winicour, J. Math. Phys. **33**, 4 (1992).
[6] David Garfinkle, Phys. Rev. D **51**, 5558 (1995).
[7] Lior M. Burko, in *The Eighth Marcel Grossmann Meeting: Proceedings*, edited by Tsvi Piran (World Scientific, Singapore, 1999), Vol. I, pp. 750–752.
[8] R. Bartnik and A. Norton, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **22**, 917 (2000).
[9] M. A. Scheel, T. W. Baumgarte, G. B. Cook, S. L. Shapiro, and S. A. Teukolsky, Phys. Rev. D **58**, 044020 (1998).
[10] M. A. Scheel, T. W. Baumgarte, G. B. Cook, S. L. Shapiro, and S. A. Teukolsky, Phys. Rev. D **56**, 6320 (1997).
[11] Peter Hübner, Phys. Rev. D **53**, 701 (1996).
[12] Jörg Frauendiener, Living Rev. Relativ. **3**, 4 (2000), http://www.livingreviews.org/Volume3/2000-4/frauendiener
[13] R. L. Marsa and M. W. Choptuik, Phys. Rev. D **54**, 4929 (1996).
[14] B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods* (Wiley, New York, 1995).
[15] B. Strand, J. Comput. Phys. **110**, 47 (1994).
[16] Carsten Gundlach, Richard H. Price, and Jorge Pullin, Phys. Rev. D **49**, 883 (1994).
[17] Carsten Gundlach, Richard H. Price, and Jorge Pullin, Phys. Rev. D **49**, 890 (1994).
[18] Lior M. Burko and Amos Ori, Phys. Rev. D **56**, 7820 (1997).
[19] H. Friedrich and G. Nagy, Commun. Math. Phys. **210**, 619 (1999).