PHYSICAL REVIEW D VOLUME 44, NUMBER 10 15 NOVEMBER 1991

Consistency of finite-difference solutions of Einstein's equations

M. W. Choptuik
Canadian Institute for Theoretical Astrophysics, 60 St. George Street, Toronto, Ontario, Canada M5S 1A1
(Received 5 July 1990)

In the past, arguments have been advanced suggesting that certain finite-difference solutions of the $3+1$ form of Einstein's equations suffer from a fundamental inconsistency. Specifically, it has been claimed that freely evolved solutions, where the constraint equations are not explicitly imposed after the initial time, will generally satisfy discrete versions of the constraints to a lower order in the basic mesh spacing h than the truncation order of the discretized evolution equations. This issue is reexamined here, and using the key observation, originally due to Richardson, that a numerical differentiation need not produce an $O(h^{p-1})$ quantity from an $O(h^p)$ one, it is argued that there should be *no* such inconsistency for convergent difference schemes. Numerical results from a study of spherically symmetric solutions of a massless scalar field minimally coupled to the gravitational field are presented in support of this claim. These results show that the expected convergence of various residual quantities can be achieved in practice.

I. INTRODUCTION

In $3+1$ numerical relativity [1,2], as with other branches of computational physics involving the solution of initial-value problems, a principle task is the conversion of a set of continuum *evolution equations* to a corresponding discrete algebraic form which can then be solved numerically. To date, finite differencing has been the main vehicle for performing this discretization, and although other techniques have been successful applied, it is clear that finite-difference methods will continue to play a major role in the development of the field. The task of differencing Einstein's equations is complicated by the fact that there are also *constraint equations* which must be satisfied. It is a basic result of the Hamiltonian treatment of general relativity [3] that the evolution equations preserve the constraints by virtue of the Bianchi identities. Thus, in principle, the constraint equations need to be explicitly solved only on the initial data surface and this approach has often been successfully adopted in numerical work, resulting in what numerical relativists term *free evolution schemes* [4–11]. However, even if the initial data supplied to such an algorithm (code) satisfy some differenced from of the constraints to machine precision, deviations from the constraints—at a much higher level than that attributable to hardware limitations—invariably develop during the evolution of the data.

The fact that such deviations appear is not in itself very surprising. As has been frequently pointed out, the Bianchi identities are differential identities, so that, in general, one can only expect them to be *approximately* satisfied by a finite-difference solution. Nevertheless, from time to time, considerable concern has been expressed about this particular feature of free-evolution schemes, both by practitioners of the art and by outside observers. Most of this concern seems to come from a general consensus that it is most desirable to construct a spacetime, slice by slice, so that on each hypersurface some form of the constraint equations are properly satisfied. There has always been a somewhat vague feeling of uncertainty about the "drifting" of the numerical data "off of the constraint surface" during the course of a free evolution and the extent to which such data can reliably be regarded as a "good" approximate solution of Einstein's equations. These worries have been further exacerbated by the empirical observation that algorithms in which discrete versions of the constraints are imposed at regular intervals tend to be less prone to numerical instabilities than free-evolution schemes [12–16].

One particularly intriguing view of the nature of the constraints in $3+1$ numerical relativity is discussed in review articles by Piran [17] and Stewart [18]. Both of these papers contain an argument which suggests that free evolution schemes generically suffer from an *inconsistency* with respect to the constraint equations. We will review this argument in more detail in the next section, but basically, it goes as follows. We focus attention on some dynamical geometrical variable g (a three-metric or extrinsic curvature component) which, roughly speaking, can be evolved either by (1) using the specific *evolution* equation for the variable, or (2) solving a *constraint* equation for the variable once all other quantities appearing in that equation have been evolved. We then imagine writing two distinct finite-difference codes which differ only in how the function g is approximated. In one code, we approximate g with a *grid function* (discrete representation) $\hat{g}$, which satisfies a differenced version of the *evolution* equation for g. We have the other code compute another grid function $\tilde{g}$, which, at all times, satisfies a discrete form of the *constraint* equation. Now, the basic parameter which will govern the accuracy of these finite-difference codes is a discretization scale h (which we will also refer to as the *mesh spacing* or *grid spacing*). Typically, in our construction of the two algorithms, we will have replaced all derivatives by finite differences which are $O(h^2)$ ("second-order") accurate, and then, at least naively, we can expect second-order convergence of both

the freely evolved grid function $\hat{g}$ and the constrained grid function $\hat{g}$—that is, we anticipate $\hat{g}=g+O(h^2)$ and $\hat{g}=g+O(h^2)$ as $h\to 0$. The inconsistency seems to arise when we ask how well the freely evolved quantity $\hat{g}$ will satisfy the constraint equation. Since $\hat{g}$ is defined on the finite-difference grid, we cannot evaluate how well it satisfies the *analytic* constraint, and it is only natural to examine how well it satisfies the same discrete version of the constraint which we use to compute $\tilde{g}$. Specifically, if we write the discrete constraint as $\tilde{H}\tilde{g}=0$, where $\tilde{H}$ is the "discrete constraint operator," then the grid function $\hat{\xi}$ defined by $\hat{\xi}\equiv\tilde{H}\hat{g}$ is a measure of the deviation of the freely evolved grid function from the constraint. Presumably, if $\hat{g}$ is consistent with the constraint equations, we will find that $\hat{\xi}=O(h^2)$—we certainly have $\xi\equiv\tilde{H}g=O(h^2)$, by our *construction* of a second-order difference scheme for the constraint equations. However, the arguments made by Piran and Stewart suggest that we will discover, quite generically, that $\hat{\xi}=O(h)$ or $\hat{\xi}=O(1)$—that is, we will find that the freely evolved quantity $\hat{g}$ satisfies the (discretized) constraint equation to *a lower order of accuracy* in h than the order of accuracy of the difference schemes used to compute either $\hat{g}$ or $\tilde{g}$. It is in this sense that free-evolution schemes have been claimed to be inconsistent with the constraint equations.

Now, as we will discuss below, this "inconsistency argument" hinges on the application of a familiar rule of thumb of numerical analysis which states that if a grid function $\hat{g}$ is an $O(h^p)$ approximation of some continuum function g then a *numerical derivative* of $\hat{g}$ will generally be an $O(h^{p-1})$-accurate approximation of the derivative of g. The basis of this rule of thumb is simply the observation that a numerical (first) derivative of a grid function invariably contains an overall factor of h^{-1}. However, it was realized a long time ago by Richardson [19] that this rule would *not* necessarily apply to grid functions which are solutions to finite-difference analogues of partial differential equations. Richardson was specifically concerned with grid functions which satisfied second-order, *centered* difference approximations, and, although his considerations apply in modified form to cases where other types of differencing are used, we will also focus on second-order, centered approximations. Richardson argued that if a grid function $\hat{u}$ satisfies an $O(h^2)$, centered-difference approximation to some partial differential equation (PDE) which has the exact solution u, then, in the limit $h\to 0$, we can expect u to have an expansion of the form

$$u=\hat{u}+h^2e_2+h^4e_4+\cdots\ , \tag{1.1}$$

where $e_2,e_4,\ldots$ are *error functions* which are *independent of the mesh spacing* h. On the basis of the (presumed) existence of such expansions, Richardson went on to introduce the well-known technique of *Richardson extrapolation.* However, he also noted that the existence of such an expansion implies, for example, that if we take a numerical first derivative of $\hat{u}$, using an $O(h^2)$ difference formula, then we will in general get an $O(h^2)$, *not an* $O(h)$, approximation of the first derivative of u.

The main purpose of this paper is to reanalyze this issue of consistency in numerical relativity in light of these old ideas of Richardson. The basic conclusion we will come to is that provided that we are sufficiently careful with our finite differencing, it is not too difficult, in principle, to construct a freely evolved differenced form of Einstein's equations which preserves a discrete form of the constraints to the same order of accuracy as the accuracy of the underlying difference scheme. Here, an important qualification is in order: the arguments presented for consistency can be expected to apply *only when we are dealing with PDE's and initial data which result in sufficiently smooth solutions.* For example, we do not expect expansions of the form (1.1) to hold for numerical solutions of model systems which admit shocks or turbulent phenomena. For such problems, which have the feature that a "wave-number cascade" concentrates power on the *smallest* resolvable scale (i.e., the grid scale h) in any numerical computation, we simply cannot usefully apply Richardson's ideas. However, it is clear that there are many interesting general-relativistic problems which *do* involve smooth solutions. One such problem is the subject of Sec. III, where we present some results from two separate codes where the expected (consistent) rate of convergence of $\hat{\xi}$ has been observed.

