Gravitational wave forms at finite distances and at null infinity

R. Gómez and J. Winicour
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 19 November 1991)

Motivated by the methodology of numerical studies of gravitational radiation, we investigate the discrepancies that arise if wave forms are observed at a finite distance as opposed to infinity. Our study is based upon scalar radiation from a spherically symmetric Einstein-Klein-Gordon system. This allows us to isolate the effects of backscattering and redshifting while avoiding more complicated effects that arise in nonspherical systems with gravitational radiation. We show that discrepancies close to 100% can arise at large observation distances $R >> M$ for sufficiently periodic systems. They are most pronounced for radiation losses between one-quarter and one-half of the initial mass. This falls within the expected regime of the spiral infall of a relativistic binary system. The predominant contribution to this discrepancy stems from a time-dependent redshift arising from radiative mass loss.

PACS number(s): 04.30.+x, 95.30.Sf

I. INTRODUCTION

This paper addresses the following question: How well can you identify gravitational radiation fields without going to infinity? A practical aspect of this question has arisen in the context of numerical relativity, where the standard approach uses a spatial grid terminated at a finite radius R [1–3]. For a system of mass $M \ll R$, it has tacitly been assumed that the wave form at the grid boundary approximates the wave form at infinity, after compensating for the $1/r$ falloff, with an error of the order of magnitude of M/R. An important result of this paper is that this is not true, even in the simplest of systems, for radiation consisting of a long wave train.

Accurate wave forms are the crucial theoretical input for the design of optimal filters for gravitational-wave antennas. The most potent sources of radiation are highly asymmetric systems, such as binary black holes. Sufficient computational power to study such systems is on the verge of becoming available with the new massively parallel machines. For this purpose, it would be expeditious to truncate the grid domain at a small radius, e.g., $R = 10M$. This would only be useful, for the purpose of antenna design, if the distant wave form could be reconstructed from the fields at the grid boundary. Such wave form extraction at $R \approx 12.5M$ has been accomplished by matching an exterior linearized solution to the numerical interior in the axially symmetric case of radiation from small pulsations of a relativistic star [3]. It is difficult to assess to what extent this method can give accurate wave forms and polarization properties in the case of highly nonlinear asymmetric sources. In this paper we use the simple model of a spherically symmetric, self-gravitating, massless scalar field to isolate how strong radiation fields affect the wave form apparent at a finite radius.

For clarification of the issues involved here, we review some standard results. In theories based upon the wave equation in four-dimensional spacetime, pure radiation fields exist only in special circumstances. Plane-wave solutions of the linear wave equation are the most notable case, where the amplitude is described by an arbitrary function of the phase fronts, i.e., for a scalar field by $\Phi(t-z)$ for a plane wave moving in the z direction. Such plane-wave solutions exist for all the rest mass zero fields of arbitrary spin and even in some nonlinear theories such as general relativity.

These plane-wave solutions are physically unrealistic in the sense that they have infinite total energy and that their sources must have infinite spatial extent. However, they serve as reliable models for the far field behavior of radiation from a compact source. In electromagnetic theory, the $O(1/r)$ part of the electric and magnetic fields possess the algebraic properties of a plane wave. For a localized distribution of charges and currents, this behavior is embodied in the peeling property [4] of Maxwell fields, which describes the asymptotic properties of the eigenvectors l^μ of the Maxwell field tensor $F_{\mu\nu}$:

$$F_{\mu\nu}l^\nu = \lambda l_\mu. \tag{1.1}$$

In the nondegenerate case (1.1) admits two independent eigenvectors. They satisfy the null vector condition $l^\mu l_\mu = 0$. The peeling theorem implies that these eigenvectors become degenerate as $r \to \infty$ and approach a unique outgoing null vector k^μ which satisfies the asymptotic eigenvalue equation

$$F_{\mu\nu}k^\nu = O(r^{-2}). \tag{1.2}$$

Here the limit $r \to \infty$ holds $t - r$ constant and defines future null infinity J^+ and the null vector k^ν defines the asymptotic propagation direction. In that limit, $rF_{\mu\nu}k^\nu$ defines the pure radiation field. In gravitational theory, the analogous peeling property describes how the four independent null eigenvectors of the Weyl curvature tensor become asymptotically fourfold degenerate as the gravitational field of an isolated source attains a local plane-wave structure at large distance along the outgoing light cone [4].

