Singularities and Horizons in the Collisions of Gravitational Waves

Thesis by

Ulvi Yurtsever

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1989

(Submitted September 29, 1988)

Acknowledgments

I would like to thank my advisor, Kip Thorne, for his continuous encouragement, his relentless patience, and his understanding throughout the course of this project. During my stay at Caltech, I have greatly enjoyed and benefited from various discussions with Ted Allen, Luc Blanchet, Sandip Chakrabarti, Sam Finn, Mike Morris, Don Page, and Wai-Mo Suen; to all of whom I would like to give my thanks. Last and most, I would like to thank Ayse Franko, for her pure love and inexhaustible friendship.

Abstract

This thesis presents a study of the dynamical, nonlinear interaction of colliding gravitational waves, as described by classical general relativity. It is focused mainly on two fundamental questions: First, what is the general structure of the singularities and Killing-Cauchy horizons produced in the collisions of *exactly plane-symmetric* gravitational waves? Second, under what conditions will the collisions of almost-plane gravitational waves (waves with large but finite transverse sizes) produce singularities?

In the work on the collisions of exactly-plane waves, it is shown that Killing horizons in any plane-symmetric spacetime are unstable against small plane-symmetric perturbations. It is thus concluded that the Killing-Cauchy horizons produced by the collisions of some exactly plane gravitational waves are nongeneric, and that generic initial data for the colliding plane waves always produce "pure" spacetime singularities without such horizons. This conclusion is later proved rigorously (using the full nonlinear theory rather than perturbation theory), in connection with an analysis of the asymptotic singularity structure of a general colliding plane-wave spacetime. This analysis also proves that asymptotically the singularities created by colliding plane waves are of inhomogeneous-Kasner type; the asymptotic Kasner axes and exponents of these singularities in general depend on the spatial coordinate that runs tangentially to the singularity in the non-plane-symmetric direction.

In the work on collisions of almost-plane gravitational waves, first some general properties of single almost-plane gravitational-wave spacetimes are explored. It is shown that, by contrast with an exact plane wave, an almost-plane gravitational wave cannot have a propagation direction that is Killing; i.e., it must diffract and disperse as it propagates. It is also shown that an almost-plane wave cannot be precisely sandwiched between two null wavefronts; i.e., it must leave behind tails in the spacetime region through which it passes. Next, the occurrence of spacetime singularities in the collisions of almost-plane waves is

investigated. It is proved that if two colliding, almost-plane gravitational waves are initially exactly plane-symmetric across a central region of sufficiently large but *finite* transverse dimensions, then their collision produces a spacetime singularity with the same local structure as in the exact-plane-wave collision. Finally, it is shown that a singularity still forms when the central regions are only approximately plane-symmetric initially. Stated more precisely, it is proved that if the colliding almost-plane waves are initially sufficiently close to being exactly plane-symmetric across a bounded central region of sufficiently large transverse dimensions, then their collision necessarily produces spacetime singularities. In this case, nothing is now known about the local and global structures of the singularities.

Table of Contents

CHAPTER 1

Introduction

According to general relativity, the evolution of gravitational and matter fields is governed by the Einstein field equations. The Einstein field equations are notoriously *nonlinear*: almost all novel features of relativistic gravitation (e.g., the formation of black holes by the gravitational collapse of stars or star clusters, and the generation of gravitational waves by compact astrophysical sources) are either direct consequences of this essential nonlinearity, or they owe to it the richness and attractiveness of their properties. Today this nonlinearity is studied by a variety of techniques, including: (i) exact solutions, where novel mathematical techniques are employed to generate new explicit solutions of Einstein's equations (usually in the presence of symmetry); (ii) global methods, where the geometry and the causal structure of a general spacetime are explored with the techniques of differential geometry and topology; (iii) perturbation theory, where Einstein's equations are analyzed via an order-by-order expansion around an explicit background solution; and (iv) numerical relativity, where Einstein's equations are integrated numerically to simulate the action of relativistic gravity in complex astrophysical processes. From mathematical relativity to relativistic astrophysics, the nonlinear nature of gravity permeates all areas of research involving the physics of gravitation.

No manifestation of the nonlinear nature of gravitation is more striking than the nonlinear coupling of gravity to itself, a fact often dramatized by the aphorism: "gravity gravitates." As an example, just like any matter field, gravitational radiation propagating through empty space generates background curvature, which in turn couples back to the gravitational waves propagating on the background. An extreme (theoretical) example of this coupling is the phenomenon of "geons," localized lumps of gravitational radiation held together (for a finite time long compared to the periods of the constituent waves) by the background curvature that the waves themselves generate (somewhat in analogy with the gluon states of quantum chromodynamics). A slightly more realistic example is provided by the interactions between gravitational waves freely propagating on an empty background spacetime; it is these interactions that constitute the subject matter of this thesis.

In recent years, a phenomenal growth of activity guided by the modern techniques of contemporary nonlinear mathematics has occurred in the search for new exact solutions to Einstein's equations. An offshoot of this growth in exact-solutions research has been the discovery of several explicit solutions to the vacuum Einstein equations describing the fully nonlinear interactions of gravitational plane waves propagating and colliding in an otherwise flat spacetime. These exact solutions are the starting point for the investigations contained in this thesis.

It should be noted at this point that in the present epoch of the Universe it is highly unlikely that collisions of gravitational waves are ever strong enough to produce significant nonlinear effects. Simple order-of-magnitude calculations show that at the time they collide with other waves, gravity waves generated by compact astrophysical sources are very likely to be either too weak in amplitude or too small in transverse size to yield any of the interesting nonlinear effects associated with the collisions of exactly-plane (infinite-size) or almost-plane (finite but very large-size) gravitational waves. It is possible in principle that in the very early Universe, near a higly inhomogeneous and anisotropic initial singularity, the nonlinear interactions of colliding gravitational waves might have played an important role, and therefore that the results of this thesis might yield some insight into the study of such inhomogeneous and anisotropic cosmologies. We do not, however, understand enough about the early Universe to know for sure whether this is the case.

Regardless of this potential application, the primary purpose of this thesis is to gain insight into the nonlinear interactions of gravity with itself through studying the problem of colliding gravitational waves as an issue of principle. Throughout the thesis, the point of view is adopted that the interactions between colliding gravitational waves constitute a convenient model problem for a deeper study of the nonlinear, dynamical nature of gravitation. From this point of view, the problem of colliding plane waves has the particular advantage that the presence of plane symmetry (unlike the presence of, for example, spherical symmetry) allows dynamical vacuum degrees of freedom into the problem, while simplifying the equations sufficiently to make an exact analytic treatment possible. Morover, the more general problem of colliding gravitational waves that are not precisely planar (by contrast with the analogous problem of gravitational collapse) has the advantage that the analysis and the results do not depend on the choice of a particular stress-energy tensor for matter fields.

In the rest of this introduction, I will try to provide some necessary background to the nonspecialist, explain in more detail some of the terms and assertions made in these opening paragraphs and in the Abstract, and present a summary of the remaining chapters of this thesis.

GRAVITATIONAL PLANE WAVES AND THEIR COLLISIONS

Consider a gravitational wave in flat Minkowski spacetime so weak that its propagation is accurately described by linearized theory. Such a wave gives rise to a spacetime metric

$$g_{\mu\nu}=\eta_{\mu\nu}+h_{\mu\nu} \qquad |h_{\mu\nu}| \ll 1 , \tag{1}$$

where $\eta_{\mu\nu}$ denotes the flat Minkowski line element. The linearized Einstein equations for the "wave field" $h_{\mu\nu}$ are just[1]

$$\Box \bar{h}_{\mu\nu} \equiv \bar{h}_{\mu\nu}{}^{,\alpha}{}_{,\alpha} = 0 , \qquad \bar{h}^{\mu\alpha}{}_{,\alpha} = 0 ,$$

$$\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2} h \, \eta_{\mu\nu} , \qquad h \equiv h_\alpha{}^\alpha . \tag{2}$$

By imposing the "transverse traceless" gauge conditions $h_{\mu 0} = h = 0$ on $h_{\mu\nu}$, as in Sec. 35.4 of Ref. 1, it is easy to construct the standard plane-wave solutions of Eqs. (2)

$$h_{xx} = h_+(t-z) , \qquad h_{yy} = -h_+(t-z) , \qquad \text{all other } h_{\mu\nu} \equiv 0 . \tag{3}$$

For this solution the metric (1) takes the form

$$g = [1 + h_+(t-z)] \, dx^2 + [1 - h_+(t-z)] \, dy^2 + dz^2 - dt^2$$

$$= [1 + h_+(u)] \, dx^2 + [1 - h_+(u)] \, dy^2 - du dv , \tag{4}$$

where we have introduced the null coordinates $u \equiv t - z$, $v \equiv t + z$. The arbitrary function $h_+(u)$ represents the linearized amplitude of the "plus"-polarization mode in the plane wave (4). The remaining "cross"-polarization mode, if present, would be represented by an additional term in Eq. (4) of the form $h_\times(u) \, dxdy$.

Although Eq. (4) represents a perfectly exact solution to the *linearized* equations (2), it nevertheless is not an exact solution of the full nonlinear Einstein equations: in general, the Ricci tensor of the metric (4) contains nonzero terms of order ${h_+}^2$. However, it is a remarkable property of the Einstein field equations that by slightly modifying the metric (4) one can build *exact* solutions that describe fully nonlinear plane gravitational waves. The simplest such solution (due to Bondi[2]) can be written in the form:[3,1]

$$g = L^2(u)\,(e^{2h(u)}dx^2 + e^{-2h(u)}dy^2) - du\,dv \;, \tag{5}$$

where the functions $L(u)$ and $h(u)$ are related by the only vacuum Einstein equation

$$L'' + (h'^2)L = 0 \;, \qquad ' \equiv \frac{d}{du} \;. \tag{6}$$

The linearized metric (4) can be recovered from (5) in the limit $L(u) \approx 1$, $h(u)=$ arbitrary but small. Like the linearized solution (4), the exact plane-wave metric (5) incorporates only the "plus"-mode of polarization; however the solutions that involve both modes of polarization are equally easy to write down (see Sec. II A of Chapter 7).

It is not very difficult to see from Eq. (6) that for a sandwich plane wave whose curvature is "sandwiched" between two null wave fronts $\{u=0\}$ (the initial wave front) and $\{u=a\}$ (the final wave front), the background factor $L(u)$ decreases from its initial value $L=1$ at $u<0$ to $L=0$ at a u value $u=f$ which, in order of magnitude, is given by $f \sim \lambda\!\!\!^{-2}/h^2 a$. Here h and $\lambda\!\!\!^{-}$ are the typical values of the amplitude [i.e., of $h(u)$] and of the wavelength [i.e., of h/h'], respectively. Consequently, the surface $\{u=f\}$ is a coordinate singularity for the metric (5) and the coordinates (u,v,x,y). In fact, this coordinate singularity is a manifestation of the *focusing effect* of the plane wave (5): Spacetime is actually flat after the wave passes (and, in particular, near $u=f$ it is flat). However the background curvature generated by the plane wave in its sandwich region $0<u<a$ focuses all null geodesics (x,y,v)=constant, causing them to converge on each other at the "focal plane" $u=f$. Since the coordinate lines of Eq. (5) are attached to these null geodesics, the distance between them is driven to zero as $u \rightarrow f$; and, thus, L goes to zero. This focusing effect, crucial for understanding the nonlinear interactions between colliding gravitational waves, is reviewed in great

detail in Sec. II of Chapter 3.

Consider now two gravitational plane sandwich waves of the form (5) that propagate and collide in flat spacetime. (Here for simplicity only parallel-linear-polarized colliding waves are considered; under this restriction there is no cross-polarization piece in the metric. The general case is treated in Chapter 7.) By applying a Lorentz transformation if necessary, one can find a coordinate system (u, v, x, y) in which the waves collide head-on, i.e., in which the metric can be written in the form[10]

$$g = L^2(u,v)\,(e^{2h(u,v)}dx^2 + e^{-2h(u,v)}dy^2) - e^{-M(u,v)}\,du dv \ , \tag{7}$$

Here, on and before the initial wavefront $\{v=0\}$ of wave 2, i.e., for $v \leq 0$, M vanishes, and L, h are equal to their values $L(u)$, $h(u)$ for wave 1; and on and before the initial wavefront $\{u=0\}$ of wave 1, i.e., for $u \leq 0$, M vanishes and L, h are equal to their values $L(v)$, $h(v)$ for wave 2 (see Fig. 1 of Chapter 4). The values of $L(u,v)$, $h(u,v)$, and $M(u,v)$ in the interaction region, where $u \geq 0$, $v \geq 0$, are to be found by solving the vacuum Einstein field equations for the metric (7) with the above initial conditions. These field equations are more complicated than Eq. (6), and the analysis and solution of the above initial-value problem are difficult. Nevertheless, using several ingenious techniques many researchers have succeeded in finding exact solutions of the form (7) that describe colliding plane waves; the first and prototypical such solution being the one discovered by Khan and Penrose.[4]

The Khan-Penrose solution describes the collision between two impulsive, plane-symmetric gravitational waves propagating in a flat background spacetime. The gravitational field generated by the collision is not only qualitatively different from the linear superposition of the two incoming fields, but in fact the spacetime curvature in the interaction region increases without bound along all timelike worldlines, and it ultimately diverges to form a curvature singularity where all observers' worldlines reach and terminate in finite proper time. The local and global structure of this solution is complicated;[5] but its physical interpretation is simple: Each of the two colliding plane waves generates a spacetime geometry in its wake which acts like an infinite, perfectly converging lens,[6] focusing any radiation field which passes through the plane wave while propagating in the opposite direction (the focusing effect discussed above). When the two plane waves collide, each of them is thus perfectly focused by the other's background geometry; diffraction effects are prevented from counterbalancing this perfect focusing by the global exact-plane-symmetry of spacetime (i.e. the infinte transverse size of both waves). As a result, while they propagate through the interaction region the amplitudes of the colliding waves grow without bound and ultimately diverge, creating a spacelike curvature singularity which bounds the interaction region in all future directions.

SINGULARITIES OR HORIZONS?
THE GENERIC OUTCOME OF COLLISIONS BETWEEN PLANE WAVES

Thanks to the work of Chandrasekhar and Xanthopoulos[7] who first discovered this phenomenon, it is now known that colliding plane waves do not always create spacelike curvature singularities with a global structure similar to that of the Khan-Penrose solution: for some choices of the incoming plane waves' waveforms $h(u)$ and $h(v)$, their collision produces a nonsingular Killing-Cauchy horizon at the points where ordinarily one would expect curvature singularities to form. Killing-Cauchy horizons are nonsingular surfaces along which the coordinates (u, v, x, y) develop *coordinate* singularities, but the curvature of spacetime remains finite and observers can pass through. The precise definition and structure of Killing-Cauchy horizons in the context of colliding plane-wave spacetimes are discussed extensively in Chapter 2. If a Killing-Cauchy horizon forms in the interaction region of a colliding plane-wave spacetime, the spacetime can be extended smoothly across it (in nonunique ways) to obtain several inequivalent, maximal solutions, which all evolve from the same initial data posed by the incoming, colliding plane waves. (There is a breakdown of predictability.) It is therefore of fundamental importance to determine (i) under what conditions on the initial data (the incoming plane waves) the collision creates singularities and under what conditions it creates horizons; (ii) what are the local structures of the singularities and horizons thus created; and (iii) whether "generic" initial data (with respect to some appropriate notion of genericity) always produce "pure" spacetime singularities without Killing-Cauchy horizons, i.e., whether any breakdowns in global predictability can occur in "generic" gravitational plane-wave collisions. The investigation of these issues occupies much of Chapters 2, 5, and 7, culminating in the fundamental conclusions that Killing-Cauchy horizons in colliding plane-wave solutions are nongeneric phenomena, and that generic initial data for colliding plane waves always produce "pure" spacetime singularities without horizons. These generic singularities are similar in global structure to the singularity of the Khan-Penrose solution.

ALMOST-PLANE GRAVITATIONAL WAVES AND THEIR COLLISIONS

It is natural to raise the issue of whether (or under what conditions) spacetime singularities can be produced by the collisions of gravitational waves which are not exactly plane-symmetric, but which have finite but very large transverse "spatial" sizes; i.e., by the collisions of *almost-plane* gravitational waves. Almost-plane gravitational waves can be visualized as the fully nonlinear analogues of the well-known Gaussian-beam solutions to the linearized wave equation (2) with "waist radius" a_o

huge compared to wavelength λ. General properties of almost-plane waves are discussed in Chapter 3. Later chapters focus attention on the structure and collisions of almost-plane waves. Attention is restricted, initially, to almost-plane waves whose initial data across a bounded "central" region are identical with the initial data of exactly-plane waves, but fall off in an arbitrary way at larger transverse distances. It is proved that if the central region of exact plane symmetry is sufficiently large, then the collision between the almost-plane waves is guaranteed to produce a spacetime singularity with the same local structure as in an exact-plane-wave collision. It is then shown that if the initial data for the two colliding almost-plane waves are not exactly plane symmetric over any region, but only *sufficiently close* to being exactly plane symmetric across a sufficiently large but bounded region of the initial surface, then their collision must still produce spacetime singularities. Although our analysis proves rigorously the existence of these general singularities, it does not give any information about either their global structure or their local asymptotic behavior.

DETAILED SUMMARY OF THE REMAINING CHAPTERS

The reimainder of this thesis consists of a series of papers, all of which have been published in or submitted to The Physical Review.

In Chapter 2 [*Physical Review D* **36**, 1662 (1987)], we show that Killing-Cauchy horizons in exactly plane-symmetric spacetimes are unstable against plane-symmetric perturbations and thence argue that generic spacetimes representing colliding plane waves are likely to have spacelike singularities without Killing-Cauchy horizons. More specifically, in this Chapter we give an explicit definition of Killing-Cauchy horizons in plane-symmetric spacetimes and we classify these horizons into two types: those which are smooth surfaces, called "type I," and those which are not smooth, called "type II." We then show that type I horizons are unstable with respect to any generic, plane-symmetric perturbation data posed on a suitable initial null boundary and evolved with arbitrarily nonlinear field equations satisfying some very general requirements; linearized gravitational perturbations constitute a special case of this instability, but fully nonlinear gravitational perturbations do not. We then consider plane-symmetric Killing-Cauchy horizons of type II, and prove that they are unstable with respect to generic, plane-symmetric perturbations satisfying linear evolution equations; a special case again is linearized gravitational perturbations.

In Chapter 3 [*Physical Review D* **37**, 2810 (1988)], we review some crucial features of the well-known exact solutions for colliding exactly plane waves and we argue that one of these features, the breakdown of "local inextendibility" can be regarded as nongeneric. We then introduce a new framework for analyzing general colliding exactly plane-wave spacetimes; we give an alternative proof of a theorem due to Tipler[8] implying the existence of singularities in all generic colliding plane-wave solutions; and we discuss the fact that the Chandrasekhar-Xanthopoulos[7] colliding plane-wave solutions are not strictly plane symmetric and thus do not satisfy the conditions and the conclusion of Tipler's theorem. Our alternative proof of Tipler's theorem emphasizes the role and the necessity of strict plane symmetry in establishing the existence of singularities in colliding plane-wave spacetimes. However, we argue on the basis of Chapter 2 that the breakdown of strict plane symmetry as exhibited in the Chandrasekhar-Xanthopoulos solutions is a nongeneric phenomenon. We then propose a definition of general *gravitational-wave spacetimes*, of which almost-plane waves are a special case; and we develop some mathematical tools for studying them. An old result of Dautcourt[9] implies that the only gravitational-wave spacetimes with a Killing propagation direction are the plane-fronted waves with parallel rays (PP waves); and we prove a new, related result, that the only gravitational-wave spacetimes with a precisely sandwiched curvature distribution are PP waves. These properties imply that almost-plane waves cannot propagate without diffraction, and that as opposed to the case for precisely plane waves, the curvature in an almost-plane-wave spacetime cannot be precisely sandwiched between two null surfaces (i.e., the wave must have tails). We also prove a "peeling theorem" for components of the Weyl curvature in general gravitational-wave spacetimes.

In Chapter 4 [*Physical Review D* **37**, 2790 (1988)], we construct an infinite-parameter family of exact solutions to the vacuum Einstein field equations describing colliding, exactly plane gravitational waves with parallel polarizations. The interaction regions of the solutions in this family are locally isometric to the interiors of those static axisymmetric (Weyl) black-hole solutions which admit both a nonsingular horizon, and an analytic extension of the exterior metric to the interior of the horizon. As a member of this family of solutions we also obtain, for the first time, a colliding plane-wave solution where both of the two incoming plane waves are purely anastigmatic, i.e., where both incoming waves have equal focal lengths.

When the colliding exactly plane waves have parallel (linear) polarizations, the mathematical analysis of the field equations in the interaction region is especially simple. Using a formulation of these field equations previously given by Szekeres,[10] in Chapter 5 [*Physical Review D* **38**, 1706 (1988)] we analyze the asymptotic structure of a general colliding parallel-polarized plane-wave solution near its singularity. We show that the metric is asymptotic to an inhomogeneous Kasner solution as the singularity is approached. We give explicit expressions which relate the asymptotic Kasner exponents along the singularity to the initial pre-collision waveforms of the two plane waves. It becomes clear from these expressions that for specific choices of initial waveforms the curvature singularity created by the colliding waves degenerates to a coordinate singularity, and that a nonsingular Killing-Cauchy horizon is thereby obtained. Our equations prove that this horizon is unstable in the full nonlinear theory against small but generic perturbations of the initial data, and that in a very precise sense, "generic" initial data always produce all-embracing, spacelike curvature singularities without Killing-Cauchy horizons. We give several examples of exact solutions which illustrate some of the asymptotic singularity structures that are discussed in the chapter. In particular, we construct a new family of exact colliding parallel-polarized plane-wave solutions, which create Killing-Cauchy horizons instead of a spacelike curvature singularity. The maximal analytic extension of one of these solutions across its Killing-Cauchy horizon results in a colliding plane-wave spacetime, in which the interior of a Schwarzschild black hole is created out of the collision between two plane-symmetric sandwich waves propagating in a cylindrical universe.

In Chapter 6 [*Physical Review D* **38**, 1731 (1988)], we consider the problem of whether (or under what conditions) singularities can be produced in the collisions of gravitational waves with finite but very large transverse sizes; i.e., in the collisions of almost-plane gravitational waves. On the basis of (nonrigorous) order-of-magnitude considerations, we discuss the outcome of the collision in two fundamentally different regimes for the parameters of the colliding waves; these parameters are the transverse sizes $(L_T)_i$, typical amplitudes h_i, typical reduced wavelengths $\lambda\!\!\!^{-}_i \equiv \lambda_i/2\pi$, thicknesses a_i, and focal lengths $f_i \sim \lambda\!\!\!^{-}_i{}^2/a_i h_i{}^2$ $(i=1,2)$ of the waves 1 and 2. For the first parameter regime where $(L_T)_1 \gg (L_T)_2$ and $h_1 \gg h_2$, we conjecture the following: (i) If $(L_T)_2 \ll \sqrt{\lambda\!\!\!^{-}_2 f_1}(\frac{h_1}{h_2})^{1/4}$, the almost-plane wave 2 will be focused by wave 1 down to a finite, minimum size, then diffract and disperse [Fig. 1(a) of Chapter 6]. (ii) If $(L_T)_2 \gg \sqrt{\lambda\!\!\!^{-}_2 f_1}(\frac{h_1}{h_2})^{1/4}$ (and if wave 1 is sufficiently anastigmatic), wave 2 will be focussed by wave 1 so strongly that it forms a singularity surrounded by a horizon, and the end result is a black hole flying away from wave 1 [Fig. 1(b) of Chapter 6]. For the second parameter regime where $(L_T)_1 \sim (L_T)_2 \equiv L_T$ and $h_1 \sim h_2$, we conjecture that if $L_T \gg \sqrt{f_1 f_2} \equiv f$, a horizon forms around the two colliding waves shortly before their collision, and the collision produces a singularity inside a black hole that is at rest in a reference frame in which $f_1 \sim f_2 \sim f$ (Fig. 2 of Chapter 6). As a first step in proving this conjecture, we give a rigorous analysis of the second regime in the singularity-forming case $L_T \gg f$. Our rigorous analysis is confined to the special situation of colliding parallel-polarized (almost-plane) gravitational waves which are exactly plane-symmetric across a region of transverse size $\gg f$, but which fall off in an arbitrary way at larger transverse distances. This analysis shows that the collision is guaranteed to produce a spacetime singularity with the same local structure as in an exact plane-wave collision, but it does not prove that the singularity is surrounded by a horizon.

In Chapter 7 [*Physical Review D*, submitted], we explore the structure of the singularities produced in the collisions of arbitrarily-polarized gravitational plane waves, and we reconsider the problem of whether (or under what conditions) singularities can be produced in the collisions of *almost-plane* gravitational waves with finite but very large transverse sizes. First we analyze the asymptotic structure of a general, arbitrarily-polarized, colliding, plane-wave spacetime near its singularity. We show that the metric is asymptotic to a generalized inhomogeneous-Kasner solution as the singularity is approached. In general, the asymptotic Kasner axes as well as the asymptotic Kasner exponents at the singularity are functions of the spatial coordinate that runs tangentially to the singularity in the non-plane-symmetric direction. It becomes clear that for specific values of these asymptotic Kasner exponents and axes the curvature singularity created by the colliding waves degenerates to a coordinate singularity, and that a nonsingular Killing-Cauchy horizon is thereby obtained. Our analysis proves that these horizons are unstable in the full nonlinear theory against small but generic plane-symmetric perturbations of the initial data, and that in a very precise and rigorous sense, "generic" initial data for colliding arbitrarily-polarized plane waves always produce all-embracing, spacelike curvature singularities without Killing-Cauchy horizons. Next we turn to the problem of colliding almost-plane gravitational waves, and by combining the results that we obtain in this Chapter and in the previous Chapters with the Hawking-Penrose singularity theorem and the Cauchy stability theorem, we prove that if the initial data for two colliding almost-plane waves are sufficiently close to being exactly plane symmetric across a sufficiently large but bounded region of the initial surface, then their collision must produce spacetime singularities. Although our analysis proves rigorously the existence of these singularities, it does not give any information about either their global structure (e.g., whether they are hidden behind an event horizon) or their local asymptotic behavior (e.g., whether they are of Belinsky-Khalatnikov-Lifshitz generic-mixmaster type).

OVERALL ACKNOWLEDGMENTS

I would like to acknowledge the support from the Caltech Physics Department in the form of Teaching Assistantships and Tuition Scholarships for the last five academic years, and in the form of a Summer Reading Grant for the first summer of my stay here. I would also like to thank Kip Thorne for providing, through NSF grant AST85-14911, Research Assistantships for the last four summers and the last academic year of my stay.

REFERENCES FOR CHAPTER 1

[1] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation*, (Freeman, San Francisco, 1973).

[2] H. Bondi, Nature (London) **179**, 1072 (1957).

[3] J. Ehlers and W. Kundt in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).

[4] K. Khan and R. Penrose, Nature (London) **229**, 185 (1971).

[5] R. A. Matzner and F. J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[6] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[7] S. Chandrasekhar and B. Xanthopoulos, Proc. R. Soc. Lond. **A410**, 311 (1987).

[8] F. J. Tipler, Phys. Rev. D **22**, 2929 (1980).

[9] G. Dautcourt, in *Conference Internationale sur les Theories Relativistes de la Gravitation*, (Pergamon, Oxford, 1964).

[10] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

CHAPTER 2

Instability of Killing-Cauchy Horizons in Plane-Symmetric Spacetimes

Published in Physical Review D **36**, 1662 (1987).

ABSTRACT

It is well known that when plane-symmetric gravitational waves collide, they produce singularities. Presently known exact solutions representing such collisions fall into two classes: those in which the singularities are spacelike, and those in which timelike singularities appear preceded by a Killing-Cauchy horizon. This paper shows that Killing-Cauchy horizons in plane-symmetric spacetimes are unstable against plane-symmetric perturbations and thence argues that generic spacetimes representing colliding plane waves are likely to have spacelike singularities without Killing-Cauchy horizons. More specifically, this paper gives an explicit definition of Killing-Cauchy horizons in plane-symmetric spacetimes and classifies these horizons into two types: those which are smooth surfaces, called "type I," and those which are singular, called "type II." It is then shown that type I horizons are unstable with respect to any generic, plane-symmetric perturbation data posed on a suitable initial null boundary and evolved with arbitrarily nonlinear field equations satisfying some very general requirements; linearized gravitational perturbations constitute a special case of this instability. Horizons of type II are shown to be unstable with respect to generic, plane-symmetric perturbations satisfying linear evolution equations; a special case again is linearized gravitational perturbations.

I. INTRODUCTION AND SUMMARY

It has been known since the early 1970's[1] that when two plane gravitational waves propagating in an otherwise flat background collide, they focus each other so strongly as to produce a spacetime singularity. Until recently all the known solutions to the Einstein field equations describing such collisions[1,2] entailed all-encompassing, spacelike singularities that could not be avoided by any observer on any timelike world line. However, recently Chandrasekhar and Xanthopoulos[3] have constructed exact solutions in which the collision produces a Killing-Cauchy horizon, which in turn (if one continues the metric through the horizon analytically) is followed by a timelike singularity that is readily avoided by almost all observers travelling on timelike world lines. On the other hand, there exist strong arguments[4] to the effect that generic colliding plane-wave spacetimes should be free of Killing-Cauchy horizons, and there are theorems[5,4] to the effect that nonflat plane-symmetric spacetimes without such horizons must be geodesically incomplete.

Hence, the question naturally arises as to which of the above outcomes of plane-wave collisions is generic (if, indeed any of them really is.) The present paper makes no attempt to formulate this question precisely (which in itself is a nontrivial task to accomplish.) However, this paper shows that the Killing-Cauchy horizons present in the recent Chandrasekhar-Xanthopoulos solutions can not be generic, because such horizons in any plane-symmetric spacetime are unstable against linear vacuum perturbations (as well as nonvacuum perturbations) that preserve the plane symmetry. It is natural to expect that the growth of these instabilities, in a generic plane-symmetric situation, will convert the horizon into an all-encompassing spacelike singularity, and that such singularities are therefore the generic outcome of plane-wave collisions. However, this paper does not make any attempt at proving this speculation rigorously.

(Independently of, and simultaneously with our proof of this instability, Chandrasekhar and Xanthopoulos discovered that the presence of a perfect fluid with (energy density) = pressure, or a null dust, in their spacetime destroys the horizon in the full nonlinear theory.)

Before turning to a detailed formulation and proof of the instability results, we illustrate them by two simple examples of plane-symmetric spacetimes with Killing-Cauchy horizons. In section II of this paper we shall classify such horizons into two classes which we call type I and type II. A simple example of a spacetime with a Killing-Cauchy horizon of type I is the plane-polarized, plane sandwich wave[6] with the metric

$$g=-dudv+F^2(u)dx^2+G^2(u)dy^2 , \tag{1.1}$$

where F, G are constant (hence g is flat) for $u \leq 0$ and

$$F(u)=(f_1-u) ,$$

$$G(u)=(f_2-u) \tag{1.2}$$

for $u \geq 1$, where $f_2 \geq f_1 > 1$. In the region $0 \leq u \leq 1$, F and G are determined by the spacetime curvature associated with the gravitational wave. The wave is sandwiched inside the region $0 \leq u \leq 1$ since this spacetime is flat not only for $u \leq 0$ but also for $u \geq 1$, as becomes evident after transforming to the global coordinate system (U,V,X,Y) given by (for $u \geq 1$)

$$x=\frac{X}{(f_1-U)} ,$$

$$y=\frac{Y}{(f_2-U)} ,$$

$$u=U\ ,$$

$$v=V+\frac{X^2}{(f_1-U)}+\frac{Y^2}{(f_2-U)} \tag{1.3}$$

in which the metric is

$$g=-dUdV+dX^2+dY^2\ , \qquad\qquad U\geq 1. \tag{1.4}$$

As is clear from the form of the metric in Eq.(1.1), the plane-wave spacetime admits the two spacelike Killing vectors $\vec{\xi}_1=\partial/\partial x$ and $\vec{\xi}_2=\partial/\partial y$ as plane-symmetry generators. In the global Minkowskian chart (U,V,X,Y) that covers the whole spacetime including the surface $\{u=U=f_1\}$ in a nonsingular fashion, these Killing vectors are given by the expressions

$$\vec{\xi}_1=\frac{\partial}{\partial x}=(f_1-U)\frac{\partial}{\partial X}-2X\frac{\partial}{\partial V}\ ,$$

$$\vec{\xi}_2=\frac{\partial}{\partial y}=(f_2-U)\frac{\partial}{\partial Y}-2Y\frac{\partial}{\partial V}\ . \tag{1.5}$$

The Killing vector $\vec{\xi}_1$ or both $\vec{\xi}_1$ and $\vec{\xi}_2$ become null on the Killing-Cauchy horizon $\mathcal{S}=\{U=f_1\}$ according to whether $f_2>f_1$ or $f_2=f_1$. (See figure 1 for the case $f_1=f_2$.) In either case they are both spacelike before the horizon ($U<f_1$) and become tangent to the horizon as U approaches f_1, one (or both) of them pointing along the null generators of the Killing-Cauchy horizon when $U=f_1$. In the case $f_1=f_2$, both $\vec{\xi}_1$ and $\vec{\xi}_2$ vanish on the null line $\mathcal{C}=\{U=f_1,X=Y=0\}$ in $\mathcal{S}$, whereas on any neighborhood of $\mathcal{C}$ in $\mathcal{S}$ at least one of $\vec{\xi}_i$ $(i=1,2)$ is nonzero. In the case $f_2>f_1$, $\vec{\xi}_1$ vanishes on the null two-plane $\mathcal{P}=\{U=f_1,X=0\}$ in $\mathcal{S}$, whereas it is nonzero on any neighborhood of $\mathcal{P}$ in $\mathcal{S}$. On the other hand, $\vec{\xi}_2$ remains spacelike and nonzero on $\mathcal{S}$ in this case.[7] Figure 1 depicts the Killing vector field $\vec{\xi}_1=\partial/\partial x$, surfaces of constant u and v, and the Killing-Cauchy horizon $\mathcal{S}=\{u=U=f_1\}$ for this example, in the case $f_1=f_2$ and in Minkowskian coordinates with the Y direction suppressed. As one sees from this figure or from Eqs.(1.3) and (1.4), the horizon $U=f_1$ is a smooth hypersurface in spacetime generated by endless null geodesics. This turns out to be the feature that distinguishes type I horizons from type II.

In section III we study the propagation of a wide class of classical fields on a plane-symmetric spacetime having a Killing-Cauchy horizon of type I as in the above example. The class of fields we work with is constrained only by the type of wave equation they satisfy and these constraints are very weak; for example, they admit linear scalar waves satisfying $\Box\phi=0$, linearized gravitational perturbations, and fields satisfying arbitrarily nonlinear evolution equations that respect the causal structure of the unperturbed background spacetime (e.g., the $\lambda\phi^4$ field theory); but not (in general) the fully nonlinear gravitational perturbations. Section III shows that when generic, plane-symmetric initial data for such fields are propagated with the corresponding field equations on a plane-symmetric spacetime with a Killing-Cauchy horizon of type I, the fields become singular as they approach the horizon.

This instability of Killing-Cauchy horizons of type I is well illustrated by the example of a linear scalar field satisfying the wave equation $\Box\phi=0$ in the above plane sandwich-wave spacetime given by Eqs. (1.1) and (1.2). The scalar wave equation

$$\Box\phi=\frac{1}{\sqrt{-|g|}}\frac{\partial}{\partial x^\alpha}\left(\sqrt{-|g|}\,g^{\alpha\beta}\frac{\partial\phi}{\partial x^\beta}\right)=0$$

in this case takes the form [Eq. (1.1)]

$$-4\phi_{,uv}-2\left(\frac{F_{,u}}{F}+\frac{G_{,u}}{G}\right)\phi_{,v}+\frac{1}{F^2}\phi_{,xx}+\frac{1}{G^2}\phi_{,yy}=0\,. \qquad (1.6)$$

For a plane-symmetric field $\phi(u,v)$ and for $u \geq 1$, this equation becomes [cf. Eq. (1.2)]

$$-2\phi_{,uv}+\left(\frac{1}{f_1-u}+\frac{1}{f_2-u}\right)\phi_{,v}=0 \tag{1.7}$$

and has the general solution

$$\phi=\frac{a(v)}{[(f_1-u)(f_2-u)]^{1/2}}+b(u)\,, \tag{1.8}$$

where a and b are functions that are uniquely determined by initial data for ϕ on the null boundary consisting of the null surfaces $u=1$ and $v=0$. Clearly, for generic initial data, a and b will be nonzero and ϕ will diverge as $u=U \rightarrow f_1$. If initial data on the surface $\{v=0\}$ and for $v \geq v_1 > 0$ on the surface $\{u=1\}$ are zero, then this initial-value problem describes the collision of a scalar plane sandwich wave with the background gravitational plane wave. In that case the solution simply is

$$\phi=a(v)\left[\frac{(f_1-1)(f_2-1)}{(f_1-u)(f_2-u)}\right]^{1/2}=a\left(V+\frac{X^2}{f_1-U}+\frac{Y^2}{f_2-U}\right)\left[\frac{(f_1-1)(f_2-1)}{(f_1-U)(f_2-U)}\right]^{1/2} \tag{1.9}$$

where $a(v)$ is equal to ϕ on the initial surface $\{u=1\}$ and vanishes for $v \geq v_1$ and for $v \leq 0$.

Geometrically, the reason for this singular behavior is simple: The symmetry of the spacetime, as embodied in the Killing vector fields $\vec{\xi}_1=\partial/\partial x$ and $\vec{\xi}_2=\partial/\partial y$, forces the plane-symmetric field to focus onto the line $\mathcal{C}$ (Fig. 1); a line to which all curves of constant v,x,y converge as $u \rightarrow f_1$; and this focussing of the waves produces a divergence in their amplitude. The proof of instability in Sec. III shows that this behavior is quite general for plane-symmetric spacetimes with type I Killing-Cauchy horizons.

