Appearance of coordinate shocks in hyperbolic formalisms of general relativity

Miguel Alcubierre*
Department of Physics and Astronomy, University of Wales, College of Cardiff, P.O. Box 913, Cardiff CF1 3YB, United Kingdom
(Received 3 September 1996)

I consider the appearance of shocks in hyperbolic formalisms of general relativity. I study the particular case of the Bona-Massó formalism with a zero shift vector and show how shocks associated with two families of characteristic fields can develop. These shocks do not represent discontinuities in the geometry of spacetime, but rather regions where the coordinate system becomes pathological. For this reason I call them ''coordinate shocks.'' I show how one family of shocks can be eliminated by restricting the Bona-Massó slicing condition $\partial_t \alpha = -\alpha^2 f(\alpha) \mathrm{tr}K$ to the case $f = 1 + k/\alpha^2$, with k an arbitrary constant. The other family of shocks cannot be eliminated even in the case of harmonic slicing $(f=1)$. I also show the results of the numerical evolution of nontrivial initial slices in the special cases of a flat two-dimensional spacetime, a flat four-dimensional spacetime with a spherically symmetric slicing, and a spherically symmetric black hole spacetime. In all three cases coordinate shocks readily develop, confirming the predictions of the mathematical analysis. Although I concentrate in the Bona-Massó formalism, the phenomena of coordinate shocks should arise in any other hyperbolic formalism. In particular, since the appearance of the shocks is determined by the choice of gauge, the results presented here imply that in *any formalism* the use of a harmonic slicing can generate shocks. [S0556-2821(97)05510-0]

PACS number(s): 04.20.Ex, 04.20.Dw

I. INTRODUCTION

In the last few years there has been a renewed interest in the study of initial-value formulations of general relativity [1–9]. This interest has been motivated mainly by the desire of rewriting the Einstein system of evolution equations in an explicit hyperbolic form, so that it can be solved numerically using modern high-resolution methods from fluid dynamics [10].

One can separate the new hyperbolic formalisms according to the way in which they treat the evolution of the lapse function α. Some formulations assume the existence of an arbitrarily prescribed gauge; i.e., the lapse is an arbitrary function of spacetime known *a priori* [8,9] (''prescribed gauge'' formalisms). Other formulations include the lapse function as part of the system of dynamical variables, and postulate for it an evolution equation that guarantees the hyperbolicity of the *whole* system, geometry plus gauge [1–7] (''hyperbolic gauge'' formalisms). The resulting formalisms remain hyperbolic for any *prescribed* shift vector.[1]

Prescribed gauge formalisms, though certainly useful theoretically, might have a limited applicability in numerical relativity simply because there is no recipe that can give us the *a priori* form of the lapse except in trivial cases (for example $\alpha = 1$). Hyperbolic gauge formalisms, on the other hand, by allowing the lapse functon to adapt itself to the evolution of the geometry while maintaining the hyperbolic structure of the system of equations, would appear to be much more promising.

Hyperbolic gauge formalisms, however, are probably more susceptible to a problem that seems to have been overlooked until now. By rewriting the whole evolution system (gauge plus geometry) in hyperbolic form, they open up the possibility of running into a well-known nonlinear effect associated with hyperbolic systems: the appearance of shocks. Here I use the term ''shock'' in a somewhat loose form to refer to a discontinuous solution that develops from smooth initial data, without worrying about the existence of weak solutions or jump conditions.

The fact that in vacuum general relativity one can have shock fronts is well known [12–15]. By shock fronts, however, one generally understands discontinuities in the curvature of spacetime present in the initial data that propagate with the speed of light. In the theory of nonlinear hyperbolic equations such solutions are not considered proper shocks, but are called instead ''contact discontinuities.'' Here, however, I will consider the existence of discontinuous solutions that arise from smooth initial data even in a flat spacetime. Clearly those solutions do not correspond to a physical discontinuity in the geometry of spacetime. Instead the discontinuities indicate regions where our coordinate system becomes pathological: the time slices can become nonsmooth, or a spatial coordinate might map a finite proper distance to an infinitesimal interval. It is for this reason that I shall refer to them as ''coordinate shocks.''

Even though modern high-resolution numerical methods can deal with the presence of shock waves, clearly the appearance of coordinate shocks is something that must be avoided. In the first place, coordinate shocks create completely artificial discontinuities in solutions that otherwise

*Present address: Max Planck Institut für Gravitationsphysik, Albert Einstein Institut, Schlaatzweg 1, D-14473 Potsdam, Germany.

[1]To date, in all these hyperbolic formalisms the shift vector is assumed to be known *a priori*. The author is aware, however, of some efforts to find evolution equations for the shift that will keep the whole system hyperbolic [11].

represent perfectly smooth geometries. Not only that, but since in general our gauge conditions are not obtained from a conservation law, we will not have an analogue of the ''weak solutions'' to such laws. This means that after a shock forms our gauge conditions will just break down, and even if the numerical solution remains well behaved, *it will not have any clear physical meaning*. In particular, as the numerical mesh is refined, the solution will not converge after the formation of the shock.

In this paper I will concentrate in one particular hyperbolic formalism of general relativity, the Bona-Massó (BM) formalism [1–4], and I will show how these coordinate shocks can and do indeed develop even in very simple situations.

II. THE BONA-MASSÓ FORMALISM

In this section I will make a brief introduction to the BM hyperbolic formalism for general relativity. I will use the most recent form of this formalism as presented in [4].

Let us start from the standard 3+1 formulation of general relativity of Arnowitt, Deser, and Misner (ADM) [16,17]. The evolution equations for the metric g_{ij} and extrinsic curvature K_{ij} are

$$(\partial_t - \mathcal{L}_\beta) g_{ij} = -2\alpha K_{ij}, \tag{1a}$$

$$(\partial_t - \mathcal{L}_\beta) K_{ij} = -\nabla_i \nabla_j \alpha + \alpha [R^{(3)}_{ij} + \mathrm{tr}K K_{ij} - 2K_{ik}K^k_j - R^{(4)}_{ij}], \tag{1b}$$

where α is the lapse function, β^k the shift vector, and where $R^{(3)}_{ij}$ and $R^{(4)}_{ij}$ represent the components of the Ricci tensor for the spatial hypersurfaces and for the full spacetime, respectively. In what follows I will restrict myself to the case of the zero shift vector. The ADM equations then reduce to

$$\partial_t g_{ij} = -2\alpha K_{ij}, \tag{2a}$$

$$\partial_t K_{ij} = -\nabla_i \nabla_j \alpha + \alpha [R^{(3)}_{ij} + \mathrm{tr}K K_{ij} - 2K_{ik}K^k_j - R^{(4)}_{ij}]. \tag{2b}$$

In order to obtain a system that is first order in space we introduce the quantities

$$A_k = \partial_k \ln \alpha, \quad D_{kij} = \frac{1}{2}\partial_k g_{ij}. \tag{3}$$

The evolution equation for K_{ij} can then be rewritten as

$$\partial_t K_{ij} + \partial_k (\alpha \lambda^k{}_{ij}) = \alpha S_{ij}, \tag{4}$$

where we have defined

$$\lambda^k{}_{ij} = D^k{}_{ij} + \frac{1}{2}\delta^k_i (A_j + 2V_j - D_{jm}{}^m) + \frac{1}{2}\delta^k_j (A_i + 2V_i - D_{im}{}^m), \tag{5}$$

with

$$V_k \equiv D_{km}{}^m - D^m{}_{mk}. \tag{6}$$

The source term S_{ij} in Eq. (4) involves only the fields themselves and not their derivatives:

$$S_{ij} = -R^{(4)}_{ij} + \mathrm{tr}K K_{ij} - 2K_{ik}K^k_j + 4D_{kmi}D^{km}{}_j + \Gamma^k{}_{km}\Gamma^m{}_{ij} - \Gamma_{ikm}\Gamma_j{}^{km} + (A^k - 2D_m{}^{km})(D_{ijk} + D_{jik}) + A_i\left(V_j - \frac{1}{2}D_{jk}{}^k\right) + A_j\left(V_i - \frac{1}{2}D_{ik}{}^k\right). \tag{7}$$