II. ANALYSIS OF CONSISTENCY IN NUMERICAL RELATIVITY

In this section we go through the previously sketched arguments in more detail, and for this, some familiarity with finite-difference techniques for initial-value problems [20] (time-dependent partial differential equations) will be helpful. However, we must first define some notation, and in order to keep the development reasonably self-contained, we review some basic notions from numerical analysis which are useful in analyzing the issue of consistency.

The principal mathematical objects we will be concerned with in our discussion are *grid functions.* As the name suggests, these are functions defined on a *grid (mesh, lattice)* and will generally satisfy finite-difference analogues of partial differential equations. A grid is a discrete set of events which typically forms a regular lattice in the coordinate space of our problem domain. The specific differential equations we consider in this paper involve, at most, dependence on one spatial coordinate $x^1\equiv x$ or r, and time, $x^0\equiv t$. (Differentiation with respect to these variables will often be denoted by a prime and an overdot, respectively.) We will always work with grids which are either (1) *uniform* (constant spacing between grid points in each coordinate direction) and *rectilinear* (aligned with the coordinate axes), or (2) the union of some small number of such grids. Specifically, a uniform, rectilinear grid, Ω^2, satisfies

$$\Omega^h\subset\{(x^0_n,x^1_j)\,|\,x^0_n=x^0_0+n\Delta x^0,x^1_j=x^1_0+j\Delta x^1\}\ , \tag{2.1}$$

where n and j assume integral values, the origin of the mesh is at (x^0_0,x^1_0), and Δx^0 and Δx^1 are of order-1 multiples of the basic mesh spacing *(grid spacing, resolution),* h. Because we will be considering hyperbolic systems, we

expect temporal *and* spatial variations in our solutions to be on the same scale, so we will keep $\lambda \equiv \Delta x^0/\Delta x^1$ fixed when we vary h. Thus $(\Delta x^0)^p$ and $(\Delta x^1)^p$ will always be $O(h^p)$.

Grid functions $(\hat{g},\tilde{g},\ldots)$ will be distinguished from continuum functions $(g,\ldots)$ by the use of a caret or tilde. We will use the former notation $(\hat{g})$ for all grid functions except those, such as $\tilde{g}$, which are associated with the constraint equations. Continuum functions, which generally satisfy differential equations, are, of course, defined at all events of the continuum. Strictly speaking, when interpreting, "mixed-mode" expressions such as "$g-\hat{g}$," we should adopt either of the following conventions: continuum functions are restricted to the mesh in some fashion, or mesh functions are in some way extended to the continuous domain. Our discussion will not be sufficiently rigorous that our conclusions will depend on which interpretation we choose, but in understanding Richardson's ideas, it does seem useful to adopt the latter viewpoint. Also, we will use an operator formalism when we write down and manipulate actual finite-difference expressions and the view that for any given resolution h any specific grid function is defined at *all* events of the continuum is more logically consistent with this formalism. The finite-difference operators we use are defined in Fig. 1; also given are the first few terms of the formal expansions (in powers of h) of these operators in terms of *differential* operators.

Now, let u be a function satisfying the differential equation

$$Lu=0\ , \tag{2.2}$$

where L is some differential operator, and let

$$\hat{L}\hat{u}=0 \tag{2.3}$$

be a finite-difference approximation of (2.2) where $\hat{L}$ is some finite-difference analogue of L. We define the *truncation error* $\hat{\tau}$ of the operator (difference scheme) $\hat{L}$ as

$$\hat{\tau}\equiv\hat{L}u\ . \tag{2.4}$$

We say the difference scheme is pth-order accurate [or pth order, or $O(h^p)$] if $\hat{\tau}=O(h^p)$. We also define the *solution error* $\hat{e}$ by

$$\hat{e}\equiv u-\hat{u}\ . \tag{2.5}$$

Let us say that the difference scheme (2.4) is *optimally convergent* if the truncation error and the solution error are *of the same order* in the mesh spacing, that is, if $O(\hat{e})=O(\hat{\tau})$. Previous discussions of the consistency of free-evolution schemes have tacitly assumed that the difference schemes used in numerical relativity *are* optimally convergent. This assumption is rather central to the issue of consistency and we will later present heuristic arguments for why we should expect it to be valid. At this point, for the sake of exposition, we will also proceed with the assumption that our difference schemes are generically optimally convergent.

We can now easily reconstruct the argument [17,18] which suggests that freely evolved finite-difference solutions are generally inconsistent with the constraint equations. Let $\mathcal{G}$ and $\mathcal{T}$ denote the sets of geometric and matter variables for some solution of Einstein's equations. For some particular geometric variable (three-metric component or extrinsic curvature component) contained in $\mathcal{G}$, let g and $\hat{g}$ denote the exact solution and the output from some $O(h^p)$ free-evolution scheme, respectively. Thus, the continuum function g satisfies the differential evolution equation

$$\dot{g}=L[\mathcal{G},\mathcal{T}]\ , \tag{2.6}$$

where L is some differential operator, while the grid function $\hat{g}$ satisfies a finite-difference equation

$$\hat{U}\hat{g}=\hat{L}[\hat{\mathcal{G}},\hat{\mathcal{T}}]\ . \tag{2.7}$$

Here $\hat{U}$ and $\hat{L}$ are $O(h^p)$ approximations of $\partial/\partial t$ and L. Now, the truncation error $\hat{\tau}$ of the discrete evolution equation is

$$\hat{\tau}\equiv\hat{U}g-\hat{L}[\mathcal{G},\mathcal{T}]\ , \tag{2.8}$$

and, by construction, is $O(h^p)$. By the assumption of optimal convergence, we have that the solution error $\hat{e}$ of the freely evolved grid function

$$\hat{e}\equiv g-\hat{g} \tag{2.9}$$

is also $O(h^p)$. We now consider one of the constraint equations, which we write schematically as

DIFFERENCE OPERATORS

$$\Delta_\pm^x f \equiv \pm\frac{f(x\pm h)-f(x)}{h}=\frac{df}{dx}+\frac{1}{24}h^2\frac{d^3f}{dx^3}+O(h^4)\Big|_{x\pm\frac{h}{2}}$$

$$\Delta_0^x f \equiv \frac{f(x+h)-f(x-h)}{2h}=\frac{df}{dx}+\frac{1}{6}h^2\frac{d^3f}{dx^3}+O(h^4)\Big|_{x}$$

$$\Delta_+^x\Delta_-^x f=\frac{f(x+h)-2f(x)+f(x-h)}{h^2}=\frac{d^2f}{dx^2}+\frac{1}{12}h^2\frac{d^4f}{dx^4}+O(h^4)\Big|_{x}$$

AVERAGING AND EXTRAPOLATION OPERATORS

$$\mu_\pm^x f \equiv \frac{f(x\pm h)+f(x)}{2}=f+\frac{1}{8}h^2\frac{d^2f}{dx^2}+O(h^4)\Big|_{x\pm\frac{h}{2}}$$

$$\chi_+^x \equiv \frac{3f(x)-f(x-h)}{2}=f-\frac{3}{8}h^2\frac{d^2f}{dx^2}+O(h^3)\Big|_{x+\frac{h}{2}}$$

Note: μ and χ have precedence over other algebraic operations. For example:

$$\mu_+^x\left(\frac{fg^2}{h}\right)\equiv\frac{(\mu_+^x f)(\mu_+^x g)^2}{\mu_+^x h}$$

FIG. 1. Definitions of the difference, averaging, and extrapolation operators used in the text. When applied to smooth functions of more than one variable, $(x\in\{x^0,x^1\})$, the definitions and expansions remain valid (with $d/dx\to\partial/\partial x^0$ or $\partial/\partial x^1, h\to\Delta x^0$ or Δx^1) provided that a uniform, rectlinear mesh is used.