These asymptotic properties unambiguously identify the pure radiation field at null infinity, i.e., the ideal limit of observations made by a sequence of antennas at farther and farther distances. In practice, with antennas at finite distances, something else is measured. In scalar terminology, we represent this as $\Phi(t, R, \theta, \phi)$. Assume that there are enough antennas to supply complete knowledge of the field in some (t, θ, ϕ) local patch of the world tube $r = R$. To account for a radial spread in antenna sites, we might consider this patch thickened between closeby radii R_1 and R_2. However, for technical simplicity, we adopt the physically equivalent view that all radial gradients of Φ are known in a patch of radius R. Let the totality of this information about Φ and its radial gradient in this patch be denoted by $\Phi_R(t, \theta, \phi)$.

In this scalar terminology, one aspect of our primary question is the following: Can we pick out the radiation part of the local information contained in $\Phi_R(t, \theta, \phi)$? Everyday experience listening to a car radio would suggest "yes." Yet on mathematical grounds this is impossible. There is no unique local prescription for splitting this information into an inductive part and a purely radiative part. In the radio example, the signal is "distorted" when the car is too close to the transmitter. Here the "distortion" is not the fault of the radio, which we hypothesize to possess perfect fidelity. The distortion is apparent because the detected signal is not the signal that was intended to be sent. The intended signal is the signal which would be received by an ideal antenna at infinity. Radio communication does of course work with negligible distortion under the intended circumstances, i.e., many wavelengths from the source where inductive effects are comparatively small. This defines the physical context for which radio communication is designed. However, this context is not invariant under boost transformations. At any given distance from the transmitter, significant apparent distortion (in addition to Doppler shifting) would be apparent to a sufficiently relativistic car.

This last feature is highlighted by the Coulomb field of a stationary charge. In the limit as the observer velocity approaches the velocity of light, this field appears to be an impulsive plane wave. The same effect arises in the case of a stationary observer and moving source. Furthermore, superimposed versions of this boost distortion cannot be transformed away. For instance, in the case of two sufficiently relativistic charged particles moving in opposite directions, with negligible acceleration and radiation, a strong impulsive signal will be detected by an observer at a finite distance in the center-of-mass frame.

The above considerations tell us that $\Phi_R(t, \theta, \phi)$ adequately describes the intended signal under some set of normal circumstances. But to decide what is normal we need information about the position and motion of the sources relative to the observer. In laboratory or terrestrial electrodynamics, there is more than ample experience and information to decide this issue. In the case of relativistic astrophysical sources, there is often a great deal of ignorance about their structure but the observational distances are so large as to guarantee the correct radiation zone interpretation.

In the context of the numerical treatment of a general-relativistic system, there are both gauge ambiguities and curvature effects that complicate this issue. It is therefore useful to first consider a scalar wave in flat spacetime. In that case, let $\Phi_R(t, \theta, \phi)$ be given globally on the world tube $r = R$ except limited from below by some finite time t_0. Is it then possible to determine the radiation field for retarded times $t - r > t_0 - R$? From a consideration of the domain of dependence determined by the given data, the answer is "no." This domain of dependence is in fact empty, the reason being that any spacetime point in the the region $r > R$ can be reached by a light ray that does not register on the world tube. In the language of the car analogy, there can be competing signals from other distant radio sources exterior to the world tube.

This competition from external sources is never completely absent. Mathematically, it corresponds to incoming radiation fields from the infinite past. Information about such fields can be included in the problem by specifying additional data on the spacelike hypersurface which extends from the world tube to infinity at time t_0. This removes the obvious problem with the domain of dependence. Now every light ray which reaches a point with $r > R$ and $t > t_0$ must pass through one of the data hypersurfaces. In many practical circumstances this competitive signal due to incoming radiation is negligible. In that case, for a charged system, the additional data at $t = t_0$ might be suitably approximated by a Coulomb or static multipole field.

Suppose again that we knew $\Phi_R(t, \theta, \phi)$ on the world tube and the appropriate data at $t = t_0$. The exterior field, including the radiation field is now uniquely determined by these data [5]. However, in general, this exterior field cannot be extended to a solution in a four-dimensional neighborhood of the world tube. This constitutes an example of an improperly posed boundary-value problem for the wave equation [6]. The standard version of the boundary-value problem for the wave equation is to specify the Cauchy data Φ and $\partial_t \Phi$ at some initial time. These Cauchy data then determine a unique solution within its domain of dependence. Furthermore, this solution depends continuously on the data. It is in this sense that the Cauchy problem is well posed. There are well-posed versions of the initial-value problem based upon data on characteristic hypersurfaces, such as a light cone, or on combinations of characteristic and spacelike hypersurfaces. But quite generally, a boundary-value problem for the wave equation is not well posed when two points on the boundary can be connected by a timelike line. The data hypersurface in the above example fails this criterion in two ways. First, there are timelike lines which connect points at $t = t_0$, $r > R$ to points on the world tube. Second, there are timelike lines connecting pairs of points on the world tube itself. Given compatible data on the world tube and on an extension of the spacelike hypersurface $t = t_0$, a unique solution does exist for some range of r in both the interior and the exterior of the world tube but not in a neighborhood of the world tube. In general, the world tube will act as a three-dimensional sheet source.