We also need an evolution equation for the D_{kij}. This we obtain, by taking the spatial derivative of the evolution equation for g_{ij},

$$\partial_t D_{kij} + \partial_k (\alpha K_{ij}) = 0. \tag{8}$$

The quantities V_k defined in Eq. (6) are very important. Their evolution equation can be obtained from Eq. (8). In order to ensure hyperbolicity, however, it is crucial to modify the resulting equation using the momentum constraints to obtain

$$\partial_t V_k = \alpha P_k, \tag{9}$$

where

$$P_k = G^0_k + A_m(K^m_k - \delta^m_k \mathrm{tr}K) + K^m_n(D_{km}{}^n - \delta^n_k D_{ma}{}^a) - 2K_{mn}(D^{mn}{}_k - \delta^n_k D_a{}^{am}), \tag{10}$$

and where $G_{\mu\nu}$ is the Einstein tensor of the spacetime. The quantities V_k are now considered independent, and Eq. (6) becomes an algebraic constraint that must be satisfied by the physical solutions.

Finally, we need evolution equations for the lapse α and its derivative A_k, i.e., we need to choose a slicing condition. In the BM formalism the following slicing condition is used:

$$\partial_t \alpha = -\alpha^2 f(\alpha) \mathrm{tr}K, \tag{11}$$

with $f(\alpha) > 0$ but otherwise arbitrary.

The complete system of evolution equations then takes the form

$$\partial_t \alpha = -\alpha^2 f(\alpha) \mathrm{tr}K, \tag{12a}$$

$$\partial_t g_{ij} = -2\alpha K_{ij}, \tag{12b}$$

and

$$\partial_t A_k + \partial_k (\alpha f \mathrm{tr}K) = 0, \tag{13a}$$

$$\partial_t D_{kij} + \partial_k (\alpha K_{ij}) = 0, \tag{13b}$$

$$\partial_t K_{ij} + \partial_k (\alpha \lambda^k{}_{ij}) = \alpha S_{ij}, \tag{13c}$$

$$\partial_t V_k = \alpha P_k. \tag{13d}$$

To study the characteristic structure of the system of Eqs. (12) and (13) we choose a fixed space direction x and consider only derivatives along that direction. It can then be

shown that the system is hyperbolic with the following structure.[2]

25 fields propagate along the time lines (zero speed). These fields are

$$\{\alpha, g_{ij}, A_{x'}, D_{x'ij}, V_i, A_x - fD_x{}^m{}_m\} \quad (x' \neq x). \tag{14}$$

10 fields propagate along the physical light cones with speeds

$$\lambda^l{}_{\pm} = \pm \alpha \sqrt{g^{xx}}. \tag{15}$$

These fields are

$$w^l{}_{ix'\pm} = K_{ix'} \pm \sqrt{g^{xx}}(D_{xix'} + \delta_i^x V_{x'}/g^{xx}) \quad (x' \neq x). \tag{16}$$

2 fields propagate with the ''gauge speeds''

$$\lambda^f{}_{\pm} = \pm \alpha \sqrt{f g^{xx}}. \tag{17}$$

They are

$$w^f{}_{\pm} = \sqrt{f}\,\mathrm{tr}K \pm \sqrt{g^{xx}}(A_x + 2V^x/g^{xx}). \tag{18}$$

III. MATHEMATICAL ANALYSIS OF THE NONLINEARITIES

Here I will try to understand the nonlinearities present in our system of equations, trying in particular to determine whether shocks can develop. From the discussion in the previous section it is clear that the system of evolution equations in the BM formalism has the structure

$$\partial_t u_i = p_i, \quad i \in \{1, \ldots, N_u\}, \tag{19a}$$

$$\partial_t v_i + \partial_x F_i = q_i, \quad i \in \{1, \ldots, N_v\}. \tag{19b}$$

The fluxes F_i that appear in the above equations have the form

$$F_i = \sum_{j=1}^{N_v} M_{ij} v_j, \tag{20}$$

where the coefficients M_{ij} are functions of the u's but not of the v's.

Let us now call λ_i the eigenvalues and $\mathbf{e}_i$ the corresponding eigenvectors of the Jacobian matrix $M_{ij} = \partial F_i / \partial v_j$. Let us also introduce the matrix $R = [\mathbf{e}_1 | \mathbf{e}_2 | \cdots | \mathbf{e}_{N_v}]$ of column eigenvectors. The eigenfields w_i are then defined by

$$\mathbf{v} = R\mathbf{w} \Rightarrow \mathbf{w} = R^{-1}\mathbf{v}. \tag{21}$$

A given eigenfield w_i is called ''linearly degenerate'' [18] if the following condition holds:

$$\frac{\partial \lambda_i}{\partial w_i} = \sum_{j=1}^{N_v} \frac{\partial \lambda_i}{\partial v_j} \frac{\partial v_j}{\partial w_i} = \nabla_v \lambda_i \cdot \mathbf{e}_i = 0. \tag{22}$$

Since in our case the λ's do not depend on the v's, it is obvious that all the eigenfields are linearly degenerate.

In the case of systems of conservation laws where the sources vanish, linear degeneracy is enough to guarantee that no shocks will form. However, when the sources are nonzero, this is not true anymore. This is easy to see if we consider for a moment the prototype of nonlinear hyperbolic equations, Burgers' equation

$$\partial_t u + u \partial_x u = 0. \tag{23}$$

If we now define

$$v := \partial_x u, \tag{24}$$

then we can rewrite Eq. (23) as the system

$$\partial_t u = -uv, \tag{25a}$$

$$\partial_t v + \partial_x (uv) = 0. \tag{25b}$$

This has precisely the form (19). The only eigenvalue turns out to be equal to u which is clearly independent of v. By the definition above the system is linearly degenerate. However, it is clearly nonlinear and will generate shocks since it is only Burgues' equation in disguise. The nonlinearities have now been buried in the sources.

Clearly the condition that must be imposed to guarantee that no shocks will develop is that a given eigenvalue λ_i should not be affected by changes in the corresponding eigenfield w_i. The condition for linear degeneracy (22) asks for the eigenvalue not to be explicitly dependent on its associated eigenfield. In the presence of sources, however, the coupling can introduce an indirect dependency. In order to study this dependency let us consider the time evolution of λ_i:

$$\dot{\lambda}_i = \partial_t \lambda_i = \sum_{j=1}^{N_u} \frac{\partial \lambda_i}{\partial u_j} \partial_t u_j = \nabla_u \lambda_i \cdot \mathbf{p}. \tag{26}$$

Now, we want this time derivative to be independent of the eigenfield w_i:

$$\frac{\partial \dot{\lambda}_i}{\partial w_i} = \frac{\partial}{\partial w_i}(\nabla_u \lambda_i \cdot \mathbf{p}) = 0. \tag{27}$$

I shall call this condition ''indirect linear degeneracy'' and I will refer to condition (22) as ''explicit'' or ''direct'' linear degeneracy.