Turn now to the second example, a spacetime with metric

$$g=-dt^2+dz^2+t^2dx^2+dy^2\ , \qquad t\leq 0 \tag{1.10a}$$

or, by putting $u=t-z\ ,v=t+z$;

$$g=-dudv+\frac{1}{4}(u+v)^2dx^2+dy^2\ , \qquad u+v\leq 0 \tag{1.10b}$$

in which the plane-symmetry-generating Killing vectors are again $\vec{\xi}_1=\partial/\partial x$ and $\vec{\xi}_2=\partial/\partial y$. In this case $\vec{\xi}_2$ is everywhere spacelike, while $\vec{\xi}_1$ becomes null on the Killing-Cauchy horizon $t=0$ but is spacelike prior to the horizon ($t<0$). This spacetime is actually flat as one sees from the coordinate transformation

$$T=t\cosh x\ ,$$

$$X=t\sinh x\ ,$$

$$Y=y\ ,$$

$$Z=z\ , \tag{1.11}$$

$$g=-dT^2+dX^2+dY^2+dZ^2\ . \tag{1.12}$$

Figure 2 depicts the Killing vector fields $\vec{\xi}_1=\partial/\partial x$, $\vec{\xi}_2=\partial/\partial y$, surfaces of constant t, and the horizon $\{t=0\}$ $(=\{T=-|X|\})$ in the Minkowskian coordinates with the Z (z) dimension suppressed. As one sees from this figure, the horizon $t=0$ is not everywhere smooth; it has a crease on the curve denoted by $\mathcal{C}$ in the figure; i.e., at $T=X=0$. This kind of nonsmooth behavior characterizes type II horizons; it shows up, for example, in the Killing-Cauchy horizons of the exact, colliding plane-wave solutions studied by Chandrasekhar and Xanthopoulos in Ref. 3 [their Eq. (124)].

Section IV of this paper studies the propagation of fields satisfying linear wave equations (e.g., scalar fields or linearized gravitational perturbations) in a plane-symmetric spacetime with a type II Killing-Cauchy horizon. When these fields are constrained to be plane symmetric and are evolved from generic initial data, they diverge as they approach the horizon. As an example, consider a scalar field satisfying $\Box\phi=0$ in the spacetime with metric (1.10). The general plane-symmetric (i.e., x,y independent) solution to

$$\Box\phi=-\frac{\partial^2\phi}{\partial t^2}-\frac{1}{t}\frac{\partial\phi}{\partial t}+\frac{\partial^2\phi}{\partial z^2}+\frac{1}{t^2}\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}=0 \tag{1.13}$$

is

$$\phi=\int_{-\infty}^{+\infty}[A(\omega)J_0(\omega t)+B(\omega)Y_0(\omega t)]e^{i\omega z}\,d\omega\,, \tag{1.14}$$

where J_0,Y_0 are the Bessel functions of the first and second kind and the functions $A(\omega)$, $B(\omega)$ are uniquely determined by initial data for ϕ on some initial t=const. surface prior to the horizon $t=0$. As we approach the horizon $t=0$, $J_0(\omega t)$ remains well behaved but $Y_0(\omega t)$ diverges logarithmically

$$Y_0(\omega t)\sim\frac{2}{\pi}\ln|\omega t|+\text{const}\,; \tag{1.15}$$

and correspondingly, unless $B(\omega)$ vanishes for all ω (a non-generic case),

$$\phi\sim E(z)\ln|t|=\frac{1}{2}E(Z)\ln|T^2-X^2| \tag{1.16}$$

for some (generically nonzero) function $E(z)$.

As in the type I case, the reason for this instability is geometrical: The $\vec{\xi}_1$ symmetry of the field and of the spacetime forces the waves to focus onto the line $\mathcal{C}$ (T=X=0), to which all curves of constant x,y,z converge as $t \to 0$ (Fig. 2); and this focussing of the field produces a divergence in its amplitude. The proof of instability in Sec. IV shows that this behavior is quite general for linear fields in plane-symmetric spacetimes with type II horizons.

In the concluding section (Sec. V) we briefly recapitulate the implications of these results for the general structure of singularities in plane-symmetric spacetimes.

Throughout this paper our notation and conventions are the same as those in Ref.8, in particular the metric has signature (−,+,+,+) and the Newman-Penrose equations are used in the "rationalized" form appropriate to that signature[9,4].

II. CLASSIFICATION

By a plane-symmetric spacetime we shall mean a maximal spacetime $(\mathcal{M},g)$ with a C^2 metric g on which there exist (i) a pair of commuting Killing vectors $\vec{\xi}_i \equiv \vec{\xi}_1, \vec{\xi}_2$, and (ii) a dense open subset at each point of which the $\vec{\xi}_i$ generate a spacelike two dimensional plane in the tangent space. If the dense open subset is equal to $\mathcal{M}$, we call $(\mathcal{M},g)$ strictly plane symmetric as no breakdowns of plane symmetry occur on $\mathcal{M}$.

By a Killing-Cauchy horizon in a plane-symmetric spacetime $(\mathcal{M},g)$ we shall mean a null, achronal, edgeless[8] three-dimensional connected (C^{1-}) surface $\mathcal{S}$ in $\mathcal{M}$ on which at least one of the Killing vectors $\vec{\xi}_i$ degenerates to a null Killing vector (which is not identically zero on $\mathcal{S}$); and whose null geodesic generators have no past endpoints in $\mathcal{M}$ and are past complete. It follows from the definition of plane symmetry that both $\vec{\xi}_i$ must be tangent to $\mathcal{S}$, and hence the Killing vector(s) which degenerates to a null vector on $\mathcal{S}$ is tangent to the null generators of $\mathcal{S}$ on $\mathcal{S}$. As the spacetime is maximal and the generators of $\mathcal{S}$ are tangent to Killing directions, we assume (without loss of generality) that the null geodesics generating $\mathcal{S}$ are also future complete in $\mathcal{M}$ (or at least in a neighborhood of $\mathcal{S}$ in $\mathcal{M}$.[8])

If $(\mathcal{M},g)$ is a spacetime with a Killing-Cauchy horizon $\mathcal{S}$ for which the above definitions are satisfied only on $I^-(\mathcal{S})\cup\mathcal{S}$, we will still regard $(\mathcal{M},g)$ as plane symmetric for it will become clear later that this is all we need to prove our results. (See the remarks following theorems 1 and 2.)

On any plane-symmetric spacetime there are local coordinate systems (u,v,x,y) (covering at least $I^-(\mathcal{S})$) such that $\vec{\xi}_i=\partial/\partial x^i \quad (x^1\equiv x, x^2\equiv y)$. By plane symmetry, in any such coordinate system a Killing-Cauchy horizon $\mathcal{S}$ in $\mathcal{M}$ will be given by an expression of the form $\{f(u,v)=\text{const.}\}$ since $\vec{\xi}_i$ are tangent to $\mathcal{S}$. Then there are two possible cases:

If there exists a local coordinate system (u,v,x,y) in which $\vec{\xi}_i=\partial/\partial x^i$ and $\mathcal{S}$ is given by $\mathcal{S}=\{f(u,v)=\text{const.}\}$ where $\vec{\nabla}f$ is a smooth, everywhere nonvanishing vector field on $\mathcal{S}$, then we will say that $\mathcal{S}$ is a Killing-Cauchy horizon of type I.

If in every local coordinate system of the above kind and for every $f(u,v)$ such that $\mathcal{S}=\{f(u,v)=\text{const.}\}$, $\vec{\nabla}f$ either vanishes or blows up at some points on $\mathcal{S}$, then we call $\mathcal{S}$ a Killing-Cauchy horizon of type II.

Clearly, the first example of a Killing-Cauchy horizon which we described in the last section [Sec. I, Eqs. (1.1)—(1.5)] is of type I since it was given by $\mathcal{S}=\{u=U=f_1\}$ and $\vec{\nabla}u=-2\partial/\partial v=-2\partial/\partial V$ is a smooth everywhere nonzero vector field on $\mathcal{S}$. On the other hand our second example [Eqs. (1.10)—(1.12)] was of type II as it was given by $\mathcal{S}=\{t=\frac{1}{2}(u+v)=0\}$ where

$$\vec{\nabla}t = -\frac{\partial}{\partial t} = -\frac{T}{(T^2-X^2)^{1/2}}\frac{\partial}{\partial T} - \frac{X}{(T^2-X^2)^{1/2}}\frac{\partial}{\partial X}$$

which blows up on $\mathcal{S}$. An alternative choice for f, $f(u,v)=t^2=\frac{1}{4}(u+v)^2$ leads to $\vec{\nabla}f=2t\vec{\nabla}t$, which vanishes on the crease line $\mathcal{C}=\{T=X=0\}\subset\mathcal{S}$. It is not possible to describe $\mathcal{S}$ globally by any $f(u,v)=0$ where $\vec{\nabla}f$ is smooth and everywhere nonzero on $\mathcal{S}$.

III. INSTABILITY OF HORIZONS OF TYPE I

Before stating our instability theorem for horizons of type I, we formulate some of our assumptions:

Assumption (A1): $(\mathcal{M},g)$ is a plane-symmetric vacuum spacetime.

Assumption (A2): There is an open subset in $\mathcal{M}$ on which g is flat.

Assumption (A2) is not true of all plane-symmetric spacetimes, but it is true of spacetimes containing nothing but plane-symmetric gravitational waves (possibly coupled with matter or electromagnetic radiation), since such spacetimes are flat before any of the waves arrive.

By (A1) we can define a canonical null tetrad on $(\mathcal{M},g)$: $\vec{l},\vec{n}$ are the null geodesic congruences everywhere orthogonal to the $\vec{\xi}_i$ and Lie parallel along $\vec{\xi}_i$; $\vec{m},\vec{m}^*$ are linearly independent linear combinations of the $\vec{\xi}_i$, normalized such that $-g(\vec{l},\vec{n})=g(\vec{m},\vec{m}^*)=1, g(\vec{m},\vec{m})=0$. Then as is shown by Szekeres,[10] it follows from the presence of only two nontrivial dimensions that we can find a local chart (u,v,x,y) with $\vec{\xi}_i=\partial/\partial x^i$ such that

$$\vec{l}=\frac{\partial}{\partial u}+P^i(u,v)\frac{\partial}{\partial x^i}\ ,$$

$$\vec{n}=R(u,v)\frac{\partial}{\partial v}+Q^i(u,v)\frac{\partial}{\partial x^i}\ ,$$

$$\vec{m}=\frac{1}{F(u,v)}\frac{\partial}{\partial x}+\frac{1}{G(u,v)}\frac{\partial}{\partial y}\ , \tag{3.1}$$

where P^i, Q^i, R are real and F, G are complex, with $F^*G-G^*F\neq 0$ throughout the region on which strict plane symmetry holds and the tetrad (3.1) and the coordinate chart (u,v,x,y) are well behaved. The commutation relations[4] for the tetrad (3.1) yield

$$RP^1{}_{,v}-Q^1{}_{,u}=\frac{4\alpha}{F}+\frac{4\alpha^*}{F^*}-\frac{R_{,u}}{R}Q^1\ ,$$

$$RP^2{}_{,v}-Q^2{}_{,u}=\frac{4\alpha}{G}+\frac{4\alpha^*}{G^*}-\frac{R_{,u}}{R}Q^2\ , \tag{3.2}$$

where α denotes the Newman-Penrose spin coefficient. We can eliminate the $P^i\partial_i$ and $Q^i\partial_i$ terms from (3.1) by a coordinate transformation of the form

$$u'=u\ ,$$

$$v'=v\ ,$$

$$x^{i'}=x^i+\zeta^i(u,v)\ , \tag{3.3}$$

if $P^i+\zeta^i{}_{,u}=Q^i+R\,\zeta^i{}_{,v}=0$. But the integrability conditions for $P^i+\zeta^i{}_{,u}=Q^i+R\,\zeta^i{}_{,v}=0$ are

$$RP^i{}_{,v}-Q^i{}_{,u}=-\frac{R_{,u}}{R}Q^i\ ,$$

which by (3.2) are equivalent to $\alpha\equiv 0$. However, it follows by standard arguments[10] using the Ricci identities[10,4] in the vacuum case that assumption (A2) guarantees $\alpha\equiv 0$

on $\mathcal{M}$ when (3.1) is suitably set in the flat region. Hence we can, by a coordinate change (3.3), put our tetrad into the form

$$\vec{l}=\frac{\partial}{\partial u} ,$$

$$\vec{n}=R(u,v)\frac{\partial}{\partial v} ,$$

$$\vec{m}=\frac{1}{F(u,v)}\frac{\partial}{\partial x}+\frac{1}{G(u,v)}\frac{\partial}{\partial y} . \tag{3.4}$$

The Newman-Penrose commutation relations for the tetrad (3.4) give zero values for the following combinations of spin coefficients

$$\kappa=\nu=\alpha=\beta=\tau=\pi=\gamma+\gamma^{*}=\rho-\rho^{*}=\mu-\mu^{*}=0 ;$$

and the field equations then imply that two of the components of the Weyl tensor vanish:

$$\Psi_1=\Psi_3=0 .$$

The other spin coefficients can also be calculated using the commutation relations. Of them we will only need the complex expansion

$$\rho=\frac{1}{2(F^{*}G-G^{*}F)}\left[F^{*}G\left(\frac{F_{,u}}{F}+\frac{G^{*}{}_{,u}}{G^{*}}\right)-G^{*}F\left(\frac{F^{*}{}_{,u}}{F^{*}}+\frac{G_{,u}}{G}\right)\right] . \tag{3.5}$$

Assumption (A3): There is a Killing-Cauchy horizon $\mathcal{S}$ of type I in $(\mathcal{M},g)$.

Assumption (A4): The metric g is analytic in a neighborhood of $\mathcal{S}$ in $\mathcal{M}$, i.e., there are admissible coordinate systems in a neighborhood of $\mathcal{S}$ in which the metric coefficients are analytic functions.

Assumption (A4) guarantees [as $g(\vec{\xi}_i, \vec{\xi}_i)$ are analytic functions near $\mathcal{S}$] that strict plane symmetry holds on a neighborhood $\mathcal{W}$ of $\mathcal{S}$ in $\mathcal{M}$, with the exception of breaking down on $\mathcal{S}$ itself.

Also note that the tetrad component $R(u,v)$ [Eq. (3.4)] is bounded and nonzero on $\mathcal{S}$ since the vanishing or divergence of R at $\mathcal{S}$ will cause curvature singularities (in Ψ_2 and Ψ_4) to appear on $\mathcal{S}$. (See, e.g., Ref. 10 and Ref. 11.) On the other hand by (A3) $\mathcal{S}$ is of the form $\{f(u,v)=\text{const}\}$ where $\vec{\nabla} f$ is smooth and everywhere nonzero on $\mathcal{S}$. Therefore by the implicit function theorem,[12] $\mathcal{S}=\{f(u,v)=\text{const}\}$ is a smooth (at least C^1) null surface and hence is generated by null geodesics without endpoints. Then, since the generators of $\mathcal{S}$ are future and past complete in $\mathcal{M}$ by assumption, by exactly the same argument as we will give in the proof of theorem 2 below, it follows that we can find a function $\hat{f}$ which is smooth, vanishes on $\mathcal{S}$, and has a smooth, null nonzero gradient everywhere in a neighborhood of $\mathcal{S}$. As $\hat{f}$ has these properties globally on all of $\mathcal{S}$, it can be chosen to be a function of only u and v. (Since $\vec{\xi}_i=\partial/\partial x^i$ are Killing and hence have zero convergence and since they become tangent to the horizon $\mathcal{S}=\{\hat{f}=0\}$, they can not be threading through every family of surfaces $\{\hat{f}=\text{const}\neq 0\}$ each of which consists of parallel null surfaces generated by complete null geodesics without endpoints.) Redefine $f\equiv\hat{f}$ since $\hat{f}=0$ on $\mathcal{S}$. Then $\mathcal{S}=\{f=0\}$ and $0=g(\vec{\nabla} f, \vec{\nabla} f)=-Rf_{,u}f_{,v}$ [by Eq. (3.4)] in a neighborhood of $\mathcal{S}$. As $R\neq 0$ on $\mathcal{S}$, this implies either $f_{,u}\equiv 0$ or $f_{,v}\equiv 0$ (but not both since $\vec{\nabla} f\neq 0$) in a neighborhood of $\mathcal{S}$, which clearly tells us that $\mathcal{S}$ is a surface of the form $\{u=\text{const}\}$ or $\{v=\text{const}\}$. We shall assume, without loss of generality, that $\mathcal{S}=\{u=f\}$ where f is a constant.

Theorem 1. Let $(\mathcal{M},g)$ be a spacetime satisfying assumptions (A1)—(A4). Let $\{\mathcal{Q}^a\}$ denote an arbitrary multi-index field (e.g., scalar, tensorial, or spinorial) defined on the spacetime, which satisfies field equations obeying the following conditions:

a) $\mathcal{Q}^a \equiv 0$ is a solution of the field equations.

b) The characteristic surfaces for the field equations are null surfaces of $(\mathcal{M}, g)$ and the evolution of $\{\mathcal{Q}^a\}$ is globally causal: if initial data for $\{\mathcal{Q}^a\}$ are zero outside a closed set $\mathcal{K}$ in an initial surface Σ, then there exists an open neighborhood $\mathcal{U}(\Sigma)$ of $\overline{D^+(\Sigma)}$ in $\mathcal{M}$ such that whenever there exists a smooth extension of the solution on $D^+(\Sigma)$ to $\mathcal{U}(\Sigma)$ it can be chosen so that $\mathcal{Q}^a = 0$ on $\mathcal{U}(\Sigma) - (J^+(\mathcal{K}) \cup J^-(\mathcal{K}))$.

c) There is a consistent characteristic initial-value formalism for the field equations for $\{\mathcal{Q}^a\}$: if $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$ is an initial null boundary consisting of three dimensional null surfaces $\mathcal{N}_1$, $\mathcal{N}_2$ intersecting in a two dimensional spacelike surface Z, then one can freely pose initial data on $\mathcal{N}$ (satisfying some constraint equations on $\mathcal{N}$.) Morover, uniqueness and local existence of solutions in $D^+(\mathcal{N})$ hold for both the general characteristic initial-value problem and for the plane-symmetric initial-value problem for $\{\mathcal{Q}^a\}$; the latter being obtained from the field equations by assuming $(\mathcal{L}_{\vec{\xi}_i} \mathcal{Q})^a \equiv 0$.

If these conditions are satisfied, then there is a null boundary $\mathcal{N}$ in $I^-(\mathcal{S})$ such that the evolution of any generic member of a class of plane symmetric initial data for $\{\mathcal{Q}^a\}$ on $\mathcal{N}$ that we will describe develops singularities on the Killing-Cauchy horizon $\mathcal{S}$.

Remarks:

(i) First note that the conditions (a), (b), (c) are universal properties of all physical fields that do not, by their stress-energy, act back on the geometry of the background spacetime; hence in particular of linearized gravitational perturbations. Although we are primarily interested in fields satisfying linear evolution equations, it is clear that inclusion of higher order terms in the equations will not affect the validity of the theorem so long as these terms respect the causal structure of the background spacetime. (Note that fully nonlinear gravitational perturbations will not, in general, have this property.[13]) Linearity is not necessary for any of the conditions (a), (b), (c).

(ii) As will be clear from the proof, the theorem will still hold if our assumptions (A1), (A2) and (A3) are valid only in an open subset of the region $I^{-}(\mathcal{S})$ whose closure in $\mathcal{M}$ contains $\mathcal{S}$.

(iii) Our only use of the vacuum assumption is in the Ricci identities involving Φ_{10} and Φ_{21}, which are ingredients in the proof[10,4] that (A2) permits setting the Newman-Penrose spin coefficient α to zero and thence permits specializing the tetrad from (3.1) to (3.4). Consequently, the theorem is also valid for a spacetime $(\mathcal{M},g)$ satisfying assumptions (A1)—(A4) with the exception that the stress-energy tensor T, instead of being zero, is assumed to only satisfy $T(\vec{l},\vec{\xi}_i)=T(\vec{n},\vec{\xi}_i)=0$ on $\mathcal{M}$, which will guarantee $\Phi_{10}=\Phi_{21}=0$.

(iv) We will formulate the genericity condition on the data for $\{\mathcal{Q}^a\}$ on $\mathcal{N}$ in the course of the proof.

(v) The reader may find it helpful, when going through the details of the proof that follows, to carry along and look at the prototype example of a Killing-Cauchy horizon of type I discussed in the introduction [Eqs. (1.1)—(1.9), and Figs. 1 and 3].

Proof of Theorem 1. By (A3) at least one of the $\vec{\xi}_i$, which we can without loss of generality assume to be $\vec{\xi}_1=\partial/\partial x$, degenerates to a null Killing vector on $\mathcal{S}$ and becomes orthogonal to $\vec{\xi}_2$ since the (unique) null direction tangent to $\mathcal{S}=\{u=f\}$ is at the same time orthogonal to all vectors tangent to $\mathcal{S}$. This implies, putting $g_{ij}=g(\vec{\xi}_i,\vec{\xi}_j)$,

$$\lim_{u\to f} g_{11}=\lim_{u\to f} g_{12}=0 . \tag{3.6}$$

On the other hand, throughout the open set $\mathcal{W}-\mathcal{S}$ on which strict plane symmetry holds we have $g(\vec{m},\vec{m})=0$, $g(\vec{m},\vec{m}^*)=1$ which read

$$\frac{1}{F^2}g_{11}+\frac{2}{FG}g_{12}+\frac{1}{G^2}g_{22}=0\ , \tag{3.7a}$$

$$\frac{1}{F^*F}g_{11}+\left(\frac{1}{FG^*}+\frac{1}{F^*G}\right)g_{12}+\frac{1}{GG^*}g_{22}=1\ . \tag{3.7b}$$

Then, at least one of $\lim_{u\to f}F$ or $\lim_{u\to f}G$ has to vanish since otherwise by Eq. (3.7a) $\lim_{u\to f}g_{22}=0$ and it is impossible to satisfy Eq. (3.7b) in a neighborhood of $\mathcal{S}$ since $\lim_{u\to f}g_{11}=\lim_{u\to f}g_{12}=0$ by Eq. (3.6). Since by (A4) $g(\vec{\xi}_i,\vec{\xi}_i)$ are analytic functions in a neighborhood of $\mathcal{S}$, it is clear that F and G are regular in a neighborhood of $\mathcal{S}=\{u=f\}$, and hence by (A4) and (A3) (namely that the Killing-Cauchy horizon $\mathcal{S}$ is of type I), we can express them as convergent power series in $(u-f)$ in a neighborhood of $\mathcal{S}$:

$$F(u,v)=\sum_{n=k}^{\infty}\alpha_n(v)(u-f)^n\ ,$$

$$G(u,v)=\sum_{n=l}^{\infty}\beta_n(v)(u-f)^n\ , \tag{3.8}$$

where $k\geq 0, l\geq 0$ and $k+l\geq 1$; $\alpha_n(v)$, $\beta_n(v)$ are (not necessarily analytic) complex functions with $\alpha_k(v)\neq 0$, $\beta_l(v)\neq 0$. Inserting Eqs. (3.8) into Eq. (3.5) we obtain that the asymptotic behaviour of ρ to leading order as $u\to f$ is given by

$$\rho\sim\frac{k+l}{u-f}\ , \qquad (u\to f) \tag{3.9}$$

where $k+l \geq 1$.

Now consider a point $p_0 \in I^-(\mathcal{S})$ lying in the region of strict plane symmetry $\mathcal{W}-\mathcal{S}$ with $u(p_0)<f$, and consider the two-surface Z_{p_0} obtained by sweeping the point p_0 with the Killing symmetry generators $\vec{\xi}_i$; i.e., let Z_{p_0} be the Killing-orbit of p_0. (See Fig. 3.) Clearly, the null geodesic generators of $\dot{J}^+(Z_{p_0})$ which have their past endpoints on Z_{p_0} will consist of those in the $\vec{l}$–direction on which $v=v(p_0)$, and those in the $\vec{n}$–direction on which $u=u(p_0)$. Since $R \neq 0$ on $\mathcal{S}=\{u=f\}$, the tangent vectors to the null geodesic generators of $\dot{J}^+(Z_{p_0})$ in the $\vec{l}$–direction which lie in the surface $\{v=v(p_0)\}$ and which are given by $R\vec{l}$ have convergence $\hat{\rho}=R\rho$ which by Eq. (3.9) diverges to $-\infty$ as $u \rightarrow f$. This guarantees[8] that every null geodesic generator of $\dot{J}^+(Z_{p_0})$ having its past endpoint on Z_{p_0} has a conjugate point to Z_{p_0} along itself on the surface $\mathcal{S}$. We now claim that this actually corresponds to the null generators of the null surface $\{v=v(p_0)\}$ converging and intersecting each other in caustics on the Cauchy horizon $\mathcal{S}$. To see this, note that outside $\mathcal{S}$ the Killing vectors $\vec{\xi}_i$ generate translations on the set of null generators of the surface $\{v=v(p_0)\}$ by generating symmetries on their past endpoints in Z_{p_0}. On the other hand, if the null surface $\{v=v(p_0)\}$ intersects the null surface $\mathcal{S}$ transversally (i.e., not tangentially), then the intersection has to be a spacelike two-surface. But this is impossible since on $\mathcal{S}$ there does not exist a pair of spacelike linearly independent Killing vectors to generate translations on the set of null generators of $\{v=v(p_0)\}$ in this spacelike two-surface. Hence $\{v=v(p_0)\}$ intersects $\mathcal{S}$ non-transversally and as the convergence $\hat{\rho}$ of its generators diverges on $\mathcal{S}$, the intersection takes place either on a spacelike curve tangent to the Killing vector $\vec{\xi}_2$ which is still spacelike on $\mathcal{S}$, in the case that only one of $\vec{\xi}_i$ (namely $\vec{\xi}_1$) becomes null; or on a single point, in the case that both $\vec{\xi}_1$ and $\vec{\xi}_2$ become null on $\mathcal{S}$ (Fig. 3).

Note that in the first case, when $\vec{\xi}_2$ is still spacelike on $\mathcal{S}$, it generates translations on the set of generators of $\{v=v(p_0)\}$ along the curve in $\mathcal{S}$ on which these null generators converge and intersect each other, while the vector $\vec{\xi}_1$ which is null on $\mathcal{S}$ has to vanish on this curve. In the second case (the case depicted in Figs. 1 and 3) both $\vec{\xi}_1$ and $\vec{\xi}_2$ have to vanish at the point in $\mathcal{S}$ on which the generators of $\{v=v(p_0)\}$ intersect each other, since they must not generate any translations on the set of these generators at that point.

Therefore there is a null two-surface $\mathcal{P}$ (a null curve $\mathcal{C}$) in $\mathcal{S}$ which is the union of all spacelike curves (points) in $\mathcal{S}$ on which generators of the surfaces $\{v=v_0\}$ converge as v_0 ranges from $-\infty$ to $+\infty$, in the case $\vec{\xi}_2$ is spacelike on $\mathcal{S}$ (in the case both $\vec{\xi}_1,\vec{\xi}_2$ are null on $\mathcal{S}$). Moreover, this two-surface $\mathcal{P}$ (curve $\mathcal{C}$) is generated by the past endless null generators of $\dot{J}^+(Z_{p_0})$. (This can be seen by noting that the local chart (u,v,x,y) is regular on $I^-(\mathcal{S})$, thus all points in $I^-(\mathcal{S})$ with $v<v(p_0)$ are outside $J^+(Z_{p_0})$ and therefore, as $\dot{J}^+(Z_{p_0})$ is edgeless,[8,14] $\mathcal{P}$ ($\mathcal{C}$) must be generated — in $\overline{J^-[\{v=v(p_0)\}]}$ — by null geodesic generators of $\dot{J}^+(Z_{p_0})$ along $\mathcal{S}$ which are past endless and which intersect the generators of $\{v=v(p_0)\}$ at their focal points on $\mathcal{S}$.) The Killing vector $\vec{\xi}_1$ (both $\vec{\xi}_1$ and $\vec{\xi}_2$) vanishes on this surface $\mathcal{P}$ (curve $\mathcal{C}$) and since by (A3) and (A4) $\vec{\xi}_1$ ($\vec{\xi}_1,\vec{\xi}_2$) is a null vector not identically vanishing on $\mathcal{S}$ whose components in some coordinate frame are analytic functions, it has to be nonzero outside $\mathcal{P}$ ($\mathcal{C}$) on any open neighborhood in $\mathcal{S}$ of $\mathcal{P}$ ($\mathcal{C}$), generating symmetries along the null generators of $\mathcal{S}$.

We now show that it is sufficient to prove the theorem only for the case where $\{Q^a\}$ is a single scalar field ϕ. Let ∂_μ , μ=1,2,3,4, denote respectively the local coordinate basis fields $\partial/\partial u$, $\partial/\partial v$, $\partial/\partial x$, $\partial/\partial y$. Then for an arbitrary multicomponent (contravariant) tensor field $\{Q^a\}$, the Lie derivative along $\vec{\xi}_i$ of the inner product of Q with the $(p,0)$ tensor basis elements is given by:

$$\vec{\xi}_i[g(Q,\partial_{\mu_1}\otimes\cdots\otimes\partial_{\mu_p})]=g(Q,\sum_{k=1}^{p}\partial_{\mu_1}\otimes\cdots\otimes\mathcal{L}_{\vec{\xi}_i}\partial_{\mu_k}\otimes\cdots\otimes\partial_{\mu_p})+$$

$$+g(\mathcal{L}_{\vec{\xi}_i}Q,\partial_{\mu_1}\otimes\cdots\otimes\partial_{\mu_p}),$$

where we have made use of the fact that $\vec{\xi}_i$ are Killing hence $\mathcal{L}_{\vec{\xi}_i}g\equiv 0$. But the first term is zero as $\mathcal{L}_{\vec{\xi}_i}\partial_\mu=[\vec{\xi}_i,\partial_\mu]\equiv 0$, thus

$$\vec{\xi}_i[g(Q,\partial_{\mu_1}\otimes\cdots\otimes\partial_{\mu_p})]=g(\mathcal{L}_{\vec{\xi}_i}Q,\partial_{\mu_1}\otimes\cdots\otimes\partial_{\mu_p}).$$

This equation tells us that each component of a multicomponent tensor field $\{Q^a\}$ in the basis frame field $(\partial_u,\partial_v,\partial_x,\partial_y)$ behaves exactly like a scalar field under Lie transport by $\vec{\xi}_i$ since (as $\vec{\xi}_i=\partial/\partial x^i$ are Killing) $\vec{\xi}_i[g(\partial_\mu,\partial_\nu)]\equiv 0$. For spinor fields, by the same argument, the components of an arbitrary spinor field in the spin basis corresponding to the null tetrad (3.1) or (3.4) will behave like scalar fields under Lie transport by the $\vec{\xi}_i$. Clearly, this is also true for the components of the arbitrary tensor or spinor field in any local basis field that is Lie parallel along the $\vec{\xi}_i$, or in the spin basis that corresponds to any null tetrad that is Lie parallel along the $\vec{\xi}_i$. Therefore, despite the obvious fact that these basis fields themselves will in general develop singularities on the Killing-Cauchy horizon $\mathcal{S}$, precisely the following arguments by which we prove the singularity result of the theorem for a scalar field ϕ will prove the same result for an arbitrary field $\{Q^a\}$ (after constructing a suitable basis field Lie parallel along the $\vec{\xi}_i$ for each such field $\{Q^a\}$) when the initial data satisfy the conditions of the theorem.

Now consider a characteristic initial value problem for the scalar field ϕ (Fig. 3) in which the initial null boundary is given by $\mathcal{N}=\mathcal{N}_1\cup\mathcal{N}_2$, $\mathcal{N}_2=\{u=u(p_0)<f\}$, $\mathcal{N}_1=\{v=v(p_0)\}$, $\mathcal{N}_1\cap\mathcal{N}_2=Z_{p_0}$, and the initial data have the form: $\phi\equiv 0$ on $\mathcal{N}_1$, and $\phi=\phi(v)$, generic, nonzero, plane symmetric ("sandwich") data on $\mathcal{N}_2$ vanishing for $v\geq v(p_1)>v(p_0)$ and for $v\leq v(p_0)$, and satisfying the constraint equations (when there are any). The well-posedness of this problem is clear from the conditions (a), (b), (c) of the theorem. By condition (c), the evolution in $D^+(\mathcal{N})$ will have the full Killing symmetries: $\mathcal{L}_{\vec{\xi}_i}\phi=\vec{\xi}_i(\phi)\equiv 0$ throughout spacetime. We formulate the following notion of genericity for the data on $\mathcal{N}$:

Initial data for ϕ on $\mathcal{N}$ of the above class are generic if the solution is nonzero somewhere on the surface $\mathcal{P}$ (curve $\mathcal{C}$) in $\mathcal{S}$. For a multicomponent field we similarly demand that the solutions evolving from generic initial data take nonzero tensor (or spinor) values at some points of the surface $\mathcal{P}$ (curve $\mathcal{C}$) on $\mathcal{S}$. Note that, by "the solution on $\mathcal{P}$ ($\mathcal{C}$) in $\mathcal{S}$" we mean the limit of the solution on $I^-(\mathcal{S})$ as the field point approaches the plane $\mathcal{P}$ (the curve $\mathcal{C}$) lying in $\mathcal{S}$. Hence, more precisely, initial data for ϕ are generic if either this limit does not exist or it exists and is nonzero somewhere in $\mathcal{P}$ ($\mathcal{C}$) on $\mathcal{S}$. If the limit does not exist, then the field ϕ is singular near the horizon $\mathcal{S}$ and the theorem is proved. □

Now let us assume that this limit does exist and the field ϕ obtained by evolving the above data on $\mathcal{N}$ is smooth in a neighborhood of $\mathcal{S}=\{u=f\}$. (This assumption will produce a contradiction thereby implying that ϕ cannot be smooth — the conclusion of our theorem.) Then, since ϕ is smooth and not identically zero on $\mathcal{P}$ ($\mathcal{C}$), it will be nonzero on some open subset in $\mathcal{S}$ intersecting $\mathcal{P}$ ($\mathcal{C}$) in the region on which $\phi\neq 0$. But as the Killing vector $\vec{\xi}_1$ ($\vec{\xi}_1,\vec{\xi}_2$) generates symmetries along the null generators of $\mathcal{S}$ everywhere near $\mathcal{P}$ ($\mathcal{C}$) except on $\mathcal{P}$ ($\mathcal{C}$) itself, and since $\vec{\xi}_1(\phi)\equiv 0$ ($\vec{\xi}_i(\phi)\equiv 0$) on $\mathcal{S}$ as this holds prior to $\mathcal{S}$ and ϕ and $\vec{\xi}_1$ ($\vec{\xi}_i$) are smooth, $\vec{\xi}_1$ ($\vec{\xi}_i$) will carry this region on which ϕ is nonzero arbitrarily down into the past along the generators of $\mathcal{S}$. But when we move a sufficiently large affine distance into the past along these generators we clearly enter the region $\overline{J^-[\{v=v(p_0)\}]}$ in which the generators along $\mathcal{P}(\mathcal{C})$ are past endless generators of $\dot{J}^+(Z_{p_0})$ and hence of $\dot{J}^+(\mathcal{N}_2)$. Therefore any neighborhood in $\mathcal{M}$ of $\mathcal{P}(\mathcal{C})$ in this region intersects a piece of $\mathcal{M}$ not contained in $J^{\pm}(\mathcal{N}_2)$ [Fig. 3]. But again by the smoothness of ϕ and as $\vec{\xi}_1(\phi)\equiv 0$ ($\vec{\xi}_i(\phi)\equiv 0$), ϕ will be nonzero at all points of $\mathcal{P}(\mathcal{C})$ in this region and thereby be nonzero in a neighborhood in $\mathcal{M}$ of any point of $\mathcal{P}(\mathcal{C})$ there, contradicting the condition (b) of the theorem. Thus the assumption that ϕ is smooth near $\mathcal{S}=\{u=f\}$ is contradictory and must be false, and the field ϕ must develop singularities on $\mathcal{S}$ proving the theorem. □

The singularity of ϕ on $\mathcal{S}$ will in most cases be of the form $\phi\neq 0$ on $\mathcal{P}(\mathcal{C})$ for $v(p_0)<v<v(p_1)$ (bounded or unbounded) whereas $\phi\equiv 0$ on $\mathcal{S}$ outside $\mathcal{P}(\mathcal{C})$, with possibly an added smooth background field on $\mathcal{S}$ which satisfies $\phi^B(p)=0\ \forall p \in \mathcal{P}(\mathcal{C})$. Thus even though the field itself might be bounded near $\mathcal{S}$, some of its derivatives will diverge on the two-surface $\mathcal{P}$ (curve $\mathcal{C}$) in $\mathcal{S}$. However, if the field equations are linear, exactly the same argument we will use in proving theorem 2 will imply (as $\vec{\xi}_1$ or $\vec{\xi}_1,\vec{\xi}_2$ vanish on the surface $\mathcal{P}$ or the curve $\mathcal{C}$ in $\mathcal{S}$) that ϕ actually diverges on the set $\mathcal{P}$ (or on $\mathcal{C}$) in $\mathcal{S}$.

IV. INSTABILITY OF HORIZONS OF TYPE II

Theorem 2. Let $(\mathcal{M},g)$ be a plane-symmetric spacetime with a Killing-Cauchy horizon $\mathcal{S}$ of type II where strict plane symmetry holds on the intersection $\mathcal{W}$ of a neighborhood of $\mathcal{S}$ with $I^-(\mathcal{S})$. Let $\{\mathcal{Q}^a\}$ denote a field satisfying an arbitrary set of evolution equations such that:

a) The equations are linear.

b) There is a consistent (noncharacteristic) initial-value formalism for the field $\{ \mathcal{Q}^a \}$ and the evolution equations it satisfies, with local existence and uniqueness holding for both the general and the plane-symmetric initial-value problems.