A boundary-value problem being improperly posed

does not necessarily mean that it is incorrect. But it gives a warning that small variations in the data might lead to large variations in the solution. Furthermore, unless the data are subject to certain constraints there might not exist solutions in a neighborhood of the boundary surface [6]. In the scalar example, these constraints imply that both Φ and $\partial_r\Phi$ cannot be independently specified on the world tube $r = R$. Compare this with the situation for the Cauchy problem. Furthermore, even Φ by itself cannot be specified freely on the world tube but must satisfy a unique continuation condition. This condition implies that knowledge of Φ on any patch of the world tube determines Φ on a larger causally related patch. It thus poses a type of functional constraint on the data. One solution of this constraint problem is that Φ be an analytic function. These bizarre mathematical conditions are unfamiliar to physicists, who for the most part deal with well-posed versions of the initial-value problem. It is not known to what extent they affect the extrapolation of radiation fields from numerically obtained data on a finite world tube.

In general relativity, there are severe complications, in addition to those mentioned above for flat space, which can potentially effect the accuracy of wave forms based upon a finite world tube. Gauge ambiguities make it unclear which components of the metric or curvature tensor to use. Some method of selecting the components transverse to the propagation direction is necessary but there is no unique means of defining this propagation direction locally. One choice might be to use one of the four null eigenvectors of the Weyl tensor. The extent to which these four eigenvectors differ could be used to measure the inherent ambiguity in the local propagation direction. However, such a scheme has not been implemented in any of the present treatments of numerical relativity. The choice of time coordinate can introduce further gauge effects. There are additional physical complications. Time-dependent versions of redshifting occur. The nonlinear gravitational self-source is noncompact and introduces backscattering which blurs the distinction between incoming and outgoing fields. Furthermore, in a case such as a binary black-hole system, there is no practical scheme for eliminating incoming waves from the initial data.

In spite of all these complications, techniques used by numerical relativists with spacelike codes seem to lead to consistent and sensible results in the range of problems where they have been applied (neutron-star oscillations, supernovae). In this context, where the nonlinearities are not too severe, confidence has developed in perhaps mathematically naive but physically practical and fruitful schemes.

On the other hand, it was precisely the challenge of these mathematical difficulties in formulating an unambiguous theory of gravitational radiation that has given rise to a successful new approach. The use of null coordinates [7] and the compactification of the points at null infinity [8] has led to a mathematically rigorous description of gravitational radiation in terms of geometrically defined quantities such as the Bondi mass and news function. These geometrical techniques have been incorporated into a numerical evolution scheme based upon the characteristic initial-value problem on a grid with compactified future null infinity J^+. This allows the calculation of radiation fields according to their geometric definition at J^+ [9].

In this paper, we use this approach to compare wave forms at J^+ with the approximate wave forms obtained by a simple prescription at the boundary of a finite spatial grid. We deal here with the radiation of a self-gravitating, spherically symmetric, massless scalar field, as described by the coupled Einstein-Klein-Gordon equations. The assumption of spherical symmetry eliminates ambiguities in the local radial direction and the context of a scalar field eliminates gauge ambiguities. Furthermore, initial incoming waves are eliminated in the test region by choosing initial data with compact support. In the nonspherical case, this remedy is not possible for the initial data describing gravitational degrees of freedom because of the constraint equations. Elimination of these effects allows isolation of the nonlinear effects of backscattering and time-dependent redshifting. Furthermore, spherical symmetry allows the luxury of a very fine grid so that these effects can be studied here without contamination by numerical noise.

II. DESCRIPTION OF THE SYSTEM

We give a brief description of the characteristic initial-value problem for a spherically symmetric, zero rest mass scalar field Φ. For more thorough treatments, see Refs. [10] and [11]. Einstein's equation

$$G_{\mu\nu} = 8\pi(\nabla_\mu\Phi\nabla_\nu\Phi - \tfrac{1}{2}g_{\mu\nu}\nabla_\alpha\Phi\nabla^\alpha\Phi) \tag{2.1}$$

is decomposed with respect to a family of outgoing null cones emanating from the central geodesic. Let u be the proper time along this geodesic, with u =const on the outgoing null cones. Let r be a luminosity distance on these null cones, so that $4\pi r^2$ is the surface area of the spheres of symmetry. Then, in this Bondi-type coordinate system, the line element becomes

$$ds^2 = e^{2\beta}du\left(\frac{V}{r}du + 2\,dr\right) - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \tag{2.2}$$

where θ and ϕ are the usual polar coordinates.