If we assume that the condition for explicit linear degeneracy holds, then the condition for indirect linear degeneracy can be reduced to

$$\nabla_u \lambda_i \cdot \frac{\partial \mathbf{p}}{\partial w_i} = \nabla_u \lambda_i \cdot \sum_{j=1}^{N_v} \frac{\partial \mathbf{p}}{\partial v_j} \frac{\partial v_j}{\partial w_i} = 0, \tag{28}$$

which can be rewritten as

$$\nabla_u \lambda_i \cdot (\mathbf{e}_i \cdot \nabla_v)\mathbf{p} = 0. \tag{29}$$

[2]Here I use the term hyperbolic in the weak sense to mean that the characteristic matrix of the system has real eigenvalues. It should be noticed that this weak form of hyperbolicity does not guarantee that the system can be diagonalized. A crucial feature of the BM formalism is that even though it is only weakly hyperbolic, it can in fact be diagonalized as long as $f > 0$.

This condition must supplement the condition for explicit linear degeneracy (22) if we want to guarantee that no shocks will develop.

Let us now apply the previous condition to the BM system of evolution equations (12) and (13). From the discussion of the previous section it is clear that, on a given spatial direction x we only have the following nontrivial eigenvalues:

$$\lambda^{l}{}_{\pm}=\pm\alpha\sqrt{g^{xx}},\quad \lambda^{f}{}_{\pm}=\pm\alpha\sqrt{fg^{xx}}. \tag{30}$$

The time derivative of $\lambda^{l}{}_{\pm}$ will then be

$$\dot{\lambda}^{l}{}_{\pm}=\pm\lambda^{l}{}_{\pm}\left[\frac{1}{\alpha}\partial_t\alpha+\frac{1}{2g^{xx}}\partial_t g^{xx}\right]$$
$$=\pm\lambda^{l}{}_{\pm}\left[\frac{1}{\alpha}\partial_t\alpha-\frac{g^{xm}g^{xn}}{2g^{xx}}\partial_t g_{mn}\right]. \tag{31}$$

Using now Eqs. (12) we find

$$\dot{\lambda}^{l}{}_{\pm}=\pm\alpha\lambda^{l}{}_{\pm}(K^{xx}/g^{xx}-f\mathrm{tr}K). \tag{32}$$

Now, from the definitions of w_l and w_f [Eqs. (16) and (18)] we can easily find that

$$\mathrm{tr}K=\frac{1}{2\sqrt{f}}(w^{f}{}_{+}+w^{f}{}_{-}) \tag{33}$$

and $(p,q\neq x)$

$$K^{xx}=g^{xx}\mathrm{tr}K+K_{pq}(g^{xp}g^{xq}-g^{xx}g^{pq})=\frac{1}{2}\left[\frac{g^{xx}}{\sqrt{f}}(w^{f}{}_{+}+w^{f}{}_{-})\right.$$
$$\left.+(g^{xp}g^{xq}-g^{xx}g^{pq})(w^{l}_{pq+}+w^{l}_{pq-})\right]. \tag{34}$$

Substituting these results back in the expression for $\dot{\lambda}^{l}{}_{\pm}$ we find

$$\dot{\lambda}^{l}{}_{\pm}=\pm\frac{\alpha\lambda^{l}{}_{\pm}}{2}\left[\frac{1}{\sqrt{f}}(1-f)(w^{f}{}_{+}+w^{f}{}_{-})+(g^{xp}g^{xq}-g^{xx}g^{pq})\right.$$
$$\left.\times(w^{l}_{pq+}+w^{l}_{pq-})\right]. \tag{35}$$

In the same way we find, for the time derivative of $\lambda^{f}{}_{\pm}$,

$$\dot{\lambda}^{f}{}_{\pm}=\pm\alpha\lambda^{f}{}_{\pm}[K^{xx}/g^{xx}-(f+\alpha f'/2)\mathrm{tr}K], \tag{36}$$

where $f'=\partial_\alpha f$. Substituting again the expression for K_{xx} and $\mathrm{tr}K$ in terms of the eigenfields we find

$$\dot{\lambda}^{f}{}_{\pm}=\pm\frac{\alpha\lambda^{f}{}_{\pm}}{2}\left[\frac{1}{\sqrt{f}}(1-f-\alpha f'/2)(w^{f}{}_{+}+w^{f}{}_{-})\right.$$
$$\left.+(g^{xp}g^{xq}-g^{xx}g^{pq})(w^{l}_{pq+}+w^{l}_{pq-})\right]. \tag{37}$$

Equations (35) and (37) are very important results. Consider first the situation for $\lambda^{f}{}_{\pm}$. If we want $\dot{\lambda}^{f}{}_{\pm}$ to be independent of $w^{f}_{\pm}$, and hence satisfy the condition for indirect linearly degeneracy, we must clearly ask for

$$1-f-\alpha f'/2=0. \tag{38}$$

This differential equation can be easily solved to give

$$f(\alpha)=1+k/\alpha^2, \tag{39}$$

with k an arbitrary constant. We must in fact take $k\geqslant 0$ in order to ensure that we will have $f>0$ for all $\alpha>0$.

We have then show that the function f must have the form (39) in order to guarantee that the eigenfields $w^{f}_{\pm}$ will not generate shocks. Notice that if we take $k=0$ the condition reduces to that of harmonic slicing, i.e., for harmonic slicing the eigenfields $w^{f}_{\pm}$ do not generate shocks.

Consider now the situation for $\lambda^{l}{}_{\pm}$. From Eq. (35) it is clear that if we want $\dot{\lambda}^{l}{}_{\pm}$ to be independent of $w^{l}_{qp\pm}$ we must have

$$g^{xp}g^{xq}-g^{xx}g^{pq}=0\quad (p,q\neq x). \tag{40}$$

This condition is very restrictive. In particular, it is impossible to satisfy with a diagonal metric. We then reach the conclusion that in the general case, the eigenfields $w^{l}_{qp\pm}$ can *always* generate shocks. Notice how this result is independent of the value of f, it will therefore remain true even in the case of harmonic slicing.

One must stress here the fact that we have not actually proved that shocks will indeed develop. Whether they do or not in any particular case should depend in a critical way on the form of the initial data. In the following sections I will consider some examples that show how coordinate shocks can indeed develop even in very simple cases.

IV. FLAT TWO-DIMENSIONAL SPACETIME

A. Evolution equations

As a first example, consider a flat two-dimensional spacetime (a ''1+1'' spacetime) with coordinates $\{t,x\}$. Notice that these coordinates do not have to correspond to the Minkowski coordinates $\{x_M,t_M\}$, so we can have a nontrivial evolution even though the spacetime is flat. Since we only have one spatial dimension, I will simplify the notation in the following way:

$$g:=g_{xx},\quad A:=A_x,\quad D:=D_{xxx},\quad K:=K_{xx}. \tag{41}$$

Notice that the variable V_x is identically zero.

The system of evolution equations (12) and (13) reduces in this case to

$$\partial_t\alpha=-\alpha^2 fK/g, \tag{42a}$$

$$\partial_t g=-2\alpha K, \tag{42b}$$

and

$$\partial_t A+\partial_x(\alpha fK/g)=0, \tag{43a}$$

$$\partial_t D+\partial_x(\alpha K)=0, \tag{43b}$$

$$\partial_t K + \partial_x(\alpha A) = \alpha/g(AD - K^2). \tag{43c}$$

The characteristic structure of this system is very simple.

There are three fields that propagate along the time lines (speed zero). These fields are

$$\{\alpha, g, A - fD/g\}. \tag{44}$$

The two remaining fields propagate with the ''gauge speeds''

$$\lambda^f{}_{\pm} = \pm\alpha\sqrt{f/g}. \tag{45}$$

They are

$$w^f{}_{\pm} = \sqrt{f}K/g \pm A/\sqrt{g}. \tag{46}$$

Notice how there are no fields propagating along the physical light cones. According to the discussion of the previous section, we should then expect shocks only when condition (39) is violated.