If these conditions are satisfied, then there exists a spacelike partial Cauchy surface Σ in $I^-(\mathcal{S})$ such that the evolution of any generic, plane-symmetric initial data for $\{ \mathcal{Q}^a \}$ on Σ results in singularities on the Killing-Cauchy horizon $\mathcal{S}$.

Remarks:

(i) As will be clear from the proof, the assumptions of the theorem need only hold on $I^-(\mathcal{S}) \cup \mathcal{S}$ in $\mathcal{M}$.

(ii) The condition of genericity for the initial data on Σ will be formulated in the proof.

(iii) When studying the proof the reader may find it helpful to carry along and look at the prototype example of a type II horizon discussed in the Introduction [Eqs. (1.10) — (1.16)].

Proof of Theorem 2. We can set up the canonical local tetrad (3.1) on $(\mathcal{M}, g)$ in which the metric will be of the form

$$g = -\frac{1}{R(u,v)} du dv + A(u,v) du^2 + B(u,v) dv^2 + M^2(u,v) dx^2 +$$

$$+ N^2(u,v) dy^2 + K(u,v) dx dy + L_i(u,v) du dx^i + J_i(u,v) dv dx^i \ , \tag{4.1}$$

where $R(u,v)$ is positive, bounded and nonzero on $\mathcal{S}$. Put ${}^{(2)}g = -du dv + RA du^2 + RB dv^2$. Find local functions $t(u,v)$, $z(u,v)$ such that $t=0$ on $\mathcal{S}$ and

$$^{(2)}g=P\,(dz^2-dt^2)\ , \tag{4.2}$$

where $P\,(>0)$ is the conformal factor. (This can be done, for example, by solving the initial value problem $\{t=0$ on $\mathcal{S}$, $^{(2)}\Box t=0\ \}$ which in general has nonunique solutions, and then finding a "conjugate" $z\,(u\,,v\,)$ such that (4.2) is satisfied.[15]) Then the metric (4.1) becomes

$$g=-\frac{1}{\hat{R}\,(t\,,z\,)}(dt^2-dz^2)+F^2(t\,,z\,)dx^2+G^2(t\,,z\,)dy^2+K\,(t\,,z\,)dxdy+$$

$$+\hat{L}_i\,(t\,,z\,)dtdx^i+\hat{J}_i\,(t\,,z\,)dzdx^i\ , \tag{4.3}$$

where $\hat{R}\,(t\,,z\,)$ is again positive and bounded on $\mathcal{S}$. In both coordinate systems (4.1) and (4.3) the Killing vectors are $\vec{\xi}_i=\partial/\partial x^i$. Our definition of t guarantees that $\mathcal{S}$ is given by $\{t=0\}$; and by choosing $\vec{\xi}_1$ to be the Killing vector that becomes null and tangent to $\mathcal{S}$ on $\mathcal{S}$, we find that $F\,(t=0,z\,)=0$. Note that $\vec{\nabla}t=-\hat{R}\,\partial/\partial t$ is a timelike vector field which blows up on $\mathcal{S}$ while at the same time becoming tangent to $\mathcal{S}$. It will not be necessary in what follows to fix the coordinate (gauge) freedom further than that of Eq.(4.3).

Now we claim that since $\mathcal{S}$ is a Killing-Cauchy horizon of type II, some null generators of $\mathcal{S}$ must have future endpoints on $\mathcal{S}$. Null generators of $\mathcal{S}$ by definition have no past endpoints; if they do not have any future endpoints either, then one can globally express $\mathcal{S}$ in the form $\{f\,(u\,,v\,)=\text{const}\}$ where $\vec{\nabla}f$ is perfectly smooth and everywhere nonzero on $\mathcal{S}$, contradicting our assumption that $\mathcal{S}$ is of type II. To see this, assume null generators of $\mathcal{S}$ have no endpoints. Take a spacelike two-dimensional section Z of $\mathcal{S}$, and take a smooth field of (spacelike) basis fields $\vec{k}_1$, $\vec{k}_2$ on Z in $\mathcal{S}$. (This can be done since by plane symmetry the spacelike sections of $\mathcal{S}$

will not have spherical topology.) Propagate $\vec{k}_i$ along null generators $\vec{l}$ of $\mathcal{S}$ by parallel transport to all of $\mathcal{S}$. Since generators of $\mathcal{S}$ are both past and future complete in $\mathcal{M}$ by definition, $\vec{k}_i$ will be smooth on $\mathcal{S}$ and will have smooth extensions to a neighborhood of $\mathcal{S}$ in $\mathcal{M}$. Construct a null vector field $\vec{n}$ on $\mathcal{S}$ satisfying $g(\vec{n},\vec{k}_i)=0$, $g(\vec{n},\vec{l})=-1$. $\vec{n}$ will be a smooth vector field on $\mathcal{S}$ and will have a smooth extension (as a null vector) to a neighborhood of $\mathcal{S}$ in $\mathcal{M}$. Then take f to be the affine parameter along geodesics in the $\vec{n}$ direction, so that $\mathcal{S}=\{f=0\}$ and $\vec{\nabla} f$ on $\mathcal{S}$ is equal to $-\vec{l}$ and hence is smooth, null and everywhere nonzero on $\mathcal{S}$. By choosing $\vec{l}$, hence $\vec{n}$ and the (now not necessarily affine) parameter f such that f is constant on a family of parallel null surfaces near $\mathcal{S}$, $\vec{\nabla} f$ will retain these properties over a neighborhood of $\mathcal{S}$ in $\mathcal{M}$. Finally, by the same argument as we gave just before the statement of theorem 1, f can be chosen to be a function of only u and v.

Therefore, there is a nonempty subset $\mathcal{C}$ of $\mathcal{S}$ which consists of the endpoints of null generators of $\mathcal{S}$. (As $\mathcal{S}$ is achronal and edgeless it is a closed set and must contain these endpoints.) Now our Killing field $\vec{\xi}_1$ becomes null and tangent to $\mathcal{S}$ on $\mathcal{S}$, pointing along its null generators. But since $\mathcal{S}$ is a Killing horizon, the convergence and shear of its null geodesic generators must identically vanish on $\mathcal{S}$, and since $\mathcal{S}$ has no edge[8,14] the only way these generators can have endpoints on $\mathcal{S}$ is by intersecting other non-neighboring geodesic generators. Therefore at any point in $\mathcal{C}$, there are at least two distinct null directions pointing to the past along two distinct generators of $\mathcal{S}$. Then, as $\vec{\xi}_1$ is smooth and parallel to these generators on $\mathcal{S}$, it has to vanish at all points in $\mathcal{C} \subset \mathcal{S}$. (This is also expected because the set $\mathcal{C}$ represents an isolated set of points with a special geometric property that would be left invariant under the action of $\vec{\xi}_1$ if it were nonzero on $\mathcal{C}$.) Thus, we have a nonempty subset $\mathcal{C}$ of $\mathcal{S}$ on which the Killing field $\vec{\xi}_1$ vanishes (that is, $\mathcal{C}$ is the bifurcation set for the Killing horizon

$\mathcal{S}$).

We now note that, as before we only need to prove the theorem in the case $\{\mathcal{Q}^a\}$ is a scalar field ϕ. As each component of a multi-index field $\{\mathcal{Q}^a\}$ in the basis frame field $(\partial_t, \partial_z, \partial_x, \partial_y)$ or in the spin basis corresponding to the tetrad (3.1) (and similarly in any local basis field Lie parallel along the $\vec{\xi}_i$ or in the spin basis corresponding to any null tetrad Lie parallel along the $\vec{\xi}_i$ so that the argument we gave in the proof of theorem 1 applies without modification) behaves like a scalar field under Lie transport by $\vec{\xi}_i$, exactly the same arguments that prove the singularity of ϕ on $\mathcal{S}$ will prove the singularity of an arbitrary field $\{\mathcal{Q}^a\}$ (by constructing a suitable basis field Lie parallel along the $\vec{\xi}_i$ for each such field $\{\mathcal{Q}^a\}$) when the initial data satisfy the conditions of the theorem.

Now consider the spacelike partial Cauchy surface $\Sigma=\{t=-c\}$ in $I^-(\mathcal{S})$ where $c>0$ is sufficiently small so that Σ lies within the region of strict plane symmetry $\mathcal{W}$ [Fig. 4]. Since $\mathcal{S}$ has past endless null generators and $\mathcal{S}=H^+(\Sigma)$, Σ has no edge, i.e., it is infinite in the Killing $\vec{\xi}_i$ directions.[8] Consider generic, plane symmetric initial data for our scalar field ϕ on Σ. We will adopt the following notion of genericity:

Plane symmetric initial data for ϕ on Σ are generic, if we can find an arbitrarily large number L and coordinate values $x=a$, $x=b$ with $b-a=L$ such that if we cutoff the data for ϕ on Σ except on the portion of Σ between $x=a$ and $x=b$ (thereby breaking the plane symmetry), then the solution $\phi^{(L)}$ to the initial value problem with data $\{\phi^{(L)}=0,\ \dot{\phi}^{(L)}=0$ *on* Σ, *except on the strip between* $x=a$ *and* $x=b$ *where they are equal to the data of* $\phi\}$ will be nonzero at least on some points of the subset $\mathcal{C}$ on $\mathcal{S}$. (Note that, even though the data for $\phi^{(L)}$ on Σ are cutoff in the x-direction, they still extend infinitely far in the other Killing (y-) direction.) In the case of a multicomponent field $\{\mathcal{Q}^a\}$, plane-symmetric initial data on Σ are called generic if there is an arbitrarily large $L=b-a$ so that the solution developing from the truncation of these initial data in the manner described above takes nonzero tensor (or spinor) values at some points on the subset $\mathcal{C}$ in $\mathcal{S}$. As before, the values of the solution $\phi^{(L)}$ at the points on the subset $\mathcal{C}$ in $\mathcal{S}$ are defined as the limiting values of the solution on $I^{-}(\mathcal{S})$ as the field points approach the set $\mathcal{C}$ in $\mathcal{S}$. Again to be more precise, we will call the initial data for ϕ on Σ generic if either this limit does not exist for $\phi^{(L)}$, or it does exist and is nonzero somewhere on the subset $\mathcal{C}$ in $\mathcal{S}$. In case this limit does not exist, the solution ϕ is clearly singular (and divergent) on the horizon $\mathcal{S}$ and the theorem is proved. Therefore, in the following we will assume that this limit does exist for $\phi^{(L)}$ and takes nonzero values somewhere in the subset $\mathcal{C}$ of $\mathcal{S}$.

But now consider the action of the symmetry group generated by $\vec{\xi}_1$, given by $\mathcal{G}_L: (x,y,z,t) \rightarrow (x+L,y,z,t)$. By assumption (b) of the theorem, if we Lie transport the initial data truncated in the manner of the preceding paragraph with the Killing vector field $\vec{\xi}_1$, then the solution will be Lie transported by $\vec{\xi}_1$. But $\vec{\xi}_1$ vanishes on $\mathcal{C}$ and $\mathcal{L}_{\vec{\xi}_1}\phi=\vec{\xi}_1(\phi)$; therefore the action of $\vec{\xi}_1$ leaves the value of $\phi^{(L)}$ on $\mathcal{C}$ invariant. However, by the linearity of the field equations, the solution for the original plane symmetric initial data will be

$$\phi=\sum_{n=-\infty}^{\infty} \mathcal{G}_L{}^n\,(\phi^{(L)})\,;$$

hence on $\mathcal{C}$, since $\mathcal{G}_L\,(\phi^{(L)})\,(\mathcal{C})=\phi^{(L)}\,(\mathcal{C})$,

$$\phi(\mathcal{C})=\sum_{n=-\infty}^{\infty} \phi^{(L)}\,(\mathcal{C})=\phi^{(L)}\,(\mathcal{C})\,(\sum_{n=-\infty}^{\infty} 1\,)\,,$$

and thus ϕ diverges on $\mathcal{C}$ as $\phi^{(L)}\,(\mathcal{C})\neq 0$ by genericity; and the theorem is proved. $\square$

V. CONCLUSIONS

We have shown the instability of Killing-Cauchy horizons in plane symmetric spacetimes to arbitrary plane symmetric perturbations satisfying reasonable genericity conditions. Although it remains to be shown that our genericity criteria follow, under suitable restrictions, from the more general and standard notions of genericity employed by mathematicians[16], it seems intuitively clear to us that they agree quite naturally with a physicist's notion of genericity. Accepting this, then it is clear that if initial data whose evolution is a plane symmetric spacetime containing a Killing-Cauchy horizon are slightly perturbed in some "generic" plane symmetric direction, the horizon will be destroyed. Therefore, we conclude that the type II Killing-Cauchy horizons present in the new Chandrasekhar-Xanthopoulos solutions[3] of colliding plane wave spacetimes are probably isolated features and will not be present in a generic colliding plane wave solution. This conclusion, as was mentioned in the introduction, is in accord with the simultaneous and independent work by Chandrasekhar and Xanthopoulos showing that null dust or a fluid with pressure = (energy density), when inserted into their spacetime, destroys the horizon.

It is intriguing to note that, despite this non-generic horizon behaviour, the Chandrasekhar-Xanthopoulos solutions are more general than the previously known exact solutions for colliding plane waves with parallel polarizations - more general in the same sense as the Kerr solution is more general than the Schwarzschild solution. Nevertheless, the previously known exact solutions for colliding plane waves possess the generic plane-symmetric causal structure (no Killing-Cauchy horizons), while the Chandrasekhar-Xanthopoulos solutions do not.

It is also interesting to note that, the occurrence of timelike singularities in a plane-symmetric spacetime would imply the existence of a Killing-Cauchy horizon if in the vicinity of such a singularity at least one of the Killing vectors which generate plane symmetry becomes timelike. Even though this is the case for the presently known solutions[3] with timelike singularities, a satisfactory argument to the effect that in any plane symmetric spacetime with sufficiently "strong" timelike curvature singularities[17] at least one of the plane symmetry generating Killing vectors must be timelike near the singularity is unavailable to the author. If such an argument could be provided (possibly with some weak assumption of genericity, e.g., under the restriction that the spacetime has no Killing symmetries other than plane symmetry), then the results of the present paper would indicate that the singularities in a "generic", plane-symmetric spacetime can not be timelike (in the sense of Penrose[18]); and this would constitute an interesting verification of the Cosmic Censorship Hypothesis[18,17] in the restricted domain of plane-symmetric spacetimes. A (possibly) stronger result which would be sufficient to reach this last conclusion rigorously would be the formulation and proof of a theorem to the effect that whenever the evolution of "generic", plane symmetric Cauchy data for the gravitational and matter fields on an initial surface Σ results in the formation of a Cauchy horizon $\mathcal{S}$ for Σ, $\mathcal{S}$ is also a Killing horizon for at least one of the plane symmetry generating Killing vectors on $D^+(\Sigma)$.

ACKNOWLEDGEMENT

The author wishes to thank Kip S. Thorne for a critical reading of the first draft of this paper and for guidance in its subsequent revision. This work was supported by the National Science Foundation under Grant No. AST 85—14126.

REFERENCES FOR CHAPTER 2

[1] K. Khan and R. Penrose, Nature **229**, 185 (1971); P. Szekeres, J. Math. Phys. **13**, 286 (1972); Y. Nutku and M. Halil, Phys. Rev. Lett. **39**, 1379 (1977).

[2] R.A. Matzner and F.J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[3] S. Chandrasekhar and B.C. Xanthopoulos, Proc. R. Soc. Lond. A (1987), (to be published).

[4] U. Yurtsever, Phys. Rev. D **37**, 2803 (1988) (Chapter 3).

[5] F.J. Tipler, Phys. Rev. D **22**, 2929 (1980).

[6] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[7] For a discussion of the relevance of the Killing-Cauchy horizon $\mathcal{S}$ to the focussing effect of the plane wave on null geodesics see Ref. 6 and section II A of Ref. 4.

[8] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973). We adopt all of our terminology from this book. In addition, we will say that a future directed causal geodesic γ which intersects a subset $\mathcal{N} \subset \mathcal{M}$ is future (past) complete *in* $\mathcal{N}$, if either γ can be extended in $\mathcal{N}$ for an infinite affine parameter in the future (past) direction, or there exists an open neighborhood $\mathcal{O}$ of $\overline{\mathcal{N}}$ and affine parameter values $\lambda_1 < \lambda_2$ ($\lambda_1 > \lambda_2$) such that $\gamma(\lambda_1) \in \mathcal{N}$ and $\gamma(\lambda_2) \in (\mathcal{O} - \overline{\mathcal{N}})$.

[9] R. Price and K.S. Thorne, Phys. Rev. D **33**, 915 (1986) (Appendix A).

[10] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

[11] U. Yurtsever, (in preparation).

[12] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups* (Springer, New York, 1983).

[13] See, for example, chapters 20 and 35 of C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973); and also S. Deser, Gen. Rel. and Grav. **1**, 9 (1970).

[14] R. Penrose, *Techniques of Differential Topology in Relativity* (Society for Industrial and Applied Mathematics, Philadelphia, 1972).

[15] Y. Choquet-Bruhat, C. DeWitt-Morette and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North Holland, Amsterdam, 1982).

[16] The subset of all plane-symmetric, generic initial data according to our definitions is presumably an open dense subset with respect to any reasonable topology (e.g., the compact-open topology or the metric topology induced from a suitable norm) on the function space of all plane-symmetric initial data, although to prove this rigorously one would probably have to study the detailed properties of the evolution equations. A more elegant approach to the notion of generic subsets in infinite dimensional topological spaces (specifically in the space of all smooth vector fields on a manifold) is discussed in chapter 7 of R. Abraham and J.E. Marsden, *Foundations of Mechanics* (Benjamin/Cummings, London, 1982).

[17] F.J. Tipler, C.J.S. Clarke and G.F.R. Ellis, in *Einstein Centenary Volume*, edited by A. Held and P. Bergmann (Plenum, New York, 1979).

[18] R. Penrose, in *General Relativity: An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

FIGURE CAPTIONS FOR CHAPTER 2

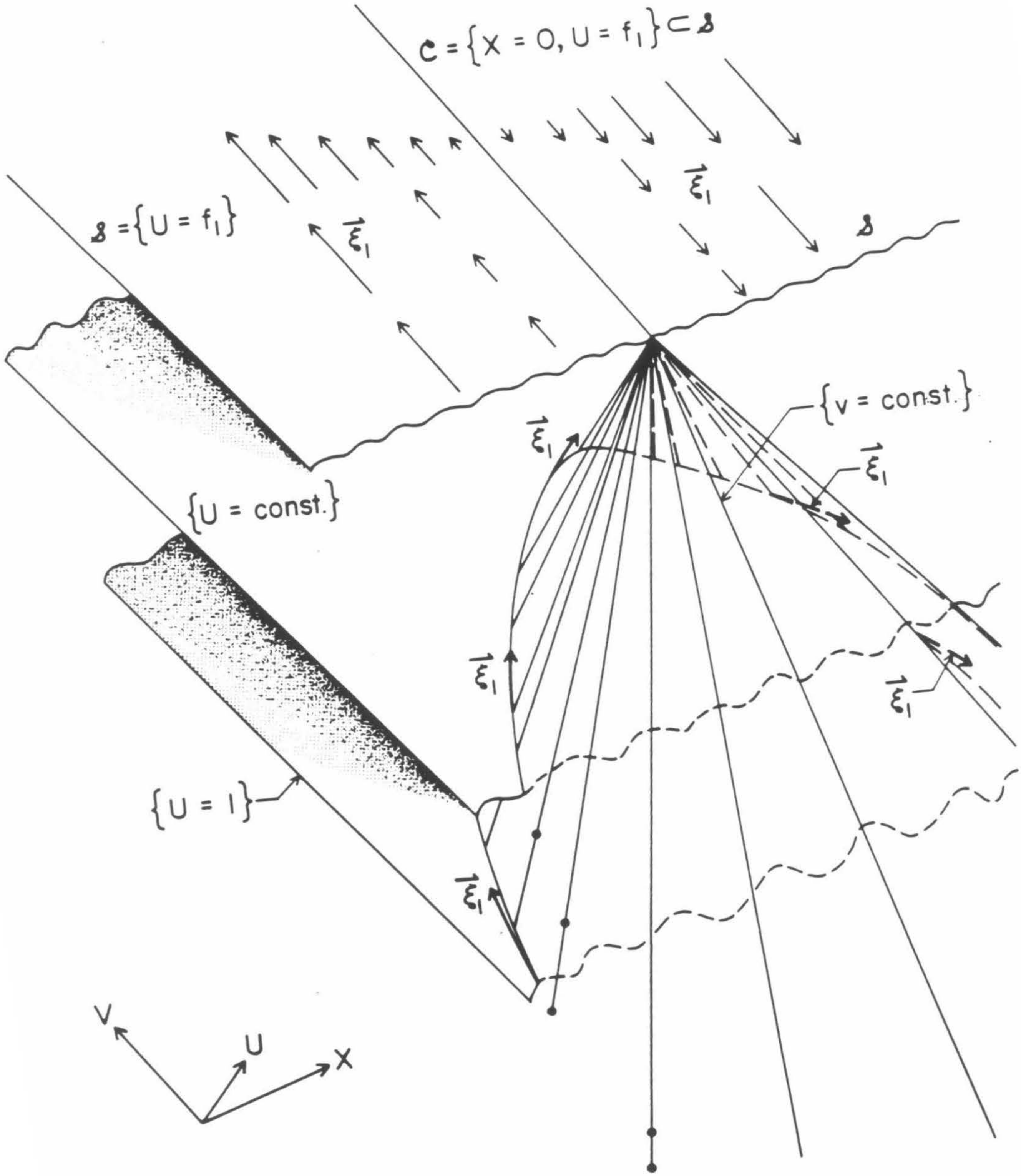
FIG. 1. The type I Killing-Cauchy horizon $\mathcal{S}$ in the Minkowskian region of the plane sandwich-wave spacetime described by Eqs.(1.1)— (1.5) with $f_1=f_2$. The Minkowskii region is given by $U \geq 1$ and the horizon $\mathcal{S}$ is located at $U=f_1=f_2$. The Y-dimension is suppressed. The Minkowskian null cone centered on the line $\mathcal{C}=\{X=0, U=f_1\}$ in $\mathcal{S}$ is (the closure of) a $\{v=\text{const}\}$ surface and has one generator in common with $\mathcal{S}$ along the line $\mathcal{C}$. The remaining generators of this cone are lines of constant v, x and y on which $u=U$ ranges from 1 to f_1. The Killing vector field $\vec{\xi}_1=\partial/\partial x$ is tangent to the intersections of $\{u=U=\text{const}\}$ surfaces with the $\{v=\text{const}\}$ cones which are obtained by rigidly translating the illustrated null cone along the line $\mathcal{C}$. On the Killing-Cauchy horizon $\mathcal{S}$, $\vec{\xi}_1$ degenerates to a null vector tangent to the null generators of $\mathcal{S}$ and vanishes on $\mathcal{C}$.

FIG. 2. The type II Killing-Cauchy horizon $\mathcal{S}$ in Minkowski space described by Eqs. (1.10)—(1.12). The Z dimension is suppressed. The horizon $\mathcal{S}$ is located at $\{t=0\}$, i.e., at $\{T=-|X|\}$ in Minkowskian coordinates. A $\{t=\text{const}<0\}$ surface lying under $\mathcal{S}$ is shown along with the orbits of the plane symmetry generating Killing vectors $\vec{\xi}_1$ and $\vec{\xi}_2$ on it. Even though it is spacelike below the horizon $\mathcal{S}=\{t=0\}$, the Killing vector $\vec{\xi}_1$ becomes null on the horizon $\mathcal{S}$ and points along its null generators, whereas the other plane symmetry generator $\vec{\xi}_2$ is everywhere spacelike. $\vec{\xi}_1$ vanishes on the line (two-plane) $\mathcal{C}$ in $\mathcal{S}$ given by $\{T=X=0\}$. The horizon $\mathcal{S}$ has a "crease" singularity on this line $\mathcal{C}$, at which the null generators of $\mathcal{S}$ have their future endpoints and onto which all lines of constant x, y and z converge as $t \rightarrow 0$.

FIG. 3. The initial value problem of theorem 1 illustrated by the example of a plane sandwich wave spacetime with a type I Killing-Cauchy horizon $\mathcal{S}$ [Eqs. (1.1)—(1.5) and Fig. 1]. As in Fig.1, the case $f_1=f_2$ is depicted with the Y-direction suppressed. The initial null boundary $\mathcal{N}$ consists of $\mathcal{N}_2$: the part of the null surface $\{u=u(p_0)=1\}$ lying above $\{v=v(p_0)\}$; and of $\mathcal{N}_1$: the piece of the null cone $\{v=v(p_0)=\text{const}\}$ lying above the surface $\{u=1\}$ with the exception of the single generator of this cone which lies in $\mathcal{S}$. $\mathcal{N}_1$ and $\mathcal{N}_2$ intersect on the spacelike two-surface Z_{p_0}. The initial data for the plane-symmetric scalar field ϕ are zero on $\mathcal{N}$ except on the dotted strip in $\mathcal{N}_2$ lying between Z_{p_0} and the line (two-surface) $v=v(p_1)$. If these data are generic, ϕ will be nonzero at some point q on the line $\mathcal{C}=\{X=0\}$ lying in the Killing-Cauchy horizon $\mathcal{S}=\{u=f_1\}$. If ϕ is smooth near $\mathcal{S}$, there will be an open neighborhood $\mathcal{A}$ in $\mathcal{S}$ around q where $\phi\neq 0$. Since the Killing vector $\vec{\xi}_1$ is null and nonzero on $\mathcal{S}$ outside the line $\mathcal{C}$, it will transport this neighborhood $\mathcal{A}$ onto an infinite strip in $\mathcal{S}$ around $\mathcal{C}$ on which $\phi\neq 0$. Sufficiently far in the past, this strip will be neighboring the single null generator of the null cone $\{v=v(p_0)\}$ along $\mathcal{C}$ that does not belong to $\mathcal{N}_1$. In that region (labelled $\mathcal{B}$ in the figure), any neighborhood of this strip in space-time will contain points that do not belong to either $J^+(\mathcal{N}_2)$ or $J^-(\mathcal{N}_2)$ and a smooth ϕ will therefore be incompatible with causal evolution.

FIG. 4. The initial value problem of theorem 2 depicted in the t—x plane with the y and z directions suppressed. The Killing-Cauchy horizon $\mathcal{S}$ on which $\vec{\xi}_1$ becomes null is given by $\{t=0\}$ and has a bifurcation singularity at $\mathcal{C}$ on which $\vec{\xi}_1$ vanishes. The spacelike initial surface $\Sigma=\{t=-c\}$ is a Killing orbit of $\vec{\xi}_1$ and sits in the open region $\mathcal{W}$ of strict plane symmetry which lies between $\mathcal{S}$ and the dashed line below Σ which also is a Killing orbit for $\vec{\xi}_1$. Plane-symmetric initial data for the linear scalar

field ϕ are posed on the initial surface Σ. When these data are generic, there will be a strip in Σ of arbitrarily large but finite extent in the x-direction which in the figure is the shaded line segment lying between the points $x=a$ and $x=b$. This strip has the property that if the initial data on Σ everywhere outside it are replaced with zero, then the solution corresponding to these truncated initial data (which, even though cutoff in the x-direction, still extend infinitely far in the other Killing y-direction) will take nonzero values somewhere on the subset $\mathcal{C}$ of the horizon $\mathcal{S}$.



$\mathcal{C} = \{X = 0, U = f_1\} \subset \mathcal{S}$
$\vec{\xi}_1$
$\mathcal{S} = \{U = f_1\}$
$\mathcal{S}$
$\{v = \text{const.}\}$
$\{U = \text{const.}\}$
$\{U = 1\}$
V
U
X

FIGURE 1

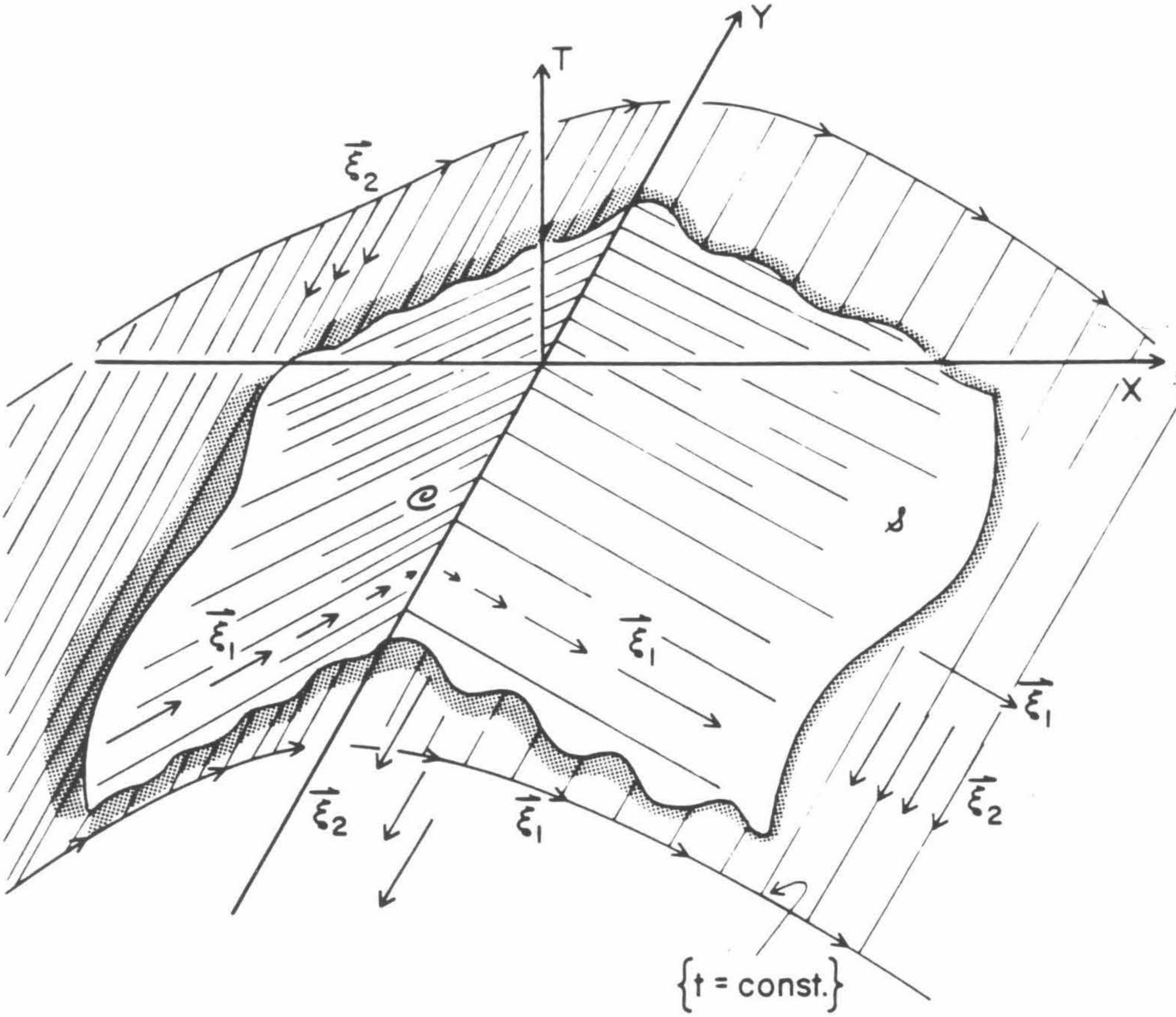


FIGURE 2

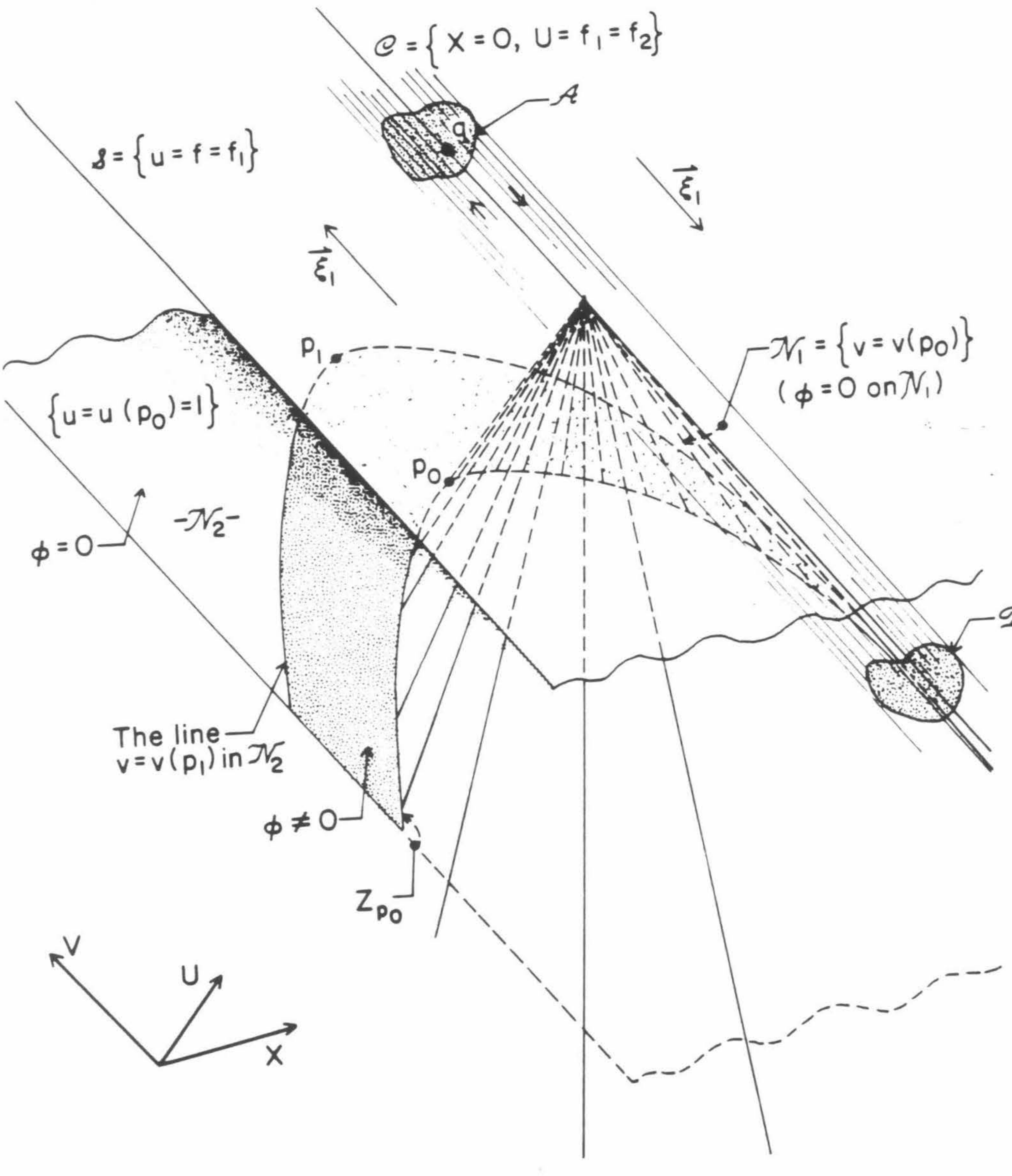
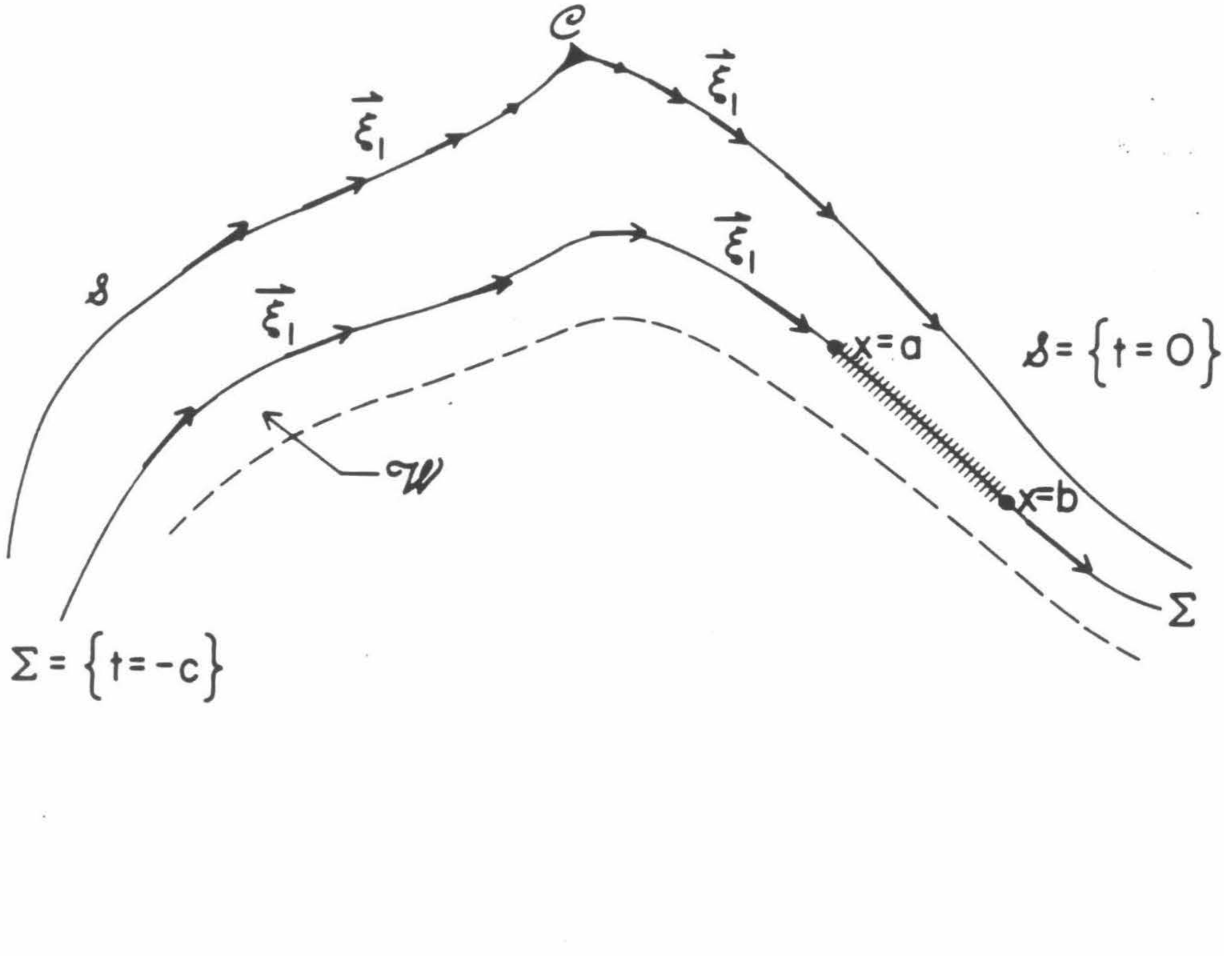

$\mathcal{C} = \{X = 0,\ U = f_1 = f_2\}$
$\mathcal{A}$
q
$\mathcal{S} = \{u = f = f_1\}$
ξ_1
ξ_1
p_1
$\mathcal{N}_1 = \{v = v(p_0)\}$
$(\phi = 0 \text{ on } \mathcal{N}_1)$
$\{u = u(p_0) = 1\}$
p_0
$\phi = 0$
$-\mathcal{N}_2-$
$\mathcal{B}$
The line
$v = v(p_1)$ in $\mathcal{N}_2$
$\phi \neq 0$
Z_{p_0}
V
U
X

FIGURE 3



$\mathcal{C}$
$\vec{\xi}_1$
$\vec{\xi}_1$
$\vec{\xi}_1$
$\vec{\xi}_1$
$\mathcal{S}$
$x=a$
$x=b$
$\mathcal{S} = \{t = 0\}$
$\mathcal{W}$
Σ
$\Sigma = \{t = -c\}$

FIGURE 4

CHAPTER 3

Colliding Almost-Plane Gravitational Waves: Colliding Plane Waves and General Properties of Almost-Plane-Wave Spacetimes

Published in Physical Review D **37**, 2803 (1988).