$$H[\mathcal{G},\mathcal{T}]=0\ , \tag{2.10}$$

and which we assume can be solved for g. We further assume, as will typically be the case, that this equation includes terms linear in the first and second (spatial) derivatives of g. Thus, for the sake of argument, we write (2.10) as

$$Hg\equiv P[\mathcal{G},\mathcal{T}]g''+Q[\mathcal{G},\mathcal{T}]g'+R[\mathcal{G},\mathcal{T}]=0\ , \tag{2.11}$$

and, using the operators defined in Fig. 1, we difference it *to second order* as

$$\tilde{H}\tilde{g}\equiv\tilde{P}[\tilde{\mathcal{G}},\tilde{\mathcal{T}}]\Delta_+^x\Delta_-^x\tilde{g}+\tilde{Q}[\tilde{\mathcal{G}},\tilde{\mathcal{T}}]\Delta_0^x\tilde{g}+\tilde{R}[\tilde{\mathcal{G}},\tilde{\mathcal{T}}]=0\ . \tag{2.12}$$

Thus, $\tilde{g}$ satisfies a discrete version of the constraint equation, and will generally differ from $\hat{g}$. Associated with $\tilde{H}$ is another truncation error $\tilde{\tau}$ defined by

$$\tilde{\tau}\equiv\tilde{H}g\ , \tag{2.13}$$

which is $O(h^2)$ by construction. Now we consider the quantity $\hat{\xi}$, which, as discussed in the Introduction, measures how well the freely evolved grid function satisfies the discrete constraint equation. Specifically,

$$\begin{aligned}\hat{\xi}&\equiv\tilde{H}\hat{g}\\&=\tilde{H}(g-\hat{e})=\tilde{\tau}-\tilde{P}[\hat{\mathcal{G}},\hat{\mathcal{T}}]\Delta_+^x\Delta_-^x\hat{e}\\&\qquad-\tilde{Q}[\hat{\mathcal{G}},\hat{\mathcal{T}}]\Delta_0^x\hat{e}+O(h^2)\ ,\end{aligned} \tag{2.14}$$

where there are additional $O(h^2)$ terms which arise if $\tilde{H}$ is nonlinear or if the coefficient functionals $\tilde{P}$, $\tilde{Q}$, and $\tilde{R}$ in (2.12) depend on $\tilde{g}$. Now, $\hat{e}$ is an $O(h^p)$ quantity. The difference operators $\Delta_+^x\Delta_-^x$ and Δ_0^x have factors h^{-2} and h^{-1}, respectively. Thus, we reason

$$\Delta_+^x\Delta_-^x\hat{e}=O(h^{p-2}) \tag{2.15}$$

and

$$\Delta_0^x\hat{e}=O(h^{p-1})\ . \tag{2.16}$$

In particular, for second-order, free-evolution schemes, $p=2$, and this argument implies that $\hat{\xi}=O(1)$. However, $\Delta_+^x\Delta_-^x$ and Δ_0^x are *operators* not algebraic quantities, and there is no justification for this last step, which amounts to an invalid application of the familiar "rule of thumb" which states that a numerical differentiation generally produces an $O(h^{p-1})$ quantity from an $O(h^p)$ one. More care is needed in our analysis, as the following simple example illustrates.

Consider the equation

$$Lu\equiv\left[\frac{d}{dx}-1\right]u=0 \tag{2.17}$$

on the half line, $x\geq 0$. With the initial condition $u(0)=1$, this equation has the solution

$$u=e^x\ . \tag{2.18}$$

Let us first consider the problem of computing an $O(h^2)$-accurate grid function $\hat{u}$, which approximates u. Using the operator notation of Fig. 1, we write down the following finite-difference version of (2.17):

$$\hat{L}\hat{u}\equiv(\Delta_+^x-\mu_+^x)\hat{u}=0\ , \tag{2.19}$$

where our grid is $\{jh, j=0,1,\ldots\}$. We can easily show that this is second-order accurate by performing Taylor-series expansions about $\bar{x}\equiv(j+\frac{1}{2})h$ (the event at which the difference scheme is centered):

$$\begin{aligned}\hat{\tau}\equiv\hat{L}u&=(\Delta_+^x-\mu_+^x)u\\&=\left[\frac{d}{dx}-1\right]\bar{u}+h^2\left[-\frac{1}{8}\frac{d^2}{dx^2}+\frac{1}{24}\frac{d^3}{dx^3}\right]\bar{u}\\&\quad+O(h^4)\\&=O(h^2)\ ,\end{aligned} \tag{2.20}$$

where $\bar{u}\equiv u(\bar{x})$. Let us adopt the usual subscript notation for specific elements of a grid function: $\hat{u}_j\equiv\hat{u}(x_j)\equiv\hat{u}(hj)$. Then the particular solution of difference equation (2.19) for $\hat{u}_0=1$ is

$$\hat{u}_j=\rho^j\ , \tag{2.21}$$

where

$$\rho\equiv\frac{1+h/2}{1-h/2}\ . \tag{2.22}$$

But

$$\begin{aligned}\rho^j&=\exp\left\{j\left[\ln\left[1+\frac{h}{2}\right]-\ln\left[1-\frac{h}{2}\right]\right]\right\}\\&=e^{x_j}+\tfrac{1}{12}h^2x_je^{x_j}+O(h^4)\ ,\end{aligned} \tag{2.23}$$

so we have

$$\hat{e}_j=u_j-\hat{u}_j=-\tfrac{1}{12}h^2x_je^{x_j}+O(h^4)\ . \tag{2.24}$$

We first note that since $\hat{e}_j$ is $O(h^2)$, the difference scheme (2.19) *is* optimally convergent. More importantly for our present discussion, however, is the *form* of the solution error. In accordance with Richardson's general arguments, the continuum solution u admits an asymptotic expansion of the form

$$u=\hat{u}+h^2e_2+h^4e_4+\cdots\ , \tag{2.25}$$

where $e_2, e_4, \ldots$ are h-independent functions with smoothness comparable to u (or more precisely, comparable to some appropriately high-order derivative of u). In fact, we can derive (2.24) in another way by *assuming* that the above expansion holds. Thus, if we solve (2.25) for $\hat{u}$, then substitute the result in (2.19) and expand about $\bar{x}$, we find (again, with the aid of the operator expansions given in Fig. 1)

$$\left[\frac{d}{dx}-1\right]\bar{u}+h^2\left[-\left[\frac{d}{dx}-1\right]\bar{e}_2+\left[\frac{1}{24}\frac{d^3}{dx^3}-\frac{1}{8}\frac{d^2}{dx^2}\right]\bar{u}\right]+O(h^4)=0\ . \tag{2.26}$$

Now, the terms in this expansion must vanish order by order in h. The order-1 term vanishes by virtue of the original differential equation (consistency of the difference scheme). The vanishing of the $O(h^2)$ term gives us a differential equation for e_2:

$$e_2' = e_2 + \left[\frac{1}{24}\frac{d^3}{dx^3} - \frac{1}{8}\frac{d^2}{dx^2} \right] u \ . \tag{2.27}$$

Since $u = e^x$, and $e_2(0) = 0$ (we use the exact initial condition in our difference scheme), this equation has the solution

$$e_2 = -\tfrac{1}{12} x e^x \ . \tag{2.28}$$

Thus we recover (2.24), as claimed.

Now, suppose that we want to numerically evaluate the second derivative of the difference solution (2.21). By the "inconsistency argument" outlined above, we expect $\Delta_+^x \Delta_-^x - \hat{u}$ to be (only) an order-1-accurate approximation of the exact second derivative, $d^2u/dx^2 = e^x$. However,

$$\Delta_+^x \Delta_-^x \hat{u} = h^{-2}(\rho - 2 + \rho^{-1})\hat{u} \tag{2.29}$$

and

$$\rho - 2 + \rho^{-1} = h^2 + \tfrac{1}{4}h^4 + O(h^6) \ . \tag{2.30}$$

Thus

$$\begin{aligned} \Delta_+^x \Delta_-^x \hat{u} &= [e^x + \tfrac{1}{12}h^2 x e^x + O(h^4)][1 + \tfrac{1}{4}h^2 + O(h^4)] \\ &= e^x + h^2(\tfrac{1}{4} + \tfrac{1}{12}x)e^x + O(h^4) \\ &= \frac{d^2u}{dx^2} + O(h^2) \end{aligned} \tag{2.31}$$

and we have, in fact, an $O(h^2)$ (optimally convergent), *not* order-1, approximation of the second derivative. Again, we can readily verify this result using (2.25) [with the specific leading-order error term given by (2.27)] and the expansion

$$\Delta_+^x \Delta_-^x = \frac{d^2}{dx^2} + h^2 \frac{1}{12}\frac{d^4}{dx^4} + O(h^4) \tag{2.32}$$

from Fig. 1.