B. Numerical simulations

Since we are dealing with a flat spacetime, the only way to obtain a nontrivial evolution is to start with a nontrivial initial slice. I will therefore consider an initial slice given in terms of Minkowski coordinates $\{x_M, t_M\}$ as

$$t_M = h(x_M). \tag{47}$$

I will assume that the dynamical spatial coordinate x coincides initially with the Minkowski spatial coordinate x_M. It is then not difficult to show that the initial metric g and extrinsic curvature K are given by

$$g = 1 - h'^2, \tag{48a}$$

$$K = -h''/\sqrt{g}. \tag{48b}$$

The initial value of D can be obtained directly from its definition in terms of g. The initial lapse is taken to be equal to 1 everywhere, which implies that $A = 0$.

In all the simulations shown here, the function $h(x)$ has a Gaussian profile:

$$h(x) = H\exp\left\{-\frac{(x - x_c)^2}{\sigma^2}\right\}, \tag{49}$$

with $\{H, \sigma, x_c\}$ constants. The particular values of $\{H, \sigma\}$ used in the simulations presented here are

$$H = 5, \quad \sigma = 10. \tag{50}$$

I have also always taken the initial perturbation to be centered around $x_c = 150$. The initial values of all the variables can be seen in Fig. 1. All the results presented below where obtained using a time step of $\Delta t = 0.125$ and a spatial increment of $\Delta x = 0.25$.

In all the simulations, the evolution proceeds at first in a similar way: The initial perturbation in g, D, and K gives rise to perturbations in α and A. These perturbations rapidly develop into two separate pulses traveling in opposite directions with a speed $\sim\sqrt{f}$. What happens later depends crucially on the form of the function $f(\alpha)$.

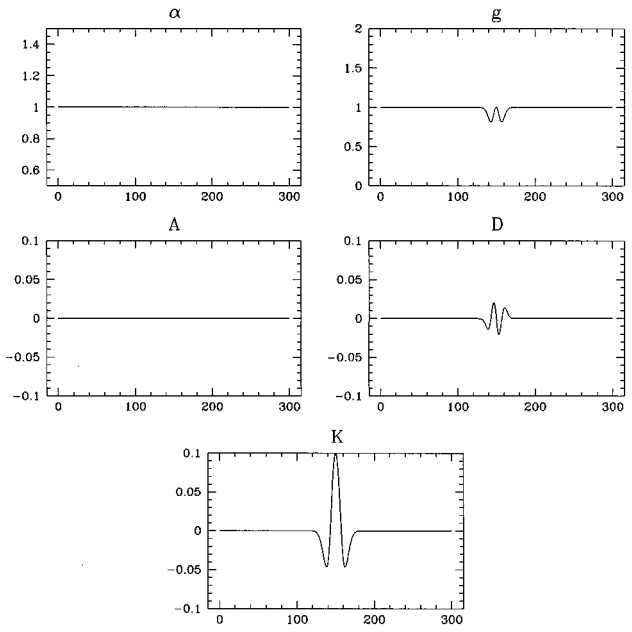

FIG. 1. Two-dimensional flat spacetime. Initial values of the dynamical variables.

For harmonic slicing ($f = 1$), the pulses remain smooth as they move away. Once the pulses are gone, the lapse, the metric, and the variables A and D return to their initial values, and the extrinsic curvature becomes 0. Figure 2 shows the values of the variables at $t = 100$.

When f is a constant larger than 1, the pulses do not remain smooth and shocks develop. In fact, we have two shocks developing in each pulse, one in front of it and one behind it. At those points, the lapse and the metric develop large gradients, while the extrinsic curvature and the vari-

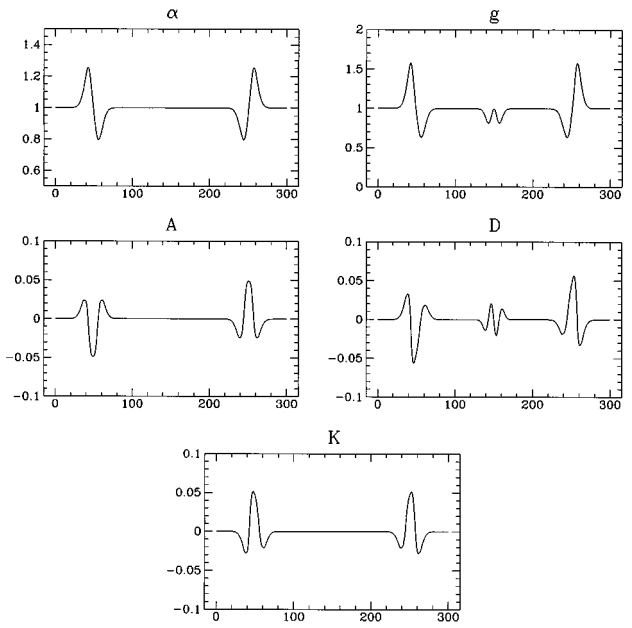

FIG. 2. Two-dimensional flat spacetime. Values of the variables at $t = 100$ for harmonic slicing ($f = 1$).

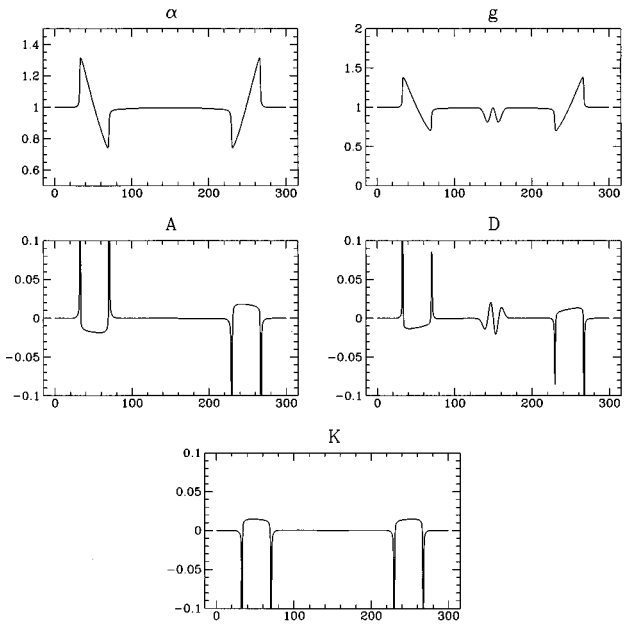

FIG. 3. Two-dimensional flat spacetime. Values of the variables at $t=75$ in the case when $f=1.69$.

ables A and D develop very tall and narrow spikes. Figure 3 shows the values of the variables at $t=75$ in the particular case when $f=1.69$.

When f is a constant smaller than 1, a single shock develops in the middle of each pulse. Apart from this, the situation is very similar to the case $f>1$. Figure 4 shows the values of the variables at $t=75$ in the particular case when $f=0.49$.

Finally, when f is of the form (39), no shocks develop in agreement with the predictions. The pulses remain smooth and move away with a speed $\sim\sqrt{1+k}$. Figure 5 shows the values of the variables at $t=70$ in the particular case when $f=1+1/\alpha^2$.

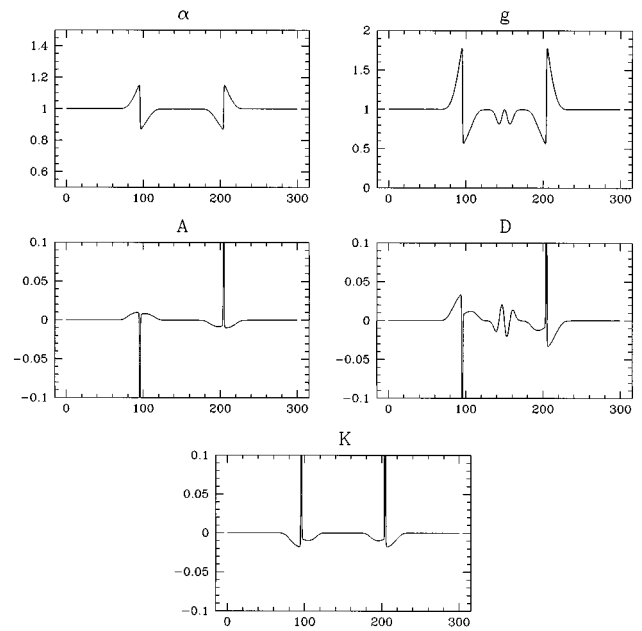

FIG. 4. Two-dimensional flat spacetime. Values of the variables at $t=75$ in the case when $f=0.49$.