In the neighborhood of the origin, we adopt, as boundary conditions on the coordinates

$$V(u,r) = r + O(r^3) \text{ and } \beta(u,r) = O(r^2), \tag{2.3}$$

so that the metric reduces to a Minkowski (null polar) form along the central world line. The resulting metric does not take an asymptotic Minkowski form in the limit $r \to \infty$ of null infinity. We set $H(u) = \beta(u,\infty)$. Then Bondi time u_B at J^+ is related to proper time u along the central geodesic by

$$\frac{du_B}{du} = e^{2H}. \tag{2.4}$$

The coordinates (u_B, r, θ, ϕ) constitute a standard Bondi

frame in the neighborhood of J^+ in which the metric does take an asymptotic Minkowski form. Bondi time u_B is appropriate for discussing asymptotic quantities such as the mass, the news function, and radiation wave forms. However, central time u is more convenient in dealing with horizons. A horizon forms at a finite central time $u = u_{\mathcal{H}}$ but at an infinite Bondi time, with the central redshift determined by Eq. (2.4).

In these coordinates, the field equations reduce to the two radial equations

$$\beta_{,r} = 2\pi r(\Phi_{,r})^2, \tag{2.5}$$

$$V_{,r} = e^{2\beta}, \tag{2.6}$$

and the scalar wave equation $\Box\Phi = 0$, which takes the form

$$2(r\Phi)_{,ur} = r^{-1}(rV\Phi_{,r})_{,r}. \tag{2.7}$$

The boundary conditions (2.3) imply that $\beta \geq 0$ and $V \leq e^{2\beta}r$, with β a monotonically increasing function of r.

The initial null data for evolution consist of $\Phi(u_0, r)$ at an initial retarded time $u = u_0$. We use the gauge freedom $\Phi \rightarrow \Phi +$ const to set $\Phi(u_0, \infty) = 0$. Furthermore, we only consider nonsingular, initial data with the asymptotic behavior

$$\Phi = \frac{Q}{r} + O\Big(\frac{1}{r^2}\Big). \tag{2.8}$$

Here $Q(u)$ is the scalar monopole moment. The hypersurface equations (2.5) and (2.6) uniquely determine $\beta(u_0, r)$ and $V(u_0, r)$ from these data. Formal evolution then proceeds by determining $\partial_u\Phi(u_0, r)$ from the radial integral of the wave equation (2.7) which gives, after integration by parts,

$$2r\Phi_{,u} = V\Phi_{,r} + \int_0^r \Big(\frac{V}{r}\Big)\Phi_{,r}dr. \tag{2.9}$$

Our algorithm for numerical evolution is based upon an identity obtained by integrating this equation along the radially incoming null geodesics [11]. It is carried out on a compactified grid in terms of the radial coordinate $x = r/(1 + r)$, where $x = 1$ at J^+.

The Bondi mass is given by the asymptotic expression

$$M = \frac{1}{2}e^{-2H}r^2\Big(\frac{V}{r}\Big)_{,r}\Big|_{r=\infty} \tag{2.10}$$

and the scalar news function by

$$N = e^{-2H}Q_{,u} = Q_{,u_B}. \tag{2.11}$$

Here the factors of e^{-2H} arise from the relation of central time to Bondi time. Alternatively, by using the field equations, they may be expressed as the radial integrals

$$M = 2\pi\int_0^\infty e^{2(\beta-H)}r^2(\Phi_{,r})^2dr \tag{2.12}$$

and

$$N = \frac{1}{2}e^{-2H}\int_0^\infty \frac{V}{r}\Phi_{,r}dr. \tag{2.13}$$

The Bondi mass loss formula is

$$M_{,u_B} = -4\pi N^2. \tag{2.14}$$

III. WAVE FORMS

In terms of $g = r\Phi$, the radiation amplitude at J^+ is $Q(u_B) = g(u_B, \infty)$, where the Bondi time u_B plays the role of the proper time used by an observer at infinity. As the counterpart of $Q(u_B)$ based upon the world tube $r = R$, we choose $q(\tau; R) = g\big(u_B(\tau, R), R\big)$, where $\tau = \tau(u_B, R)$ is the proper time on the world tube. These times are related by

$$\frac{\partial\tau}{\partial u_B} = e^{\beta-2H}\sqrt{\frac{V}{R}}, \tag{3.1}$$

as follows from (2.2) and (2.4). This gives the relative redshift between observers on the world tube and at J^+. Since $V \sim e^{2H}r$, and $\beta \sim H$, there is no redshift between τ and u_B in the limit $R \rightarrow \infty$. An asymptotic expansion gives

$$\frac{\partial\tau}{\partial u_B} = \sqrt{1 - \frac{2M(u_B)}{R}} + O\Big(\frac{1}{R^3}\Big). \tag{3.2}$$