ABSTRACT

It is well known that when two *precisely* plane-symmetric gravitational waves propagating in an otherwise flat background collide, they focus each other so strongly as to produce a curvature singularity. This paper is the first of several devoted to *almost*-plane gravitational waves and their collisions. Such waves are more realistic than plane waves in having a finite but very large transverse size. In this paper we review some crucial features of the well-known exact solutions for colliding plane waves and we argue that one of these features, the breakdown of "local inextendibility" can be regarded as nongeneric. We then introduce a new framework for analyzing general colliding plane-wave spacetimes; we give an alternative proof of a theorem due to Tipler implying the existence of singularities in all generic colliding plane-wave solutions; and we discuss the fact that the recently constructed Chandrasekhar-Xanthopoulos colliding plane-wave solutions are not strictly plane-symmetric and thus do not satisfy the conditions and the conclusion of Tipler's theorem. Our alternative proof of Tipler's theorem emphasizes the role and the necessity of strict plane symmetry in establishing the existence of singularities in colliding plane-wave spacetimes. However, we argue on the basis of previous work that the breakdown of strict plane symmetry as exhibited in the Chandrasekhar-Xanthopoulos solutions is a nongeneric phenomenon. We then propose a definition of general *gravitational-wave spacetimes*, of which almost-plane waves are a special case; and we develop some mathematical tools for studying them. An old result of Dautcourt implies that the only gravitational-wave spacetimes with a Killing propagation direction are the plane-fronted waves with parallel rays (PP waves); and we prove a new, related result, that the only gravitational-wave spacetimes with a precisely sandwiched curvature distribution are PP waves. These properties imply that almost-plane waves

cannot propagate without diffraction, and that as opposed to the case for precisely plane waves, the curvature in an almost-plane-wave spacetime cannot be precisely sandwiched between two null surfaces (i.e., the wave must have tails). We also prove a "peeling theorem" for components of the Weyl curvature in general gravitational-wave spacetimes.

I. INTRODUCTION AND OVERVIEW

This is the first of a series of papers describing work aimed at understanding the nonlinear interaction of colliding gravitational waves in general relativity. It has been known since the early 1970s, from work on exact solutions of the Einstein field equations, that when two gravitational plane waves propagating in an otherwise flat spacetime collide, they interact so strongly as to eventually cause a curvature singularity to develop in the future of the collision plane. It is natural to ask whether this singularity is an artifact of the unphysical idealization that the waves are precisely planar and thus extend infinitely far in the "transverse" directions, or whether a singularity would still be produced if the waves were transversely finite but had arbitrarily large "size"— i.e., if they were "almost-plane waves." And if there is a regime in which spacetime singularities are guaranteed to be produced as a result of almost-plane-wave collisions, what are the conditions on the colliding almost-plane waves which characterize this regime? This paper is the first in a series whose ultimate goal is to answer these and related questions.

This first paper in the series lays foundations for the subsequent papers by reviewing (briefly) old results and presenting some new ones on colliding exact plane-wave spacetimes, and by introducing the concept of a gravitational wave (GW) spacetime—of which almost-plane waves are a special case—and proving some theorems about GW spacetimes which imply several important properties of almost-plane waves. More specifically, we note the following.

In Sec. II we briefly review some global properties of the exact solutions for the so-called plane fronted waves with parallel rays ("PP waves"), and for plane waves. The principal purpose of this section is to introduce the reader to our terminology and viewpoint on issues that will be central to the rest of this paper and to future papers in the series.

In Sec. III we turn attention to colliding exact plane waves. We begin, in Sec. III A, by briefly reviewing the properties of some exact solutions to Einstein's equations representing plane-wave collisions, and we discuss a peculiar property of these colliding plane-wave spacetimes: the fact that, even after one has maximally extended them in a global sense, they are not "locally inextendible" if one uses the standard notion of local inextendibility (Sec. 3.1 of Ref. 1). We elucidate this peculiarity by introducing a new notion of "generic local inextendibility," which these spacetimes satisfy. In Sec. III B we give an alternative proof of a theorem due to Tipler[2] which predicts that collisions of exact plane waves must produce singularities. Our alternative proof of Tipler's theorem emphasizes the role and the necessity of strict plane symmetry (a concept we shall define with care) in establishing the existence of singularities in colliding plane-wave solutions and in more general plane-symmetric spacetimes. The importance of strict plane symmetry becomes clear when, following Chandrasekhar and Xanthopoulos,[3] one notices that in contrast with the usual case where they produce spacelike spacetime singularities, some colliding plane waves can produce Killing-Cauchy horizons on which strict plane symmetry breaks down and thereby can avoid the conclusion of Tipler's theorem. (In a previous paper[4] we have shown that Killing-Cauchy horizons in plane-symmetric spacetimes are unstable against plane-symmetric perturbations, and thence that any generic colliding plane-wave solution will be devoid of such horizons. In accordance with this result but independently of it, Chandrasekhar and Xanthopoulos[5] have recently discovered that the Killing-Cauchy horizons in their colliding gravitational plane-wave spacetimes are destroyed and are replaced by spacelike singularities, when the colliding plane waves are coupled with plane symmetric null fluids propagating along with the waves. Thus, the assumption of strict plane symmetry required in the proof of Tipler's theorem is probably satisfied by all but a set of measure zero of colliding plane-wave spacetimes).

In Sec. IV we introduce the concept of a "gravitational-wave (GW) spacetime," and we use the Newman-Penrose[6] and the characteristic initial value formalisms[7–9] to prove several theorems about GW spacetimes. These theorems have important implications for almost-plane waves (which are special cases of GW spacetimes).

Section IV A is devoted to a careful definition of GW spacetimes and associated discussion. Roughly speaking a GW spacetime is a solution to the vacuum Einstein field equations which is flat prior to the arrival of a curvature disturbance (gravitational wave), but may or may not settle back down into flatness afterward. This section also introduces a class of "standard" coordinate systems and "standard" null tetrads to be used in studying GW spacetimes.

In Sec. IV B we discuss a previous theorem of Dautcourt[10] which directly implies that any GW spacetime possessing a null Killing vector field pointing along the propagation direction–i.e., a spacetime which represents a gravitational wave propagating in a perfectly diffraction-free manner, with no change in its wave form—must be a PP-wave spacetime. Since PP waves are always infinitely large in transverse extent, this result implies that almost-plane waves (which have finite transverse "size") must always exhibit diffraction.

In Sec. IV C we present a "peeling-off"-type theorem about the behavior of the Weyl curvature quantities associated to a standard tetrad on a general GW spacetime. A discussion of this theorem is given preceding its proof.

In Sec. IV D we introduce the characteristic initial-value formalism of Penrose,[7] Muller zum Hagen and Seifert,[8] and Friedrich[9] which we will need to prove the theorem of Sec. IV D. We give a brief review of this formalism in a form that is appropriate to our conventions and notation and we emphasize those aspects relevant to our purposes.

Section IV E is devoted to another theorem about GW spacetimes: a proof that any GW spacetime that not only begins flat before the wave arrives but also returns to perfect flatness after the wave passes (i.e., any precisely "sandwich" GW spacetime), must actually be a PP wave spacetime. Since all PP waves are infinitely large in transverse extent, this theorem implies that almost-plane waves must always leave "tails" behind, in any region of space through which they have propagated.

In Sec. V we briefly recapitulate the principal conclusions of this paper.

Throughout this paper (with the exception of Sec. III B) we will deal with purely gravitational (vacuum) waves; for Einstein-Maxwell plane waves and for plane waves coupled with fluid motions, results similar to those of Secs. II—III hold with appropriate modifications.[5]

Throughout this paper we use, without explanation, terminology and concepts from Hawking and Ellis.[1] Our mathematical conventions and notation are those of Hawking and Ellis,[1] and Misner, Thorne, and Wheeler.[11] In particular we adopt the metric signature (–,+,+,+) and use the "rationalized" Newman-Penrose equations appropriate to this signature. These equations are listed in the Appendix. Our terminology and general usage of the Newman-Penrose formalism are in accordance with those of Chandrasekhar[12] after the proper conversion from his (+,–,–,–) metric signature to ours.

II. EXACT PLANE WAVE AND PP-WAVE SPACETIMES: A REVIEW INTRODUCING OUR TERMINOLOGY AND NOTATION

A plane fronted (PP) wave with parallel rays[13] $(\mathcal{M},g)$ is a spacetime where one can introduce a global coordinate chart (U,V,X,Y): $\mathcal{M}\rightarrow R^4$ in which the metric takes the form

$$g=dX^2+dY^2+h(U,X,Y)dU^2-dU\,dV\ , \tag{2.1}$$

where $h(U,X,Y)$ is C^2 and satisfies

$$\frac{\partial^2 h}{\partial X^2}+\frac{\partial^2 h}{\partial Y^2}=0\ . \tag{2.2}$$

In such a spacetime, $\partial/\partial V$ is parallel [i.e., $\nabla(\partial/\partial V)\equiv 0$] and is in general the only Killing vector field on $(\mathcal{M},g)$. The special case

$$h(U,X_i)=h_{ij}(U)X_iX_j \qquad (i,j=1,2)\ , \tag{2.3}$$

where $h_{ij}(U)$ is a symmetric matrix with $h_{ii}(U)=0$, defines the plane-wave[13] spacetimes with their five-dimensional group of isometries.

Note that, when $h(U,X,Y)=0$, except for $0<U<a$, the PP-wave spacetime represents an exact "sandwich wave" for which spacetime is flat for $U\leq 0$ and $U\geq a$. Note also that whatever may be h, the propagation direction $\partial/\partial V$ is Killing, so the PP wave propagates without diffraction. The PP waves must be of infinite extent in the spacelike X,Y directions because of Eq. (2.2), even though they are not in general plane symmetric. In fact, we will show in Sec. IV that neither of the above solitonlike properties of PP waves (flatness after the passage of the wave, and diffraction-free propagation) can hold true for almost-plane gravitational waves of finite transverse extent.

For a plane polarized plane wave in the "Kerr-Schild"-type chart[13] (U,V,X,Y), the function h takes the form

$$h(U,X_i)=h(U)(X^2-Y^2)\,.$$

When $h(U)=0$ for $U \geq a$ and for $U \leq 0$ (i.e., for a sandwich plane wave), it is also useful to introduce the "Rosen-type" chart[13] (u,v,x,y), which is defined on the open domain $\{F(U)G(U)\neq 0\}$ of $\mathcal{M}$ by

$$X=xF(u)\,,\quad Y=yG(u)\,,\quad U=u\,,$$

$$V=v+x^2FF'+y^2GG'\,, \tag{2.4}$$

where F and G are the unique C^4 solutions to the equations

$$\frac{F''}{F}=h\,,\quad \frac{G''}{G}=-h\,, \tag{2.5}$$

with initial conditions $F(0)=G(0)=1$, $F'(0)=G'(0)=0$, and $F(U)=G(U)=1$ for $U \leq 0$. In this local coordinate system the metric is

$$g=F^2(u)dx^2+G^2(u)dy^2-du\,dv\,, \tag{2.6}$$

and the plane symmetry generators are given by $\vec{\xi}_i=\partial/\partial x^i$ on the domain of the (u,v,x,y) chart, with $(i=1,2)$ and $x^1=x$, $x^2=y$.

The field equations (2.5) imply that, for a sandwich wave (2.6), in the domain $\{u \geq a\}$ where $h=0$ and where the spacetime is flat,

$$F(u)=\frac{F(a)}{a-f_1}(u-f_1)\,,$$

$$G(u)=\frac{G(a)}{a-f_2}(u-f_2)\,, \tag{2.7}$$

where, because of the field equations (2.5), $f_1 > a$, and $f_2 \in [-\infty, a] \cup [f_1, +\infty]$. These metric functions display, as we shall see, the focusing effect of the plane sandwich wave (2.6) on the null geodesics propagating in the u direction. We call the case $f_1 = f_2$ the anastigmatic case and the generic case $f_1 \neq f_2$ the astigmatic case. We also denote the null surfaces (wave fronts) $\{u=0\}$ and $\{u=a\}$ by $\mathcal{N}$ and $\mathcal{N}'$, respectively.

To see the focusing effect of the plane wave on null geodesics (discussed in greater detail, e.g., in Ref. 13), consider, for an arbitrary value of v_0, the null surface $\{v=v_0\}$ generated by null geodesics on which u is an affine parameter and along which x, y, and v are constant. In the Minkowskian region $I^-(\mathcal{N})$, these null geodesics generate a standard, flat, Minkowskian null surface; namely they generate the null surface $\{v=V=v_0\}$. On the other hand, assuming for simplicity that the plane wave is anastigmatic and using Eqs. (2.4) and (2.7), it is easily seen that in the other Minkowskian region $I^+(\mathcal{N}')$ lying to the future of the wave, the null surface $\{v=v_0\}$ is a Minkowskian null cone C_Q centered at the point Q which in the (U,V,X,Y) coordinates is given by $Q=(V=v_0, U=f_1, X=Y=0)$. In other words, after they pass through the spacetime curvature sandwiched between the wave fronts $\mathcal{N}$ and $\mathcal{N}'$ of the plane wave, the initially parallel (shear-free and convergence-free) null geodesics generating the surface $\{v=v_0\}$ are focused along the null generators of the Minkowskian null cone C_Q, to the point Q in $I^+(\mathcal{N}')$. Moreover, it is easy to see that the null generators of the surface $\{v=v_0\}$ constitute one-half of the null generators of the achronal boundary[1] $\dot{J}^+(Z_p)$ which have their past end points on Z_p. Here p is any point in $I^-(\mathcal{N})$ with $v(p)=v_0$, and Z_p is the spacelike two-surface generated by p under the action of plane symmetry. The single null generator of the null cone C_Q which runs parallel to (and thus does not intersect) the plane wave is the single past endless generator of

$\dot{J}^+(Z_p)$. Similarly, in the general *astigmatic* case, one-half of the generators of $\dot{J}^+(Z_p)$ which have their past endpoints on Z_p generate the null surface $\{v=v_0\}$, and after passing through the plane sandwich wave these generators are focused onto a spacelike curve lying in the null plane $\{U=f_1\}$. This spacelike curve is given by $\{U=f_1, X=0, V=v_0+Y^2/(f_1-f_2)\}$ in the (U,V,X,Y) coordinate system. Along the null plane $\{U=f_1\}$, which we will henceforth denote by $\mathcal{S}$, there is a one-parameter family of null generators of $\dot{J}^+(Z_p)$ which do not have past end points and which all run parallel to the plane wave.

Similar conclusions apply for the null generators of the achronal boundaries $\dot{J}^+(p)$ where $p \in I^-(\mathcal{N})$ is a point sufficiently far away from the wave (before the wave's arrival). However, in this case the null generators are focused to a point (or a spacelike curve) lying beyond the surface $\mathcal{S}$, i.e., at $U>f_1$. [13]

The plane symmetry generated by the Killing vectors $\vec{\xi}_i$ breaks down on the null surface $\mathcal{S}$; that is, in the tangent space at any point on this surface $\mathcal{S}$, the Killing vectors $\vec{\xi}_i$ generate a subspace which is *not* a two-dimensional spacelike plane (see Sec. III B of this paper). This breakdown of "strict" plane symmetry on $\mathcal{S}$ (Sec. III B) allows the null generators of the achronal boundary $\dot{J}^+(Z_p)$ to intersect each other at points in $\mathcal{S}$. In the *anastigmatic* case, the $\vec{\xi}_i$ degenerate on $\mathcal{S}$ to null Killing vectors that are proportional to $\partial/\partial V$ and that vanish on the line $X=Y=0$; hence, in this case the $\vec{\xi}_i$ span a one-dimensional null line at each point on $\mathcal{S}-\{X=Y=0\}$. In the *astigmatic case*, $\vec{\xi}_1$ degenerates on $\mathcal{S}$ to a null Killing vector that is proportional to $\partial/\partial V$ and that vanishes along the two-surface $X=0$, while $\vec{\xi}_2$ is still spacelike on $\mathcal{S}$, generating symmetries along the spacelike line to which null generators of $\dot{J}^+(Z_p)$ are focused. In this case, the $\vec{\xi}_i$ span a two-dimensional *null* plane at each point on $\mathcal{S}-\{X=0\}$.

A further consequence of these focusing properties is the fact that plane-wave spacetimes are not globally hyperbolic,[13] even though they are geodesically complete and satisfy stable causality. Any partial Cauchy surface Σ in $I^-(\mathcal{S})$ cannot intersect $\mathcal{S}$; hence $\mathcal{S}$ is a future Cauchy horizon for Σ. This was to be expected, since strict plane symmetry on Σ will be preserved throughout the domain of dependence of Σ, which therefore cannot include $\mathcal{S}$. [The past Cauchy horizon for a partial Cauchy surface that intersects the wave will be the time reversed analogue of $\mathcal{S}$ lying in $I^-(\mathcal{N})$.]

III. COLLIDING EXACT PLANE WAVES

A. Review of exact solutions and a new viewpoint on the breakdown of local inextendibility

The first results on exact solutions of the vacuum Einstein equations representing colliding plane impulsive gravitational waves with parallel, linear polarizations were obtained by Khan and Penrose,[14] and Szekeres.[15] Later Nutku and Halil[16] obtained the generalization of these solutions to arbitrary relative linear polarizations for the two incoming impulsive waves. Szekeres,[17] and Chandrasekhar and Xanthopoulos[17] have obtained similar results for Einstein-Maxwell waves. The main result of these exact solutions (see Fig. 1) is that the future of the collision surface (region IV) is bounded by a curvature singularity in future directions. Surprisingly, the singularity extends over to the past endless null generators of the surfaces $\mathcal{N}_1$ and $\mathcal{N}_2$ which would lie in the respective Cauchy horizons of the single plane-wave spacetimes if the collision were not taking place.[18,14] A curious result, therefore, is that the spacetime pictured in Fig. 1 is maximally extended, i.e., the points of $\mathcal{S}_2$ and $\mathcal{S}_1$ touching the surfaces $\mathcal{N}_1$ and $\mathcal{N}_2$ cannot be added to the spacetime even though these points do not represent local curvature singularities–i.e., even though there are timelike curves running into and terminating on these singularities which are completely contained within the flat regions II and III. In other words the maximally extended colliding plane-wave spacetime is not C^k-locally inextendible although it is globally C^k-inextendible (maximal) for any $k \geq 1$ (for definitions, see Sec. 3.1 of Ref. 1).

The first new result of this paper is to point out that the above failure of local inextendibility is not generic in the following sense.

We define a spacetime $(\mathcal{M},g)$ to be *generically* C^k-locally inextendible if there exists *no* open set $\mathcal{U}$ in $\mathcal{M}$ with the following properties.

(i) $\mathcal{U}$ has noncompact closure in $\mathcal{M}$.

(ii) $(\mathcal{U}, g \mid_{\mathcal{U}})$ has a C^k-extension $(\mathcal{U}',g')$ in which the image of $\mathcal{U}$ has compact closure.

(iii) There is a point $q \in \mathcal{U}$ and an open subset $\mathcal{O}$ of $T_q\mathcal{M}$ consisting of timelike vectors, such that for any vector $X \in \mathcal{O}$ the geodesic $\{\gamma_X(t)=\exp_q(tX), 0 \leq t < t_1\}$ from q is inextendible in $\mathcal{M}$, and t_1 is finite. Here t_1 is the smallest upper bound of all $t' \in R$ such that $\gamma_X(t)=\exp_q(tX)$ is defined and contained in $\mathcal{U}$ for all $t \in [0,t')$.

With this definition, the maximally extended colliding plane-wave spacetime of Fig. 1 is not only globally inextendible; it is also generically C^k-locally inextendible for any $k \geq 1$.

Note that the key idea behind the usual definition of local inextendibility[1] is this: One identifies as locally *extendible*, among others, those (maximal) spacetimes that possess purely topological singularities, i.e., singularities which do not involve unbounded curvature but which nevertheless cannot be removed without destroying the topological manifold structure of the spacetime. Such a singularity may be modeling the global influence of some essentially nonsingular matter distribution (e.g., a cosmic string), or may be the result of some unusual "cutting and pasting" employed in the construction of the spacetime (such as those that appear in the covering space of the two-dimensional Minkowski space with the origin removed). In all such cases our definition would identify these spacetimes as generically locally inextendible; i.e., the presence of their "noncurvature" singularities will be *directly* felt only by those freely falling observers which constitute a set of measure zero. Only the spacetimes in which such topological singularities are unavoidable for a "finite fraction" of all freely falling observers (e.g., because of the focusing of causal geodesics onto these singularities) will fail to satisfy generic local inextendibility. For example, if the topological singularities of the colliding plane-wave spacetimes which we have described above were to lie *beyond* the respective Cauchy horizons of the colliding waves instead of lying *on* them, then these spacetimes would fail to be generically locally inextendible. However, as our discussion of Tipler's theorem in the next section will make clear, this is not a possible outcome of generic plane-wave collisions. Thus, except possibly for a set of measure zero, all colliding plane-wave spacetimes will satisfy generic local inextendibility.

Another important property of the above examples of colliding plane-wave solutions is that they are globally hyperbolic, since neither of the Cauchy horizons $\mathcal{S}_1$, $\mathcal{S}_2$ is contained in $\mathcal{M}$. In particular, the singularities present in these spacetimes are "not timelike" in the sense of Penrose;[19] that is, the singular points are part of an achronal future c boundary for $(\mathcal{M}, g)$.

We should also remark that recently Chandrasekhar and Xanthopoulos[3] have obtained new exact solutions describing colliding gravitational impulsive-shock waves with nonparallel polarizations, in which the interaction region is bounded by a

Killing-Cauchy horizon instead of by a spacelike singularity, and in which a timelike singularity appears when the solution is analytically extended beyond this horizon. However, as we will also discuss in the next section, it is shown in Ref. 4 that such Killing-Cauchy horizons in any plane-symmetric spacetime are unstable against purely plane-symmetric perturbations. Therefore, it is reasonable to expect that the spacetimes resulting from "generic" plane-wave collisions will always involve spacelike curvature singularities with the same global structure as the solutions we have discussed above, regardless of the relative configuration of the incoming polarizations and wave forms.

B. A general framework for studying colliding plane-wave spacetimes and an alternative proof of Tipler's theorem on their singularities

The global structure of plane-symmetric spacetimes (e.g., plane waves and colliding plane waves) is nontrivial when they possess Killing-Cauchy horizons on which their plane symmetry breaks down. When discussing such spacetimes from a general standpoint some care is needed. In this section we introduce a brief framework for analyzing general plane-wave and colliding plane-wave spacetimes. This framework is based on some intuitively plausible definitions and constructions which make precise the basic notions that one needs in such a general discussion. We conclude the section with an important application of this framework: a discussion and an alternative proof of Tipler's Theorem[12] on singularities of colliding plane-wave spacetimes.

We will call a maximal (see Sec. 3.1 of Ref. 1) spacetime $(\mathcal{M},g)$ *plane symmetric* if there exists a pair of commuting Killing vectors $\vec{\xi}_1,\vec{\xi}_2$ on $\mathcal{M}$, and an open dense subset of $\mathcal{M}$ at every point of which $\vec{\xi}_1,\vec{\xi}_2$ span a spacelike two-dimensional subspace in the tangent space. So as to exclude cylindrical symmetry, we assume that the orbits of $\vec{\xi}_i$ $(i=1, 2)$ are homeomorphic to R^1. If the open dense subset is all of $\mathcal{M}$, i.e., if no breakdowns of plane symmetry occur, then we say $(\mathcal{M},g)$ is *strictly plane symmetric*.

In the strictly plane-symmetric region of any plane-symmetric spacetime there exist *standard null tetrads* constructed as follows: Since $\vec{\xi}_i$ are Killing and span a spacelike two-plane at each point, there exist precisely two null geodesic congruences everywhere orthogonal to $\vec{\xi}_i$. Let $\vec{l},\vec{n}$ be tangent vector fields to these congruences normalized so that $g(\vec{l},\vec{n})=-1$. Let $\vec{m}$, $\vec{m}^*$ be two linearly independent complex null linear combinations of $\vec{\xi}_i$, which are complex conjugate, satisfy $g(\vec{m},\vec{m}^*)=1$, and vary smoothly over the region of strict plane symmetry. Then $(\vec{l},\vec{n},\vec{m},\vec{m}^*)$ is a null tetrad which is locally regular although it will not in general cover all of $\mathcal{M}$. We will call the tetrad $(\vec{l},\vec{n},\vec{m},\vec{m}^*)$, together with the additional requirement that $\vec{l},\vec{n}$ are Lie parallel along $\vec{\xi}_i$ (which we can obviously impose since $\vec{\xi}_i$ are Killing and commute), a *standard tetrad*.

We will say that a plane-symmetric nonflat spacetime is a *plane wave* if in a standard tetrad either $\Psi_0=\Psi_2\equiv 0$ or $\Psi_4=\Psi_2\equiv 0$. Note that this property is independent of the choice of the standard tetrad which is unique up to tetrad rotations of type III (Sec. 7.3 of Ref. 12).

We have seen in the last section that single plane sandwich wave solutions are not strictly plane symmetric, as focusing causes the breakdown of plane symmetry along a null hypersurface $\mathcal{S}$ in $I^+(\mathcal{N}')$, where $\mathcal{N}$, $\mathcal{N}'$ are the past and future wave fronts. Now consider a spacetime representing the collision between two plane waves moving in opposite directions. A plane-symmetric spacetime $(\mathcal{M},g)$ will be said to model *colliding plane waves* if there exist two null surfaces $\mathcal{N}_1$, $\mathcal{N}_2$ in $\mathcal{M}$, intersecting in a spacelike two-surface Z, such that in any standard tetrad $\Psi_4=\Psi_2\equiv 0$ but $\Psi_0\neq 0$

on $I^{-}(\mathcal{N}_1)$, $\Psi_0=\Psi_2\equiv0$ but $\Psi_4\neq0$ on $I^{-}(\mathcal{N}_2)$, and $\Psi_0,\Psi_4\neq0$ on $I^{+}(Z)$. Figure 2 depicts such a spacetime.

In the specific colliding plane-wave solutions reviewed above,[15,16,18] the collision produces a spacetime singularity. That this is a rather general outcome of plane-wave collisions is shown by a theorem of Tipler.[2] However, a key requirement for the proof of Tipler's theorem is that strict plane symmetry holds *throughout* the colliding plane-wave spacetime. Since this notion of strict plane symmetry is crucial to the discussion that we will give in the next few paragraphs of this section, we first present a restatement and an alternative proof of Tipler's theorem which emphasize the requirement of strict plane symmetry explicitly.

Theorem 1 ("Tipler's theorem"[2]). Let $(\mathcal{M},g)$ be a strictly plane-symmetric spacetime with a C^2 metric g, with the following properties.

(i) Null convergence[1] holds on $\mathcal{M}$: $R_{ab}K^aK^b\geq0$ for any null vector K.

(ii) There exists a point p at which either at least one of $(\Psi_0,\sigma,\Phi_{00})$ is nonzero or at least one of $(\Psi_4,\lambda,\Phi_{22})$ is nonzero, in some standard tetrad on $\mathcal{M}$.

(iii) Through the above point $p\in\mathcal{M}$, there exists a partial Cauchy surface Σ which intersects each null geodesic generator of $\dot{J}^{\pm}(p)$ and which is noncompact in the spacelike direction orthogonal to $\vec{\xi}_i$.

Then, $(\mathcal{M},g)$ is *not* null geodesically complete.

Proof. Fix the standard tetrad mentioned in property (ii). Then since $\vec{l},\vec{n}$ are geodesic and hypersurface orthogonal, we can arrange that the following Newman-Penrose spin coefficients vanish (cf. the Appendix)

$$\kappa=\nu=\varepsilon+\varepsilon^*=\rho-\rho^*=\mu-\mu^*=0\ .$$

Now assume, without loss of generality, that it is one of $(\Psi_0, \sigma, \Phi_{00})$ that is nonzero at p (otherwise interchange the role of $\vec{l}$ and $\vec{n}$ in the argument, and accordingly interchange the roles of the spin quantities).

Let Z_p denote the orbit of p under the action of the Killing symmetry group generated by $\vec{\xi}_1, \vec{\xi}_2$. Then by plane symmetry, properties (ii) and (iii) hold at every point $q \in Z_p$.

Now assume, in contradiction of the theorem's conclusion, that $(\mathcal{M}, g)$ is null geodesically complete. Consider $\dot{J}^{\pm}(Z_p)$. These achronal boundaries are generated by null geodesics which by (iii) and because of strict plane symmetry all have their past (future) end points on Z_p, and which are everywhere orthogonal to Z_p and hence (since $\vec{\xi}_i$ are Killing) to $\vec{\xi}_1, \vec{\xi}_2$. Thus $\dot{J}^{\pm}(Z_p)$ are generated by integral curves of $\vec{l}$ and $\vec{n}$ that start off from Z_p.

It is shown by Tipler in Ref. 2 that as a result of the assumptions (i) and (ii) above [and of the Ricci identities (A5) and (A6)], any null geodesic γ_q parallel to $\vec{l}$ and passing through any point q in Z_p will have a conjugate point to Z_p along itself at some affine distance f from q. If we now fix our time orientation so that the conjugate point lies to the future of q, then every null generator of $\dot{J}^{+}(Z_p)$ parallel to $\vec{l}$ has a conjugate point to Z_p along itself at an affine distance $u = f > 0$; and f is independent of the null generator. The noncompactness of the partial Cauchy surface in property (iii) guarantees that the null geodesic generators of $\dot{J}^{+}(Z_p)$ parallel to $\vec{l}$ cannot intersect (except on Z_p) those parallel to $\vec{n}$, and consequently since $\dot{J}^{+}(Z_p)$ has no boundary (proposition 6.3.1 of Ref. 1), the submanifold $\dot{J}_l^{+}(Z_p) - Z_p$ of $\dot{J}^{+}(Z_p)$ generated by null geodesics parallel to $\vec{l}$ has no boundary.

We construct the map

$$\phi : Z_p \times (0,f\,] \rightarrow \dot{J}_l^{\,+}(Z_p) - Z_p \quad ,$$

given by

$$\phi : (q\,,u\,) \rightarrow \gamma_q(u\,) \in \dot{J}_l^{\,+}(Z_p) - Z_p \quad .$$

Claim: ϕ is a diffeomorphism.

That ϕ is onto is obvious since all points on $\dot{J}_l^{\,+}(Z_p)$–Z_p are on null geodesics γ from Z_p and for $u > f$ $\gamma_q(u\,)$ does not belong to $\dot{J}_l^{\,+}(Z_p)$–Z_p for any $q \in Z_p$ (see Chap. 4 of Ref. 1). That ϕ is one to one is an immediate consequence of the strict plane symmetry holding at every point of $\mathcal{M}$, which strict plane symmetry prevents different null generators γ_q and $\gamma_{q'}$ $(q \neq q\,')$ from intersecting each other. That ϕ and ϕ^{-1} are smooth is clear by construction.

Thus,

$$[\dot{J}_l^{\,+}(Z_p) - Z_p\,] \cong Z_p \times (0,f\,] \cong R^2 \times (0,f\,]\,.$$

Here, the symbol $\cong$ denotes "is diffeomorphic to." But $R^2 \times (0,f\,]$ has a boundary which is diffeomorphic to R^2, and therefore we obtain a contradiction to the proposition 6.3.1 of Ref. 1.

Therefore, the assumption that $(\mathcal{M},g\,)$ is null geodesically complete must be false–a conclusion that proves the theorem. □

Tipler's theorem implies, as a specific application, that if the spacetime produced by the collision of two plane waves is strictly plane symmetric–as is the case in the classic examples (Refs. 14—16), then the collision must produce a singularity (null geodesic incompleteness). We have argued at length in Sec. III of Ref. 4 that in a spacetime which represents the collision between an exact plane gravitational wave and a plane wave of *any* physical field belonging to some restricted class, the breakdown of strict plane symmetry is incompatible with global causality. Therefore, strict plane symmetry is a *natural* restriction to impose on colliding plane-wave spacetimes. However, the fully nonlinear gravitational field does *not* belong to the class of fields for which the arguments of Ref. 4 are valid; consequently these arguments do not *prove* that colliding plane-wave spacetimes are strictly plane symmetric. In fact, just as spacetimes containing a single plane wave fail (beyond the Cauchy horizon $\mathcal{S}$) to be strictly plane symmetric, so also *some* colliding plane-wave spacetimes possess (Killing-)Cauchy horizons at which strict plane symmetry breaks down. Examples are the Chandrasekhar-Xanthopoulos[3] solutions. Tipler's theorem cannot be applied to such spacetimes.

On the other hand, as is suggested by calculations of Chandrasekhar and Xanthopoulos (Ref. 5) and proved by the author (Ref. 4), all such Killing-Cauchy horizons which break strict plane symmetry are unstable against plane-symmetric perturbations. Morover, as was shown by Chandrasekhar and Xanthopoulos[5] for special cases, it is plausible (though not yet proved in general) that the (full nonlinear) growth of these instabilities always destroys the Killing-Cauchy horizon, thereby making the spacetime strictly plane symmetric. If this is the case, then *all colliding plane-wave spacetimes whose causal structures are stable against plane-symmetric perturbations are strictly plane symmetric, and Tipler's theorem implies that they also are all singular.*

It is interesting in revealing the depth of Tipler's theorem to note that for a single plane-wave spacetime the only conditions of the theorem that do not hold are strict plane symmetry and the existence of the partial Cauchy surface satisfying the requirements in (iii). As we argued in Sec. II these conditions are intimately related and presumably imply each other in the generic case.[4] The partial Cauchy surface condition is used to guarantee that all generators of $\dot{J}^+(Z_p)$ have past end points on Z_p; whereas the strict plane symmetry is used to show that the map ϕ is a diffeomorphism, which is the vital step in our proof of Tipler's theorem. In fact, we will use this aspect of Tipler's theorem in a future paper to produce a qualitative argument for the existence of singularities in colliding almost-plane-wave spacetimes when the relevant parameters belong to a certain regime.

Also note that (Fig. 2) Tipler's theorem [simply by taking the point p as an arbitrary point in $I^+(\mathcal{N}_2) \cap I^-(\mathcal{N}_2')$] implies that the points on the past endless generators of $\dot{J}^+(Z_p)$ which would lie in the Cauchy horizon $\mathcal{S}$ will become singular, and consequently $\mathcal{S}$ will be cut off completely from the colliding plane-wave spacetime, a result that is not obvious from the analytical structure of the known exact solutions.[18]

IV. GRAVITATIONAL-WAVE (GW) SPACETIMES

In this section we turn attention to general solutions to the vacuum Einstein equations which represent a single "gravitational wave" propagating in an otherwise flat space. Plane-wave and PP-wave spacetimes are simple examples of such solutions; and we frequently will refer to them for comparison and motivation while discussing more general gravitational-wave (GW) spacetimes. Our primary interest in studying GW spacetimes is to learn about "almost-plane waves"–GW spacetimes that in some suitable sense are of "finite spatial extent," representing a transversly bounded curvature disturbance carrying finite "energy" and propagating in an otherwise flat spacetime. (We will define almost-plane waves more precisely in paper 2 of this series.) Clearly almost-plane waves cannot be plane symmetric, since they have an amplitude that must satisfy suitable falloff conditions at large "transverse" distances.[20] We will see in Sec. IV B, by a theorem of Dautcourt,[10] that relaxing the assumption of plane symmetry on a such a spacetime forces it to have no Killing vectors in general and hence leaves little hope for an exact solution. Indeed, one can already guess that for a nonplanar gravitational wave the linear and nonlinear effects of diffraction and backscattering might cause the wave to evolve as it propagates, thereby preventing the existence of a Killing propagation vector. However, it is by no means clear whether the nonlinearity of the field equations can make possible the existence of localized, nondispersive, solitonlike solutions. Dautcourt's result shows that it cannot.

The plan of this section is as follows.

Section IV A is devoted to a careful definition of GW spacetimes and associated discussion. Roughly speaking a GW spacetime is a solution to the vacuum Einstein field equations which is flat prior to the arrival of a curvature disturbance (gravitational wave), but may or may not settle back down into flatness afterward. This section also introduces a class of "standard" coordinate systems and "standard" null tetrads to be used in studying GW spacetimes.