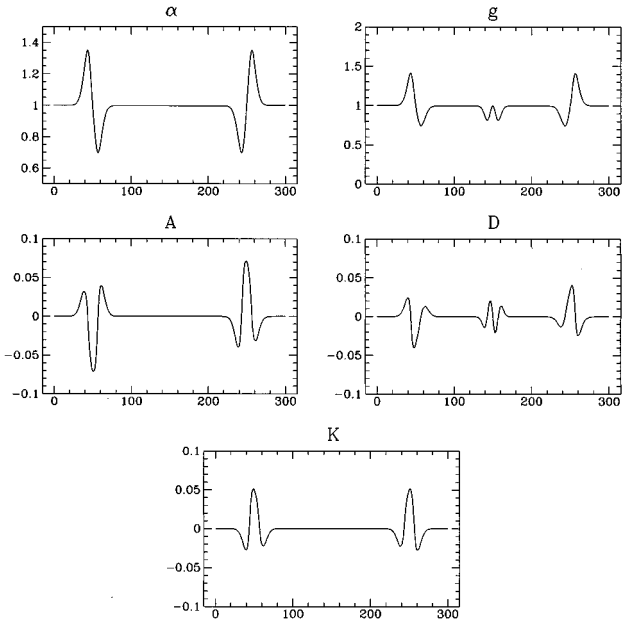

FIG. 5. Two-dimensional flat spacetime. Values of the variables at $t=70$ in the case when $f=1+1/\alpha^2$.

As a final comment, it should be mentioned that I have performed similar simulations with many different values of the amplitude H and width σ of the Gaussian profile of the initial slice. When f is not of the form (39) shocks apparently always develop, though at different times. The crucial feature that seems to determine the time of shock formation is the maximum absolute value of the extrinsic curvature K. For small values of K, shocks take a long time to appear, whereas for large values they develop very rapidly.

V. SPHERICALLY SYMMETRIC VACUUM SPACETIME

A. Evolution equations

As a second example, consider a spherically symmetric four-dimensional vacuum spacetime. Let us introduce the coordinate system $\{t,r,\theta,\phi\}$. The only independent dynamical variables will then be

$$\{\alpha,g_{rr},g_{\theta\theta},A_r,D_{rrr},D_{r\theta\theta},K_{rr},K_{\theta\theta},V_r\}. \tag{51}$$

The system of evolution equations (12) and (13) reduces now to

$$\partial_t\alpha=-\alpha^2 f\,\mathrm{tr}K, \tag{52a}$$

$$\partial_t g_{rr}=-2\alpha K_{rr}, \tag{52b}$$

$$\partial_t g_{\theta\theta}=-2\alpha K_{\theta\theta}, \tag{52c}$$

and

$$\partial_t A_r+\partial_r(\alpha f\,\mathrm{tr}K)=0, \tag{53a}$$

$$\partial_t D_{rrr}+\partial_r(\alpha K_{rr})=0, \tag{53b}$$

$$\partial_t D_{r\theta\theta}+\partial_r(\alpha K_{\theta\theta})=0, \tag{53c}$$

$$\partial_t K_{rr}+\partial_r(\alpha\lambda^r_{rr})=\alpha S_{rr}, \tag{53d}$$

$$\partial_t K_{\theta\theta} + \partial_r(\alpha \lambda^r_{\theta\theta}) = \alpha S_{\theta\theta}, \tag{53e}$$

$$\partial_t V_r = \alpha P_r. \tag{53f}$$

with

$$\lambda^r_{rr} = A_r + 2V_r - 2D_{r\theta\theta}/g_{\theta\theta}, \tag{54a}$$

$$\lambda^r_{\theta\theta} = D_{r\theta\theta}/g_{rr}, \tag{54b}$$

and

$$S_{rr} = K_{rr}(2K_{\theta\theta}/g_{\theta\theta} - K_{rr}/g_{rr}) + A_r(D_{rrr}/g_{rr} - 2D_{r\theta\theta}/g_{\theta\theta}) \tag{55a}$$

$$+ 2D_{r\theta\theta}/g_{\theta\theta}(D_{rrr}/g_{rr} - D_{r\theta\theta}/g_{\theta\theta}) + 2A_r V_r, \tag{55b}$$

$$S_{\theta\theta} = K_{rr}K_{\theta\theta}/g_{rr} - D_{rrr}D_{r\theta\theta}/{g_{rr}}^2 + 1, \tag{55c}$$

$$P_r = -2/g_{\theta\theta}[A_r K_{\theta\theta} - D_{r\theta\theta}(K_{\theta\theta}/g_{\theta\theta} - K_{rr}/g_{rr})]. \tag{55d}$$

We also have the following algebraic constraint that must be satisfied by the physical solutions:

$$V_r = 2D_{r\theta\theta}/g_{\theta\theta}. \tag{56}$$

The characteristic structure of this system turns out to be the following.

Five fields propagate along the time lines (speed zero). These fields are

$$\{\alpha, g_{rr}, g_{\theta\theta}, V_r, A_r - f{D_r}^m{}_m\}. \tag{57}$$

Two fields propagate along the physical light cones with speeds

$$\lambda^l{}_\pm = \pm \alpha/\sqrt{g_{rr}}. \tag{58}$$

These fields are

$$w^l{}_\pm = \sqrt{g_{rr}}K_{\theta\theta} \pm D_{r\theta\theta}. \tag{59}$$

Two fields propagate with the ''gauge speeds''

$$\lambda^f{}_\pm = \pm \alpha\sqrt{f/g_{rr}}. \tag{60}$$

They are

$$w^f{}_\pm = \sqrt{f g_{rr}}\,\mathrm{tr}K \pm (A_r + 2V_r). \tag{61}$$

Notice how we now have both fields propagating with the speed of light and fields propagating with the gauge speed. We should then expect to see two different types of shocks forming. In particular, shocks produced by the $w^l{}_\pm$ fields can be expected always, even for harmonic slicing.

B. Numerical simulations for a flat spacetime

Again, since we are dealing with flat spacetime, the only way to obtain a nontrivial evolution is to start with a nontrivial initial slice. I will therefore consider an initial slice given in terms of Minkowski coordinates $\{r_M, t_M\}$ as

$$t_M = h(r_M). \tag{62}$$

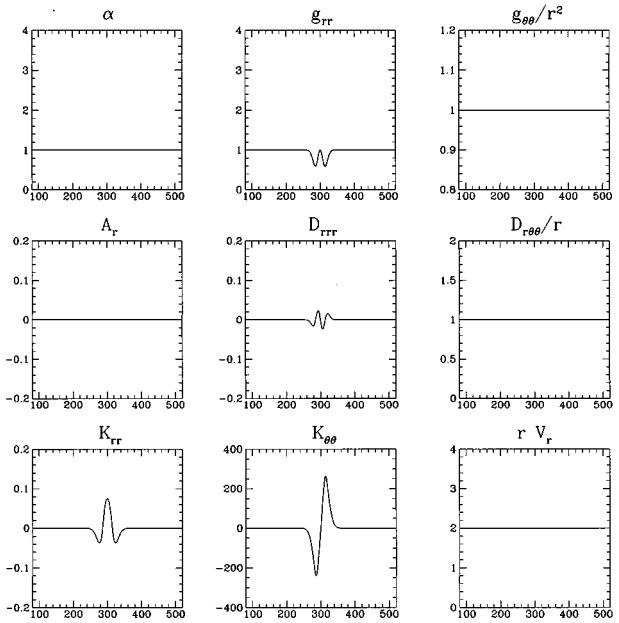

FIG. 6. Spherically symmetric flat spacetime. Initial values of the dynamical variables.