The fact that the "numerical second derivative" of our finite-difference solution $\hat{g}$ is second-order accurate may be a little surprising. However, we see that the result follows rather directly if we adopt (2.25) as an *ansatz* and observe that finite-difference operators have asymptotic expansions in terms of differential operators. It was precisely this type of result which Richardson anticipated when he wrote [19] *(emphasis* provided) "... the errors of the integral [solution of a difference equation] and *of any differential expressions derived from it,* due to using the simple central differences ... instead of differential coefficients, are of the form"

$$h^2 f_2(x,y,z) + h^4 f_4(x,y,z) + h^6 f_6(x,y,z) + \cdots \ .$$

Let us return to the question of consistency of free-evolution schemes in $3+1$ numerical relativity. Provided that the following type of expansion holds for the geometric variable g:

$$g = \hat{g} + \sum_{q=p}^{\infty} h^q \hat{e}_q \ , \tag{2.33}$$

where the $\hat{e}_q$ are *smooth* functions, *all* of the quantities on the right-hand side of (2.14) are $O(h^p)$. Specifically, the suspect terms become

$$\Delta_+^x \Delta_-^x \hat{e} = h^p \hat{e}_p'' + O(h^{p+r}) \ , \tag{2.34}$$

$$\Delta_0^x \hat{e} = h^p \hat{e}_p' + O(h^{p+r}) \ , \tag{2.35}$$

for some $r > 0$. In particular, if $p = 2$, so that we are using a second-order accurate free-evolution scheme, then $\hat{\xi}$ should *also* be $O(h^2)$. However, even if $p = 1$, which is the best that can be expected for many schemes which have been used in numerical relativity, then $\hat{\xi}$ should be $O(h)$. Only if the underlying difference scheme is inconsistent with the original set of Einstein-matter equations ($\hat{\tau} \not\to 0$ as $h \to 0$) should a $\hat{\xi}$ of order 1 be observed.

Heuristically then, we are arguing the following. If we use a (free-evolution) difference scheme that produces grid functions which admit asymptotic expansions such as (2.33) (which we call Richardson expansions), then we should *not* observe any basic inconsistency of the type we have discussed. Furthermore it seems plausible to *expect* expandability (as we will refer to that property attained by grid functions which have Richardson expansions) for the solutions of carefully constructed difference schemes. To understand this, we observe that for any set of PDE's and a given finite-difference discretization of that set of PDE's, we can imagine writing down Richardson expansions for all of the grid functions and then carrying out a development along the lines of (2.25)–(2.28). This will involve expanding various finite-difference operators and deriving a hierarchy (in h) of additional systems of PDE's for the various error functions appearing in the expansions. In order to rigorously establish expandability, we would have to demonstrate that these additional systems are well posed, and yield unique solutions when supplied with appropriate boundary data [21,22]. However, a key point is that the *form* (and hence *nature*) of these auxiliary systems is apt to be similar to (and at least as complex as) the form of the *original* set of equations. (For example, if we discretize a "wave equation," and our finite-difference solutions of this equation exhibit wave like behavior, then surely the various error functions satisfy some other "wave equations.") Thus, if we are willing to accept the existence and uniqueness of the solution of the continuum equations, it seems reasonable for us to expect that our grid functions will be expandable provided, of course, that they are convergent.

Finally, even though we may not be able to *prove, a priori,* the proposition that a given finite-difference solution of Einstein's equations will be expandable (from which consistency follows almost immediately), we can always perform "numerical experiments" designed to test the proposition. Moreover, these experiments are, in principle, quite easy to carry out. They involve little more than performing several computations which use the same initial data, but different values of h. Such experiments are the subject of the next section, where we

first discuss two $O(h^2)$ difference schemes for a model problem in numerical relativity and then show that, at least empirically, the schemes *are* consistent, as we have argued they should be.

III. SOME CONSISTENT NUMERICAL SOLUTIONS

The results described in the remainder of this paper come from a numerical study [23] of one of the simplest conceivable general-relativistic systems. The model consists of a single massless scalar field minimally coupled to the gravitational field with the additional restriction of spherical symmetry. For several reasons, this system is particularly useful for providing clear evidence of the consistent behavior of finite-difference schemes which we posited in the preceding section. First, due to the spherical symmetry, the gravitational field has no "true" (radiative) dynamical degrees of freedom. Provided appropriate choices of coordinate systems are made, this means that we can construct difference schemes which are either *completely constrained* (no use of the geometric evolution equations) or *completely freely evolved* (no use of the constraint equations, except at $t=0$). Second, the high degree of symmetry renders the model quite tractable computationally so that, using reasonable amounts of computer time, it is possible to investigate the behavior of our finite-difference solutions over a considerable range of values for h. Finally, the simplicity of the equations of motion, the spherical symmetry, and the fact that solutions of the model equations are asymptotically flat, conspire to make it relatively easy to design numerical boundary conditions which produce convergent finite-difference solutions, *on a bounded computational domain.* By contrast, in more generic (less symmetric) calculations, it is generally the case that we expect to recover the exact solution only in the limit that $h\rightarrow 0$ and $R\rightarrow\infty$ *simultaneously,* where R is the radius of the computational domain, and this is largely due to the use of boundary conditions which are asymptotically ($R\rightarrow\infty$) correct. We take advantage of these features of the model to set up numerical experiments where we can directly and unambigously measure the rate of residual quantities such as $\hat{\xi}$.

The full description of the differential equations for the model and the difference techniques used to generate the numerical solutions is somewhat involved and has been documented in detail elsewhere [23]. Here, a schematic view of the various equations will be given, along with a discussion of those numerical points which bear directly on the issue of consistency. (We note at this point that this model has also been studied recently in considerable detail by Christodoulou [24] (analytic work) and Goldwirth and Piran [25] (numerical work based on the method of characteristics)).

The manifestly covariant field equations for the model are easily derived from the action

$$I=\int dV\left[\frac{1}{16\pi}R-\tfrac{1}{2}\phi_{,\mu}\phi^{,\mu}\right], \tag{3.1}$$

where R is the Ricci scalar and ϕ is the scalar field, in the standard fashion, with the familiar results

$$G_{\mu\nu}=8\pi T_{\mu\nu}\ , \tag{3.2}$$

$$T_{\mu\nu}=\phi_{,\mu}\phi_{,\nu}-\tfrac{1}{2}g_{\mu\nu}\phi_{,\mu}\phi^{,\mu}\ , \tag{3.3}$$

$$\phi^{;\mu}{}_{;\mu}=\frac{1}{\sqrt{-g}}(\sqrt{-g}\,\phi^{,\mu})_{,\mu}=0\ . \tag{3.4}$$

Here g is the determinant of the four-metric $g_{\mu\nu}$ and we have adopted the units and sign conventions of Misner, Thorne, and Wheeler [26]. Hamiltonian equations of motion for the "true" Arnowitt-Deser-Misner (ADM) variables (g_{ij},π^{ij}) may also be derived [27,28] from (3.1). Here, we adopt the $3+1$ approach which has been used so extensively in numerical relativity, and specialize it to the case of spherical symmetry. Using the standard spherical-polar coordinates $x^i=(r,\theta,\varphi)$, the three-metric and extrinsic curvature tensors can both be written in diagonal form, each with two independent components:

$$g_{ij}=\mathrm{diag}(a(r,t)^2,r^2b(r,t)^2,r^2b^2\sin^2\theta)\ , \tag{3.5}$$

$$K^i{}_j=\mathrm{diag}(K^r{}_r(r,t),K^\theta{}_\theta(r,t),K^\theta{}_\theta)\ . \tag{3.6}$$

Generally, $(a,K^r{}_r)$ and $(b,K^\theta{}_\theta)$ constitute dynamical pairs of variables. Introducing the remaining two degrees of coordinate freedom via the lapse function $\alpha(r,t)$, and the shift vector $\beta^i=(\beta(r,t),0,0)$, the most general spherically symmetric space-time metric is

$$ds^2=(-\alpha^2+a^2\beta^2)dt^2+2a^2\beta\,dt\,dr+a^2dr^2+r^2b^2d\Omega^2\ , \tag{3.7}$$

where $d\Omega^2$ is the (flat) metric on the two-sphere.