In Sec. IV B we discuss a previous theorem of Dautcourt[10] which directly implies that any GW spacetime possessing a null Killing vector field which points along the propagation direction–i.e., possessing a gravitational wave which propagates in a perfectly diffraction free manner, indefinitely preserving its wave form–must be a PP-wave spacetime. Since PP waves are always infinitely large in transverse extent (Sec. II), this result implies that almost-plane waves (which have finite transverse "size") must always exhibit diffraction.

In Sec. IV C we discuss, present, and prove a "peeling-off"-type theorem about the behavior of the Weyl curvature quantities associated with a standard tetrad on a general GW spacetime.

In Sec. IV D we introduce the characteristic initial-value formalism of Penrose,[7] Muller zum Hagen and Seifert,[8] and Friedrich[9] which we will need to prove the theorem of Sec. IV D. We give a brief review of this formalism in a form that is appropriate to our conventions and notation, and we emphasize those aspects relevant to our purposes.

Section IV E is devoted to another theorem about GW spacetimes: a proof that any GW spacetime that not only begins flat before the wave arrives but also returns to perfect flatness after the wave passes (i.e., any precisely "sandwich" GW spacetime), must actually be a PP-wave spacetime. Since all PP waves are infinitely large in transverse extent (Sec. II), this theorem implies that almost-plane waves must always leave "tails" behind in any region of space through which they have propagated.

A. Definition of a GW spacetime

Definition. A gravitational-wave (GW) spacetime is a geodesically complete (hence maximal), vacuum spacetime $(\mathcal{M},g)$ with a C^2 metric g, satisfying the following conditions.

(i) $\mathcal{M}$ is diffeomorphic to R^4.

(ii) There exist two nonintersecting, null, achronal three-dimensional C^2 submanifolds (without edge) $\mathcal{N}$ and $\mathcal{N}'$, whose null geodesic generators have no past or future end points in $\mathcal{M}$, and which satisfy $\mathcal{N}\subset I^-(\mathcal{N}')$.

(iii) $(\mathcal{M},g)$ is flat on $I^-(\mathcal{N})$.

(iv) There exists a noncompact partial Cauchy surface through every point $p\in\mathcal{M}$.

(v) g is C^∞ outside $\mathcal{N}$ and $\mathcal{N}'$.

Remark. The differentiability class of $\mathcal{M}$ is assumed C^∞. It can be shown, using the characteristic initial-value formalism which we will outline in Sec. IV D, that there exist spacetimes satisfying all the above conditions except geodesic completeness. Completeness cannot be proved for these spacetimes because of the local nature of the existence theorems; nevertheless, in view of its mathematical naturalness and the relatively unimportant role it will play in what follows, we retain the assumption of completeness. We also remark that, by appealing to Christodoulou's recent theorems[21] proving the global existence of solutions to the initial-value problem for the vacuum Einstein equations with "small" initial data, it seems physically plausible (and in fact extremely likely) that for sufficiently "weak" gravitational waves–not a serious restriction for our purposes–the completeness condition will indeed be satisfied.

On any GW spacetime there exist local coordinate systems (u,v,x,y) for which $\mathcal{N}$, $\mathcal{N}'=\{u=0\},\{u=a\}$ and in which we can find a local null tetrad with the form

$$\vec{l}=R\left[\frac{\partial}{\partial u}+A\frac{\partial}{\partial v}\right], \qquad \vec{n}=\frac{\partial}{\partial v},$$

$$\vec{m}=\hat{M}\frac{\partial}{\partial x}+\hat{N}\frac{\partial}{\partial y}+\omega\frac{\partial}{\partial v}, \tag{4.1}$$

where $R(u,v,x,y)$, $A(u,v,x,y)$, and $\omega(u,v,x,y)$ are real and $\hat{M}(u,v,x,y)$, $\hat{N}(u,v,x,y)$ are complex functions. (A proof and detailed discussion will be given in a future paper. [20]) We call such a local chart and tetrad a "standard coordinate system" and its "associated standard tetrad." If the GW spacetime has a (null) Killing propagation direction, we also require a standard coordinate system

(u, v, x, y) to satisfy $\partial/\partial v$=Killing vector, but we drop the requirement that $\mathcal{N}' = \{u = a\}$. Note that for both the general case and for a GW spacetime with a Killing propagation direction, neither the standard charts nor the standard tetrads are uniquely defined; in both cases a large amount of coordinate and tetrad transformation freedom remains in the choice of these charts and tetrads. For example, for a sandwich plane-wave spacetime (Sec. II), the "Kerr-Schild"-type chart and the "Rosen"-type chart are both standard coordinate systems.

B. The only diffraction-free GW spacetimes are PP waves

In a short paper[10] published in 1964, Dautcourt classified all vacuum spacetimes possessing a null Killing vector. According to his classification, such spacetimes either are PP waves or are certain solutions of Petrov type II or I. Furthermore, his solutions of Petrov type II or I have the property that their curvature-invariant $R_{abcd}R^{abcd}$ is nonzero on a region that extends into $I^{-}(\mathcal{N})$ for any null surface $\mathcal{N}$ satisfying property (ii) above, and diverges on a three-dimensional timelike hypersurface.[10] Obviously, these type-I or type-II solutions cannot be gravitational-wave spacetimes according to our definition above. Therefore, as we have stated earlier, the only GW spacetimes with a Killing propagation direction (the only diffraction-free GW spacetimes) are PP waves; and this in turn implies that almost-plane waves must always exhibit diffraction.

C. A "peeling"-type property of general GW spacetimes

We now prove a "peeling"-type result about the behavior of the curvature tensor in a general GW spacetime.

Theorem 2. Let $(\mathcal{M},g)$ be a gravitational-wave spacetime with wave fronts $\mathcal{N}$, $\mathcal{N}'$; thus $(\mathcal{M},g)$ is flat on $I^-(\mathcal{N})$ where $\mathcal{N}$ is the past wave front. Then, there exists a collection of open sets $\{U_\alpha\}$, $U_\alpha \subset I^+(\mathcal{N})$, such that $\overline{\cup_\alpha U_\alpha} \supset \mathcal{N}$, and on each U_α, $\Psi_1 = \Psi_2 = \Psi_3 = \Psi_4 = 0$ in any standard chart and tetrad.

Remarks. This result implies that any generic observer through whom the curvature disturbance of the GW spacetime passes will first feel only the Ψ_0 component of the Weyl tensor in any standard tetrad. Only later, and in a "sudden" (i.e., nonanalytic, shocklike) fashion, the other components Ψ_1, Ψ_2, Ψ_3, and Ψ_4 (which represent back-scattered curvature) will appear in the measured gravitational field. Hence, if we trace the history of the observer's measurements backwards in time, the quantities $\Psi_1, \Psi_2, \Psi_3, \Psi_4$ will "peeloff" (not necessarily in that order) before the quantity Ψ_0 vanishes and the disturbance is turned off.

Proof. In any standard chart (u, v, x, y), the surface $\mathcal{N}$ is given by $\mathcal{N} = \{u=0\}$ and the standard tetrad is of the form of Eq. (4.1):

$$\vec{l} = R\left[\frac{\partial}{\partial u} + A\frac{\partial}{\partial v}\right] \qquad (R \neq 0)$$

$$\vec{n} = \frac{\partial}{\partial v}\,, \qquad \vec{m} = \hat{M}\frac{\partial}{\partial x} + \hat{N}\frac{\partial}{\partial y} + \omega\frac{\partial}{\partial v}\,.$$

Since the metric is C^2, and the spacetime is flat on $I^-(\mathcal{N})$, all curvature quantities vanish on $\mathcal{N} = \{u=0\}$. Now assume, in contradiction to the theorem's conclusion, that there is no set of neighborhoods $\{U_\alpha\}$, $U_\alpha \subset I^+(\mathcal{N})$ satisfying $\overline{\cup_\alpha U_\alpha} \supset \mathcal{N}$ such that on each U_α , $\Psi_1=\Psi_2=\Psi_3=\Psi_4\equiv 0$. Let $\{V_\alpha\}$ be the collection of all open sets in $I^+(\mathcal{N})$ on which $\Psi_i\equiv 0$, $i=1,2,3,4$. Then the complement $\mathcal{O}$ of $\overline{\cup_\alpha V_\alpha}$ in $\mathcal{M}$ is an open set intersecting $\mathcal{N}$, and $\mathcal{O}\cap I^+(\mathcal{N})$ does not contain any open neighborhoods on which $\Psi_i\equiv 0$. Thus $\mathcal{O}\cap I^+(\mathcal{N})\subset \mathrm{supp}(\Psi_1,\Psi_2,\Psi_3,\Psi_4)$; in fact, Int $\mathrm{supp}(\Psi_1,\Psi_2,\Psi_3,\Psi_4)=\mathcal{O}\cap I^+(\mathcal{N})$, and therefore $[\mathcal{O}\cap I^+(\mathcal{N})]\cap\{p\in\mathcal{M}, (\Psi_1,\Psi_2,\Psi_3,\Psi_4)(p)\neq 0\}$ is a nonempty open set whose closure intersects $\mathcal{N}$ in the closure of an open set $\mathcal{W}$ in $\mathcal{N}$. Then, it follows from the repeated application of the argument below to noncharacteristic surfaces in a neighborhood of $\mathcal{N}$, that there exists at least one open subset $\mathcal{W}$ of $\mathcal{N}$ and an open neighborhood $\mathcal{U}'$ around it, for which at least one of $\Psi_1,\Psi_2,\Psi_3,\Psi_4$ is nonzero at any point in $\mathcal{U}'\cap I^+(\mathcal{N})$.

Now in general Ψ_0 will be nonzero on $\mathcal{U}'$. Then, perform a type-II tetrad rotation[12] with a local function b to make $\Psi_0'\equiv 0$ on $\mathcal{U}'$. [This can be done since $(\Psi_1,\Psi_2,\Psi_3,\Psi_4)\neq 0$ at any point.] The rotated tetrad will be of the form

$$\vec{l}'=R\frac{\partial}{\partial u}+RA'\frac{\partial}{\partial v}+P'\frac{\partial}{\partial x}+Q'\frac{\partial}{\partial y},$$

$$\vec{n}'=\vec{n}=\frac{\partial}{\partial v}, \qquad \vec{m}'=\hat{M}\frac{\partial}{\partial x}+\hat{N}\frac{\partial}{\partial y}+\omega'\frac{\partial}{\partial v}.$$

Henceforth we will omit primes over the quantities belonging to the new tetrad.

Now in this new tetrad $\Psi_0\equiv 0$ on $\mathcal{U}'$. But then, the Bianchi identities give us

$$-D\Psi_1=-3\kappa\Psi_2+2(\varepsilon+2\rho)\Psi_1,$$

$$-D\Psi_2=-\delta^*\Psi_1-2\kappa\Psi_3+3\rho\Psi_2+2(\pi-\alpha)\Psi_1,$$

$$-D\Psi_3=-\kappa\Psi_4-\delta^*\Psi_2-2(\varepsilon-\rho)\Psi_3+3\pi\Psi_2-2\lambda\Psi_1,$$

$$-D\Psi_4=-\delta^*\Psi_3-(4\varepsilon-\rho)\Psi_4+(4\pi+2\alpha)\Psi_3-3\lambda\Psi_2\,,$$

which, when written in terms of partial derivatives with respect to the coordinates according to the tetrad above, and when the spin coefficients and metric components are regarded as known functions, yields us a system of first-order linear partial differential equations for Ψ_1, Ψ_2, Ψ_3, Ψ_4.

By the C^2-ness of g, $\Psi_1=\Psi_2=\Psi_3=\Psi_4=0$ on $\mathcal{N}$, and hence also on $\mathcal{W}\subset\mathcal{N}$. Since $R=R'\neq 0$ at any point, the surface $\mathcal{W}$ is a noncharacteristic surface for the above system of equations, given locally by $\{u=0\}$. Since the coefficients are smooth [at least C^3 since g is C^4 on $I^+(\mathcal{N})\cap I^-(\mathcal{N}')$], any C^1 solution $(\Psi_1,\Psi_2,\Psi_3,\Psi_4)$ of the system above is uniquely determined in some neighborhood $\mathcal{V}'$ of $\mathcal{W}$ by Holmgren's uniqueness theorem extended to nonanalytic equations (John[22] and Smoller[23]). (In fact, one can safely assume g to be piecewise analytic thus eliminating the need for such an extended uniqueness theorem: that piecewise analyticity, by the Cauchy-Kovalewski theorem,[22] implies that the unique C^1 solution of the above system is also analytic on its domain of uniqueness.) But as we clearly see from the above system of equations, $\Psi_1=\Psi_2=\Psi_3=\Psi_4=0$ is a C^1 solution in any neighborhood of $\mathcal{W}$, of the above initial-value problem. Therefore it is the unique solution in some neighborhood $\mathcal{V}'$ of $\mathcal{W}$, and $\Psi_1=\Psi_2=\Psi_3=\Psi_4\equiv 0$ in that neighborhood $\mathcal{V}'$.

But since $\Psi_0\equiv 0$ on $\mathcal{U}'$, $\mathcal{U}'\cap\mathcal{V}'$ is a flat neighborhood of $\mathcal{W}$ and hence in the original tetrad, on $\mathcal{U}'\cap\mathcal{V}'$, $\Psi_1=\Psi_2=\Psi_3=\Psi_4\equiv 0$. This contradicts our assumption about $\mathcal{W}$ and $\mathcal{U}'$, since $(\mathcal{U}'\cap\mathcal{V}')\cap I^+(\mathcal{N})$ is nonempty and is contained in $\mathcal{U}'\cap I^+(\mathcal{N})$. This contradiction proves the theorem 2. □

D. Review of the characteristic initial-value formalism

Our next result about GW spacetimes is a uniqueness theorem similar to that of Dautcourt: Whereas Dautcourt's theorem (Sec. IV B above) says that the only diffraction-free GW spacetimes are PP waves, our next theorem (Sec. IV E below) says that the only sandwich GW spacetimes are PP waves. Since the proof of the theorem makes extensive use of the characteristic initial-value formalism as developed by Penrose,[7] Muller zum Hagen and Seifert,[8] Friedrich,[9] and others, we first give in this subsection a brief review of this formalism, emphasizing those aspects that are relevant to our purposes. We follow Friedrich[9] quite closely, though with entirely different conventions.

We assume that we are given a "spacetime $\mathcal{M}$ with boundary," where the boundary $\partial\mathcal{M}\equiv S$ consists of two null surfaces $\mathcal{N}_1$, $\mathcal{N}_2$ intersecting and terminating in the past directions on a spacelike two-dimensional submanifold $Z=\mathcal{N}_1\cap\mathcal{N}_2$; $\partial\mathcal{M}\equiv S=\mathcal{N}_1\cup\mathcal{N}_2\cup Z$. Here the geometry on the boundary $\partial\mathcal{M}$ is to be understood as the geometry given by the limit of the metric g which lives in the open interior of $\mathcal{M}$; this limiting metric defines smooth tensor fields on the manifolds without boundary: $\mathrm{Int}\mathcal{N}_1$, $\mathrm{Int}\mathcal{N}_2$, and Z. We describe the situation in Fig. 3.

We will now outline the construction of a local coordinate system and tetrad on $\mathcal{M}$, which are particularly well suited for the discussion of the initial-value problem. We will call them Friedrich's tetrad and coordinate system.[9] They are constructed as follows.

On Z choose coordinates $x^3, x^4 \equiv x^A$ $(\equiv x, y)$.

On $\mathcal{N}_1$ choose a function $u \geq 0$ which vanishes on Z and which is the affine parameter along integral curves of $\vec{e}_1 \equiv \vec{l}$, the null geodesic generators of $\mathcal{N}_1$. Let Z_{u_0} be the two-dimensional submanifold $\{u=u_0\}$ in $\mathcal{N}_1$. Choose on Z complex vector fields $\vec{e}_3$, $\vec{e}_4=\vec{e}_3^{\,*}$ with $g(\vec{e}_3,\vec{e}_4)=1$, $g(\vec{e}_3,\vec{e}_3)=0$ which are tangent to Z. Propagate $\vec{e}_3,\vec{e}_4$ onto $\mathcal{N}_1$ in the following manner: Construct $\vec{e}_3{}'$, $\vec{e}_4{}'$ as the parallel transports of $\vec{e}_3,\vec{e}_4$ along $\vec{e}_1$. At any point in $\mathcal{N}_1$, $\vec{e}_3$ $(\vec{e}_4)$ lies in the intersection of the $\vec{e}_3{}'\wedge\vec{e}_1$ $(\vec{e}_4{}'\wedge\vec{e}_1)$ plane with the two-surface Z_{u_0} through that point, and $g(\vec{e}_3,\vec{e}_4)=1$, $g(\vec{e}_3,\vec{e}_3)=g(\vec{e}_4,\vec{e}_4)=0$.

Choose a coordinate $u\geq 0$ on $\mathcal{M}$ coinciding with u on $\mathcal{N}_1$, such that $\forall$ u_0, $u=u_0$ is a null hypersurface in $\mathcal{M}$. Put $\vec{e}_2\equiv-\vec{\nabla}u$ on $\mathcal{M}$. Parallel transport $\vec{e}_1$, $\vec{e}_3$, $\vec{e}_4$ from $\mathcal{N}_1$ to all of $\mathcal{M}$ along integral curves of $\vec{e}_2$. Choose a function $v\geq 0$ on $\mathcal{M}$ and functions x^A on $\mathcal{M}$ $(A=3,4)$ such that (i) x^A coincide with x^A on Z, (ii) x^A are constant along null generators of $\mathcal{N}_1$ and null generators of the $\{u=u_0\}$ hypersurfaces [and hence $\vec{e}_2(x^A)=0$], (iii) $v=0$ on $\mathcal{N}_1$, and (iv) $\{u,v,x^A\}$ form a coordinate system such that $\vec{e}_2\equiv\partial/\partial v$.

As a result of these constructions, we have

$$\vec{l}\equiv\vec{e}_1=\frac{\partial}{\partial u}+U\frac{\partial}{\partial v}+X^A\frac{\partial}{\partial x^A},$$

$$\vec{n}\equiv\vec{e}_2=\frac{\partial}{\partial v}, \quad \vec{m}\equiv\vec{e}_3=\omega\frac{\partial}{\partial v}+\xi^A\frac{\partial}{\partial x^A}, \tag{4.2}$$

$$\vec{m}^*\equiv\vec{e}_4=\omega^*\frac{\partial}{\partial v}+\xi^{*A}\frac{\partial}{\partial x^A}$$

as a null tetrad on $\mathcal{M}$. We also have $\mathcal{N}_1=\{v=0\}$, $\mathcal{N}_2=\{u=0\}$, and $Z=\{u=v=0\}$. Moreover,

$$U=X^A=\omega=0 \quad \text{on } \mathcal{N}_1 \ (v=0),$$

$$\kappa=\varepsilon=0 \quad \text{on } \mathcal{N}_1 \ (v=0), \tag{4.3}$$

while on the whole spacetime $\mathcal{M}$

$$\nu=\gamma=\tau=\pi-(\alpha+\beta^*)=\mu-\mu^*=0 \quad \text{on } \mathcal{M} . \tag{4.4}$$

We now formulate the fundamental theorems of the characteristic initial-value formalism.

An initial data set is a set of complex- and real-valued functions

$$U, X^A, \omega, \xi^A, \mu, \beta, \alpha, \lambda, \rho, \varepsilon, \sigma, K, \Psi_0, \Psi_1, \Psi_2, \Psi_3, \Psi_4$$

on $S\equiv\partial\mathcal{M}=\mathcal{N}_1\cup\mathcal{N}_2\cup Z$. A reduced initial data set is a set of complex- and real-valued functions $\mu, \rho, \sigma, \lambda, \pi, \xi^A$ on Z such that $g^{AB}=\xi^A\xi^{*B}+\xi^B\xi^{*A}$ is a positive definite metric on Z, and complex-valued functions Ψ_4 on $\mathcal{N}_2$ and Ψ_0 on $\mathcal{N}_1$. It is assumed that the initial and reduced initial data sets satisfy Eqs. (4.3) and (4.4).

Theorem 3. Let an initial data set on S satisfy all the constraint equations obtained by restricting the vacuum Einstein field equations onto the initial surface S. Then this initial data set uniquely determines, in some neighborhood of S, a vacuum spacetime $(\mathcal{M},g)$ with boundary $\partial\mathcal{M}=S$ and with the data on S coinciding with the restrictions to S of the spin quantities on $\mathcal{M}$ in some suitable null tetrad and coordinate system on $\mathcal{M}$.

Theorem 4. A *reduced initial data set* on S uniquely determines, by using the constraint equations, an *initial data set* on S that satisfies the constraints.

For the proof of theorem 3, see Refs. 9 and 8. In our proof of theorem 5 in the next section, we will need the intermediate steps of the proof of theorem 4. Therefore, we sketch here an outline of this proof, following Friedrich.[9]

Proof of theorem 4 .

(1) To find the initial data on Z from the reduced initial data, first use the commutation relations

$$(\alpha-\beta^*)\xi^A+(\beta-\alpha^*)\xi^{*A}=(\xi^B\xi^{*A}{}_{,B}-\xi^{*B}\xi^A{}_{,B}) \quad \text{on } Z \tag{4.5}$$

and $\pi=\alpha+\beta^*$ [Eqs. (4.4)] to find α,β,π on Z. All other initial data on Z are known from the reduced initial data and Eqs. (4.3) and (4.4) (since $Z\subset\mathcal{N}_1$), except Ψ_1, Ψ_2, and Ψ_3 which are found from the following Ricci identities restricted to Z:

$$\delta^*\sigma-\delta\rho=\rho(\alpha^*+\beta)-\sigma(3\alpha-\beta^*)+\Psi_1\,, \tag{4.6}$$

$$\delta^*\beta-\delta\alpha=\mu\rho-\lambda\sigma+\alpha\alpha^*+\beta\beta^*-2\alpha\beta+\Psi_2\,, \tag{4.7}$$

$$\delta^*\mu-\delta\lambda=\mu(\alpha+\beta^*)+\lambda(\alpha^*-3\beta)+\Psi_3\,. \tag{4.8}$$

(2) To find the initial data on $\mathcal{N}_1$, proceed as follows: First use Ψ_0 on $\mathcal{N}_1$ given by the reduced initial data, and the following commutation relations and Ricci identities, restricted to $\mathcal{N}_1$,

$$\xi^A{}_{,u}=-\rho^*\xi^A-\sigma\xi^{*A}\,, \tag{4.9}$$

$$-D\rho=-\rho_{,u}=\rho^2+\sigma\sigma^*\,,$$

$$-D\sigma=-\sigma_{,u}=\sigma(\rho+\rho^*)-\Psi_0\,. \tag{4.10}$$

Use these to integrate, onto $\mathcal{N}_1$, by ordinary differential equations along null generators of $\mathcal{N}_1$, the reduced initial data $\{\ \xi^A,\rho,\sigma \text{ on } Z\ \}$. Then use the Ricci identities on $\mathcal{N}_1$,

$$\Psi_1=\delta^*\sigma-\delta\rho-\rho(\alpha^*+\beta)+\sigma(3\alpha-\beta^*)\,, \tag{4.11a}$$

$$-D\,\alpha = -\alpha_{,u} = \rho\alpha + \beta\sigma^* + \rho(\alpha+\beta^*)\,,$$

$$-D\,\beta = -\beta_{,u} = (2\alpha+\beta^*)\sigma + \rho^*\beta - \delta^*\sigma + \delta\rho$$

$$+\rho(\alpha^*+\beta) - \sigma(3\alpha-\beta^*)\,, \tag{4.11b}$$

to determine α,β,Ψ_1 on $\mathcal{N}_1$ from ω,ξ^A,ρ,σ on $\mathcal{N}_1$ which are known from the preceding step and from Eqs. (4.3). Similarly, use the Ricci identities on $\mathcal{N}_1$,

$$\Psi_2 = \delta^*\beta - \delta\alpha - (\rho\mu - \sigma\lambda) - \alpha\alpha^* - \beta\beta^* + 2\alpha\beta\,, \tag{4.12a}$$

$$-D\,\lambda = -\lambda_{,u} = -\delta^*(\alpha+\beta^*) + \rho\lambda + \sigma^*\mu + (\alpha+\beta^*)^2$$

$$+\alpha^2 - \beta^{*2}\,,$$

$$-D\,\mu = -\mu_{,u} = -\delta(\alpha+\beta^*) + \rho^*\mu + \sigma\lambda + |\alpha+\beta^*|^2$$

$$-(\alpha+\beta^*)(\alpha^*-\beta) - \delta^*\beta + \delta\alpha + (\rho\mu - \sigma\lambda)$$

$$+\alpha\alpha^* + \beta\beta^* - 2\alpha\beta\,, \tag{4.12b}$$

to determine μ, λ, and Ψ_2 on $\mathcal{N}_1$ by integrating, by ordinary differential equations (ODEs) along $\mathcal{N}_1$, the reduced initial data on Z. Finally, use the Ricci and Bianchi identities on $\mathcal{N}_1$,

$$\Psi_3 = \delta^*\mu - \delta\lambda - \mu(\alpha+\beta^*) - \lambda(\alpha^* - 3\beta)\,, \tag{4.13}$$

$$-D\,\Psi_4 = -\Psi_{4,u} = -\delta^*\Psi_3 + \rho\Psi_4 + (6\alpha+4\beta^*)\Psi_3$$

$$-3\lambda\Psi_2\,, \tag{4.14}$$

to determine Ψ_3 and Ψ_4 on $\mathcal{N}_1$.

(3) To find the initial data on $\mathcal{N}_2$, proceed as follows: Use the commutation relations and the Ricci identities on $\mathcal{N}_2$,

$$U_{,v}=-(\varepsilon+\varepsilon^*)+\pi\omega+\pi^*\omega^* ,$$

$$X^A{}_{,v}=\pi\xi^A+\pi^*\xi^{*A} , \tag{4.15a}$$

$$\omega_{,v}=-\pi^*+\mu\omega+\lambda^*\omega^* ,$$

$$\xi^A{}_{,v}=\mu\xi^A+\lambda^*\xi^{*A} , \tag{4.15b}$$

$$\Delta\beta=\beta_{,v}=\beta\mu+\alpha\lambda^* ,$$

$$-\Delta\alpha=-\alpha_{,v}=-\beta\lambda-\mu\alpha+\delta^*\mu-\delta\lambda-\mu(\alpha+\beta^*)$$

$$-\lambda(\alpha^*-3\beta) , \tag{4.15c}$$

to determine X^A, ξ^A, β, α, ω, and π on $\mathcal{N}_2$. This should be done *after* finding λ and μ on $\mathcal{N}_2$ from Ψ_4 on $\mathcal{N}_2$, by using the following Ricci identities on $\mathcal{N}_2$:

$$\Delta\mu=\mu_{,v}=\mu^2+\lambda\lambda^* ,$$

$$\Delta\lambda=\lambda_{,v}=2\mu\lambda-\Psi_4 . \tag{4.16}$$

(Note that Ψ_4 on $\mathcal{N}_2$ is given by reduced initial data.) Next find Ψ_3, Ψ_2, ρ, and σ on $\mathcal{N}_2$ as follows: First find ρ and σ on $\mathcal{N}_2$ by integrating the ODEs on $\mathcal{N}_2$ which follow from the Ricci identities

$$\Delta\sigma=\sigma_{,v}=\mu\sigma+\lambda^*\rho ,$$

$$-\Delta\rho=-\rho_{,v}=-(\rho\mu+\sigma\lambda)+\delta^*\beta-\delta\alpha-(\mu\rho-\lambda\sigma)$$

$$-\alpha\alpha^* - \beta\beta^* + 2\alpha\beta\,. \tag{4.17}$$

Then use the Ricci identities

$$\Psi_3 = \delta^* \mu - \delta\lambda - \mu(\alpha+\beta^*) - \lambda(\alpha^* - 3\beta)\,, \tag{4.18a}$$

$$\Psi_2 = \delta^* \beta - \delta\alpha - (\mu\rho - \lambda\sigma) - \alpha\alpha^* - \beta\beta^* + 2\alpha\beta\,, \tag{4.18b}$$

on $\mathcal{N}_2$ to compute Ψ_3 and Ψ_2 on $\mathcal{N}_2$. Finally, use Eqs. (4.15a) and the Ricci and Bianchi identities on $\mathcal{N}_2$,

$$\Delta\varepsilon = \varepsilon_{,v} = \pi^* \alpha + \pi\beta - \Psi_2\,, \tag{4.19}$$

$$\Delta\kappa = \kappa_{,v} = \pi^* \rho + \pi\sigma - \Psi_1\,, \tag{4.20}$$

$$\Psi_1 = \delta^* \sigma - \delta\rho - \rho(\alpha^* + \beta) + \sigma(3\alpha - \beta^*)\,, \tag{4.21}$$

$$-\Delta\Psi_0 = -\Psi_{0,v} = -\delta\Psi_1 + 3\sigma\Psi_2 - \mu\Psi_0 - 2\beta\Psi_1\,, \tag{4.22}$$

to determine ε, κ, Ψ_1, Ψ_0, and U on $\mathcal{N}_2$.

The uniqueness statement in the theorem now follows straightforwardly, since we have only integrated ordinary differential equations to determine the initial data on S from the reduced initial data on S. □

E. The only sandwich GW spacetimes are PP waves

Theorem 5. Let $(\mathcal{M},g)$ be a gravitational-wave spacetime with wavefronts $\mathcal{N}_2$, $\mathcal{N}_2'$; hence $(\mathcal{M},g)$ is flat on $I^-(\mathcal{N}_2)$. If $(\mathcal{M},g)$ is also flat on $I^+(\mathcal{N}_2')$, and if the fundamental theorems of the characteristic initial-value formalism hold globally (rather than just locally) on $\mathcal{M}$, then $(\mathcal{M},g)$ is a PP-wave spacetime.

Remark. A gravitational wave of this type, which leaves spacetime precisely flat both before and after its passage, is called a sandwich wave. This theorem then says that the only sandwich gravitational waves are PP waves.

Proof. We assume, as stated in the theorem, that $I^+(\mathcal{N}_2')$ (Fig. 3) is flat and that Theorem 3 holds globally on $\mathcal{M}$; and we seek to show that $(\mathcal{M},g)$ is a PP-wave spacetime.

Choose a null surface $\mathcal{N}_1$ which intersects $\mathcal{N}_2$ transversely in a spacelike two-submanifold Z (and $\mathcal{N}_2'$ in Z') (Fig. 3). Then, that portion of the spacetime which lies to the future of the initial null boundary $S=\mathcal{N}_1\cup\mathcal{N}_2\cup Z$ is uniquely determined by the reduced initial data it induces on S. From here on, we will only be interested in this region $I^+(S)$ of the spacetime $(\mathcal{M},g)$ together with the boundary of this region $S=\partial I^+(S)$, and we will denote the spacetime region with boundary, $I^+(S)\cup S$, by the same symbol $\mathcal{M}$, where $\partial\mathcal{M}=S$. The following proof will show that if $(\mathcal{M},g)$ is a precisely sandwich GW spacetime as defined above, then the reduced initial data induced on the null boundary S is PP-wave reduced initial data. This is sufficient to prove theorem 5, since the location of the transverse null surface $\mathcal{N}_1$ is arbitrary.

Before proceeding with the proof, we observe that the flatness of $I^+(\mathcal{N}_2')$ and $I^-(\mathcal{N}_2)$ requires $\Psi_0=\Psi_1=\Psi_2=\Psi_3=\Psi_4=0$ on $\mathcal{N}_2$, on Z, and on $\mathcal{N}_1\cup I^+(\mathcal{N}_2')$. That is, all curvature quantities (in any tetrad) vanish on the null boundary S *except* on that portion of $\mathcal{N}_1$ lying between $\mathcal{N}_2$ and $\mathcal{N}_2'$ (Fig. 3). We also note that, in general there is some coordinate freedom in choosing Friedrich's coordinate system and tetrad on the spacetime $\mathcal{M}$ with (null) boundary $S=\partial\mathcal{M}=\mathcal{N}_1\cup\mathcal{N}_2\cup Z$. In the following, we will use this gauge freedom in the choice of Friedrich's chart, coupled with the freedom to choose the transverse null surface $\mathcal{N}_1$ (the choice of which is completely arbitrary), to construct *a specific* null boundary S and *a specific* Friedrich-type coordinate chart on the spacetime $\mathcal{M}$ with boundary S. These choices for S and for the Friedrich-type coordinate chart on $\mathcal{M}=I^{+}(S)$ will be particularly well suited for studying the exactly sandwich GW spacetime $(\mathcal{M},g)$.

The gauge freedom in the choice of a Friedrich-type coordinate system can be decomposed into two different types of coordinate transformations. The transformations of the first type are generated by the successive application of two transformations: (i) transformations of the form $u'=\alpha(x^A)u$ on $\mathcal{N}_1$ which give, upon extending u' uniquely as a null coordinate, $u'=u'(x_A,u,v)$ on $\mathcal{M}$, where α is an arbitrary function on Z which is extended to S by keeping it constant along the null geodesic generators of $\mathcal{N}_1$ and $\mathcal{N}_2$; (ii) transformations of the form $v'=v'(u,v,x^A)$ which are so adjusted that when $x^{A\prime}=x^{A\prime}(x^A,u,v)$ on $\mathcal{M}$ are obtained from the x^A on $\mathcal{N}_1$ by keeping them constant along the new integral curves of $\vec{e}_1{}'=-\vec{\nabla}u'$, we have $\partial/\partial v'=-\vec{\nabla}u'$ and $v'=0$ on $\mathcal{N}_1$ in the new primed coordinate system. The second type of coordinate transformations generating the gauge freedom in the choice of Friedrich's chart are given by

$$v'=v+v'(u,x^A)\ ,\ \ u'=u\ ,\ \ x^{A\prime}=x^{A\prime}(u,x^A)\ ,$$

where all primed quantities are arbitrary functions of their arguments. Since $x^{A\prime}$ (and x^A) are to be constant on null generators of $\mathcal{N}_1$ and $\mathcal{N}_2$, these transformations reduce to

$$v'=v+v'(u,x^A)\ ,\ \ x^{A\prime}=x^{A\prime}(x^A)\ ,\ \ u'=u\ .$$

And if $\mathcal{N}_1$ and $\mathcal{N}_2$ are fixed, since we have $v'=0$ on $\mathcal{N}_1$, these transformations further reduce to

$$v'=v\ ,\ \ x^{A\prime}=x^{A\prime}(x^A)\ ,\ \ u'=u\ .$$

Both types of coordinate transformations induce tetrad rotations, since the old $\vec{n},\vec{l},\vec{m},\vec{m}^*$ in the new coordinate system will not be of the form (4.2), and tetrad rotations are necessary to bring them back into the form (4.2) in the $(u',v',x^{A'})$ chart.

In the next two paragraphs we will use, as we have indicated before, the above coordinate and tetrad freedom coupled with the freedom in the choice of $\mathcal{N}_1$, to bring Friedrich's coordinate system and associated tetrad into a form that meshes nicely with the structure of our sandwich GW spacetime.

We begin with the freedom to choose $\mathcal{N}_1$. We shall choose $\mathcal{N}_1$ so that it is given by $\hat{v}=0$, where $\hat{v}$ is a Minkowskian null coordinate in $I^+(\mathcal{N}_2')$. That is, $d\hat{v}$ is a parallel null 1-form in the flat region $I^+(\mathcal{N}_2')$, or in other words on $I^+(\mathcal{N}_2')$ there is a coordinate system $(\hat{v},u',x^{A'})$ in which the metric is

$$
\begin{aligned}
g = {} & dx^{3'}\otimes dx^{3'} + dx^{4'}\otimes dx^{4'} \\
& -\frac{1}{2}(du'\otimes d\hat{v} + d\hat{v}\otimes du') \qquad \text{on } I^+(\mathcal{N}_2') .
\end{aligned}
$$

Note that this choice of $\mathcal{N}_1$ on $I^+(\mathcal{N}_2')$ completely fixes it everywhere in spacetime including the region between $\mathcal{N}_2$ and $\mathcal{N}_2'$, because there exist precisely two null surfaces passing through *any* spacelike two-surface. In other words, $\mathcal{N}_1$ is extended in the past directions *beyond* the spacelike two-surface Z' (Fig. 3) as that null surface, which together with $\mathcal{N}_2'$ constitutes the unique pair of null surfaces through Z'.