Asymptotically, the redshift of the proper time on the world tube is a time-dependent version of the familiar Schwarzschild result. In order to synchronize the two time coordinates we set $\tau_0 = u_{B0} = 0$ on the initial null cone.

Our purpose is to test how accurately $q(\tau; R)$ serves as a substitute for $Q(u_B)$. This test is in the spirit of how wave forms are calculated in the spacelike codes. It has been constructed so as to be automatically satisfied in the linear weak-field case throughout any region where the wave is purely outgoing. In that case $g(u_B, r)$ is independent of r and $\tau = u_B$ so that $q(\tau; R) = Q(u_B)$.

In order to provide some orientation, first consider the weak-field initial data

$$g(u_0, r) = \begin{cases} 0.1\,r^2(1-r)^2 & \text{for } r \leq 1, \\ 0 & \text{for } r \geq 1, \end{cases} \tag{3.3}$$

which describe a pulse-shaped wave profile of compact support. Outside $r = 1$ the wave is purely outgoing and the $q(\tau; R)$ should be identical for all R. This is apparent in Fig. 1, which overlays graphs of $q(\tau; R)$, for several radii, and of the wave form $Q(u_B) = q(\tau; \infty)$ in terms of a common time parameter τ. The graphs for $R > 1$ are indistinguishable whereas those for the internal region display interference between the incoming and outgoing parts of the wave. This emphasizes how misleading the wave form at a finite radius can be when incoming waves are present. In all of the ensuing examples, we deal only with regions where there is no initial incoming wave. It should also be noted that none of the profiles have a tail; i.e., they are of finite duration. This results from the absence of backscattering associated with the validity of Huygen's principle in the weak-field regime, combined with the compact support of the initial pulse.

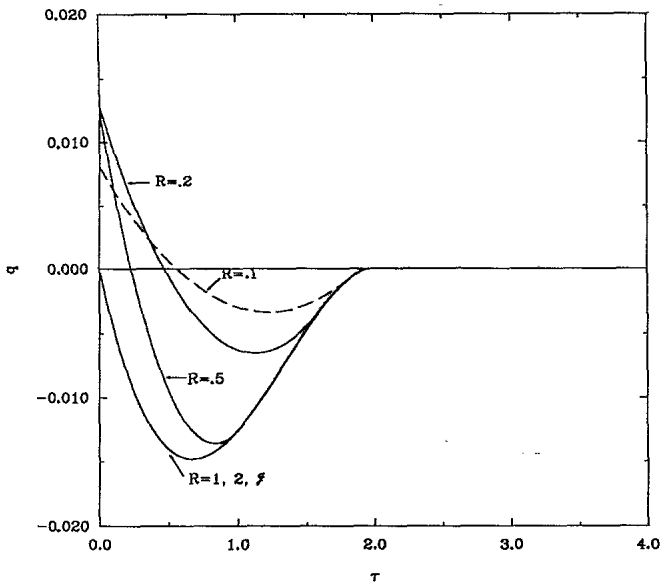

FIG. 1. Comparison of profiles at representative values of R and at J^+ in the linear regime. The profiles in the exterior region are indistinguishable.

In the strong-field examples to be discussed next, there is no clean way to decompose the wave into incoming and outgoing parts. But to make the test clear-cut, we will only consider initial data with support $r < 1$ and test radii $R \geq 1$ at which $q(\tau; R)$ measures only the outgoing radiation and the backscattering. In order to quantify the discrepancy between the wave form q and the radiative wave form Q we use the l_2 norm

$$||f||^2 = \int_0^\infty f(\tau)^2 d\tau \tag{3.4}$$

and the figure of merit

$$\mathcal{E} = \frac{||q - Q||}{||Q||}. \tag{3.5}$$

As above, τ represents the proper time measured along the world tube of radius R (or u_B in the case of Q). Since the l_2 norm is independent of basis, the same figure of merit applies to the error in the Fourier transform of the signal. Gravitational-wave antennas actually detect tidal displacements, which arise from the time derivatives of the signal. Similarly, we can compare such signals in terms of

$$\mathcal{E}' = \frac{||q_{,\tau} - Q_{,\tau}||}{||Q_{,\tau}||}, \tag{3.6}$$

in which case $||Q_{,\tau}||^2 = 4\pi\Delta M$, in terms of the total mass radiated.