The precise forms of the $3+1$ versions of the equations of motion, (3.2) and (3.4), depend strongly on the *specific* coordinate system we use. We would like to argue that, in spite of this fact, our choice of coordinates will have little (if any) impact on whether or not we can construct consistent free-evolution difference schemes. In order to provide empirical support for this conjecture, we perform two sets of numerical experiments using two distinct coordinate systems. Both of these coordinate systems which we use have been employed in previous numerical studies of spherically symmetric, general relativistic systems—most notably the investigations of Shapiro and Teukolsky [15,16] involving the coupling of distributions of collisionless particles to the gravitational field.

In the $3+1$ formalism, a coordinate system is specified by giving (1) a *slicing* condition, which determines the lapse function, and (2) a condition which determines the shift-vector component β, and hence fixes the radial coordinate labeling. We first describe the two slicing conditions we use. In both cases, the lapse function is determined through a condition on the mean extrinsic curvature K:

$$K\equiv K^i{}_i=K^r{}_r+2K^\theta{}_\theta\ . \tag{3.8}$$

The first condition is *maximal slicing* [6–11,14–17,23,29,30]

$$K=0 \tag{3.9}$$

which, when imposed on each slice of a spacetime ($\partial K/\partial t=0$), constrains the lapse to satisfy a linear,

homogeneous, elliptic equation. The second condition, resulting in what is known as *polar slicing* [16,23,31,32], is defined by

$$K = K^r{}_r \ . \tag{3.10}$$

In this case, the lapse satisfies a linear, homogeneous, parabolic equation. (It should be noted that due to the spherical symmetry, the slicing equations as well as the constraint equations are *ordinary* differential equations, so the terms "elliptic" and "parabolic" are to be understood to imply "two-point boundary-value problem" and "initial-value problem," respectively). Clearly, both of these choices allow an extrinsic curvature component to be eliminated from the $3+1$ equations.

Two different conditions are also used to fix our radial coordinates through the specification of the shift-vector component β. Our first choice requires that

$$a = b \tag{3.11}$$

at all events, resulting in *isotropic* (sometimes called *isothermal* [23,30]) coordinates and a first-order equation for β which is solved at each time step. We use this condition in conjunction with maximal slicing, giving what we will call the *maximal-isotropic* coordinate system. Our other choice of radial coordinates follows from the demand that

$$b = 1 \tag{3.12}$$

at all events. This yields *radial* or *areal* [31] coordinates—in such a system, a two-sphere of radius r has proper surface area $4\pi r^2$. When combined with polar slicing, it can be shown that β must vanish everywhere and these *polar-radial* coordinates may be thought of as a natural extension of the standard Schwarzschild labeling for a vacuum spacetime to a nonvacuum spherical system. The simplification of the Einstein–scalar-field equations induced by this choice is quite remarkable; this fact has been used to advantage in analytic studies of the model [27,28]. Particularly noteworthy are the facts that the Hamiltonian constraint (considered as an equation for a) is nonlinear, but *first order* and the momentum constraint (considered as an equation for $K^r{}_r$) is *algebraic*.

Now, some of these features of the polar-radial coordinate system caused concern when we were setting up the numerical experiments designed to demonstrate consistency of a free-evolution scheme. In particular, since the constraint equations in this coordinate system do *not* involve second spatial derivatives of the geometric variables, a skeptic could argue that our computations in this coordinate system were not truly representative of a generic calculation in numerical relativity. Thus, in designing the numerical experiments, we decided to monitor *additional* residual quantities (quantities which should vanish as $h \to 0$) which *did* involve taking numerical second derivatives of a numerical solution.

For example, through the introduction of auxiliary variables [23,33], the equations of motion for the geometric variables (in either of the coordinate systems we have adopted) may be written schematically as

$$\begin{aligned} \dot{q} &= \beta q' + \gamma p' + R \ , \\ \dot{p} &= \beta p' + \delta q' + S \ , \end{aligned} \tag{3.13}$$

where $(q,p) \in \{(a_q a', K^r{}_r), (b_q b', K^\theta{}_\theta)\}$ and a_q, b_q, γ, δ, R, and S are functions of the geometric and matter variables, r and t. It is this form of the evolution equations which we finite difference to generate the results described below. Now, using the notation of the preceding section, one of the additional residual quantities we can compute is a measure of how well a *constrained* grid function $\tilde{g}$ satisfies the discrete evolution equation which is exactly satisfied by the freely evolved grid function $\hat{g}$. From the structure of (3.13) it is clear that if g is a three-metric component, then we will have to evaluate a numerical second derivative of $\tilde{g}$ in order to compute such a residual, and, of course, the inconsistency argument implies that such a quantity is also of order 1 if $\tilde{g}$ is $O(h^2)$. Again, our aim here is to ensure that our numerical experiments, as much as possible, give an unambigouous indication of whether there is a fundamental "consistency problem" with our finite-difference solutions.

We now briefly discuss the basic approach we take in finite differencing the equations of motion for our model problem. The most notable feature of our difference schemes is that they are constructed to be *precisely* second-order accurate [$O(h^2)$ truncation error] in both space and time. As we argued in the preceding section, it is not that we feel it necessary to use $O(h^2)$ methods in order to demonstrate consistency. Rather, our approach of using $O(h^2)$ approximations everywhere in our schemes, and then establishing that quantities such as $\hat{\xi}$ are indeed $O(h^2)$, is once again principally motivated by our desire to produce clear evidence for the consistency of our particular difference solutions.

The sets of PDE's we are studying are hyperbolic, and we write them in first-order (in time) form before we finite difference them. As an aid in constructing $O(h^2)$ schemes, we use a "staggered " mesh [20] (see Fig. 2), which facilitates the centering of various spatial and temporal finite-difference expressions. The $3+1$ evolution equations for the geometric variables are, by construction, in first-order form. The wave equation (3.4) for the scalar field can be placed in this form by introducing aux-

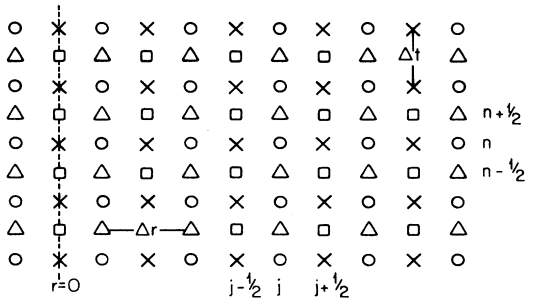

FIG. 2. Structure of the finite-difference mesh used in the algorithms discussed in the text. The various grid functions used are defined on the four constituent submeshes as follows: $G_\bigcirc : \{a^n_j, b^n_j, \alpha^n_j, \phi^n_j\}$, $G_\triangle : \{K^{r\,n+1/2}_{rj}, K^{\theta\,n+1/2}_{\theta j}, \Pi^{n+1/2}_j\}$, $G_\times : \{\beta^n_{j+1/2}, \Phi^n_{j+1/2}\}$, $G_\square : \{\beta^{n+1/2}_{j+1/2}\}$.

iliary variables

$$\Phi \equiv \phi' \tag{3.14}$$

and

$$\Pi \equiv \frac{a}{\alpha}(\dot{\phi}-\beta\phi') \ . \tag{3.15}$$

The advantage of this particular choice of variables is that the resulting pair of evolution equations

$$\dot{\Phi}=\left[\beta\Phi+\frac{\alpha}{a}\Pi\right]' , \tag{3.16}$$

$$\dot{\Pi}=\frac{1}{r^2b^2}\left[r^2b^2\left[\beta\Pi+\frac{\alpha}{a}\Phi\right]\right]'+2\left[\alpha K^{\theta}{}_{\theta}-\beta\frac{(rb)'}{rb}\right]\Pi \ , \tag{3.17}$$

do not involve time derivatives of the kinematical functions α and β. As can be seen from Fig. 2, the computational mesh can be regarded as the union of four component meshes, each of which is uniform in both space and time. Having adopted this grid structure, it was frequently necessary to use averaging, and sometimes extrapolation, operations in order to produce second-order schemes, particularly in the maximal-isotropic coordinate system. However, given the basic strategy of keeping the difference equations precisely centered, and provided the "correct" forms of the continuum equations were used [23], the actual differencing techniques were quite straightforward. For example, Eq. (3.16) was differenced as

$$\begin{aligned}\Delta_+^t\Phi^n_{j+1/2}=&\tfrac{1}{2}\Delta_0^r(\chi^t_+\beta^{n+1/2}_{j+1/2}\Phi^{n+1}_{j+1/2}+\beta^n_{j+1/2}\Phi^n_{j+1/2})\\&+\Delta_+^r\left[\chi^t_+\left[\frac{\alpha}{a}\right]^n_j\Pi^{n+1/2}_j\right]\end{aligned} \tag{3.18}$$

in the maximal-isotropic system.