I will assume that the dynamical radial coordinate r coincides initially with the Minkowski radial coordinate r_M. It is then not difficult to show that the initial metric $\{g_{rr}, g_{\theta\theta}\}$ and extrinsic curvature $\{K_{rr}, K_{\theta\theta}\}$ are given by

$$g_{rr} = 1 - h'^2, \tag{63a}$$

$$g_{\theta\theta} = r^2, \tag{63b}$$

$$K_{rr} = -h''/\sqrt{g_{rr}}, \tag{63c}$$

$$K_{\theta\theta} = -rh'/\sqrt{g_{rr}}. \tag{63d}$$

The initial values of $\{D_{rrr}, D_{r\theta\theta}, V_r\}$ can be obtained directly from their definitions in terms of the metric. The initial lapse is taken to be equal to 1 everywhere, which implies that $A_r = 0$.

In all the simulations shown here, the function $h(r)$ has a Gaussian profile:

$$h(r) = H\exp\left\{-\frac{(r - r_c)^2}{\sigma^2}\right\}, \tag{64}$$

with $\{H, \sigma, r_c\}$ constants. The particular values of $\{H, \sigma\}$ used in the simulations presented here are

$$H = 15, \quad \sigma = 20, \tag{65}$$

and I have taken the initial perturbation to be centered around $r_c = 300$. The initial values of all the variables can be seen in Fig. 6. The results presented below were obtained using a time step of $\Delta t = 0.1$ and a spatial increment of $\Delta x = 0.2$.

In all the simulations the evolution proceeds at first in a similar way: the initial perturbations in $\{g_{rr}, D_{rrr}, K_{rr}, K_{\theta\theta}\}$ give rise to perturbations in $\{\alpha, g_{\theta\theta}, A_r, D_{r\theta\theta}, V_r\}$. These perturbations develop into

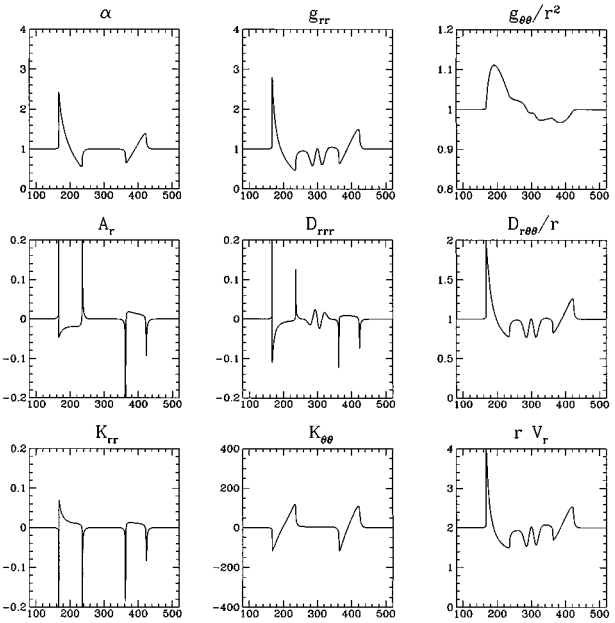

FIG. 7. Spherically symmetric flat spacetime. Values of the variables at $t=70$ in the case when $f=1.69$.

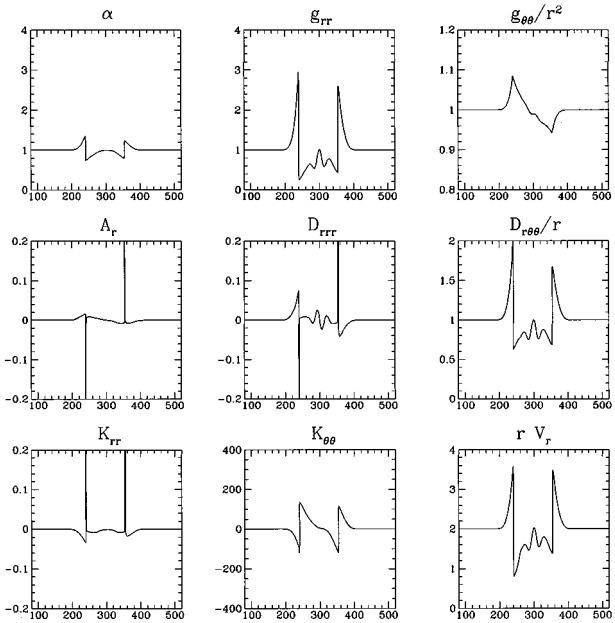

FIG. 8. Spherically symmetric flat spacetime. Values of the variables at $t=70$ in the case when $f=0.49$.

separate pulses traveling in opposite directions with a speed $\sim\sqrt{f}$. The pulses are not symmetric any more since clearly the in-going and out-going directions are not equivalent.

Consider first the case of $f>1$. As the evolution proceeds, shocks develop in both pulses. These shocks are similar to those found in the 1+1 case: two shocks develop in each pulse, one in front of it and one behind it. At those points $\{\alpha, g_{rr}, D_{r\theta\theta}, K_{\theta\theta}, V_r\}$ develop large gradients, while $\{A_r, D_{rrr}, K_{rr}\}$ develop tall and narrow spikes. The angular metric component $g_{\theta\theta}$ in contrast develops sharp corners. Figure 7 shows the values of the variables at $t=70$ in the particular case when $f=1.69$.

When $f<1$, we again find results that are similar to the 1+1 case: a single shock develops in each pulse. Again, at the shock $\{\alpha, g_{rr}, D_{r\theta\theta}, K_{\theta\theta}, V_r\}$ develop large gradients, $\{A_r, D_{rrr}, K_{rr}\}$ develop spikes, and $g_{\theta\theta}$ develops sharp corners. Figure 8 shows the values of the variables at $t=70$ in the particular case when $f=0.49$.

The most interesting case is that of harmonic slicing ($f=1$). In contrast to the 1+1 case, shocks still develop here. The shocks, however, have a different structure indicative of their different origin: the variables $\{A_r, D_{rrr}, K_{rr}\}$ now develop large gradients, while $\{\alpha, g_{rr}, D_{r\theta\theta}, K_{\theta\theta}, V_r\}$ develop sharp spikes. The angular metric component $g_{\theta\theta}$ also seems to develop a large gradient, though this gradient is less sharp than that found in other variables. This is easy to understand geometrically: any discontinuity in $g_{\theta\theta}$ must necessarily be accompanied by an infinite value of g_{rr} (we must jump a finite radial distance in an infinitesimal interval). The shocks are clearly visible in the in-going pulse, but do not seem to be present in the out-going pulse. Figure 9 shows the values of the variables at $t=70$ for harmonic slicing.

Again, I have performed similar simulations for different amplitudes and widths of the Gaussian profile of the initial slice. The shocks are always there. The only exception seems to be that for very small amplitudes of the initial perturbation the pulses moving in the in-going direction may not have enough time to develop before they reach the origin. Also, for the case of harmonic slicing, shocks in the out-going direction never seem to form, no matter how large the initial perturbation might be.

C. Numerical simulations for a black hole spacetime

In all the previous examples I have restricted myself to a flat spacetime. Since this is a very special case one might think that the shocks that we have found are just an artifact of the flatness. To show that this is not the case, I will now consider a spherically symmetric black hole spacetime.