Now the scaling freedom in u, i.e., the freedom of coordinate transformations of the first kind, is fixed by the arrangement that the wave front $\mathcal{N}_2'$ coincides with the null surface $\{u=a\}$. (Then the coordinate v is constructed as usual from $-\vec{\nabla}u$, using u and some choice of coordinates x^A on Z and thereby on $\mathcal{M}$.) We are then left with the following coordinate freedom of the second type:

$$v'=v\ ,\ x^{A\prime}=x^{A\prime}(x^A)\ ,\ u'=u\ .$$

We fix this remaining freedom totally by noting that, since Z' is a two-dimensional spacelike hypersurface in flat spacetime contained in $\{\hat{v}=0\}$, the induced metric on it is flat, and hence we can arrange $x^{A\prime}=x^{A\prime}(x^A)$ on Z in such a manner that at $\{u=a\}$ on $\mathcal{N}_1$:

$$\xi^{A\prime}(u=a)=\xi^B(u=a)\frac{\partial x^{A\prime}}{\partial x^B}=\frac{1+i}{2} \qquad \text{for } A=3\ ,$$

$$=\frac{1-i}{2} \qquad \text{for } A=4\ . \tag{4.23}$$

(Note that we are leaving the tetrad vectors $\vec{m}$, $\vec{m}^*$ fixed during the above arrangement of coordinates.) Then $g^{AB}(u=a)=\delta^{AB}$. But using Eqs. (4.10), since $Z'\subset\mathcal{N}_1$, this gives

$$\beta-\alpha^*=0 \qquad \text{on } Z'\ . \tag{4.24}$$

Now, note that we have two coordinate systems covering $I^+(\mathcal{N}_2')$: $(u',\hat{v},x^{A\prime})$ and (u,v,x^A), where the first one is Minkowskian. As we will argue later, $\mathcal{N}_2'$ (as well as $\mathcal{N}_2$) is a flat null surface in Minkowski spacetime, and therefore by construction one can find a Minkowskian coordinate system $(u'',v'',x^{A\prime\prime})$ on $\overline{I^+(\mathcal{N}_2')}$ [rotating $(u',\hat{v},x^{A\prime})$ by a Lorentz transformation, if necessary] such that $\partial/\partial v=\partial/\partial v''$ on Z'. But then on Z', $\partial/\partial v=\partial/\partial v''=\vec{n}$; and since $\vec{m}$ is tangent to Z', $\nabla_{\vec{m}}\vec{n}$ on Z' does not depend on the extension of $\vec{n}$ from Z' to $\mathcal{M}$. In particular, $\partial/\partial v''$ is an extension of $\vec{n}\,|_{Z'}$ to $\mathcal{M}$; and hence

$$(\nabla_{\vec{m}}\vec{n})\,|_{Z'}=\left[\nabla_{\vec{m}}\frac{\partial}{\partial v''}\right]|_{Z'}=0\ ,$$

since $\partial/\partial v''$ is a parallel vector field on $I^+(\mathcal{N}_2')$ and the metric is C^2. But $\nabla_{\vec{m}}\vec{n}=(\alpha^*+\beta)\vec{n}-\lambda^*\vec{m}^*-\mu\vec{m}$. Thus, on Z' we have $\alpha^*+\beta=0$, which by Eq. (4.24) gives $\alpha=\beta=0$ on Z'. Therefore, in this way we fix the above remaining coordinate freedom so that on $\mathcal{N}_1$, $\alpha(u=a)=\beta(u=a)=0$, $\xi^3(u=a)=(1+i)/2$, $\xi^4(u=a)=(1-i)/2$, and $g^{AB}(u=a)=\delta^{AB}$. Note that, with this procedure we also fix the remaining freedom for tetrad transformations of type III: $\vec{l}\to\vec{l}$, $\vec{n}\to\vec{n}$, $\vec{m}\to e^{i\theta}\vec{m}$, $\vec{m}^*\to e^{-i\theta}\vec{m}^*$, where θ is a function which depends only on x^A.

This completes our specialization of Friedrich's coordinate system and tetrad. In the next paragraph we shall derive the special values of the spin coefficients associated with this tetrad.

Now, since $\mathcal{N}_1$ is a flat null surface in Minkowski space for $u\geq a$, its null geodesic generators have no shear or convergence; and hence, since on $\mathcal{N}_1$ the null generators are tangent to $\vec{l}$, we have $\rho=\sigma=0$ for $u\geq a$ on $\mathcal{N}_1$. But by the Ricci identities (A20) and (A21) on $\mathcal{M}$,

$$\Delta\sigma=\sigma_{,v}=\mu\sigma+\lambda^*\rho \qquad \text{for } u\geq a\ ,$$

$$-\Delta\rho=-\rho_{,v}=-(\rho\mu+\sigma\lambda) \qquad \text{for } u\geq a\ .$$

These imply, by the uniqueness theorem for ODE,

$$\rho=\sigma=0 \qquad \text{for } u\geq a$$

on all of $\mathcal{M}$. Applying the Ricci identities (A8) and (A9) on $\mathcal{N}_1$ for $u\geq a$, and using the same arguments as in the last few equations, we obtain

$$\pi=\alpha=\beta=0 \qquad \text{for } u\geq a$$

on all of $\mathcal{M}$. (Here we have used the fact that, by the choice of the coordinates x^A, we have $\alpha=\beta=0$ at $u=a$.) Similarly, we obtain

$$\varepsilon=\kappa=0 \qquad \text{for } u \geq a \ ,$$

on all of $\mathcal{M}$. Since $\mathcal{N}_2$ and $\mathcal{N}_2'$ are nonsingular null surfaces whose null geodesic generators have no end points in $\mathcal{M}$ (and $\mathcal{M}$ is complete), we have

$$\lambda=\mu=0 \qquad \text{on } \mathcal{N}_2 \text{ and } \mathcal{N}_2' \ .$$

Then using Eqs. (A11) and (A12) on $\mathcal{M}$ for $u \geq a$ we obtain

$$\lambda=\mu=0 \qquad \text{for } u \geq a$$

on all of $\mathcal{M}$.

Now, having completed the construction of our specific Friedrich-type coordinate system and its associated tetrad and the specific null boundary $S=\mathcal{N}_1 \cup \mathcal{N}_2 \cup Z$ on which our sandwich GW spacetime induces a characteristic initial data set, we return to the proof that the sandwich GW spacetime $(\mathcal{M},g)$ is actually a PP wave. Clearly, by theorems 3 and 4, there is a one to one correspondence between vacuum *sandwich* GW spacetimes, and the reduced initial data sets they induce on the null boundary S, expressed in the coordinate system and tetrad constructed above. We will call such reduced initial data, which correspond to sandwich GW spacetimes, "good reduced initial data." Note that this condition of "goodness" on a reduced initial data set is equivalent to the demand that the spacetime which develops uniquely from it according to theorem 4 is flat on $I^+(\mathcal{N}_2')$ and $I^-(\mathcal{N}_2)$.

It is not hard to prove, using Eqs. (4.5)–(4.22), that any good reduced initial data set on S is completely determined by giving Ψ_0 on $\mathcal{N}_1$, between $u=0$ and $u=a$.

Therefore, the set of good initial data is in one to one correspondence with a certain subset of the set of all C^2 functions Ψ_0 on $\mathcal{N}_1$, which vanish for $u=0$ and $u \geq a$. In the following paragraphs, we will prove that any good reduced initial data set is necessarily a PP-wave reduced initial data set, which will prove the theorem.

We begin by noting that a PP-wave metric in the Kerr-Schild coordinates is associated with the null tetrad

$$\vec{l}'=2\left(\frac{\partial}{\partial U}+h(U,X,Y)\frac{\partial}{\partial V}\right),$$

$$\vec{n}'=\frac{\partial}{\partial V}, \qquad \vec{m}'=\frac{1+i}{2}\frac{\partial}{\partial X}+\frac{1-i}{2}\frac{\partial}{\partial Y},$$

in which the only nonzero spin quantities are κ' and $\Psi_0'=-\delta'\kappa'$, and in which $\delta'^{*}\kappa'=0$. When we transform this coordinate system and tetrad into Friedrich's form [Eq. (4.2)], the only nonzero spin quantities are ρ, σ, and $\Psi_0=\Psi_0'$, where $\delta^{*}\Psi_0=0$. Therefore the PP-wave reduced initial data will consist of (i) ξ^A,ρ,σ on Z with $\mu=\lambda=\pi=0$ on Z, (ii) $\Psi_4=0$ on $\mathcal{N}_2$, and (iii) Ψ_0 on $\mathcal{N}_1$ with $\delta^{*}\Psi_0=0$ and $\Psi_0=0$ for $u=0$, $u \geq a$.

A necessary condition for the reduced initial data induced from Ψ_0 to be good is that, when the Eqs. (4.10), (4.9), and (4.11b) are solved with initial conditions (4.23) and ρ, σ, α, $\beta=0$ at $u=a$, and when Eqs. (4.12b) are then solved for μ, λ with initial conditions μ, $\lambda=0$ at $u=a$, one then obtains, at $u=0$ (on Z):

$$\mu(0)=\lambda(0)=0,$$

$$\Psi_2(0)=[\delta^{*}\beta-\delta\alpha-(\mu\rho-\lambda\sigma)-\alpha\alpha^{*}-\beta\beta^{*}$$

$$+2\alpha\beta]|_{u=0}=0, \tag{4.25}$$

$$\Psi_1(0)=[\delta^* \sigma-\delta\rho-\rho(\alpha^* +\beta)+\sigma(3\alpha-\beta^*)]|_{u=0}=0 .$$

We claim that these conditions can only be satisfied if $\delta^* \Psi_0 \equiv 0$ on $\mathcal{N}_1$.

Since the proof of this claim is rather long, we will only outline in this paragraph the main steps. First define $A \equiv \delta\rho-\delta^* \sigma$. A satisfies, on $\mathcal{N}_1$,

$$DA=A_{,u}=-(2\rho+\rho^*)A+\sigma A^* -\delta^* \Psi_0 ,$$

$$A=0 \qquad \text{on } u=a \text{ in } \mathcal{N}_1 . \tag{4.26}$$

Now using theorem 2 and the Bianchi identities, this gives $A=0$ in some neighborhood U_a of $\{u=a\}$ in $\mathcal{N}_1$. By theorem 2 and the Goldberg-Sachs theorem,[24] $\lambda=0$ in some neighborhood of $\{u=0\}$ and $\{u=a\}$ in $\mathcal{N}_1$, and this gives $\alpha=\beta=0$ at $u=0$. Again by the Bianchi identities this implies $\delta^* \Psi_0=0$ on some neighborhood of $\{u=0\}$ in $\mathcal{N}_1$. It then follows from Eqs. (4.26) using standard arguments for ordinary differential equations [specifically, using an energy-type inequality, which involves a positive-definite expression depending on $|DA|$ and $|\Psi_0|^2$ and which is obtained from Eqs. (4.26), (4.10), (4.9), and (4.25)] that if $\delta^* \Psi_0 \neq 0$ at any point on $\mathcal{N}_1$, $|DA|$ and thence $A=\delta\rho-\delta^* \sigma$ are nonzero at $u=0$. But this contradicts Eq. (4.25). Therefore $\delta^* \Psi_0 \equiv 0$ on $\mathcal{N}_1$ for any "good" Ψ_0 on $\mathcal{N}_1$ and the claim is proved. [To understand this claim more intuitively, first note that we still have some freedom left in the choice of the null surface $\mathcal{N}_1$, even though we have restricted it to be a flat Minkowskian surface on $I^+(\mathcal{N}_2')$. This freedom consists of (i) rotating the surface $\mathcal{N}_1$ by Lorentz transformations applied in the flat region $I^+(\mathcal{N}_2')$, and (ii) translating $\mathcal{N}_1$ linearly in $I^+(\mathcal{N}_2')$. Thus, even if the fact $\Psi_0 \neq 0$ on $\mathcal{N}_1$ *were compatible* with Eqs. (4.25) for a *particular* choice of the surface $\mathcal{N}_1$, we could readjust the orientation of $\mathcal{N}_1$ by using the above freeedom in such a way that with the new choice of $\mathcal{N}_1$, Eqs. (4.25) would be violated.]

It is easily seen that the initial data set associated with a good reduced initial data set induced from a Ψ_0 with $\delta^* \Psi_0 \equiv 0$ on $\mathcal{N}_1$ has the following form: (i) on $\mathcal{N}_1$

$$\Psi_0=0 \qquad \text{for } u \geq a \text{ and } u=0 ,$$

$$\rho, \sigma=0 \qquad \text{for } u \geq a ,$$

$$\delta^* \Psi_0 \equiv 0 ,$$

$$\varepsilon=\kappa=\mu=\lambda=\alpha=\beta=\pi=0 ,$$

$$\Psi_1=\Psi_2=\Psi_3=\Psi_4=0 ,$$

$$X^A=U=\omega=0 ,$$

while ξ^A are found by Eq. (4.9). (ii) On Z

$$\varepsilon=\kappa=\mu=\lambda=\alpha=\beta=\pi=0 ,$$

$$\Psi_0=\Psi_1=\Psi_2=\Psi_3=\Psi_4=0 ,$$

$$X^A=\omega=U=0 ,$$

while ξ^A ,ρ,σ are nonzero. (iii) On $\mathcal{N}_2$

$$\mu=\lambda=\pi=\alpha=\beta=\varepsilon=\kappa=0 ,$$

$$\Psi_0=\Psi_1=\Psi_2=\Psi_3=\Psi_4=0 ,$$

$$U=X^A=\omega=0 ,$$

while ξ^A ,ρ,σ are nonzero but independent of v. (iv) On the whole spacetime,

$\tau=\nu=\gamma=0$.

Now we are ready to show that good reduced initial data corresponding to Ψ_0 on $\mathcal{N}_1$ with $\delta^* \Psi_0 \equiv 0$ are PP-wave reduced initial data. To prove this, it is enough to prove that the spacetime which uniquely develops from the initial data above (which are induced from reduced initial data with $\delta^* \Psi_0 \equiv 0$) is a PP-wave.

To find the spacetime that develops from these initial data, just put any of the quantities

$$\xi^A, X^A, \omega, U, \mu, \lambda, \kappa, \varepsilon, \alpha, \beta, \pi, \rho, \sigma,$$

$$\Psi_0, \Psi_1, \Psi_2, \Psi_3, \Psi_4(v, u, x^3, x^4)$$

at (v, u, x^3, x^4) equal to their values at $(v=0, u, x^3, x^4)$; in other words, just transport identically every quantity on $\mathcal{N}_1$ along integral curves of $\vec{n}=\partial/\partial v$, independently of v. Clearly the resulting spacetime will be vacuum (the Ricci and Bianchi identities are trivially checked) and will induce the above initial data on $S=\mathcal{N}_1 \cup \mathcal{N}_2 \cup Z$. Morover, by uniqueness (theorem 3), it will be the unique vacuum spacetime developing from the above initial data. Clearly the vector $\vec{n}=\partial/\partial v$ is a Killing vector for this spacetime (and is also parallel), and the resulting spacetime is flat on $I^+(\mathcal{N}_2')$ and $I^-(\mathcal{N}_2)$. Hence the spacetime is a PP wave.

This completes the proof of theorem 5. □

V. CONCLUSIONS

We have reviewed in this paper the general structure of exact colliding plane-wave solutions of the vacuum Einstein equations; and we have argued on the basis of previous work, both by the author and largely by others, that those solutions whose causality structures are stable against plane-symmetric perturbations will involve all-embracing spacelike curvature singularities bounding the spacetime in the future of the collision plane. We have given a detailed qualitative review of the well-known focusing effect of plane waves in both single and colliding plane-wave spacetimes, and by discussing and giving an alternative proof of a singularity theorem originally discovered by Tipler,[2] we have described how this focusing property makes inevitable the occurrence of singularities in *generic* plane-wave collisions. We have carefully stressed the subtle aspects of Tipler's singularity theorem and emphasized the reason for its inapplicability to single plane-wave spacetimes and to colliding plane-wave solutions which possess Killing-Cauchy horizons.[3,4]

We have defined and analyzed general gravitational-wave spacetimes and we have seen that the PP-wave solutions–a particular family of GW spacetimes–satisfy strong uniqueness theorems, much like the Kerr-Newman family which satisfies the well-known black-hole uniqueness results. We have pointed out the insight that these results give into the structures of almost-plane waves, which constitute a special case of gravitational-wave spacetimes. In particular we have seen that almost-plane waves must always exhibit diffraction, since by the classification theorem of Dautcourt[10] the only diffraction-free GW spacetimes are PP waves; and we have seen that almost-plane waves must leave behind "tails" in any region of space through which they have propagated, since the only GW spacetimes with a precisely "sandwiched" curvature distribution are the PP waves.

ACKNOWLEDGMENTS

The author wishes to thank Kip S. Thorne for a critical reading of the preliminary versions of this manuscript and for his guidance in its subsequent revision. This work was supported by the National Science Foundation under Grant No. AST 85-14911.

APPENDIX: NEWMAN-PENROSE EQUATIONS IN RATIONALIZED FORM

The Newman-Penrose equations as originally formulated[6] were based on the metric signature (+,–,–,–). When one adopts, instead, the signature (–,+,+,+), they assume the following "rationalized" form.

Commutation relations:

$$\Delta D - D\Delta = -(\gamma+\gamma^*)D - (\varepsilon+\varepsilon^*)\Delta + (\tau^*+\pi)\delta + (\tau+\pi^*)\delta^* , \tag{A1}$$

$$\delta D - D\delta = -(\alpha^*+\beta-\pi^*)D - \kappa\Delta + (\rho^*+\varepsilon-\varepsilon^*)\delta + \sigma\delta^* , \tag{A2}$$

$$\delta\Delta - \Delta\delta = \nu^* D - (\tau-\alpha^*-\beta)\Delta - (\mu-\gamma+\gamma^*)\delta - \lambda^*\delta^* , \tag{A3}$$

$$\delta^*\delta - \delta\delta^* = (\mu-\mu^*)D + (\rho-\rho^*)\Delta + (\beta^*-\alpha)\delta + (\alpha^*-\beta)\delta^* . \tag{A4}$$

Ricci identities:

$$\delta^{*} K - D\,\rho = (\rho^{2} + \sigma\sigma^{*}) + \rho(\varepsilon + \varepsilon^{*}) - \kappa^{*}\tau$$

$$-\kappa(3\alpha + \beta^{*} - \pi) - \Phi_{00}\,, \tag{A5}$$

$$\delta\kappa - D\,\sigma = \sigma(\rho + \rho^{*} + 3\varepsilon - \varepsilon^{*})$$

$$-\kappa(\tau - \pi^{*} + \alpha^{*} + 3\beta) - \Psi_{0}\,, \tag{A6}$$

$$\Delta\kappa - D\,\tau = (\tau + \pi^{*})\rho + (\tau^{*}\pi)\sigma + (\varepsilon - \varepsilon^{*})\tau$$

$$-(3\gamma + \gamma^{*})\kappa - \Psi_{1} - \Phi_{01}\,, \tag{A7}$$

$$\delta^{*}\varepsilon - D\,\alpha = (\rho + \varepsilon^{*} - 2\varepsilon)\alpha + \beta\sigma^{*} - \beta^{*}\varepsilon - \kappa\lambda$$

$$-\kappa^{*}\gamma + (\varepsilon + \rho)\pi - \Phi_{10}\,, \tag{A8}$$

$$\delta\varepsilon - D\,\beta = (\alpha + \pi)\sigma + (\rho^{*} - \varepsilon^{*})\beta - (\mu + \gamma)\kappa$$

$$-(\alpha^{*} - \pi^{*})\varepsilon - \Psi_{1}\,, \tag{A9}$$

$$\Delta\varepsilon - D\,\gamma = (\tau + \pi^{*})\alpha + (\tau^{*} + \pi)\beta - (\varepsilon + \varepsilon^{*})\gamma$$

$$-(\gamma + \gamma^{*})\varepsilon + \tau\pi - \nu\kappa - \Psi_{2} + \Lambda - \Phi_{11}\,, \tag{A10}$$

$$\delta^{*}\pi - D\,\lambda = (\rho\lambda + \sigma^{*}\mu) + \pi^{2} + (\alpha - \beta^{*})\pi - \nu\kappa^{*}$$

$$-(3\varepsilon - \varepsilon^{*})\lambda - \Phi_{20}\,, \tag{A11}$$

$$\delta\pi - D\,\mu = (\rho^{*}\mu + \sigma\lambda) + \pi\pi^{*} - (\varepsilon + \varepsilon^{*})\mu$$

$$-\pi(\alpha^{*} - \beta) - \nu\kappa - \Psi_{2} - 2\Lambda\,, \tag{A12}$$

$$\Delta\pi - D\nu = (\pi+\tau^*)\mu + (\pi^*+\tau)\lambda + (\gamma-\gamma^*)\pi$$

$$-(3\varepsilon+\varepsilon^*)\nu - \Psi_3 - \Phi_{21}\,, \tag{A13}$$

$$\delta^*\nu - \Delta\lambda = -(\mu+\mu^*)\lambda - (3\gamma-\gamma^*)\lambda$$

$$+(3\alpha+\beta^*+\pi-\tau^*)\nu + \Psi_4\,, \tag{A14}$$

$$\delta^*\sigma - \delta\rho = \rho(\alpha^*+\beta) - \sigma(3\alpha-\beta^*) + (\rho-\rho^*)\tau$$

$$+(\mu-\mu^*)\kappa + \Psi_1 - \Phi_{01}\,, \tag{A15}$$

$$\delta^*\beta - \delta\alpha = (\mu\rho-\lambda\sigma) + \alpha\alpha^* + \beta\beta^* - 2\alpha\beta$$

$$+\gamma(\rho-\rho^*) + \varepsilon(\mu-\mu^*) + \Psi_2 - \Lambda - \Phi_{11}\,, \tag{A16}$$

$$\delta^*\mu - \delta\lambda = (\rho-\rho^*)\nu + (\mu-\mu^*)\pi + \mu(\alpha+\beta^*)$$

$$+\lambda(\alpha^*-3\beta) + \Psi_3 - \Phi_{21}\,, \tag{A17}$$

$$\Delta\mu - \delta\nu = (\mu^2+\lambda\lambda^*) + (\gamma+\gamma^*)\mu - \nu^*\pi$$

$$+(\tau-3\beta-\alpha^*)\nu - \Phi_{22}\,, \tag{A18}$$

$$\Delta\beta - \delta\gamma = (\tau-\alpha^*-\beta)\gamma + \mu\tau - \sigma\tau - \sigma\nu - \varepsilon\nu^*$$

$$-\beta(\gamma-\gamma^*-\mu) + \alpha\lambda^* - \Phi_{12}\,, \tag{A19}$$

$$\Delta\sigma - \delta\tau = (\mu\sigma+\lambda^*\rho) + (\tau+\beta-\alpha^*)\tau$$

$$-(3\gamma-\gamma^*)\sigma - \kappa\nu^* - \Phi_{02}\,, \tag{A20}$$

$$\delta^*\tau - \Delta\rho = -(\rho\mu^*+\sigma\lambda) + (\beta^*-\alpha-\tau^*)\tau$$

$$+(\gamma+\gamma^*)\rho+\nu\kappa+\Psi_2+2\Lambda\ , \tag{A21}$$

$$\delta^*\gamma-\Delta\alpha=(\rho+\varepsilon)\nu-(\tau+\beta)\lambda+(\gamma^*-\mu^*)\alpha$$

$$+(\beta^*-\tau^*)\gamma+\Psi_3\ . \tag{A22}$$

Bianchi identities (in vacuum):

$$\delta^*\Psi_0-D\,\Psi_1=-3\kappa\Psi_2+2(\varepsilon+2\rho)\Psi_1$$

$$+(\pi-4\alpha)\Psi_0\ , \tag{A23}$$

$$\delta^*\Psi_1-D\,\Psi_2=-2\kappa\Psi_3+3\rho\Psi_2+2(\pi-\alpha)\Psi_1$$

$$-\lambda\Psi_0\ , \tag{A24}$$

$$\delta^*\Psi_2-D\,\Psi_3=-\kappa\Psi_4-2(\varepsilon-\rho)\Psi_3+3\pi\Psi_2$$

$$-2\lambda\Psi_1\ , \tag{A25}$$

$$\delta^*\Psi_3-D\,\Psi_4=-(4\varepsilon-\rho)\Psi_4+(4\pi+2\alpha)\Psi_3$$

$$-3\lambda\Psi_2\ , \tag{A26}$$

$$\delta\Psi_1-\Delta\Psi_0=(4\gamma-\mu)\Psi_0-(4\tau+2\beta)\Psi_1$$

$$+3\sigma\Psi_2\ , \tag{A27}$$

$$\delta\Psi_2-\Delta\Psi_1=\nu\Psi_0+2(\gamma-\mu)\Psi_1-3\tau\Psi_2$$

$$+2\sigma\Psi_3\ , \tag{A28}$$

$$\delta\Psi_3-\Delta\Psi_2=2\nu\Psi_1-3\mu\Psi_2-2(\tau-\beta)\Psi_3$$

$$+\sigma\Psi_4\ , \tag{A29}$$

$$\delta\Psi_4 - \Delta\Psi_3 = 3\nu\Psi_2 - (2\gamma + 4\mu)\Psi_3$$

$$-(\tau - 4\beta)\Psi_4\ . \tag{A30}$$

REFERENCES FOR CHAPTER 3

[1] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).

[2] F. J. Tipler, Phys. Rev. D **22**, 2929 (1980).

[3] S. Chandrasekhar and B. Xanthopoulos, Proc. R. Soc. London (to be published).

[4] U. Yurtsever, Phys. Rev. D **36**, 1662 (1987) (Chapter 2).

[5] S. Chandrasekhar and B. Xanthopoulos, Proc. R. Soc. London A (to be published).

[6] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

[7] R. Penrose, Gen. Relativ. Gravit. **12**, 225 (1980).

[8] H. Muller zum Hagen and H.-J. Seifert, Gen. Relativ. Gravit. **8**, 259 (1977).

[9] H. Friedrich, Proc. R. Soc. London **A375**, 169 (1981); H. Friedrich and M. Stewart, Max Planck Institute report 1983 (unpublished).

[10] G. Dautcourt in *Conference Internationale sur les Theories Relativistes de la Gravitation* (Pergamon, Oxford, 1964).

[11] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[12] S. Chandrasekhar, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[13] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[14] K. Khan and R. Penrose, Nature (London) **229**, 185 (1971).

[15] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

[16] Y. Nutku and M. Halil, Phys. Rev. Lett. **39**, 1379 (1977).

[17] P. Bell and P. Szekeres, Gen. Relativ. Gravit. **5**, 275 (1974); S. Chandrasekhar and B. Xanthopoulos, Proc. R. Soc. London **A398**, 223 (1985).

[18] R. A. Matzner and F. J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[19] R. Penrose, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[20] U. Yurtsever (in preparation).

[21] D. Christodoulou, Commun. Math. Phys. (to be published).

[22] F. John, *Partial Differential Equations* (Springer, NewYork, 1982).

[23] J. Smoller, *Shock Waves and Reaction-Diffusion Equations* (Springer, NewYork, 1983).

[24] F. A. E. Pirani, in *Lectures on General Relativity*, Brandeis Summer Institute in Theoretical Physics (Prentice-Hall, Englewood Cliffs, New Jersey, 1965), Vol. 1.

FIGURE CAPTIONS FOR CHAPTER 3

FIG. 1. Geometry of the maximally extended colliding plane-wave spacetime in the anastigmatic case. The null surfaces $\mathcal{N}_1$ and $\mathcal{N}_2$ represent the wave front(s) ($\mathcal{N}_1$, $\mathcal{N}_1'$ and $\mathcal{N}_2$, $\mathcal{N}_2'$) of the incoming colliding plane waves. Since the plane waves depicted in the above figure are impulsive, for both waves the past and future wave fronts are identified ($\mathcal{N}_1=\mathcal{N}_1'$, $\mathcal{N}_2=\mathcal{N}_2'$) and hence are indistinguishable. Because of the focusing of each wave by the other wave, the wave fronts $\mathcal{N}_1$ and $\mathcal{N}_2$ are represented by null cones in the future of the collision plane. The flat regions II and III lie under the cones $\mathcal{N}_2$, $\mathcal{N}_1$, respectively. The respective Cauchy horizons $\mathcal{S}_1$ and $\mathcal{S}_2$ of the incoming plane waves are completely cutoff from the spacetime, except for those points which lie on the common generators of $\mathcal{S}_1$ and $\mathcal{S}_2$ with the null cones $\mathcal{N}_2$ and $\mathcal{N}_1$, respectively. However, these points are also singular since they do not possess a regular spacetime neighborhood which is isolated from the curvature singularity. Note that the coordinate system can always be arranged, by a Lorentz transformation, so that the collision is headon as in the above figure. This figure, which was drawn by R. Penrose and published in Ref. 18 by R. Matzner and F. Tipler, is reproduced here by the kind permissions of Penrose, Matzner, and Tipler.

FIG. 2. Colliding plane sandwich waves.

FIG. 3. Characteristic initial-value problem.

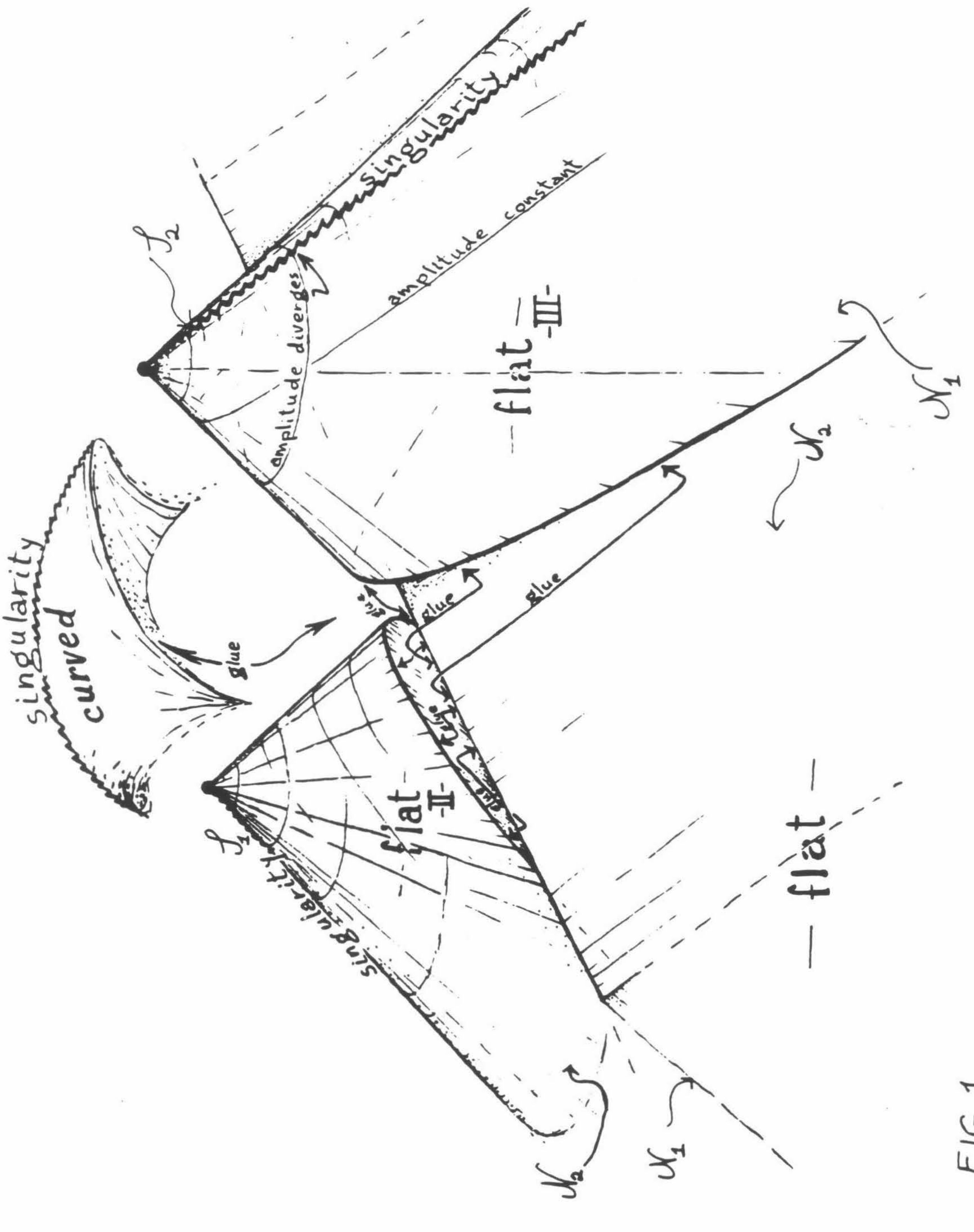
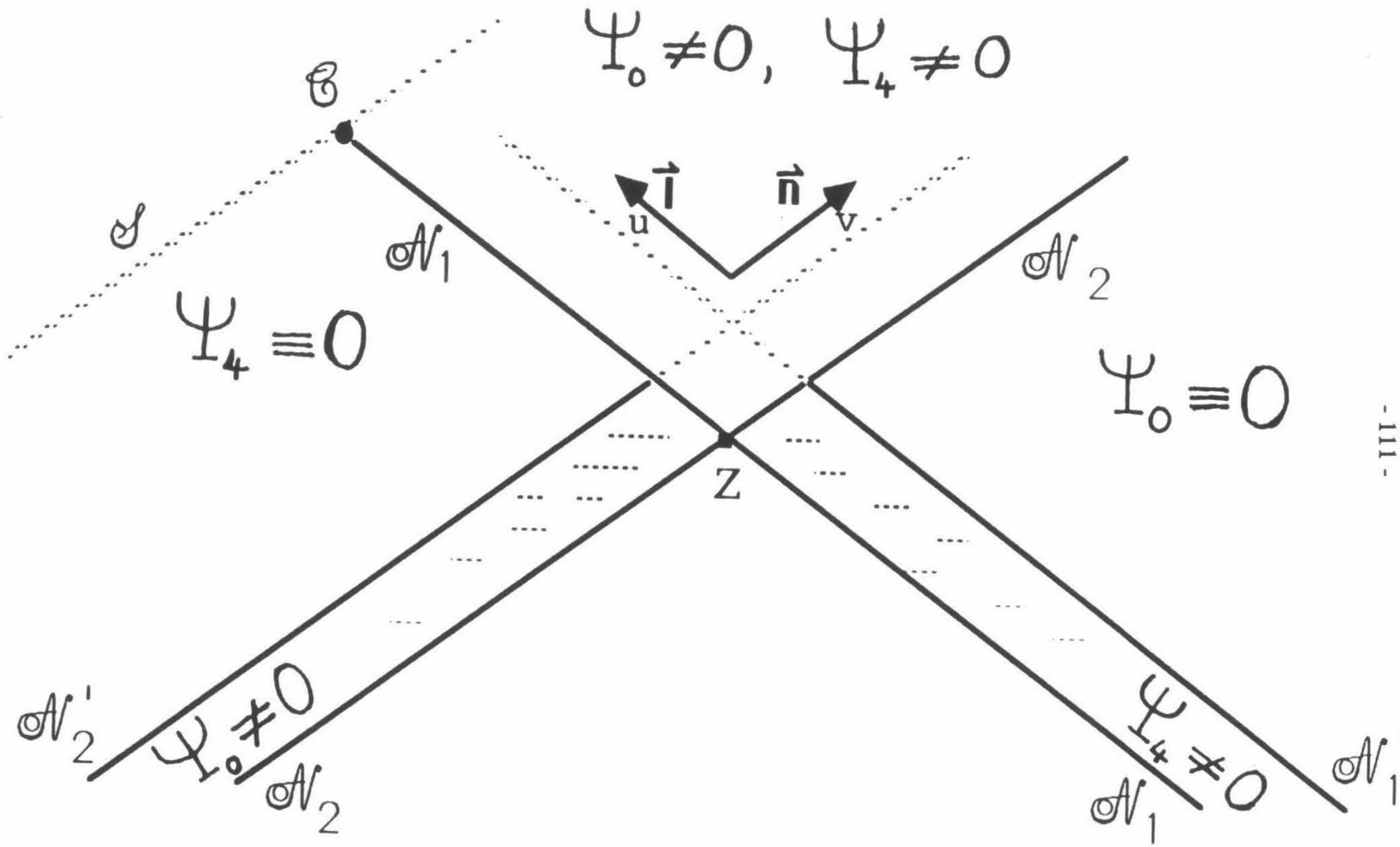

singularity
curved
glue
singularity
singularity
amplitude diverges
amplitude constant
flat
-II-
flat
-III-
glue
glue
— flat —

FIG.1



$\Psi_0 \neq 0, \ \Psi_4 \neq 0$
$\mathscr{C}$
$\mathscr{S}$
$\vec{l}$
$\vec{n}$
u
v
$\mathscr{N}_1$
$\mathscr{N}_2$
$\Psi_4 \equiv 0$
$\Psi_0 \equiv 0$
Z
$\mathscr{N}_2'$
$\Psi_0 \neq 0$
$\mathscr{N}_2$
$\Psi_4 \neq 0$
$\mathscr{N}_1$
$\mathscr{N}_1'$

FIG. 2

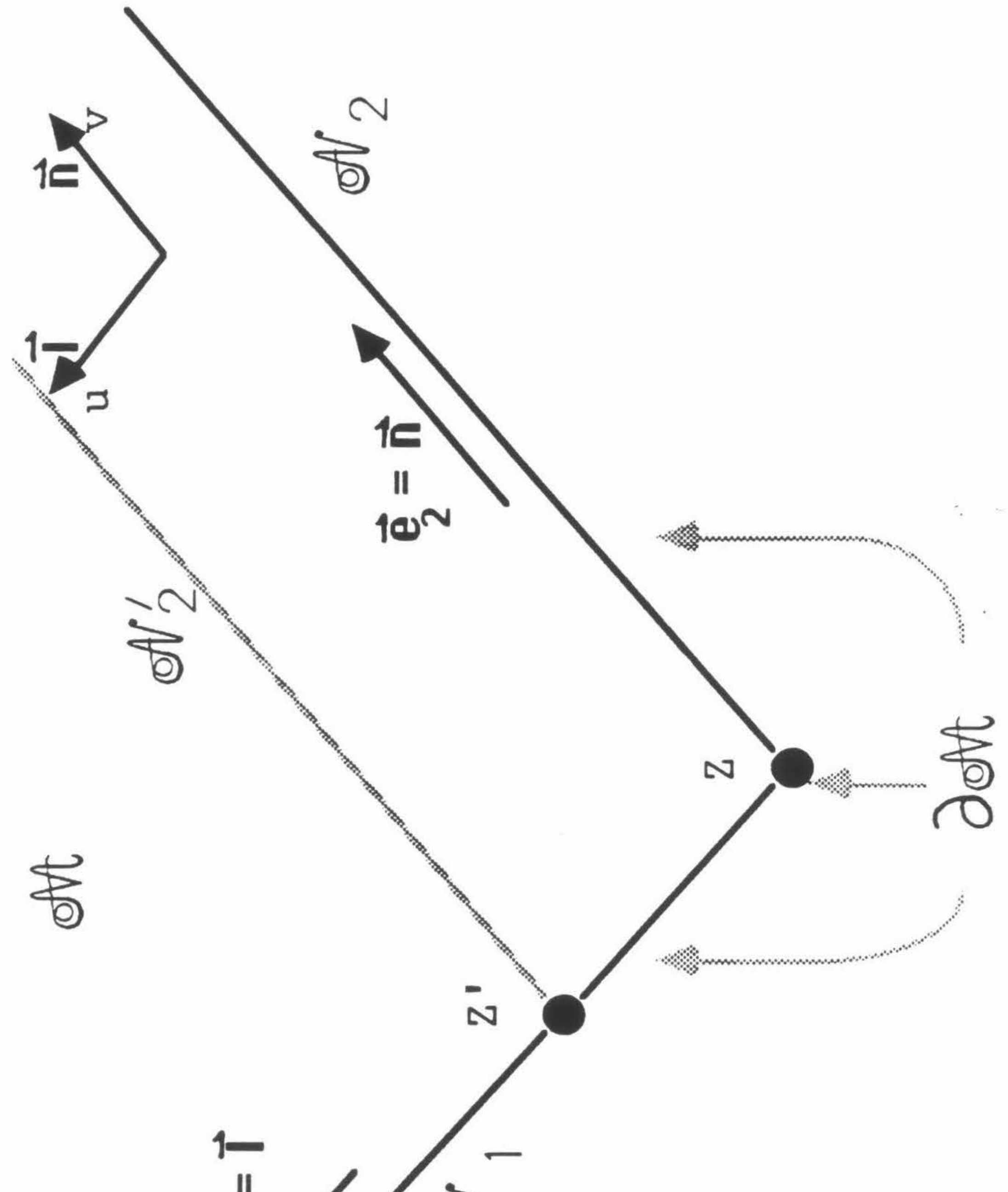


FIG.3

CHAPTER 4

A New Family of Exact Solutions for Colliding Plane Gravitational Waves

Published in Physical Review D **37**, 2790 (1988).