Finally, we maintain a uniform time step in both coordinate systems by using part of the coordinate freedom embodied in the specification of the lapse function α. Basically, since both of the slicing conditions are linear and homogeneous in α, the lapse is only determined up to an overall normalization which can be chosen so that the maximum local speed of signal propagation in the numerical solution domain remains constant [23]. The usual practice in numerical relativity [17] involves choosing a new Δt at each time step to ensure that a Courant-Friedrichs-Levy [20] stability criterion is maintained. The resulting time variability in Δt makes the construction of a second-order accurate scheme considerably more involved.

The two programs (one for each coordinate system) designed to solve the difference systems sketched above were each constructed to perform two calculations in tandem using the same initial data and grid structure. In one calculation, the pair of geometric variables, which we denote (a,K), was freely evolved, giving as output the mesh functions (a^E,K^E). The other calculation produced (a^C,K^C) which satisfied discrete versions of the Hamiltonian and momentum constraints. Note that in this section we use a superscript E or C (rather than a caret or tilde) to denote freely evolved and constrained grid functions, respectively. Generally, the equations for the lapse and shift were solved using the constrained quantities (a^C,K^C) as "coefficient" functions; however, calculations done using the freely evolved functions in this capacity yielded essentially indistinguishable results.

We write the difference equations for the geometric variables schematically in the form

$$\mathcal{O}^E a^E=\mathcal{O}^E K^E=\mathcal{O}^C a^C=\mathcal{O}^C K^C=0 \ , \tag{3.19}$$

where, for brevity of notation, the definition of $\mathcal{O}^E$ and $\mathcal{O}^C$ depends on what is being operated upon. Then, in order to investigate the consistency of our difference schemes, the programs compute the quantities

$$\mathcal{O}^E a^C, \ \mathcal{O}^E K^C, \ \mathcal{O}^C a^E, \mathcal{O}^C K^E, \ a^E-a^C, \text{ and } K^E-K^C$$

at all points of the relevant subgrids. Each of these quantities is a residual. If there is an inconsistency, then some, or all, of these functions should be of order 1 or $O(h)$. The basic idea of the numerical experiment is to monitor a measure of the sizes of these residuals, as a function of time, and attempt to determine their orders by repeating the process on a sequence of grids with mesh spacings

$$h^l \equiv 2^{-l}h_0, \ l=0,1,\ldots ,$$

where h_0 is the spacing of the coarsest mesh. We denote by $\mathcal{O}$ the l_2 norm of any of the residual quantities, $r\in\{\mathcal{O}^E a^C,\mathcal{O}^E K^C,\ldots\}$:

$$\mathcal{Q}(t)\equiv\mathcal{Q}(n\Delta t)\equiv N^{-1}\left[\sum_{j=1}^{N} r_j^n r_j^n\right]^{1/2} , \tag{3.20}$$

where N is the total number of mesh points in the radial direction. The level -l *convergence factor* of such a quantity is defined simply as

$$\frac{\mathcal{Q}^l}{\mathcal{Q}^{l+1}} \ . \tag{3.21}$$

Then, any order-1 quantity, for example, should have a convergence factor of 1 as $h\to 0$, while an $O(h^2)$ quantity will have a convergence factor of 4.

In a particular trial using the polar-radial coordinate system, initial data were constructed by specifying the initial value of ϕ,

$$\phi(r,0)=(1.0\times 10^{-3})\exp\left[-\left[\frac{(r-50.0)}{5.0}\right]^2\right] , \tag{3.22}$$

and demanding that the solution be *time symmetric* [26] about $t=0$. This means that the functions Π, $K^r{}_r$, and $K^{\theta}{}_{\theta}$ all vanish on the initial slice. The remaining variables were then determined from the coordinate condition equations and constraint equations. $O(h^2)$ Taylor-series methods were used to determine the values of Π at $t=-\frac{1}{2}\Delta t$ in accordance with the staggered structure of the mesh; the corresponding values of $K^r{}_r$ were then cal-

culated from the momentum constraint. The outer edge of all numerical grids was at $r=100.0$ and the total mass of the solution was 3.9×10^{-3}. The initial data and number of time steps were chosen so that there was no interaction of the scalar-field pulses (one outgoing and one ingoing) with either grid boundary before the runs terminated. This was done so that, as much as possible, the results would be unaffected by boundary effects.

Grids with

$$\Delta r=2\Delta t=\frac{1}{2^l},\quad 0\le l\le 4 \tag{3.23}$$

were used, and all runs stopped at $t=25.0$. Thus the coarsest grid was 100×50 (radial points $\times$ temporal points), and the finest 1600×800. The l_2 norms of the residual quantities were output every 2^l steps.

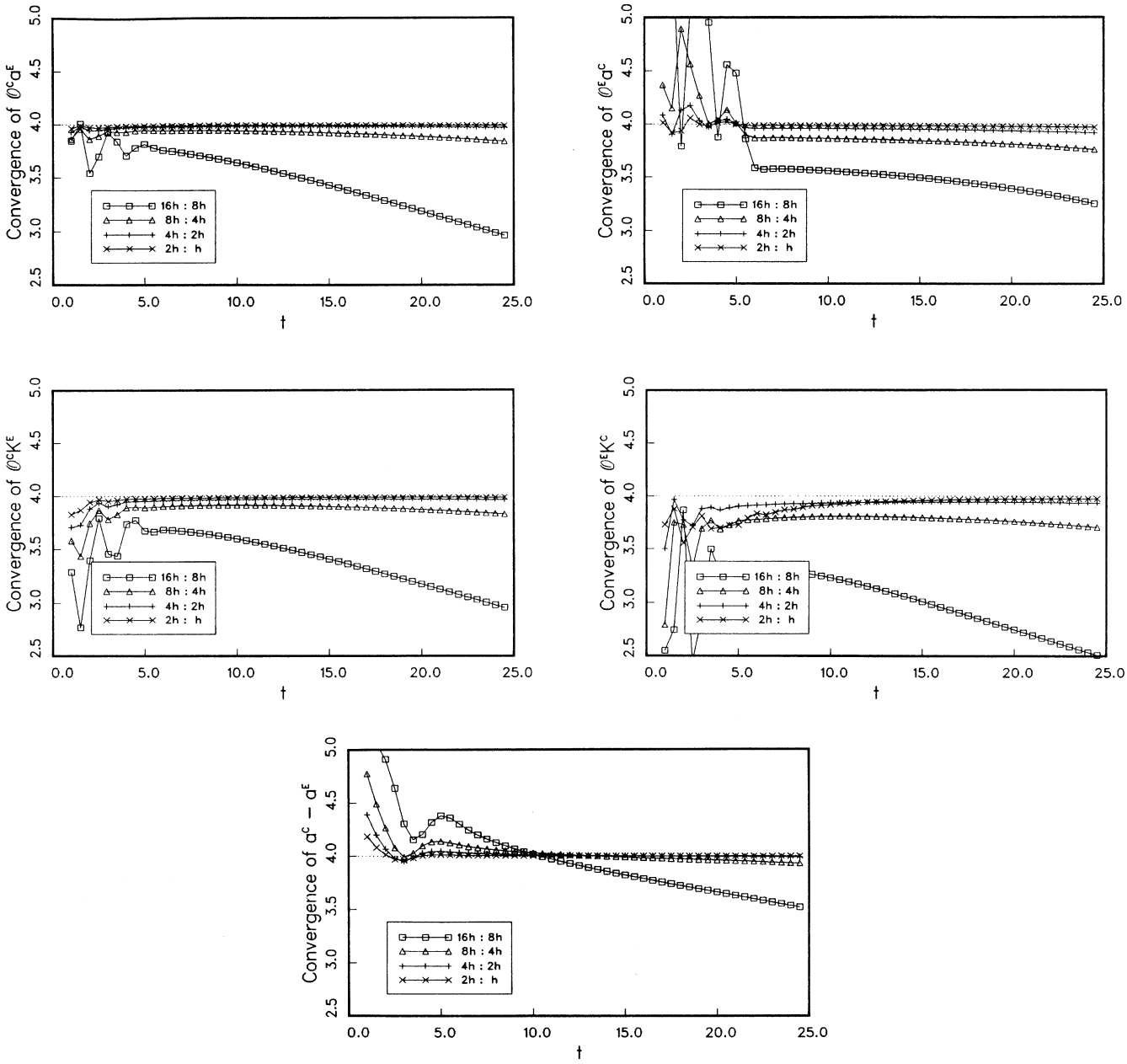

FIG. 3. Convergence factors of various residual quantities defined in the text and computed in the polar-radial coordinate system. Five geometrically related scales of discretization were used. The expected behavior as $h\to0$ for all of these quantities is the constant function 4. Note the decreasing amplitude of the initial "transient" behavior on finer meshes. The transients are probably largely due to the preparation of the initial geometric data in a fully constrained form.