The strategy here is to choose initial data characterized by two parameters, representing amplitude and wavelength, and to investigate the resulting wave forms over a comprehensive range. We have already seen that $\mathcal{E}$ is small for small amplitudes, in accord with the weak-field limit. It is also small at high amplitudes for which a black hole forms very rapidly. This stems from a rigorous version of the no-hair theorem [12], which establishes that the scalar field must vanish in the limit $\tau \to \infty$, $r > 2M_f$. As a result, in the region of greatest observational interest $r \geq 10M$, the field is zero initially by construction and, in the high amplitude regime, it never builds up any appreciable amplitude before the interior region collapses to a black hole. Systems in this regime would not be readily detectable because of their extreme redshift. We will focus our attention on the intermediate amplitude region of greatest physical relevance. We choose initial data of the form

$$g(u_0, r) = \begin{cases} \Lambda \sin[2\pi N(r-1)] & \text{for } r \leq 1, \\ 0 & \text{for } r \geq 1, \end{cases} \tag{3.7}$$

where Λ controls the amplitude and N controls the wavelength.

We look first at the long-wavelength case, taking $N = \frac{1}{2}$. The table below summarizes the pertinent results for observers at $R = 1$ and at $R = 10M_i$ in the amplitude range $\Lambda = 0.06$ (noncollapsing) to $\Lambda = 0.12$:

Λ	M_i	M_f	$\mathcal{E}(R=1)$	$\mathcal{E}(R=10M_i)$
0.065	0.112	0	0.158	0.142
0.08	0.158	0.020	0.251	0.163
0.10	0.219	0.115	0.319	0.144
0.12	0.275	0.22	0.355	0.120

For this range of parameters, the maximum discrepancy in the wave form at $R = 1$ occurs for $\Lambda = 0.12$ but at $R = 10M_i$ it occurs for $\Lambda = 0.1$. At larger observational distances, the falloff of $\mathcal{E}$ is quite accurately described by a $1/R$ dependence, as can be already seen in comparing the values in the table for $R = 1$ and $R = 10M_i$.

For $\Lambda = 0.08$, Fig. 2 displays the profiles seen by an observer at $R = 1$ and another at J^+, overlayed on the same proper time scale τ, for ($\Lambda = 0.08$). The profiles at $R = 1$ and J^+ look very similar; however, we can note some differences, which can be attributed to a redshift of the signal at J^+ with respect to that seen at $R = 1$. As seen from the graph, this redshift causes the signal at J^+ to become out of phase with respect to the signal seen by the nearby observer. Note that considerable backscatter occurs in the region interior to $R < 1$, where the incoming pulse is transformed into an outgoing signal with double pulse shape. However, there is only small backscattering in the region exterior to $R > 1$. In the worst case, the peak amplitudes at $R = 1$ and at J^+ differ only by 10%.

The discrepancy in the time derivative of the profiles seen by the two observers, as measured by $\mathcal{E}'$, is higher than the error in the profiles themselves. For $\Lambda = 0.08$, $\mathcal{E}' = 0.39$ at $R = 10M_i$ as opposed to $\mathcal{E} = 0.163$. Again, $\mathcal{E}'$ is quite accurately described by a $1/R$ falloff at larger distances.

Figure 2 suggests that larger discrepancies might result if the system were to emit a long wave train. When the radiation is emitted in a single pulse, any phase difference is initially zero, by construction, and goes to zero as the pulse decays. In the case of a wave train, the phase difference could accumulate over many cycles. Such a wave train would arise, for example, in the case of the spiral infall of a binary system.

We explore this possibility by using higher values of N

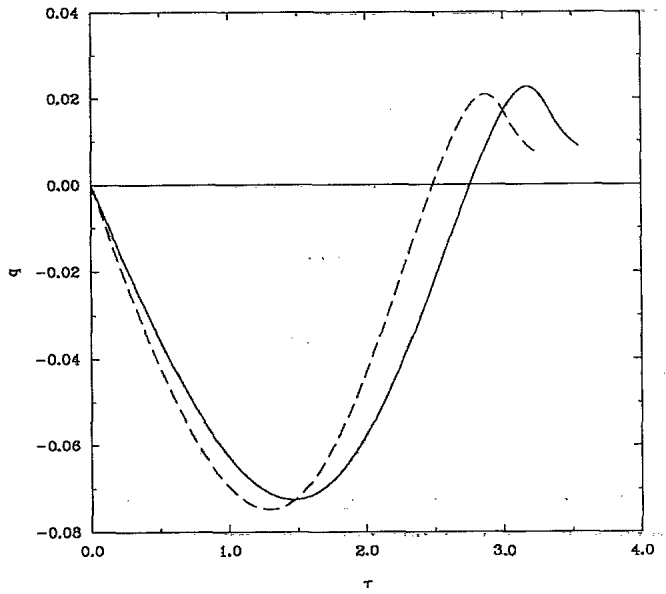

FIG. 2. Comparison of profiles at $R = 1$ (dashed line) and J^+ (solid line) for compact initial data.