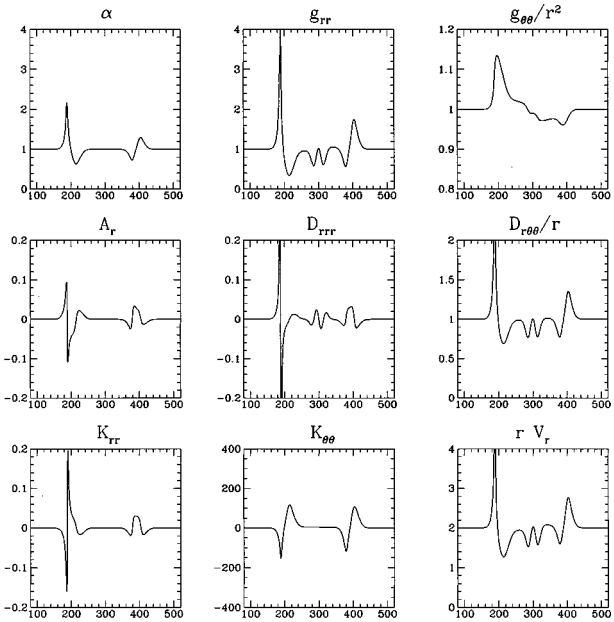

FIG. 9. Spherically symmetric flat spacetime. Values of the variables at $t=70$ in the case when $f=1$.

To find adequate initial data I start from a Schwarzschild slice with spatial metric

$$dl^2=\frac{1}{1-2M/r_s}dr_s^2+r_s^2d\Omega^2, \tag{66}$$

where r_s is the Schwarzschild radial coordinate and $d\Omega^2=d\theta^2+\sin^2\theta d\phi^2$.

In order to eliminate the singularity at $r_s=2M$, I will define a new radial coordinate r that measures proper distance along the slice. The coordinates r_s and r will be related by

$$r=\eta(r_s)+M\ln\left[\frac{r_s+\eta(r)}{r_s-\eta(r_s)}\right], \tag{67}$$

with

$$\eta(r_s)=(r_s^2-2Mr_s)^{1/2}. \tag{68}$$

Notice that even though Eq. (67) can not be inverted analytically to find $r_s(r)$, it can easily be inverted numerically to arbitrarily high accuracy.

The new metric will now have the form

$$dl^2=dr^2+[r_s(r)]^2d\Omega^2. \tag{69}$$

It is easy to see that the Schwarzschild slice has zero extrinsic curvature, so our initial data will be

$$g_{rr}=1, \tag{70a}$$

$$g_{\theta\theta}=r_s^2, \tag{70b}$$

$$K_{rr}=0, \tag{70c}$$

$$K_{\theta\theta}=0. \tag{70d}$$

Now, if we use this initial data directly we will not see any shocks develop. This is known since the BM formalism has been used before to solve this problem and no shocks have been observed [3]. The reason why shocks do not develop is that they are a consequence of transport and as such they should only develop when we have wave propagation, either in the form of real gravitational waves, or in the form of pure gauge waves. The static black hole problem has no gravitational waves, and the initial data given above does not give rise to gauge waves either.

In order to introduce gauge waves into our problem, I will consider an initial slice given in terms of Schwarzschild time t_s in the following way:

$$t_s=h(r). \tag{71}$$

It is not difficult to show that the new slice will have the metric components

$$g_{rr}=1-(\alpha_s h')^2, \tag{72a}$$

$$g_{\theta\theta}=r_s^2, \tag{72b}$$

where α_s is the Schwarzschild lapse function

$$\alpha_s=(1-2M/r_s)^{1/2}. \tag{73}$$

The components of the extrinsic curvature for this slice can now be shown to be

$$K_{rr}=-\{\alpha_s h''+\alpha_s' h'[2-(\alpha_s h')^2]\}/\sqrt{g_{rr}}, \tag{74a}$$

$$K_{\theta\theta}=-\alpha_s^2 r_s h'/\sqrt{g_{rr}}. \tag{74b}$$

As before, the initial values of $\{D_{rrr},D_{r\theta\theta},V_r\}$ can be obtained directly from their definitions in terms of the metric. The initial lapse is again taken to be equal to 1 everywhere, which implies that $A_r=0$.

For the function $h(r)$ I will again use a Gaussian:

$$h(r)=H\exp\left\{-\frac{(r-r_c)^2}{\sigma^2}\right\}, \tag{75}$$

with $\{H,\sigma,r_c\}$ constants. In order to see the development of the shocks clearly, I will consider simulations where the center of our perturbation r_c is out in the wave zone.

All the simulations I have carried out proceed in a similar way. At the throat of the wormhole we find what we expect for a black hole spacetime: the lapse collapses and the metric component g_{rr} grows rapidly. Out in the wave zone, the disturbance behaves in the same way as it did in flat spacetime: the initial perturbations in $\{g_{rr},D_{rrr},K_{rr},K_{\theta\theta}\}$ give rise to perturbations in $\{\alpha,g_{\theta\theta},A_r,D_{r\theta\theta},V_r\}$, these then develop into separate pulses traveling in opposite directions with a speed $\sim\sqrt{f}$.

In all cases, the traveling pulses develop shocks that have very similar characteristics to those that we found in the flat case. Here I will only show the results found in the case of harmonic slicing $f=1$. The particular values of $\{H,\sigma\}$ used in this simulation are

$$H=5, \quad \sigma=5. \tag{76}$$

I have also taken the initial perturbation to be centered around $r_c=50$, and the mass of the black hole to be $M=1$. The results presented here were obtained using a time step of $\Delta t=0.025$ and a spatial increment of $\Delta x=0.05$. The initial values of all the variables can be seen in Fig. 10.

Figure 11 shows the values of the variables at $t=15$. Notice how around the throat the lapse and the angular metric component $g_{\theta\theta}$ have collapsed, while the radial metric component g_{rr} has grown to a very large value. The interesting region for our purposes, however, is away from the throat. We can clearly see the two pulses resulting from our initial perturbation. The pulse moving inwards has developed a shock: the variables $\{A_r,D_{rrr},K_{rr}\}$ have developed large gradients, while $\{\alpha,g_{rr},D_{r\theta\theta},K_{\theta\theta},V_r\}$ have developed sharp spikes. The angular metric component $g_{\theta\theta}$ has also developed a large gradient.

VI. DISCUSSION

I have introduced a general approach to the study of shock development in hyperbolic systems of equations with sources. I have shown that the usual condition of explicit linear degeneracy (direct linear degeneracy) must be supplemented with a new condition which I have called ''indirect linear degeneracy'' in order to guarantee that no shocks will develop.

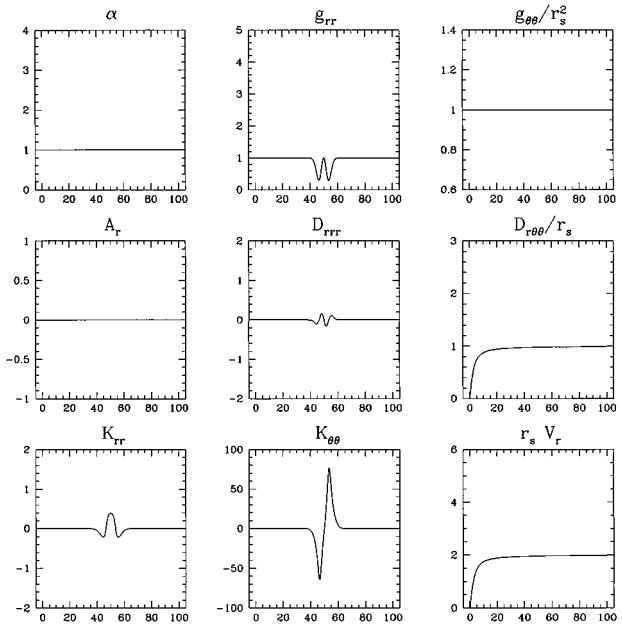

FIG. 10. Spherically symmetric black hole spacetime. Initial values of the dynamical variables.