The computed convergence factors as defined by (3.21) and (3.22) are shown in Fig. 3. (In these plots, h is the finest radial spacing, $\frac{1}{16}, K \equiv K^r{}_r$ and the symbols mark actual data points.) The dotted line in each graph indicates the ideal convergence factor of 4. Although the convergence rates deteriorate with time on the coarser grids (presumably due to non-negligible higher-order terms) there can be little doubt that all of the deviations are $O(h^2)$ quantities, so there is no inconsistency. Note that there is no graph of the convergence of $K^E - K^c$. As noted above, in this coordinate system, the momentum constraint is algebraic, and linear in $K^r{}_r$, so that, in this case, $\mathcal{O}^C K^E$ and $K^E - K^C$ are identical.

A similar experiment was performed to demonstrate consistency in the maximal-isotropic coordinate system. Initial data were determined in the fashion described previously. The initial configuration of the scalar field was

$$\phi(r,0) = (1.0\times 10^{-2}) \exp\left[-\left[\frac{(r-20.0)}{5.0} \right]^2 \right], \qquad (3.24)$$

the solution was time symmetric, and had an ADM mass of 6.6×10^{-2}. Again, the same data were evolved on a series of grids with geometrically related mesh spacings

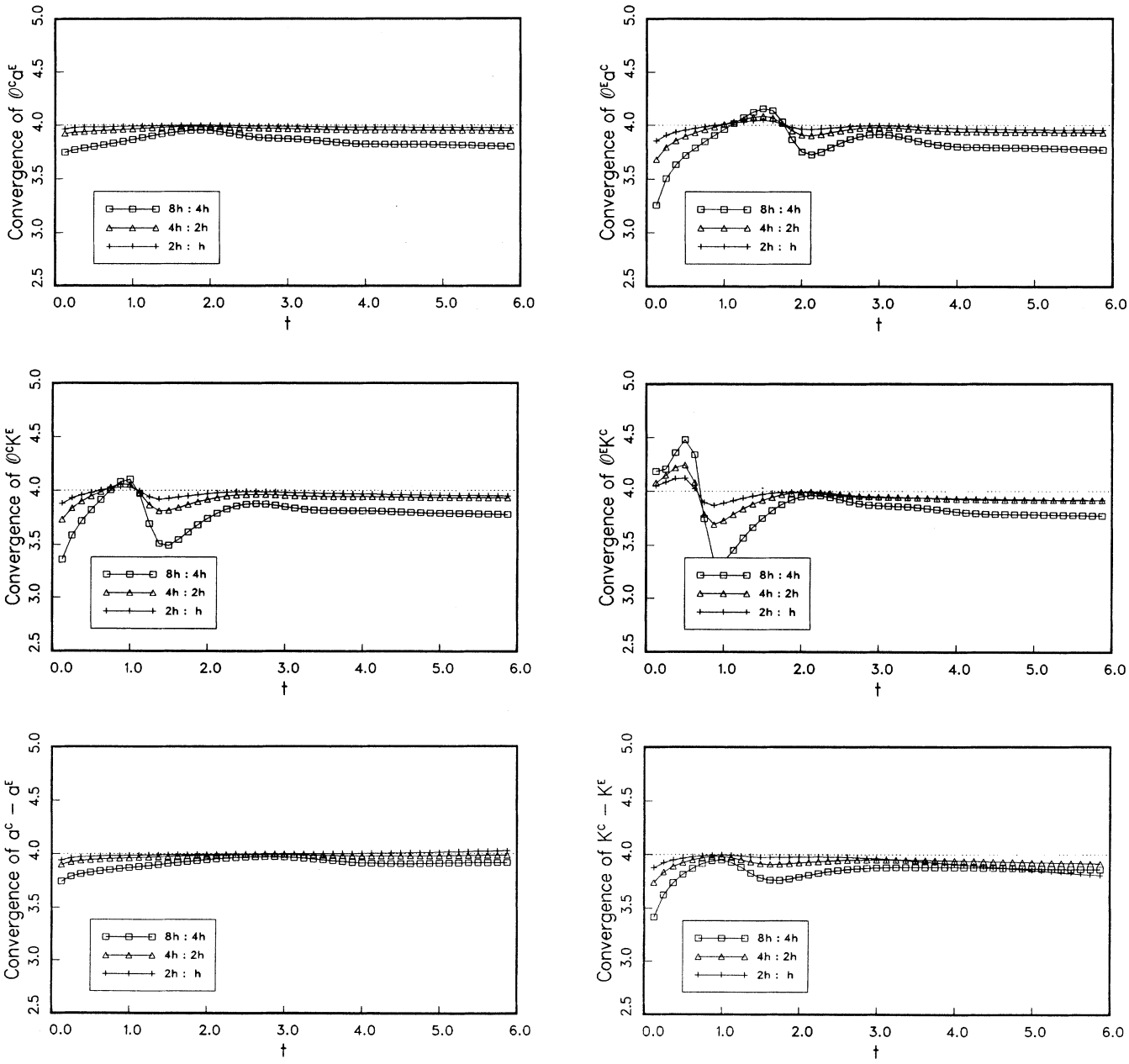

FIG. 4. Convergence factors of the various residual quantities in maximal-isotropic coordinates. Four different mesh spacings were used.