ABSTRACT

We construct an infinite-parameter family of exact solutions to the vacuum Einstein field equations describing colliding gravitational plane waves with parallel polarizations. The interaction regions of the solutions in this family are locally isometric to the interiors of those static axisymmetric (Weyl) black-hole solutions which admit both a nonsingular horizon, and an analytic extension of the exterior metric to the interior of the horizon. As a member of this family of solutions we also obtain, for the first time, a colliding plane-wave solution where both of the two incoming plane waves are purely anastigmatic, i.e., where both incoming waves have equal focal lengths.

I. INTRODUCTION

As a result of the revolutionary new techniques introduced in the last decade, there now exists an extensive collection of powerful tools to generate exact solutions for the Einstein field equations in the stationary axisymmetric case.[1] More recently, there have been successful attempts to employ these same techniques in the study of solutions with two commuting spacelike Killing vectors, i.e., in the study of plane-symmetric solutions to Einstein equations. These recent investigations have produced a rich arsenal of new exact solutions for plane-symmetric spacetimes; among these are many new solutions describing both colliding purely gravitational plane waves and colliding plane waves coupled with matter or radiation.[2]

Historically, the work on exact solutions for colliding plane waves has followed two distinct paths of development: On the one hand, the problem can be formulated as a characteristic initial-value problem for a system of nonlinear hyperbolic partial differential equations in two variables. This system involves the metric coefficients (and in the nonvacuum case the components of the matter fields) in a coordinate system where the two plane-symmetry-generating Killing vectors are equal to two members of the coordinate basis frame, so that the unknown variables are functions of the retarded and advanced time coordinates u and v only. The initial data for the metric coefficients (and the matter fields) are posed on the initial null boundary consisting of intersecting null surfaces $\mathcal{N}_1$ and $\mathcal{N}_2$, the past wavefronts of the two incoming waves (Fig. 1). The integration of this initial-value system to obtain the metric coefficients in the interaction region bounded by $\mathcal{N}_1 \cup \mathcal{N}_2$ is very difficult in general, in fact no general expression has been found for the solution in the generic case of colliding plane waves with nonparallel polarizations. However, in a paper of great ingenuity, Szekeres[3] was able to reduce the integration of arbitrary initial data for incoming gravitational plane waves with *parallel polarizations*, to the evaluation of a one-dimensional integral followed by two quadratures (see also Ref. 4 for another viewpoint). Despite this feat, however, the functions to be subjected to these elementary operations of integration and quadrature turned out to be very complex for general initial data. Consequently, exact solutions which were expressible in closed analytic form could only be obtained using this approach for a few very special incoming wave forms.

A very different and innovative alternative to the above approach for obtaining exact solutions of colliding plane waves was pioneered by the work of Khan and Penrose.[5] The idea is simply to work backward in time: (i) look for solutions to the field equations which have two commuting spacelike Killing vectors $\vec{\xi}_1$, $\vec{\xi}_2$, (ii) express the solutions in a coordinate system (u,v,x,y) where u,v are null coordinates and $\vec{\xi}_i$ are given by $\partial/\partial x^i$, and (iii) see whether it is possible to extend these solutions across the null surfaces $\mathcal{N}_1=\{u=0\}$ and $\mathcal{N}_2=\{v=0\}$ in such a way that the extension still satisfies the field equations, and that the extended metric in regions II and III (Fig. 1) describes single plane waves propagating in the appropriate directions. This technique of generating exact solutions for colliding plane waves was elevated into an art form over the recent years by the work of Chandrasekhar and Xanthopoulos,[2] who have obtained not only many new solutions describing colliding plane waves with parallel polarizations coupled with matter sources, but have also obtained new exact solutions for colliding plane waves with nonparallel polarizations, which display several unexpected novel features.[6,7] It is this technique that we use in the present paper to construct our solutions; consequently we shall describe it in more detail in the subsequent sections. Here we just remark that, as it is possible in principle to use different prescriptions for extending the metric beyond the interaction region, the alternative approach we just described will in general yield several different colliding plane-wave solutions which all have the same geometry in the interaction region I, but which for each different extension describe different incoming wave forms in the regions II and III (Fig. 1). This is in contrast with the direct method where one integrates the initial data posed by the incoming plane waves and obtains a unique colliding plane-wave spacetime. The reason for this behavior is that the same solution in the interaction region may evolve from several inequivalent sets of initial data, whereas the outcome from the direct method of integrating given initial data is constrained to be unique by the well-known uniqueness results for hyperbolic systems.

For the solutions constructed in this paper, the metric in the interaction region of the colliding plane-wave spacetimes is obtained from the interiors of the static, axisymmetric "distorted black hole" (Weyl) solutions which possess an interior. Every Weyl solution of this kind has a pair of commuting spacelike Killing vectors defined throughout its interior region. The simplest example of such Weyl solutions is the Schwarzschild spacetime. The construction by which we build our colliding plane-wave spacetimes is described in detail in the next section (Sec. II) for the Schwarzschild metric, along with a discussion of the properties of the resulting colliding plane-wave solution. Then in Sec. III we discuss the generalization of this construction to the infinite-parameter family of Weyl solutions which satisfy our regularity requirements; this generalization yields a corresponding infinite-parameter family of colliding plane-wave spacetimes. In Sec. IV two specific examples of spacetimes in this family are described briefly. The first of these examples is generated from one of the simplest nonspherical Weyl solutions in our family; this Weyl solution can be interpreted as the interior metric of a Schwarzschild black hole distorted by a static, quadrupolar matter distribution outside the horizon. The second example describes a colliding plane-wave spacetime where both of the two incoming plane waves are purely anastigmatic, i.e., where both incoming waves have equal focal lengths.[8,9] In Sec. V we recapitulate our conclusions by briefly listing both the new features and the drawbacks of the solutions that we have constructed. We also discuss some open questions and suggestions for future research on the issues raised by the present work.

It is not the purpose of this paper to discuss either the physical interpretation of colliding plane wave solutions or the significance of these solutions for general relativity in a wider context. The reader is referred to Refs. 3, 5, 10, 7, 9, and 4 and the extensive literature cited therein for a detailed exposition of these issues.

II. THE SOLUTION OBTAINED FROM THE SCHWARZSCHILD METRIC

We first write the Schwarzschild metric inside the horizon (i.e., for $r<2M$) as

$$g=-\left(\frac{1}{\frac{2M}{r}-1}\right)dr^2+\left(\frac{2M}{r}-1\right)dt^2$$

$$+r^2(d\theta^2+\sin^2\theta\, d\phi^2)\ . \tag{2.1}$$

Clearly, in this interior region where $r<2M$ the commuting Killing vectors $\partial/\partial t$ and $\partial/\partial\phi$ are both spacelike. We therefore introduce, new coordinates (x,y,u,v) tuned to the plane symmetry generated by these Killing vectors, by the following transformation (again for $r<2M$):

$$t=x\ ,\ \ \phi=(1+y/M)\ ,$$

$$\theta=\frac{\pi}{2}+(v-u)\ ,\ \ r=M[1-\sin(u+v)]\ . \tag{2.2}$$

In this new coordinate system the metric (2.1) takes the form

$$g = -4M^2[1-\sin(u+v)]^2 du\, dv$$

$$+\left(\frac{1+\sin(u+v)}{1-\sin(u+v)}\right) dx^2$$

$$+[1-\sin(u+v)]^2 \cos^2(u-v) dy^2 , \tag{2.3}$$

which explicitly displays the plane-symmetry generating, commuting, spacelike Killing vectors $\vec{\xi}_1=\partial/\partial x$ and $\vec{\xi}_2=\partial/\partial y$. We take the spacetime region $\{u \geq 0, v \geq 0, -\infty < x < +\infty, -\infty < y < +\infty\}$ with the metric (2.3) on it as the interaction region I of our colliding plane-wave solution. Note that, even though this interaction region is locally isometric to the region

$$J=\{r < \min [\, M(1+\cos\theta)\, , M(1-\cos\theta)\,]\, ,$$

$$-\infty < t < +\infty,\ 0 \leq \phi < 2\pi\}$$

of the Schwarzschild spacetime (this region J is depicted in Fig. 2), we will in effect have changed the topology of the underlying manifold from $S^2 \times R^2$ to R^4 by means of (i) extending the metric (2.3) across the surfaces $\{u=0\}$, $\{v=0\}$ (nonanalytically) in the manner described below, and (ii) by applying the coordinate transformation (2.2) in which y and $v-u$ are not regarded as periodic whereas ϕ and θ are. More specifically, by our non-analytic extension we shall eliminate the (coordinate) singularities of the (u, v, x, y) chart at $v-u=2\pi k \pm \pi/2$ (where k is any integer) that would show up in the maximal *analytic* extension, and thereby we shall transform the topology from $S^2 \times R^2$ to $S^1 \times R^3$. Subsequently, since $\partial/\partial\phi = M\, \partial/\partial y$ is Killing, the resulting metric on $S^1 \times R^3$ can be lifted to the covering space R^4 as described by the coordinate change (2.2), and this yields us the metric (2.3) defined on R^4.

We extend the metric (2.3) across the wave fronts $\{u=0\}$ and $\{v=0\}$ by the Penrose prescription[5,2] $u/a \to (u/a)H(u/a)$, $v/b \to (v/b)H(v/b)$ where $H(x)$ denotes the Heaviside step function and we have introduced two length scales a and b into the problem by putting $u \equiv u'/a$, $v \equiv v'/b$ where $ab=4M^2$, and we have redefined u' as u and v' as v. Thereby we obtain the following final metric for our colliding plane-wave spacetime:

$$g=-\left\{1-\sin\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]\right\}^2 du dv+\frac{1+\sin\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]}{1-\sin\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]}dx^2$$

$$+\left\{1-\sin\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]\right\}^2\cos^2\left[\frac{u}{a}H\left(\frac{u}{a}\right)-\frac{v}{b}H\left(\frac{v}{b}\right)\right]dy^2 . \qquad (2.4)$$

The geometry of this spacetime is depicted in Fig. 3, which describes a two-dimensional subspace given by $\{x=\text{const}, y=\text{const}\}$. (Actually the geometry is more subtle than this two-dimensional projection indicates; see Refs. 11 and 9.) A curvature singularity is present at $(u/a)+(v/b)=\pi/2$; it corresponds to the curvature singularity of the interior Schwarzschild spacetime at $r=0$. The extended spacetime consists of four regions where the metric is analytic: region I, where $u>0$, $v>0$, is the interaction region in which the metric is given by Eq. (2.3); regions II and III, where $u>0$, $v<0$ and $u<0$, $v>0$ respectively, represent the two incoming plane waves; region IV, where $u<0$, $v<0$, is the flat Minkowskian region representing the spacetime before the arrival of either wave. The only vector fields that are Killing vectors on the whole spacetime are $\partial/\partial x$ and $\partial/\partial y$ (and their constant linear combinations), whereas there exist two more R-linearly independent (i.e., linearly independent over the reals) spacelike Killing vectors in the interaction region I [Eq. (2.3)]; these extra Killing vectors correspond to the generators of spherical symmetry for the interior Schwarzschild metric (2.1). These vector fields cannot be extended as Killing vectors to the rest of the spacetime [Eq. (2.4)]. For the generalized solutions that we describe in the next section, $\partial/\partial x$ and $\partial/\partial y$ (and their constant linear combinations) are the only Killing vectors in the interaction region I, since the isometry group of the distorted, static, axisymmetric Weyl solutions is in general two dimensional. The solution (2.4) and also its generalizations described in Sec. III represent colliding plane waves with parallel polarizations, since the $x-y$ part of the metric [Eq. (2.3)] in the interaction region I is in diagonal form at all points; or equivalently since the Killing vectors $\partial/\partial x$ and $\partial/\partial y$ are hypersurface orthogonal throughout the spacetime.

According to Eq. (2.4), the metric in region II is

$$
\begin{aligned}
g_{\mathrm{II}} = &-[1-\sin(u/a)]^2 du\, dv \\
&+\left(\frac{1+\sin(u/a)}{1-\sin(u/a)}\right) dx^2 \\
&+[1-\sin(u/a)]^2 \cos^2(u/a) dy^2 ,
\end{aligned}
\tag{2.5}
$$

which entails a curvature singularity at the null surface $\{u=\pi a/2\}$. The metric g_{III} in region III is obtained by replacing u/a with v/b in Eq. (2.5) and similarly displays a curvature singularity at the null surface $\{v=\pi b/2\}$. Note that, in the most famous of the solutions for colliding plane waves[5,2,3] the corresponding null surfaces are also singular, but they do not represent curvature singularities. Instead, in those solutions, these surfaces correspond to the (nonsingular) focal planes (or Killing-Cauchy horizons[7]) of the respective incoming plane waves, and they become singular in the colliding plane-wave spacetime only because of the topological effect caused by the focusing of the plane wave moving in the opposite direction.[5,11,9] In the present case, however, these null surfaces are contained *within* the incoming plane sandwich waves, by contrast with the famous solutions where they are located in the flat regions lying to the future of the curvature disturbances associated with the incoming waves. Hence, for the solution (2.4) [see Eqs. (2.13)–(2.14) below], the curvature quantity Ψ_0 or Ψ_4 representing the radiative part of the Weyl tensor diverges on these surfaces. Physically, this could be considered a serious drawback of the solution (2.4), we expect a realistic spacetime representing a single gravitational wave propagating in empty space to be free of singularities of the above kind. However, it is possible to circumvent this difficulty by cutting off the gravitational radiation in each incoming plane wave along two null surfaces $\{u=u_c\}$ and $\{v=v_c\}$, where we can choose u_c and v_c to be arbitrarily close to $\pi a/2$ and $\pi b/2$, respectively. This results in the colliding plane-wave spacetime depicted in Fig. 4, where the metric in the regions denoted by I, II, III, and IV is exactly the same as the metric in the regions denoted by the same symbols in the original solution (2.4). Across the surfaces $\{u=u_c\}$ and $\{v=v_c\}$ the metric is C^1 but not C^2, making these surfaces shock fronts across which the curvature quantities Ψ_0 or Ψ_4 suffer jump discontinuities without delta-function contributions. (The structure of the field equations for a plane wave makes it possible to introduce such shocks at any desired null surface $\{u=\text{const}\}$; see, for example, Ref. 9.) The geometry in the regions denoted by IIa and IIIb in Fig. 4 is flat, and the surfaces $\{u=u_f\}$ and $\{v=v_f\}$ (where u_f and v_f are slightly larger than $\pi a/2$ and $\pi b/2$) correspond to the focal planes of the respective plane waves. These planes would be nonsingular if the collision were not taking place; the singularities at these focal planes are solely due to the topological effect of the focusing of the wave moving in the opposite direction.[11,9] The physics of this new solution in the interaction region is determined to an arbitrarily large extent by the metric (2.3) in the region I; even though the metric in regions Ia and Ib is not determined by Eqs. (2.3) or (2.4), by choosing u_c and v_c arbitrarily close to $\pi a/2$ and $\pi b/2$ it is possible to make the regions Ia and Ib arbitrarily small. Hence the colliding plane-wave solution (2.4) describes arbitrarily well the collision of the more "realistic" plane waves illustrated in Fig. 4.

We now turn to the proof of our implicit assertion that the metric (2.4) is indeed a genuine solution (in the sense of distributions) to the vacuum Einstein field equations. For this purpose, and also for spelling out the geometric structure of the solution (2.4) more clearly, we will find it useful to introduce the following null tetrad on our colliding plane-wave spacetime:

$$\vec{l}=2e^{M}\frac{\partial}{\partial u}\,,\quad \vec{n}=\frac{\partial}{\partial v}\,,$$

$$\vec{m}=N_1\frac{\partial}{\partial x}+N_2\frac{\partial}{\partial y}\,, \tag{2.6}$$

where

$$N_1=\frac{1+i}{2}e^{(U-V)/2}\,,$$

$$N_2=\frac{1-i}{2}e^{(U+V)/2}\,, \tag{2.7}$$

and M, U and V are functions of u and v only. The tetrad (2.6)—(2.7) gives rise to the metric

$$g=-e^{-M}du\,dv+e^{V-U}dx^2+e^{-(U+V)}dy^2\,. \tag{2.8}$$

Thus, the tetrad coefficients $M\,,U\,,V$ for the colliding plane-wave solution (2.4) are given by

$$M=-2\ln\left\{1-\sin\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]\right\}\,, \tag{2.9a}$$

$$U=-\ln\frac{1}{2}\left\{\cos\left[\frac{2u}{a}H\left(\frac{u}{a}\right)\right]\right.$$

$$\left.+\cos\left[\frac{2v}{b}H\left(\frac{v}{b}\right)\right]\right\} \tag{2.9b}$$

$$V=\ln\cos\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]$$

$$-\ln\cos\left[\frac{u}{a}H\left(\frac{u}{a}\right)-\frac{v}{b}H\left(\frac{v}{b}\right)\right]$$

$$-2\ln\left\{1-\sin\left[\frac{u}{a}H\left(\frac{u}{a}\right)+\frac{v}{b}H\left(\frac{v}{b}\right)\right]\right\}\,. \tag{2.9c}$$

The vacuum field equations for the metric (2.8) are[3,4]

$$2(U_{,uu}+M_{,u}U_{,u})-U_{,u}{}^2-V_{,u}{}^2=0\,, \tag{2.10a}$$

$$2(U_{,vv}+M_{,v}U_{,v})-U_{,v}{}^2-V_{,v}{}^2=0\,, \tag{2.10b}$$

$$U_{,uv}-U_{,u}U_{,v}=0\,, \tag{2.10c}$$

$$V_{,uv}-\frac{1}{2}(U_{,u}V_{,v}+U_{,v}V_{,u})=0\,, \tag{2.10d}$$

where the integrability condition for the first two equations is satisfied by virtue of the last two and yields the remaining field equation

$$M_{,uv}-\frac{1}{2}(V_{,u}V_{,v}-U_{,u}U_{,v})=0\ . \tag{2.11}$$

Therefore it is sufficient to solve Eqs. (2.10c) and (2.10d) first and to obtain M by quadrature from the first two equations (2.10a) and (2.10b) later, since Eq. (2.11) as well as the integrability condition for Eqs. (2.10a) and (2.10b) are automatically satisfied as a result of Eqs. (2.10c) and (2.10d).

We now proceed to verify that the field equations (2.10) and (2.11) are satisfied (in the sense of distributions) by our colliding plane-wave solution (2.4).

The field equations hold in the interaction region I (Fig. 3), since in this region (2.4) reduces to the metric (2.3), which is locally isometric to the interior Schwarzschild metric and thus is obviously vacuum.

In order to show that the field equations are satisfied in regions II and III, it is clearly sufficient to verify Eqs. (2.10) and (2.11) for the metric g_{II} given by Eq. (2.5), since the metric g_{III} in region III is locally isometric to g_{II} under the interchange $u \longleftrightarrow v$ [which incidentally is also a discrete isometry for the metric (2.3) in the interaction region]. This can be verified directly by substituting U,V and M for $u>0$, $v<0$ from Eqs. (2.9) into the left side of Eqs. (2.10) and (2.11); the result is easily shown to vanish. A more elegant approach, however, is to note that (i) verifying the field equations in regions II or III is equivalent to verifying the field equations for the *analytically extended* interaction region metric (2.3) *at* the null surfaces $\{u=0\}$, $\{v=0\}$; and (ii) the field equations for the metric (2.3) clearly hold at these null surfaces, because these equations hold *throughout* the analytically extended spacetime region covered by the (u,v,x,y) chart, and because this region contains the null surfaces $\{u=0\}$ and $\{v=0\}$ as nonsingular hypersurfaces.

The field equations hold in region IV since the metric in this region is flat.

To show that the field equations hold (in the distribution sense) on the boundaries between regions I and II and between regions I and III (Fig. 3), it is again sufficient to consider only the I—II boundary because of the $u \longleftrightarrow v$ symmetry of the problem. Now, since all field equations hold identically throughout region I and region II, they can only fail to hold on the boundary I—II if there are contributions to the left-hand side of Eqs. (2.10)—(2.11) which are nonzero only on this boundary and which are zero everywhere else. It is seen easily from the structure of the functions U, V, M displayed in Eqs. (2.9) that such contributions must involve δ-functions supported on the I—II boundary. However, as M, U, and V are functions of the arguments $(u/a)H(u/a)$ and $(v/b)H(v/b)$, the only way δ-function contributions can arise is by a two-times differentiation of U, M, or V with respect to either u or v, but not by a differentiation of the form $\partial_u \partial_v$. Therefore the last two field equations (2.10c) and (2.10d) as well as the integrability condition Eq. (2.11) automatically hold on any of the boundaries. On the boundary I—II, the first field equation (2.10a) holds trivially since this boundary is given by $\{v=0\}$ and Eq. (2.10a) contains only double u derivatives and thus cannot introduce $\delta(v)$ terms. The second equation (2.10b), however, can introduce $\delta(v)$ terms on the I—II boundary through the derivative $2U_{,vv}$. But a short calculation reveals that all $\delta(v)$ terms introduced by the differentiation $U_{,vv}$ are proportional to $\sin(2v/b)$, and thus they vanish on the I—II boundary on which $v=0$. This completes the proof that the field equations hold, in the sense of distributions, on the I—II boundary as well as on the boundary I—III between regions I and III.

The boundaries between regions II and IV and between regions III and IV (Fig. 3) are treated similarly. By the same arguments as above, and since on the II—IV boundary we have $u=0$, it is enough to show the nonexistence of $\delta(u)$ terms on the II—IV boundary. Such terms could only be introduced by the first field equation (2.10a), and the second field equation (2.10b) holds trivially on the II—IV boundary. Equation (2.10a) can introduce $\delta(u)$ terms only through the second derivative $U_{,uu}$; however, all the terms in this derivative, which involve delta functions, turn out to be proportional to $\sin(2u/a)$ and thus they vanish on the II—IV boundary on which $u=0$.

By either of the above boundary arguments, the field equations hold on the two-plane $\{u=v=0\}$. Moreover, since the coordinate system (u,v,x,y) regularly covers a neighborhood of the two-plane $\{u=v=0\}$, and since the metric coefficients in the (u,v,x,y) chart are continuous on the whole spacetime including this plane, no "conical-type" singularity can be present on the spacelike two-plane $\{u=v=0\}$.

In order to elucidate further the physics of our colliding plane-wave solution, we conclude this section with a brief discussion of the behavior of the spacetime curvature associated with the metric (2.4). The Newman-Penrose curvature quantities in the null tetrad (2.6)—(2.7) are given by[3,4]

$$\Psi_0=2ie^{2M}(M_{,u}V_{,u}+V_{,uu}-V_{,u}U_{,u}) , \tag{2.12a}$$

$$\Psi_1=0 , \tag{2.12b}$$

$$\Psi_2=-e^{M}M_{,uv} , \tag{2.12c}$$

$$\Psi_3=0 , \tag{2.12d}$$

$$\Psi_4=\frac{i}{2}[(U_{,v}-M_{,v})V_{,v}-V_{,vv}] . \tag{2.12e}$$

Substituting M, U and V from Eqs. (2.9) in the above equations, we straightforwardly obtain the following information about the behavior of the curvature quantities on our colliding plane-wave spacetime (2.4).

All nonzero curvature quantities in the interaction region I (Fig. 3) diverge towards the singularity $\{(u/a)+(v/b)=\pi/2\}$. The asymptotic behaviours of Ψ_0 , Ψ_2 and Ψ_4 near the singularity are all of the form $[(\pi/2)-(u/a)-(v/b)]^{-n}$ where $n=10$ for Ψ_0 , $n=6$ for Ψ_2, and $n=2$ for Ψ_4.

In region II (Fig. 3) the only nonzero curvature quantity is

$$\Psi_0=\frac{12i}{a^2}\,\frac{1}{[1-\sin(u/a)]^5} \tag{2.13}$$

whereas in region III the only nonzero curvature quantity is

$$\Psi_4=-\frac{3i}{b^2}\,\frac{1}{[1-\sin(v/b)]}\;. \tag{2.14}$$

In region IV all curvature quantities vanish.

On the I–II and I–III boundaries (Fig. 3) Ψ_2 has jump discontinuities which are finite but which diverge towards the singularity:

$$[\Psi_2]_{\text{I–II}}=-\frac{2}{ab}\,\frac{1}{[1-\sin(u/a)]^3}\;, \tag{2.15a}$$

$$[\Psi_2]_{\text{I–III}}=-\frac{2}{ab}\,\frac{1}{[1-\sin(v/b)]^3}\;. \tag{2.15b}$$

There are no δ-function contributions to the discontinuity of Ψ_2 along these boundaries. Along the I—II boundary Ψ_0 is continuous, whereas Ψ_4 has a jump

$$[\Psi_4]_{\text{I–II}}=-\frac{3i}{b^2}\,\frac{1}{[1-\sin(u/a)]} \tag{2.16a}$$

and also has a δ-function singularity of the form

$$-\frac{i}{b^2}\,\frac{1}{\cos(u/a)}\,\delta(v/b)\,. \tag{2.16b}$$

Along the I—III boundary Ψ_4 is continuous, whereas Ψ_0 has a jump

$$[\Psi_0]_{\text{I–III}}=\frac{12i}{a^2}\,\frac{1}{[1-\sin(v/b)]^5} \tag{2.17a}$$

and also has a δ-function singularity of the form

$$\frac{4i}{a^2}\,\frac{1}{\cos(v/b)[1-\sin(v/b)]^4}\,\delta(u/a)\,. \tag{2.17b}$$

Along the II–IV and III–IV boundaries (Fig. 3) Ψ_2 (being identically zero across these boundaries) is continuous. Across the II—IV boundary Ψ_4 (being zero) is continuous, whereas Ψ_0 has a jump

$$[\Psi_0]_{\text{II–IV}}=\frac{12i}{a^2}\,, \tag{2.18a}$$

and also has a δ-function singularity of the form

$$\frac{4i}{a^2}\delta\left(\frac{u}{a}\right)\,. \tag{2.18b}$$

Along the III—IV boundary Ψ_0 (being zero) is continuous, whereas Ψ_4 has a jump

$$[\Psi_4]_{\text{III–IV}}=-\frac{3i}{b^2}\,, \tag{2.19a}$$

and also has a δ-function singularity of the form

$$-\frac{i}{b^2}\delta\left(\frac{v}{b}\right)\,. \tag{2.19b}$$

III. THE SOLUTIONS OBTAINED FROM THE WEYL METRICS

The most general static axisymmetric spacetime with a regular axis has the metric[12]

$$g=-e^{2\psi}dt^2+e^{-2\psi}\rho^2d\phi^2+e^{2(\gamma-\psi)}(d\rho^2+dz^2)\ , \tag{3.1}$$

where (t,z,ρ,ϕ) are the cylindrical (Weyl) coordinates, and ψ and γ are functions of ρ and z only. The vacuum Einstein field equations for the metric (3.1) are

$$\psi_{,\rho\rho}+\frac{1}{\rho}\psi_{,\rho}+\psi_{,zz}=0\ , \tag{3.2a}$$

$$\gamma_{,\rho}=\rho(\psi_{,\rho}{}^2-\psi_{,z}{}^2)\ , \tag{3.2b}$$

$$\gamma_{,z}=2\rho\psi_{,\rho}\psi_{,z}\ , \tag{3.2c}$$

where Eq. (3.2a) is the integrability condition for the last two equations (3.2b) and (3.2c). The regularity of the axis $\rho=0$ requires that $\gamma=0$ at $\rho=0$. Thus, any solution $\psi(\rho,z)$ to Eq. (3.2a) uniquely determines a solution of the form (3.1) to the vacuum Einstein equations. For the Schwarzschild solution, ψ and γ are given by

$$\psi^S(\rho,z)=\frac{1}{2}\ln\left(\frac{\alpha-1}{\alpha+1}\right)\ , \tag{3.3a}$$

$$\gamma^S(\rho,z)=\frac{1}{2}\ln\left(\frac{\alpha^2-1}{\alpha^2-\mu^2}\right)\ , \tag{3.3b}$$

where

$$\mu=\frac{1}{2M}\{[\rho^2+(z-M)^2]^{1/2}-[\rho^2+(z+M)^2]^{1/2}\}\ , \tag{3.4}$$

and

$$\alpha=\frac{1}{2M}\{[\rho^2+(z-M)^2]^{1/2}+[\rho^2+(z+M)^2]^{1/2}\}\ . \tag{3.5}$$

The Schwarzschild coordinates (t, r, θ, ϕ) are related to the Weyl coordinates by

$$\rho=r\left(1-\frac{2M}{r}\right)^{1/2}\sin\theta\ ,\quad z=(r-M)\cos\theta\ ,$$

$$\phi=\phi\ ,\quad t=t\ . \tag{3.6}$$

The horizon $\{r=2M\}$ of the Schwarzschild spacetime corresponds to the surface $\{\rho=0, -M\leq z\leq M\}$ in Weyl coordinates. Note, however, that neither the horizon $\{r=2M\}$ nor the interior region where $r<2M$ is covered smoothly by the Weyl coordinate system.

In order to isolate those Weyl solutions which, like Schwarzschild spacetime, possess a nonsingular horizon and an interior region, we will find it convenient to define new metric functions $\hat{\psi}$ and $\hat{\gamma}$ by

$$\hat{\psi}\equiv\psi-\psi^S\ ,\quad \hat{\gamma}\equiv\gamma-\gamma^S\ . \tag{3.7}$$

Since the field equation (3.2a) for ψ is linear, it is satisfied in exactly the same form by the function $\hat{\psi}$. On the other hand, the field equations satisfied by $\hat{\gamma}$ as obtained from Eqs. (3.2b), (3.2c), and (3.3) are given by

$$\hat{\gamma}_{,\rho}=\rho\left(\hat{\psi}_{,\rho}{}^2-\hat{\psi}_{,z}{}^2+\frac{2}{\alpha^2-1}(\alpha_{,\rho}\hat{\psi}_{,\rho}-\alpha_{,z}\hat{\psi}_{,z})\right)\ , \tag{3.8a}$$

$$\hat{\gamma}_{,z}=2\rho\left(\hat{\psi}_{,\rho}\hat{\psi}_{,z}+\frac{1}{\alpha^2-1}(\alpha_{,z}\hat{\psi}_{,\rho}+\alpha_{,\rho}\hat{\psi}_{,z})\right)\ , \tag{3.8b}$$

where $\alpha(\rho,z)$ is defined by Eq. (3.5). (The mass M that enters into the definition of γ^S and ψ^S can be chosen arbitrarily, and in particular can be set equal to 1. However, we retain this free parameter M since it will be helpful when introducing length scales into the colliding plane-wave solutions that we are going to build shortly.) To solve the field equation (3.2a) satisfied by $\hat{\psi}$, we introduce spherical polar coordinates ν and η defined by

$$\rho=\nu\sin\eta\ ,\quad z=\nu\cos\eta\ .$$

The general solution of Eq. (3.2a) can now be written in terms of Legendre polynomials $P_k(x)$ (Ref. 13):

$$\hat{\psi}(\nu,\eta)=\sum_{k=0}^{\infty}(d_k\nu^k+c_k\nu^{-k-1})P_k(\cos\eta)\ . \tag{3.9}$$

For the time being, the coefficients d_k and c_k are simultaneously included in the above expression, because both asymptotic flatness and regularity of the horizon are irrelevant restrictions for our purposes. However, the terms involving the Legendre functions of the second kind, $Q_k(\cos\eta)$, are left out of the sum (3.9), since we assume that the axis on which $\cos\eta=\pm 1$ is nonsingular throughout spacetime. This assumption, together with the regularity condition that we impose below, will guarantee that the spacetime admits a nonsingular horizon which is located at $r=2M$ in the Schwarzschild-type coordinate system (3.6). Combining Eq. (3.1) with Eqs. (3.7), (3.3)—(3.5), and (3.6), we obtain the following general Weyl metric written in the Schwarzschild-type coordinates (t,r,θ,ϕ):

$$g=-\left(1-\frac{2M}{r}\right)e^{2\hat{\psi}(r,\theta)}dt^2+e^{-2\hat{\psi}(r,\theta)}r^2\sin^2\theta\, d\phi^2$$

$$+e^{2[\hat{\gamma}(r,\theta)-\hat{\psi}(r,\theta)]}\left(\frac{dr^2}{1-\frac{2M}{r}}+r^2d\theta^2\right) . \tag{3.10}$$

The functions $\hat{\psi}(r,\theta)$ and $\hat{\gamma}(r,\theta)$ in the above metric are calculated from the formulas [cf. Eqs. (3.8)—(3.9)]

$$\hat{\psi}(\rho,z)=\sum_{k=0}^{\infty}[d_k(\rho^2+z^2)^{k/2}+c_k(\rho^2+z^2)^{-(k+1)/2}]P_k\left(\frac{z}{\sqrt{\rho^2+z^2}}\right) , \tag{3.11}$$

$$\hat{\gamma}(\rho,z)=\int_C\rho\left[\left(\hat{\psi}_{,\rho}{}^2-\hat{\psi}_{,z}{}^2+\frac{2}{\alpha^2-1}(\alpha_{,\rho}\hat{\psi}_{,\rho}-\alpha_{,z}\hat{\psi}_{,z})\right)d\rho\right.$$

$$\left.+2\left(\hat{\psi}_{,\rho}\hat{\psi}_{,z}+\frac{1}{\alpha^2-1}(\alpha_{,z}\hat{\psi}_{,\rho}+\alpha_{,\rho}\hat{\psi}_{,z})\right)dz\right] \tag{3.12}$$

by substituting for ρ and z their expressions (3.6) as functions of r and θ. In Eq. (3.12) the function $\alpha(\rho,z)$ is defined by Eq. (3.5), and the line integral is evaluated on *any* contour C that starts on the axis $\rho=0$ (where $\hat{\gamma}$ vanishes), and that ends at the point (ρ,z) where $\hat{\gamma}$ is to be computed.[13]

Since we are interested in solutions with two spacelike Killing vectors, we now turn to the characterization of those Weyl solutions in the family (3.10)—(3.12) which possess an "interior region" in which $\partial/\partial t$ is spacelike. As we have noted before, the Weyl coordinates (t,ρ,z,ϕ) cannot cover the interior region regularly even if such a region exists. However, as the form of the metric (3.10) indicates clearly, the Schwarzschild-type coordinates (t,r,θ,ϕ) [which are defined formally by Eqs. (3.6)] will cover the interior region $r<2M$, whenever this region exists as a spacetime region with a well-defined metric (3.10). Moreover, these coordinates will cover the interior region $r<2M$ regularly, apart from the usual singularities associated with spherical coordinates. It is also clear that the interior region $r<2M$ will have a well-defined metric (3.10) if and only if the functions $\hat{\psi}(r,\theta)$ and $\hat{\gamma}(r,\theta)$ given by Eqs. (3.11) and (3.12) are well defined for $r<2M$. We now claim that in order for these functions $\hat{\psi}$ and $\hat{\gamma}$ be well defined for $r<2M$, it is necessary and sufficient that in Eqs. (3.9) and (3.11) all of the coefficients c_k vanish. This restriction on the general solution (3.9) is necessary for the existence of the interior region $r<2M$, because the expression

$$\rho^2+z^2 \equiv r^2-2Mr+M^2\cos^2\theta$$

assumes negative values at some points in the region $r<2M$; therefore we have to eliminate any term involving the product of an half-odd-integer (integer) power of ρ^2+z^2 with an even-indexed (odd-indexed) Legendre polynomial P_k from Eqs. (3.9) and (3.11). The sufficiency of the above condition for the existence of a well-defined metric (3.10) on the region $r<2M$ will become clear after the following discussion. We also note that the above restriction on the general solution (3.9) guarantees not only the existence of the interior region $\{r<2M\}$, but also the existence and regularity of the "horizon" $\{r=2M\}$.

We now have the following infinite-parameter family of interior Weyl solutions, defined on the region $\{r<2M\}$ where both of the two commuting Killing vectors $\partial/\partial t$ and $\partial/\partial\phi$ are spacelike:

$$g=\left(\frac{2M}{r}-1\right)e^{2\hat{\psi}(r,\theta)}dt^2+e^{-2\hat{\psi}(r,\theta)}r^2\sin^2\theta d\phi^2$$

$$+e^{2[\hat{\gamma}(r,\theta)-\hat{\psi}(r,\theta)]}\left(r^2d\theta^2-\frac{dr^2}{\frac{2M}{r}-1}\right) \quad \text{for } r<2M , \tag{3.13}$$

where

$$\hat{\psi}(r,\theta)=\sum_{k=0}^{\infty} d_k (r^2-2Mr+M^2\cos^2\theta)^{k/2}$$

$$\times P_k\left[\frac{(r-M)\cos\theta}{(r^2-2Mr+M^2\cos^2\theta)^{1/2}}\right], \tag{3.14}$$

$$\hat{\psi}(\rho,z)=\sum_{k=0}^{\infty} d_k (\rho^2+z^2)^{k/2} P_k\left[\frac{z}{\sqrt{\rho^2+z^2}}\right], \tag{3.15}$$

and where $\hat{\gamma}(r,\theta)$ is computed by inserting Eq. (3.15) into Eq. (3.12), using Eq. (3.5), and substituting for ρ and z their expressions (3.6) as functions of r and θ. In Eqs. (3.13)—(3.15) we have combined Eq. (3.11) with Eq. (3.6) to obtain Eq. (3.14).