I have applied this condition of indirect linear degeneracy to the BM hyperbolic formalism of general relativity in the case of a zero shift vector. My analysis has shown how two distinct families of characteristic fields can give rise to shocks. Numerical simulations have confirmed these predictions in the simple cases of a flat two-dimensional spacetime, a flat four-dimensional spacetime with spherically symmetric slices, and a spherically symmetric black hole spacetime.

The appearance of shocks that develop from smooth initial data in vacuum general relativity comes as a great surprise. These shocks, however, do not represent discontinuities in the geometry of spacetime, but indicate instead

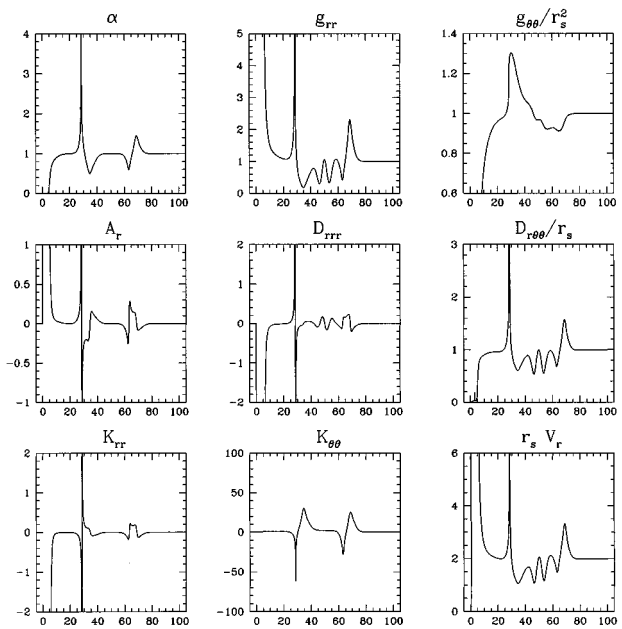

FIG. 11. Spherically symmetric black hole spacetime. Values of the variables at $t=15$ in the case when $f=1$.

regions where our coordinate system becomes pathological. It is for this reason that I refer to them as ''coordinate shocks.''

Of the two families of coordinate shocks found, one can be completely eliminated by choosing a BM gauge function $f(\alpha)$ of the form

$$f(\alpha)=1+k/\alpha^2, \tag{77}$$

with $k\geqslant 0$ an arbitrary constant. For $k>0$, however, this form of the function f will not be very useful in spacetimes with large curvatures. The reason for this is easy to see. Even though the condition will prevent the formation of shocks, it implies an evolution equation for the lapse of the form

$$\partial_t\alpha=-(\alpha^2+k)\mathrm{tr}K. \tag{78}$$

Clearly, in a region where the lapse has collapsed to a very small value we will have

$$\partial_t\alpha\simeq -k\mathrm{tr}K. \tag{79}$$

If $\mathrm{tr}K>0$, there is nothing to prevent the lapse from becoming negative (this can in fact happen very easily in black hole simulations). We are then led to the conclusion that the only value of f that will prevent the first family of shocks from developing without carrying the risk of leading to a negative lapse is $f=1$, i.e., harmonic slicing.

The second family of shocks, on the other hand, is independent of the form of f and arises even for harmonic slicing. This is a very unexpected result. After all, this is precisely the slicing used to prove the theorems of existence and uniqueness of solutions in general relativity [19–21]. Since at a shock the differential equations break down, one would expect the theorems to forbid such solutions. We must remember, however, that these theorems can only be proved *locally*, they cannot therefore rule out a shock that develops after a *finite time*.

It must be stressed that the violation of indirect linear degeneracy is not a sufficient condition for the development of shocks. The choice of initial data will have a crucial effect in whether or not shocks actually develop. In particular, since shocks are a consequence of transport, they should only develop when we have wave propagation, either in the form of real gravitational waves, or in the form of pure gauge waves as was shown in the examples presented here. Of course, in the simple cases considered in this paper one can easily find initial data that does not produce shocks: for the flat spacetimes one can just take a flat initial slice, while for a black hole spacetime we can start from an unperturbed Schwarzschild slice. In the more general case, however, it might be difficult to find such benign initial data, or even to prove that it exists at all.

One more important point should be made here. Since the shocks that I have found arise in the case of a zero shift vector, they must necessarily indicate a breakdown of the slicing condition. That is, the shocks represent places where the spatial hypersurfaces become nonsmooth. Since the presence of a shift vector can not alter the geometry of these hypersurfaces, the shocks found here must appear for *any shift condition*. A given shift might eliminate the discontinuities in some components of the spatial metric, but it cannot

eliminate the shocks completely: at least some of the dynamical quantities will remain nonsmooth for all possible shift choices.

Although in this paper I have concentrated in the BM hyperbolic formalism, the mathematical tools developed can easily be applied to any other hyperbolic formalism of general relativity. One should expect the phenomena of coordinate shocks to also arise in any such formalism. In fact, since all formalisms must have the same physical solutions, the results of this paper imply that in *any formalism* the use of a harmonic slicing will generate shocks for at least some initial conditions. Clearly, the search for gauge conditions and/or restrictions on the initial data that can prevent the development of coordinate shocks is a problem that must be addressed if hyperbolic formalisms are to become an important tool in the study of both theoretical and numerical relativity.

ACKNOWLEDGMENTS

The author wishes to thank Gabrielle D. Allen, Carles Bona, Joan Massó, and Bernard F. Schutz for many useful discussions and comments.

[1] C. Bona and J. Massó, Phys. Rev. D **40**, 1022 (1989).
[2] C. Bona and J. Massó, Phys. Rev. Lett. **68**, 1097 (1992).
[3] C. Bona, J. Massó, E. Seidel, and J. Stela, Phys. Rev. Lett. **75**, 600 (1995).
[4] C. Bona, J. Massó, E. Seidel, and J. Stela (unpublished).
[5] Y. Choquet-Bruhat and T. Ruggieri, Commun. Math. Phys. **89**, 269 (1983).
[6] Y. Choquet-Bruhat and J. W. York, C. R. Acad. Sci. Paris A **321**, 1089 (1995).
[7] A. Abrahams, A. Anderson, Y. Choquet-Bruhat, and J. W. York, Phys. Rev. Lett. **75**, 3377 (1995).
[8] H. Friedrich, Class. Quantum Grav. **13**, 1451 (1996).
[9] S. Frittelli and O. A. Reula, Phys. Rev. Lett. **25**, 4667 (1996).
[10] R. J. LeVeque, *Numerical Methods for Conservation Laws* (Birkhäuser Verlag, Basel, 1992).
[11] C. Bona and J. Massó (private communication).
[12] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[13] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957).
[14] A. Papapetrou and H. Treder, in *Recent Developments in General Relativity*, volume dedicated to L. Infeld in connection with his 60th birthday (Pergammon, Oxford and PWN-Polish Scientific Publishers, Warsaw, 1962).
[15] Y. Choquet-Bruhat, Ann. Inst. Henri Poincaré, A **8**, 327 (1968).
[16] R. Arnowitt, S. Deser, and C. W. Misner, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962)
[17] J. W. York, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge University Press, Cambridge, England, 1979).
[18] P. D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, Regional Conference Series in Applied Mathematics **11** (SIAM, Philadelphia, 1973).
[19] Y. Choquet-Bruhat, in *Gravitation: An Introduction to Current Research* [16].
[20] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).
[21] R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago, 1984).