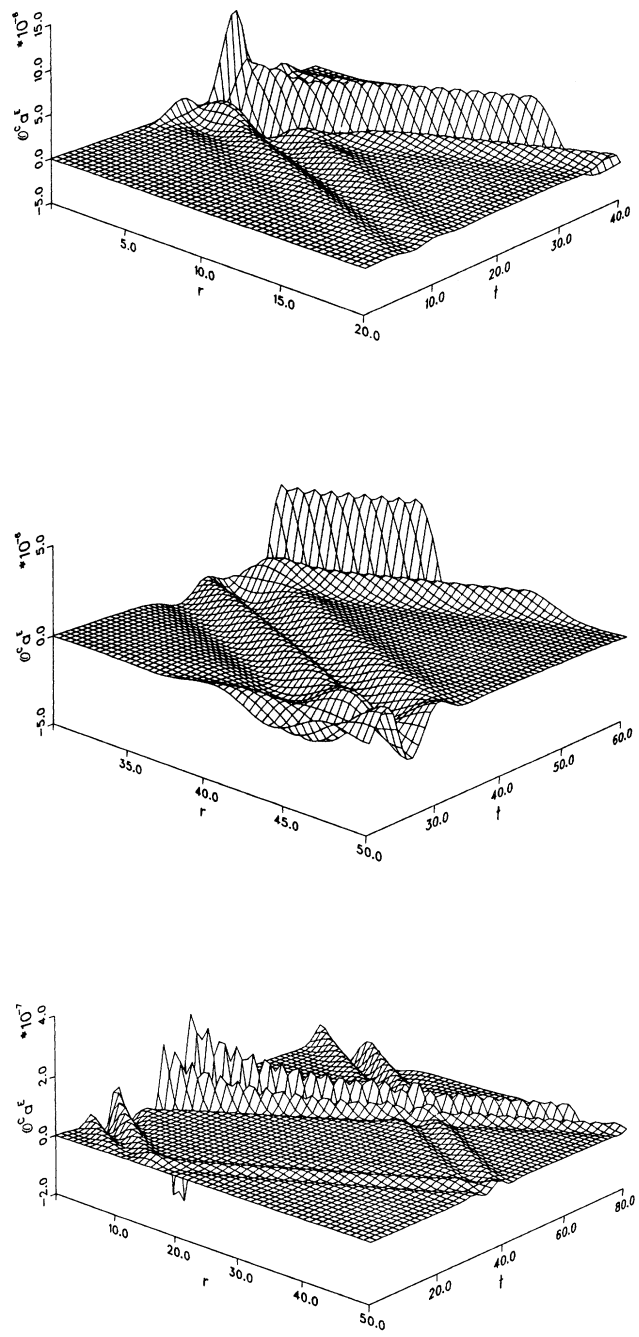

FIG. 5. Surface plots of the quantity $\hat{\xi} \equiv \mathcal{C}^C a^E$ in the two coordinate systems: (a) interior region in the polar-radial system, (b) exterior region in the polar-radial system, and (c) exterior region in the maximal-isotropic system. Features visible in (a) are due to the originally ingoing pulse of scalar radiation which "self-reflects" at $r=0$, then travels outward. In (b), which has half the vertical scale of (a), the dominant feature for early t is due to the outgoing pulse which suffers a small amount of "artificial" reflection at the outer boundary of the numerical domain. At late times, the disturbance described in (a) is propagating from left to right. In (c), which (unlike the radial axis labels might suggest) covers $r \geq 5.0$, features corresponding to those in both (a) and (b) appear.

and outer radii of 50.0. Four grids were used with

$$\Delta r = 4\Delta t = \frac{1}{2^{l+1}}, \quad 0 \leq l \leq 3 \tag{3.25}$$

and the runs ended at $t \approx 6$. The results of this test are shown in Fig. 4. In this set of plots, the finest radial spacing h is $\frac{1}{16}$ and K denotes $K^{\theta}{}_{\theta}$. Again, it seems clear that there is no inconsistency in either the constrained or freely evolved quantities. The only problems seem to be with $\mathcal{C}^E K^C$ and $K^C - K^E$. The convergency factor of the former appears to tend to a value somewhat less than 4, while that of the latter begins to deteriorate at the highest level of refinement. This suggests that there may be some problem with the way the momentum constraint was solved, and this needs further investigation.

Finally, surface plots of some of the actual residual quantities monitored in the experiments described above are shown in Fig. 5. Particularly notable are the facts that regions of high deviation are correlated with large values of the scalar field and that the deviations themselves are smooth functions. There is certainly no evidence of "exponential" growth [18] of the residuals—in fact, the residuals actually die off quite rapidly at late times in the maximal-isotropic system.

IV. DISCUSSION

The results we have presented provide strong support for the claim that properly constructed, freely evolved finite-difference solutions of Einstein's equations should satisfy the constraint equations to the truncation order of the difference scheme. Again, we stress that the behavior we have observed can be considered a direct consequence of the general existence of asymptotic expansions such as (2.32) for the solution of our difference equations. There seems to be no fundamental reason why we should not expect similar behavior for more general difference solutions in numerical relativity, including those with less symmetry or those involving stronger gravitational fields. Thus, although the standard procedure [4–9] of monitoring the degree to which freely evolved quantities fail to satisfy the constraints is not unreasonable [18], that procedure is really measuring a form of truncation error which cannot, in general, provide a direct *quantitative* estimate of the solution error. A better way to estimate the solution error is to make a comparison of the difference solutions obtained using two or more distinct mesh spacings and the same initial data. This can be done with the finite-difference solution of *any* set of differential equations, so, from the point of view of estimating the accuracy of a numerical solution of the field equations, there is nothing particularly special about the fact that general relativity has constraints [8].

Finally, the demonstration of consistency in the model system provides an excellent example of the following fact—the manner in which a difference solution changes with respect to the scale of the discretization will generally contain a great deal of useful information, particularly

about the error in the approximation. Although the basic techniques of convergence testing ("the Livermore test") are well known and well used in numerical relativity, we wish to emphasize that Richardson's ideas go beyond the basic observation that the error in our grid functions should scale as some power of h. Perhaps more than anything, Richardson's pioneering work taught us that, under ideal circumstances, the errors in our finite-difference solutions are no less computable than the approximate solutions themselves are.

ACKNOWLEDGMENTS

It is a pleasure to thank W. G. Unruh for many useful discussions, a critical reading of the manuscript, and his support and encouragement during the course of this research. I would also like to thank S. A. Teukolsky, and D. F. Chernoff for suggesting improvements to the paper. This work was supported by the Natural Sciences and Engineering Research Council of Canada and NSF PHY 86-03284.

[1] *Sources of Gravitational Radiation,* edited by L. Smarr (Cambridge University Press, Cambridge, England, 1979).
[2] *Gravitational Radiation,* edited by N. Deruelle and T. Piran (North-Holland, Amsterdam, 1983).
[3] R. Arnowitt, S. Deser, and C. W. Misner, in *Gravitation—An Introduction to Current Research,* edited by L. Witten (Wiley, New York, 1962).
[4] J. Centrella, Astrophys. J. **241**, 875 (1980).
[5] J. Centrella, in *The Origin and Evolution of Galaxies,* edited by B. J. T. Jones and J. E. Jones (D. Reidel, Boston, 1983), pp. 41–64.
[6] T. Nakamura, Prog. Theor. Phys. **65**, 1876 (1981).
[7] T. Nakamura, Suppl. Prog. Theor. Phys. **70**, 202 (1981).
[8] T. Nakamura, K. -I. Maeda, S. Miyama, and M. Sasaki, Prog. Theor. Phys. **63**, 1229 (1980).
[9] T. Nakamura, K. -I. Maeda, S. Miyama, and M. Sasaki, in *Proceedings of the Second Marcel Grossmann Meeting on Recent Developments in General Relativity,* Trieste, Italy, 1979, edited by R. Ruffini (North-Holland, Amsterdam, 1981), pp. 675–677.
[10] T. Nakamura and H. Sato, Prog. Theor. Phys. **67**, 1396 (1982).
[11] T. Piran, J. Comput. Phys. **35**, 254 (1980).
[12] K. R. Eppley, Ph.D. thesis, Princeton University, 1975.
[13] K. R. Eppley, in *Sources of Gravitational Radiation* [1], pp. 275–291.
[14] C. R. Evans, in *Dynamical Spacetimes and Numerical Relativity,* Proceedings of the Workshop, Philadelphia, Pennsylvania, 1985, edited by J. Centrella (Cambridge University Press, Cambridge, England, 1986), p. 3.
[15] S. L. Shapiro and S. A. Teukolsky, Astrophys. J. **235**, 199 (1980).
[16] S. L. Shapiro and S. A. Teukolsky, Astrophys. J. **298**, 35 (1986).
[17] T. Piran, in *Gravitational Radiation* [2], pp. 203–256.
[18] J. M. Stewart, in *Classical General Relativity,* edited by W. B. Bonnor, J. N. Isham, and M. A. H. MacCallum (Cambridge University Press, Cambridge, England, 1984).
[19] L. F. Richardson, Philos. Trans. R. Soc. **210**, 307 (1910).
[20] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems,* 2nd Ed. (Wiley, New York, 1957).
[21] H.-O. Kreiss and J. Oliger, *Methods for the Approximate Solution of Time Dependent Problems,* Garp Publications Series No. 10 (Unipub, Lanham, MD, 1973).
[22] G. I. Marchuk, *Methods of Numerical Mathematics* (Springer-Verlag, New York, 1982).
[23] M. Choptuik, Ph.D. thesis, University of British Columbia, 1986.
[24] D. Christoudolou, Commun. Math. Phys. **105**, 337 (1986).
[25] D. S. Goldwirth and T. Piran, Phys. Rev. D **36**, 3575 (1987).
[26] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[27] B. K. Berger, D. M. Chitre, V. E. Moncrief, and Y. Nutku, Phys. Rev. D **5**, 2467 (1972).
[28] W. G. Unruh, Phys. Rev. D **14**, 870 (1976).
[29] T. Piran, in *Proceedings of the Second Marcel Grossmann Meeting on Recent Developments in General Relativity* [9].
[30] J. R. Wilson, in *Sources of Gravitational Radiation* [1].
[31] J. M. Bardeen and T. Piran, Phys. Rep. **96**, 205 (1983).
[32] R. F. Stark and T. Piran, Phys. Rev. Lett. **55**, 891 (1985).
[33] L. Smarr, Ann. N.Y. Acad. Sci. **302**, 569 (1977).