in the initial data given in Eq. (3.7). As might be expected, the amplitude parameter must be set lower than in the long-wavelength case if appreciable scalar radiation is to escape before formation of a black hole. This in turn has the effect of lowering the initial mass of the system, so that an observer at a fixed distance has a larger value of R/M_i.

First consider the choice $N = 10$ and $\Lambda = 0.002$, for which $M_i = 0.045$ and $M_f = 0.026$. There is a large discrepancy, $\mathcal{E} = 0.66$, between the profiles observed at $R = 25M_i$ and at J^+. At this radius, $\mathcal{E}' = 0.69$ is now very close to $\mathcal{E}$. These profiles are compared in Fig. 3. As before, the amplitudes are in close agreement, which indicates very little backscatter in the intervening region.

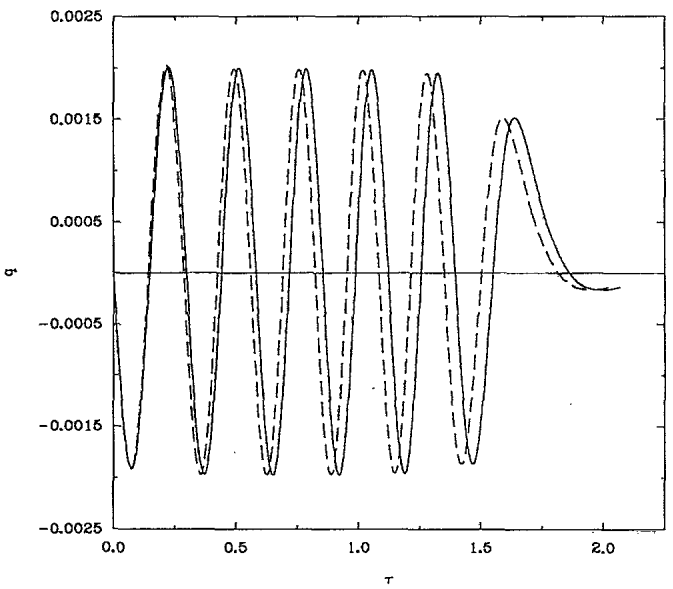

FIG. 3. Comparison of profiles at $R = 25M$ (dashed line) and J^+ (solid line) for compact initial data.

However, there is considerable phase shifting which is the prime source of the discrepancy between the wave forms.

Next, going to shorter wavelengths, we choose $N = 20$ and $\Lambda = 9.5 \times 10^{-4}$, for which $M_i = 0.041$ and $M_f = 0.029$, the discrepancy in the wave form at $R = 25M_i$ is now $\mathcal{E} = 0.91$. At further distances, we again find that $\mathcal{E}$ falls off as $1/R$ so that, for this case, $\mathcal{E} \approx 0.23$ at $R = 100M_i$.

Continuing in this manner, we find large discrepancies at $R = 100M_i$ by increasing N to 100. In that case, choosing $\Lambda = 1.3 \times 10^{-4}$, we find $M_i = 0.020$, $M_f = 0.016$, and $\mathcal{E} = 0.80$.

High-accuracy computer simulations for $N >> 100$ would require too much computer time to be practical. However the trend we have already seen for large N can be understood and extrapolated in terms of a rough model in which the wave forms observed at J^+ and at R are approximated by monochromatic waves with equal amplitudes but slightly different frequencies ω and $\omega+\delta\omega$, respectively, up to some time T when most of the radiation has been emitted, prior to the formation of a horizon. Referring to Eq. (3.2),

$$\delta\omega/\omega \approx M(\tau)/R, \tag{3.8}$$

in the region $R >> M$. Setting $Q = A \sin\omega\tau$, this implies

$$q - Q \approx (AM\omega\tau/R)\cos\omega\tau, \tag{3.9}$$

for time scales $\omega\tau < R/M$. By averaging the high-frequency terms in the integrals involved in (3.5), this gives a wave form discrepancy $\mathcal{E} \approx M\omega T/\sqrt{3}R$, over the time interval from 0 to T. Here M represents some average mass during the collapse, which we take to be $M = (M_i + M_f)/2$. For a quasiperiodic system which emits n cycles before forming a black hole, $\omega T \approx 2\pi n$. Consequently, $\mathcal{E} \approx 2\pi nM/\sqrt{3}R$. For the previously discussed cases with $N = 10$ and $N = 20$ (for which $n = 6$ and $n = 8$, respectively), this gives better than 10% agreement with the value of $\mathcal{E}$ from the numerical evolution. For the previous case with $N = 100$ (for which $n = 24$), there is 2.5% agreement between the formula and the numerical value obtained for $\mathcal{E}$. This provides strong evidence that this formula gives a reliable estimate of $\mathcal{E}$ in the large N regime.

Thus large wave-form discrepancies, $\mathcal{E} \approx 1$, can arise at any radius for a system of sufficiently high frequency, i.e., $n \approx R/M$. For quasiperiodic signals, the discrepancy in the tidal forces, as measured in terms of the time derivative of the field by $\mathcal{E}'$, is approximately equal to $\mathcal{E}$.

IV. DISCUSSION

We have shown that a large discrepancy can arise between the apparent wave form at a finite distance and the true wave form at null infinity in the process of black-hole formation. For a long wave train, a good approximation to this discrepancy is $\mathcal{E} \approx 3.6nM/R$, where n is the number of cycles in the radiated wave. This can be of the order of 100% at arbitrarily large observation distances $R \gg M$ for sufficiently high frequencies. These

observation distances satisfy all the criterion for the distant wave zone. The discrepancy is most pronounced for radiation losses between one-quarter and one-half of the initial mass. For the spiral infall of a binary black-hole system, with $n \approx 10$, the required observation distance required for 1% accuracy would be $R \approx 1000M$. The predominant contribution to this discrepancy stems from a time-dependent version of the redshift effect. The contribution from backscattering is less significant in the long-wave-train case.

For short pulses of radiation there tends to be a larger discrepancy $\mathcal{E}'$ in the time derivative of the wave form, which is related to observable tidal forces. (In the case of gravitational radiation, tidal forces actually stem from the second time derivative of the metric wave form.) Our tests were set up so that there was no initial incoming radiation in the observation region. Otherwise, interference from the incoming signal leads to further distortion of the outgoing wave form at a finite distance. However, extensive numerical experimentation shows that initial data with a $1/r$ falloff do not alter the results significantly. Such a falloff describes data with purely outgoing radiation, up to curvature effects.

Although our study was based upon spherically symmetric scalar waves, the same underlying physical effects apply to gravitational wave forms. However, in the case of gravitational radiation, there are more nonlinear sources of backscattering and there are ambiguities associated with gauge freedom and with the choice of a local propagation direction. Furthermore, except for initial data which are approximately Newtonian there is no feasible way to eliminate incoming radiation fields. These factors indicate that null infinity should be taken seriously in numerical studies of radiation.

ACKNOWLEDGMENTS

We would like to thank A. Abrahams and C. Evans for helpful discussions. This work was supported by NSF Grant No. PHY-8803073. Computer time was provided by the Pittsburgh Supercomputing Center under Grant No. PHY860023P.

[1] J. M. Bardeen and T. Piran, Phys. Rep. **96**, 205 (1983).
[2] J. L. Anderson and D. W. Hobill, in *Dynamical Spacetimes and Numerical Relativity*, edited by J. M. Centrella (Cambridge University Press, Cambridge, England, 1986).
[3] A. M. Abrahams and C. R. Evans, Phys. Rev. D **42**, 2585 (1990).
[4] Cf. R. Penrose and W. Rindler, *Spinors and Space-Time* (Cambridge University Press, Cambridge, England, 1986), Vol. 2.
[5] F. John, *Partial Differential Equations*, 4th ed. (Springer-Verlag, New York, 1982).
[6] F. John, Commun. Pure Appl. Math. **2**, 209 (1949).
[7] H. Bondi, M. G. J. van der Burgh, and A. W. K. Metzner, Proc. R. Soc. London **270**, 103 (1962).
[8] R. Penrose, Phys. Rev. Lett. **10**, 66 (1963).
[9] R. A. Isaacson, J. S. Welling, and J. Winicour, J. Math. Phys. **24**, 1824 (1983).
[10] D. Christodoulou, Commun. Math. Phys. **105**, 337 (1986).
[11] R. Gómez and J. Winicour, J. Math. Phys. (to be published).
[12] D. Christodoulou, Commun. Math. Phys. **109**, 613 (1987).