To see that the functions $\hat{\psi}(r,\theta)$ and $\hat{\gamma}(r,\theta)$ defined by Eqs. (3.14), (3.15), (3.12), (3.5), and (3.6) are well defined and real for $r<2M$, note the following facts.

(i) The Legendre polynomials $P_{2n}(x)$ are polynomials in x^2 of order n. Hence, for all even k the expression

$$(r^2-2Mr+M^2\cos^2\theta)^{k/2}$$

$$\times P_k\left[\frac{(r-M)\cos\theta}{(r^2-2Mr+M^2\cos^2\theta)^{1/2}}\right]$$

is real, well defined, and finite for $r<2M$, even at the points where $(r^2-2Mr+M^2\cos^2\theta)$ is zero or negative. Similarly, $P_{2n+1}(x)$ is equal to the product of x with a polynomial in x^2 of order n, and $x^{2n+1}=x(x^2)^n$. Hence, also for all odd k the above expression is real, well defined, and finite for $r<2M$.

(ii) The integral (3.12) can be put into the form

$$\hat{\gamma}(\rho,z)=\int_C \Bigg[\left(\hat{\psi}_{,\rho}{}^2-\hat{\psi}_{,z}{}^2+\frac{2}{\alpha^2-1}(\alpha_{,\rho}\hat{\psi}_{,\rho}-\alpha_{,z}\hat{\psi}_{,z}) \right) d(\frac{1}{2}\rho^2)$$

$$+2\left(\rho\hat{\psi}_{,\rho}\hat{\psi}_{,z}+\frac{1}{\alpha^2-1}(\alpha_{,z}\rho\hat{\psi}_{,\rho}+\rho\alpha_{,\rho}\hat{\psi}_{,z}) \right) dz \tag{3.16}$$

where α is defined in Eq. (3.5) and the contour C is as in Eq. (3.12). Moreover, by Eq. (3.15) and because of the relation

$$\rho^2+(z\pm M)^2=[r-M(1\mp\cos\theta)]^2\ , \tag{3.17}$$

all of the expressions $\alpha_{,z}$, $\rho\alpha_{,\rho}$, $\alpha_{,\rho}\hat{\psi}_{,\rho}$, and $\alpha_{,z}\hat{\psi}_{,z}$ as well as the expressions $\rho\hat{\psi}_{,\rho}$, $\hat{\psi}_{,z}$, $\hat{\psi}_{,\rho}{}^2$, and $\hat{\psi}_{,z}{}^2$ which appear in Eq. (3.16) are well defined and real throughout the region $r<2M$.

(iii) By Eq. (3.17) and because of the fact that $\alpha^2-1\neq 0$ at all points in the interior region $r<2M$, all improper integrals that are involved in the evaluation of $\hat{\gamma}(\rho,z)$ [Eq. (3.16)] are convergent; and thus the integral (3.16) yields $\hat{\gamma}$ as a well defined and real function of ρ^2 and z.

Now that we have an infinite-parameter family of interior solutions described by Eqs. (3.13)—(3.15), we can turn to the construction of the corresponding family of colliding plane-wave spacetimes. This construction proceeds in exact parallel to Sec. II, where we constructed the colliding plane-wave solution (2.4) starting from the interior Schwarzschild metric (2.1). In fact, the interior Schwarzschild solution is the special case of the family of solutions (3.13)—(3.15) for which all of the parameters d_k are zero.

We build our infinite-parameter family of colliding plane-wave solutions by the following steps.

(i) We apply the coordinate transformation (2.2) to the generalized interior metric (3.13) whose metric coefficients are defined by Eqs. (3.14) and (3.15).

(ii) We introduce two length scales a and b into the resulting metric by defining $u=u'/a$ and $v=v'/b$ where $ab=4M^2$. We then redefine u' as u and v' as v. We also redefine our parameters d_k so that the new d_k are equal to the dimensionless quantities $M^k d_k$.

(iii) We then extend the resulting interaction-region metric across the wave fronts $\{u=0\}$ and $\{v=0\}$ by the Penrose prescription; i.e., we replace u/a by $(u/a)H(u/a)$, and v/b by $(v/b)H(v/b)$.

(iv) The resulting metric on the interaction region is locally isometric to the generalized interior metric (3.13). However, as a result of the above extension and the coordinate transformation (2.2), we change the topology of our solution from $S^2 \times R^2$ [which is the topology of the manifold on which the metric (3.13) is defined], to R^4 (which is the topology of our maximal colliding plane wave spacetime, see Sec. II for details).

For each choice of the parameters $\{d_k\}$, the above construction yields a unique colliding plane-wave solution. In the following equations we describe the metric of this solution in the interaction region (Fig. 1); the complete expression for the metric on the maximally extended spacetime is obtained by replacing each u/a by $(u/a)H(u/a)$ and each v/b by $(v/b)H(v/b)$ in Eqs. (3.18), (3.19), and (3.22) below:

$$g_{\mathrm{I}}=-e^{2[\hat{\gamma}(u,v)-\hat{\psi}(u,v)]}\left[1-\sin\left(\frac{u}{a}+\frac{v}{b}\right)\right]^2 du\, dv+e^{2\hat{\psi}(u,v)}\,\frac{1+\sin\left(\dfrac{u}{a}+\dfrac{v}{b}\right)}{1-\sin\left(\dfrac{u}{a}+\dfrac{v}{b}\right)}\,dx^2$$

$$+e^{-2\hat{\psi}(u,v)}\left[1-\sin\left(\frac{u}{a}+\frac{v}{b}\right)\right]^2\cos^2\left(\frac{u}{a}-\frac{v}{b}\right)dy^2\,, \tag{3.18}$$

where

$$\hat{\psi}(u,v)=\sum_{k=0}^{\infty}d_k\left[\sin^2\left(\frac{u}{a}+\frac{v}{b}\right)-\cos^2\left(\frac{u}{a}-\frac{v}{b}\right)\right]^{k/2}$$

$$\times P_k\left(\frac{\sin\left(\frac{u}{a}+\frac{v}{b}\right)\sin\left(\frac{u}{a}-\frac{v}{b}\right)}{\left[\sin^2\left(\frac{u}{a}+\frac{v}{b}\right)-\cos^2\left(\frac{u}{a}-\frac{v}{b}\right)\right]^{1/2}}\right)\,, \tag{3.19}$$

$$\hat{\psi}(\rho,z)=\sum_{k=0}^{\infty}d_k(\frac{1}{4}ab)^{-k/2}(\rho^2+z^2)^{k/2}P_k\left(\frac{z}{\sqrt{\rho^2+z^2}}\right)\,, \tag{3.20}$$

and $\hat{\gamma}(u,v)$ is evaluated (i) by inserting Eq. (3.20) into the integral given in Eq. (3.16) where the contour C is as in Eq. (3.12) and where the function $\alpha(\rho,z)$ is given by

$$\alpha(\rho,z)=\frac{1}{\sqrt{ab}}\{[\rho^2+(z-\tfrac{1}{2}\sqrt{ab})^2]^{1/2}$$

$$+[\rho^2+(z+\tfrac{1}{2}\sqrt{ab})^2]^{1/2}\}\,, \tag{3.21}$$

and (ii) by formally substituting

$$\rho^2\equiv-\frac{ab}{4}\cos^2\left(\frac{u}{a}+\frac{v}{b}\right)\cos^2\left(\frac{u}{a}-\frac{v}{b}\right)\,, \tag{3.22a}$$

$$z\equiv-\frac{\sqrt{ab}}{2}\sin\left(\frac{u}{a}+\frac{v}{b}\right)\sin\left(\frac{u}{a}-\frac{v}{b}\right)\,, \tag{3.22b}$$

into $\hat{\gamma}(\rho,z)$, which is a smooth function of ρ^2 and z.

The functions $\hat{\gamma}(u,v)$ and $\hat{\psi}(u,v)$ are smooth functions throughout the interaction region I (Fig. 3), and generically, the metric (3.18) has a curvature singularity at $(u/a)+(v/b)=\pi/2$.

The proof that the colliding plane-wave spacetimes constructed above are genuine solutions (in the sense of distributions) to the vacuum Einstein equations is provided by exactly the same arguments with which we have shown the solution (2.4) of Sec. II to be a genuine vacuum solution in the distribution sense. The crucial observation to note in this regard is that the metric function $U(u,v)$ [Eq. (2.8)] for any of the solutions in the above family (3.18)—(3.22) is given by precisely the same expression [Eq. (2.9b)] as the corresponding function for the solution (2.4) of Sec. II.

For completeness, we conclude this section by describing the interaction region I of our solutions in an alternative coordinate system defined by

$$u'/a=\sin(u/a)\ ,\ \ v'/b=\sin(v/b)\ ,$$

$$x'=x\ \ ,\ \ y'=y\ . \tag{3.23}$$

In the following, we omit the primes over the new coordinate functions. The interaction-region metric for the family of solutions (3.18)—(3.22) is expressed below in the new coordinates (3.23). The extension of the metric beyond the interaction region is again accomplished by the substitutions $u/a\rightarrow(u/a)H(u/a)$ and $v/b\rightarrow(v/b)H(v/b)$, and these substitutions result in a colliding plane-wave spacetime globally isometric to the corresponding spacetime (3.18)—(3.22) [even though the coordinate transformation (3.23) does not hold outside region I]. Thus, the following region-I expressions produce exactly the same family of colliding plane-wave solutions as above, written in the new coordinates (3.23):

$$g_{\mathrm{I}}=-e^{2(\hat{\gamma}-\hat{\psi})}\frac{\left[1-\frac{u}{a}\left(1-\frac{v^2}{b^2}\right)^{1/2}-\frac{v}{b}\left(1-\frac{u^2}{a^2}\right)^{1/2}\right]^2}{\left(1-\frac{u^2}{a^2}\right)^{1/2}\left(1-\frac{v^2}{b^2}\right)^{1/2}}du\,dv$$

$$+e^{2\hat{\psi}}\frac{1+\frac{u}{a}\left(1-\frac{v^2}{b^2}\right)^{1/2}+\frac{v}{b}\left(1-\frac{u^2}{a^2}\right)^{1/2}}{1-\frac{u}{a}\left(1-\frac{v^2}{b^2}\right)^{1/2}-\frac{v}{b}\left(1-\frac{u^2}{a^2}\right)^{1/2}}dx^2 \tag{3.24}$$

$$+e^{-2\hat{\psi}}\left[1-\frac{u}{a}\left(1-\frac{v^2}{b^2}\right)^{1/2}-\frac{v}{b}\left(1-\frac{u^2}{a^2}\right)^{1/2}\right]^2\left[\left(1-\frac{u^2}{a^2}\right)^{1/2}\left(1-\frac{v^2}{b^2}\right)^{1/2}+\frac{uv}{ab}\right]^2 dy^2\,,$$

where

$$\hat{\psi}(u,v)=\sum_{k=0}^{\infty}d_k\left\{\left[\frac{u}{a}\left(1-\frac{v^2}{b^2}\right)^{1/2}+\frac{v}{b}\left(1-\frac{u^2}{a^2}\right)^{1/2}\right]^2-\left[\left(1-\frac{u^2}{a^2}\right)^{1/2}\left(1-\frac{v^2}{b^2}\right)^{1/2}+\frac{uv}{ab}\right]^2\right\}^{k/2}$$

$$\times P_k\left(\frac{\frac{u^2}{a^2}-\frac{v^2}{b^2}}{\left\{\left[\frac{u}{a}\left(1-\frac{v^2}{b^2}\right)^{1/2}+\frac{v}{b}\left(1-\frac{u^2}{a^2}\right)^{1/2}\right]^2-\left[\left(1-\frac{u^2}{a^2}\right)^{1/2}\left(1-\frac{v^2}{b^2}\right)^{1/2}+\frac{uv}{ab}\right]^2\right\}^{1/2}}\right),$$

(3.25)

and $\hat{\gamma}(u,v)$ is evaluated (i) by inserting Eq. (3.20) into the integral given in Eq. (3.16), where the contour C is as in Eq. (3.12) and where $\alpha(\rho,z)$ is given by Eq. (3.21), and (ii) by formally substituting

$$\rho^2\equiv-\frac{1}{4}ab\,[1-(u^2/a^2)-(v^2/b^2)]^2\,, \tag{3.26a}$$

$$z \equiv -\frac{1}{2}\sqrt{ab}\,[(u^2/a^2)-(v^2/b^2)]\,, \tag{3.26b}$$

in $\hat{\gamma}(\rho,z)$ which is a smooth function of ρ^2 and z. The curvature singularity, which, in the generic case, constitutes an achronal future c-boundary in the interaction region of the solution (3.24), is located at

$$\{(u/a)[1-(v^2/b^2)]^{1/2}+(v/b)[1-(u^2/a^2)]^{1/2}=1\}$$

in the new coordinate system (3.23).

IV. EXAMPLES

As we have noted before, when all parameters d_k are zero, the general solution (3.18)—(3.22) reduces to the solution (2.4) of Sec. II that was obtained from the interior Schwarzschild metric. From now on, we will denote by the symbol $\{d_k\}$ the unique colliding plane wave solution (3.18)—(3.22) which corresponds to a given choice of the parameters d_k. When all d_k are zero except for the parameter d_0, the solution $\{d_k\}=\{d_0,0,0,0,...\}$ is again equal to (2.4), except in this case the mass $M=\sqrt{ab}/2$ (and the boost-invariant product $\sqrt{ab}$ of the characteristic wavelengths) is rescaled by a factor e^{-d_0} corresponding to a monopolar distortion of the solution (2.4).

To illustrate the evaluation of the function $\hat{\gamma}(u,v)$ by Eqs. (3.20)—(3.22), we write down below the functions $\hat{\psi}_2(u,v)$ and $\hat{\gamma}_2(u,v)$ corresponding to the solution $\{0,0,1,0,0,...\}$, where all d_k are zero except for $d_2=1$. The metric in the interaction region of this solution is locally isometric to an interior Weyl metric (3.13); this Weyl solution can be interpreted as the interior metric of a Schwarzschild black hole distorted by a static, quadrupolar matter distribution outside the horizon:

$$\hat{\psi}_2(u,v) = \frac{1}{2}\left[3\sin^2\left(\frac{u}{a}+\frac{v}{b}\right)\sin^2\left(\frac{u}{a}-\frac{v}{b}\right)+\cos^2\left(\frac{u}{a}-\frac{v}{b}\right)-\sin^2\left(\frac{u}{a}+\frac{v}{b}\right)\right], \quad (4.1)$$

$$\hat{\gamma}_2(u,v) = \frac{4}{a^2b^2}(q^2-8qz^2) - \frac{2}{\sqrt{ab}}[I_1(q;z^2-M^2, 2(z^2+M^2), (z-M)^2, (z^2-M^2)^2)$$

$$+I_1(q;z^2-M^2, 2(z^2+M^2), (z+M)^2, (z^2-M^2)^2)]$$

$$-\frac{4}{\sqrt{ab}}z\,[(z-M)I_2(q;z^2-M^2, 2(z^2+M^2), (z-M)^2, (z^2-M^2)^2)$$

$$+(z+M)I_2(q;z^2-M^2, 2(z^2+M^2), (z+M)^2, (z^2-M^2)^2)], \quad (4.2)$$

where

$$M \equiv \frac{1}{2}\sqrt{ab},$$

$$q \equiv -\frac{ab}{4}\cos^2\left(\frac{u}{a}+\frac{v}{b}\right)\cos^2\left(\frac{u}{a}-\frac{v}{b}\right),$$

$$z \equiv -\frac{\sqrt{ab}}{2}\sin\left(\frac{u}{a}+\frac{v}{b}\right)\sin\left(\frac{u}{a}-\frac{v}{b}\right), \quad (4.3)$$

$$I_1(q;a,b,c,d) \equiv \int_0^q \frac{s\,ds}{\sqrt{s+c}\,(s+a+\sqrt{s^2+bs+d})},$$

$$I_2(q;a,b,c,d) \equiv \int_0^q \frac{ds}{\sqrt{s+c}\,(s+a+\sqrt{s^2+bs+d})}.$$

We now turn to our second example of a colliding plane-wave solution in the family (3.18)—(3.22): a solution which describes colliding purely anastigmatic plane waves.[7–9] According to Eqs. (3.18)—(3.22), the metric g_{II} on the region II (Fig. 3) of a solution $\{d_k\}$ is given by

$$g_{\mathrm{II}} = -e^{[\hat{\gamma}_{\mathrm{II}}(u) - \hat{\psi}_{\mathrm{II}}(u)]}[1-\sin(u/a)]^2 du\, dv$$

$$+e^{2\hat{\psi}_{\mathrm{II}}(u)}\left[\frac{1+\sin(u/a)}{1-\sin(u/a)}\right] dx^2 \tag{4.4}$$

$$+e^{-2\hat{\psi}_{\mathrm{II}}(u)}[1-\sin(u/a)]^2\cos^2(u/a)dy^2\,,$$

where

$$\hat{\psi}_{\mathrm{II}}(u) = \sum_{k=0}^{\infty} d_k (2t-1)^{k/2} P_k\left[\frac{t}{\sqrt{2t-1}}\right]\,, \tag{4.5a}$$

$$t \equiv \sin^2(u/a)\,. \tag{4.5b}$$

The metric (4.4)—(4.5) describes the geometry of one of the two incoming colliding plane waves in the solution $\{d_k\}$. As before, the metric g_{III} in region III (Fig. 3) (which describes the remaining incoming wave) is obtained by replacing u/a by v/b in the above equations.

Unfortunately, the polynomials $(2t-1)^{k/2}P_k(t/\sqrt{2t-1})$ are not orthogonal polynomials with respect to any weight function, since they fail the Darboux-Christoffel test (Ref. 14, Sec. 8.90). However, we shall construct one particular infinite sequence $(\bar{d}_k)$ [Eqs. (4.11) below], such that for the corresponding solution $\{\bar{d}_k\}$ the function $\hat{\psi}_{\mathrm{II}}(u)$ [Eqs. (4.5)] has the right asymptotic behavior as $u \to \pi a/2$ to make the incoming plane wave in region II [Eq. (4.4)] purely anastigmatic.[7–9] Clearly, because of the $u \leftrightarrow v$ symmetry of our solutions, with this choice of the parameters d_k the other incoming plane wave (which is represented by the metric g_{III} on region III) will also be purely anastigmatic. Also note that Eqs. (4.11) represent only one particular solution in our family of solutions for which the incoming plane waves are purely anastigmatic; the details of the construction below will make it clear that infinitely many different solutions $\{d_k\}$ with the same property can be found in the family (3.18)–(3.22).

To proceed, consider the following function $f(t)$, defined on the interval $(-1,1)$:

$$f(t)=\ln(1-t) \qquad \text{for } t\geq 0\,,$$

$$f(t)=\ln(1+t) \qquad \text{for } t\leq 0\,. \tag{4.6a}$$

Since $f(t)$ is even, there is an expansion

$$f(t)=\sum_{k=0}^{\infty}\hat{d}_k P_{2k}(t)\,. \tag{4.6b}$$

Since $f(t)$ is square integrable on $(-1,1)$, this expansion converges absolutely everywhere on $(-1,1)$ with the exception of the point $t=0$ at which $f(t)$ is not C^1. In fact,

$$\hat{d}_k=\frac{4k+1}{2}\int_{-1}^{1}f(t)P_{2k}(t)dt$$

$$=(4k+1)\int_{0}^{1}\ln(1-t)P_{2k}(t)dt\,. \tag{4.7}$$

Now consider the solution $\{\bar{d}_k\}$, where $\bar{d}_k$ are defined by

$$\bar{d}_k=\hat{d}_{k/2} \qquad \text{for } k \text{ even}\,,$$

$$\bar{d}_k=0 \qquad \text{for } k \text{ odd}\,. \tag{4.8}$$

Then, the function $\hat{\psi}_{\rm II}(u)$ [Eqs. (4.5)] for this solution $\{\bar{d}_k\}$ is given by

$$\hat{\psi}_{\rm II}(u)=\sum_{k=0}^{\infty}\hat{d}_k(2t-1)^k P_{2k}\left(\frac{t}{\sqrt{2t-1}}\right), \tag{4.9}$$

where $t \equiv \sin^2(u/a)$. However, for $t\in(0,1)$ and for all $k \geq 1$ we have[14]

$$|\,(2t-1)^k P_{2k}\left(\frac{t}{\sqrt{2t-1}}\right)\,| < 1 = \sup_{t\in(0,1)} |P_{2k}(t)|\ .$$

Therefore, the series (4.9) converges absolutely to a continuous function on the interval (0,1). We can write

$$\sum_{k=0}^{\infty}\hat{d}_k(2t-1)^k P_{2k}\left(\frac{t}{\sqrt{2t-1}}\right)=\sum_{k=0}^{\infty}\hat{d}_k P_{2k}(t)$$

$$+\sum_{k=1}^{\infty}\delta_k(1-t)^k\ ,$$

where the second series is convergent being the difference of two convergent series on (0,1). Hence, by Eq. (4.9) and Eqs. (4.6)

$$\hat{\psi}_{\rm II}(u)=\ln\left(1-\sin^2\frac{u}{a}\right)+S(u)\ ,$$

where

$$\lim_{u\to\pi a/2} S(u)=0\ .$$

We thus obtain the following for the asymptotic behavior of $\hat{\psi}_{\rm II}(u)$ as $u\to\pi a/2$

$$e^{2\hat{\psi}_{\rm II}(u)} \underset{u\to\pi a/2}{\sim} \left(1-\sin\frac{u}{a}\right)^2\left(1+\sin\frac{u}{a}\right)^2, \tag{4.10a}$$

$$e^{-2\hat{\psi}_{\rm II}(u)} \underset{u\to\pi a/2}{\sim} \left(1-\sin\frac{u}{a}\right)^{-2}\left(1+\sin\frac{u}{a}\right)^{-2} . \qquad (4.10b)$$

The asymptotic behavior of the function $\hat{\psi}_{\rm III}(v)$ [which is the counterpart of $\hat{\psi}_{\rm II}(u)$ in region III of the solution $\{\bar{d}_k\}$] will have the analogous form near the focal plane $\{v=\pi b/2\}$. From Eqs. (4.10) we conclude, by inspecting the metric $g_{\rm II}$ (and $g_{\rm III}$) given by Eq. (4.4), that for our solution $\{\bar{d}_k\}$ both incoming plane waves are purely anastigmatic, i.e., for both incoming plane waves the metric coefficients g_{xx} and g_{yy} vanish simultaneously on the respective focal planes $\{u=\pi a/2\}$ and $\{v=\pi b/2\}$.

The coefficients $\bar{d}_k$ for the solution $\{\bar{d}_k\}$ can be calculated explicitly using Eqs. (4.7) and (4.8). This gives[14]

$$\bar{d}_{2k}=\hat{d}_k=-\frac{4k+1}{2^{2k}}\sum_{l=0}^{k}\frac{(-1)^l(4k-2l)!}{(2k-2l+1)l!(2k-l)!(2k-2l)!}$$

$$\times[\psi(2k-2l+2)-\psi(1)] , \qquad (4.11a)$$

where

$$\psi(x)=\frac{d}{dx}[\ln\Gamma(x)]$$

is Euler's psi function,[14] and

$$\bar{d}_{2k+1}=0 \qquad (4.11b)$$

for any $k\geq 0$.

V. CONCLUSIONS

The infinite-parameter family of colliding plane-wave solutions we have constructed in Sec. III have the following new features.

(i) The interaction regions of our solutions are locally isometric to interior Weyl black-hole solutions. An observer who enters the interaction region will not be able to distinguish, by local measurements that she performs completely within the interaction region, the geometry of the surrounding spacetime from the geometry in the interior of a black hole.

(ii) The metric functions of our solutions have oscillatory forms in a suitable coordinate system [Eqs. (3.18)—(3.22)].

(iii) By constructing an infinite series expansion for the function $\hat{\psi}(u',v)$ [Eq. (3.19)], we have built a colliding plane-wave solution in our family for which both of the two incoming colliding plane waves are purely anastigmatic, i.e., for which both incoming waves have equal focal lengths (Sec. IV).

On the other hand, our solutions suffer from the following drawbacks, some of which are common to all presently known exact solutions for colliding plane waves.

(i) As with the famous Khan-Penrose solution,[5] so also here, there are δ-function contributions to the Riemann curvature (gravitational shock waves) on the boundaries between the adjacent regions (Fig. 3); i.e., the metric is not C^1. The reason for this discontinuous behavior is the particular prescription that we use to extend the metric beyond the interaction region. It is clear, from the form of our metric as described by Eqs. (3.18)—(3.22), that no finite sum (3.19) will eliminate the δ-function shocks across the boundaries as long as we use the Penrose prescription for extending the metric beyond the interaction region. Since infinite sums of the form we have discussed in Sec. IV will in general converge only in the mean (i.e., in the L^2 sense), we cannot reliably employ infinite series expansions of the form (3.19) to construct smoother wave forms.

(ii) Except for the characteristic wavelengths a and b which can be freely adjusted by scale transformations, the two incoming colliding plane waves in our solutions have exactly the same functional form. The reasons for this are the $u \longleftrightarrow v$ symmetry of the metric (3.18) in the interaction region, and the $u \longleftrightarrow v$ symmetry of the Penrose prescription for extending the metric beyond the interaction region.

(iii) The incoming colliding plane waves in our solutions are not of sandwich type in general; i.e., the spacetime regions II and III (Fig. 3) representing the incoming waves are not flat near the respective focal planes of these waves. As we have discussed in detail in Sec. II, this property is responsible for the curvature singularities that are present at the focal planes of our solutions. The technique of "cutting off" the incoming waves just before their focal planes, which we have discussed in Sec. II, successfully avoids this difficulty from a physical viewpoint; however, it appears exceedingly difficult to determine whether the resulting solution (e.g., the solution depicted in Fig. 4) can be expressed in closed form as a member of the family of solutions (3.18)—(3.22) that we have constructed.

We conclude by listing some open questions which suggest directions for further research on some of the issues that we have raised in this paper.

(i) Are there different prescriptions for extending the metric (3.18) beyond the interaction region which could resolve some of the drawbacks in our solutions listed above?

(ii) Can the technique of using static axisymmetric black-hole metrics to generate parallel-polarized colliding plane-wave spacetimes be generalized to stationary axisymmetric solutions? Such a generalization presumably would yield an infinite-parameter family of solutions representing colliding plane waves with nonparallel polarizations.

ACKNOWLEDGMENTS

The author wishes to thank Kip S. Thorne for a critical reading of the first draft of this paper, and Sandip Chakrabarti for stimulating discussions which initiated this research. This work was supported by the National Science Foundation under Grant No. AST 85-14911.

REFERENCES FOR CHAPTER 4

[1] B.K. Harrison, Phys. Rev. Lett. **41**, 1197 (1978); G. Neugebauer, J. Phys. A **12**, L67 (1979); C. Hoenselaers, W. Kinnersley, and B.C. Xanthopoulos, Phys. Rev. Lett. **42**, 481 (1979); J. Math. Phys. **20**, 2530 (1979); D. Kramer and G. Neugebauer, Phys. Lett. A **75**, 259 (1980); W. Dietz and C. Hoenselaers, Proc. R. Soc. London **A382**, 221 (1982).

[2] S. Chandrasekhar and V. Ferrari, Proc. R. Soc. London **A396**, 55 (1984); S. Chandrasekhar and B.C. Xanthopoulos, ibid. **A398**, 223 (1985); **A402**, 37 (1985); **A402**, 205 (1985); **A403**, 189 (1986).

[3] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

[4] U. Yurtsever, Phys. Rev. D **38**, 1706 (1988) (Chapter 5).

[5] K. Khan and R. Penrose, Nature (London) **229**, 185 (1971).

[6] S. Chandrasekhar and B.C. Xanthopoulos, Proc. R. Soc. London A (to be published).

[7] U. Yurtsever, Phys. Rev. D **36**, 1662 (1987).

[8] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[9] U. Yurtsever, Phys. Rev. D **37**, 2803 (1988) (Chapter 3).

[10] F.J. Tipler, Phys. Rev. D **22**, 2929 (1980); Kip S. Thorne, in *Nonlinear Phenomena in Physics*, proceedings of the 1984 Latin American School of Physics, edited by F. Claro (Springer, Berlin, 1985).

[11] R. A. Matzner and F.J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[12] H. Weyl, Ann. Phys. **54**, 117 (1917); L. Mysak and P. Szekeres, Can. J. Phys. **44**, 617 (1966).

[13] S. Chakrabarti, Ph. D. thesis, University of Chicago, 1985.

[14] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1980).

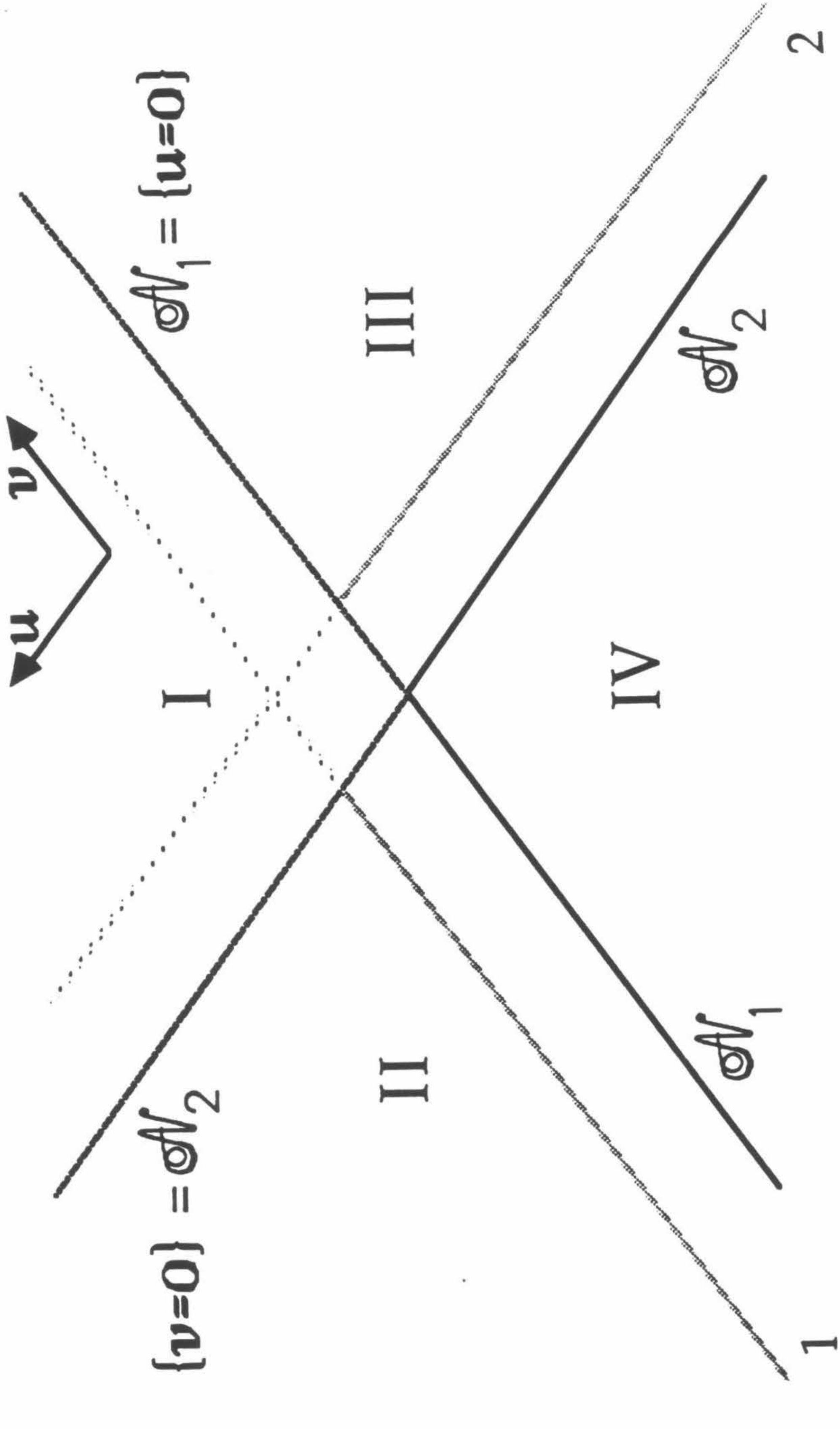
FIGURE CAPTIONS FOR CHAPTER 4

FIG. 1. The two-dimensional geometry of the characteristic initial-value problem for colliding plane waves. The null surfaces $\mathcal{N}_1=\{u=0\}$ and $\mathcal{N}_2=\{v=0\}$ are the past wave fronts of the incoming plane waves 1 and 2. Initial data corresponding to waves 1 and 2 are posed, respectively, on the upper portions of the surfaces $\mathcal{N}_2$ and $\mathcal{N}_1$ that are adjacent to the interaction region I. The geometry in region IV is flat, and the geometry in regions II and III is given by the metric describing the incoming waves 1 and 2, respectively. The geometry of the interaction region I is uniquely determined by the solution of the above initial-value problem.

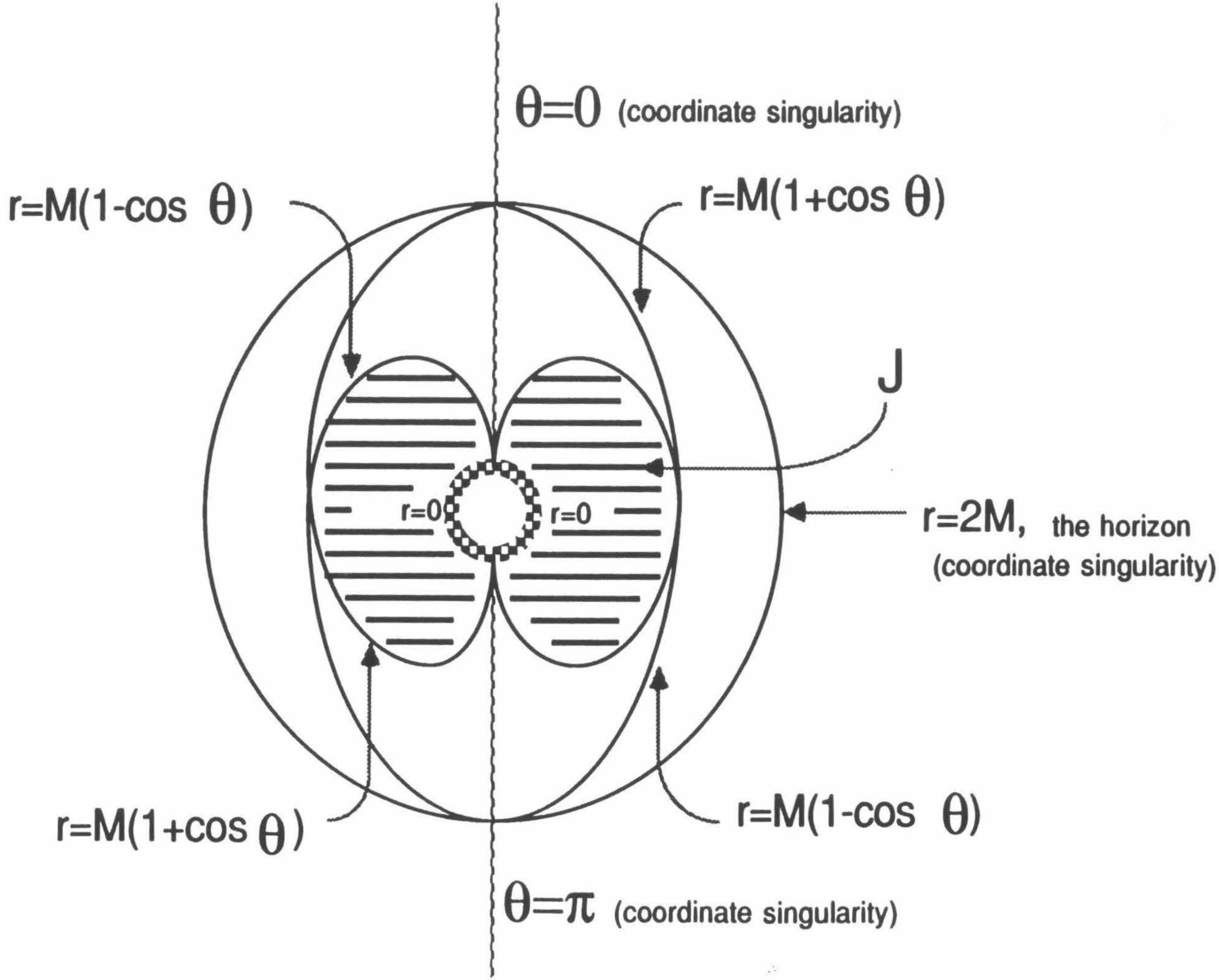
FIG. 2. The region J in Schwarzschild spacetime to which the interaction region of the colliding plane-wave solution (2.4) is locally isometric. This region J is shown shaded in this figure which is drawn in a $\{t=\text{const}\}$, $\{\phi=0,\pi\}$ plane. As explained in the text, the geometry in region J is extended nonanalytically beyond the null surfaces $r=M(1+\cos\theta)$ and $r=M(1-\cos\theta)$, which correspond to the wave fronts $\{u=0\}$ and $\{v=0\}$, respectively. Consequently, all coordinate singularities are avoided and the Schwarzschild metric on the shaded region J is lifted from $S^2\times R^2$ to R^4, on which the final metric (2.4) is defined.

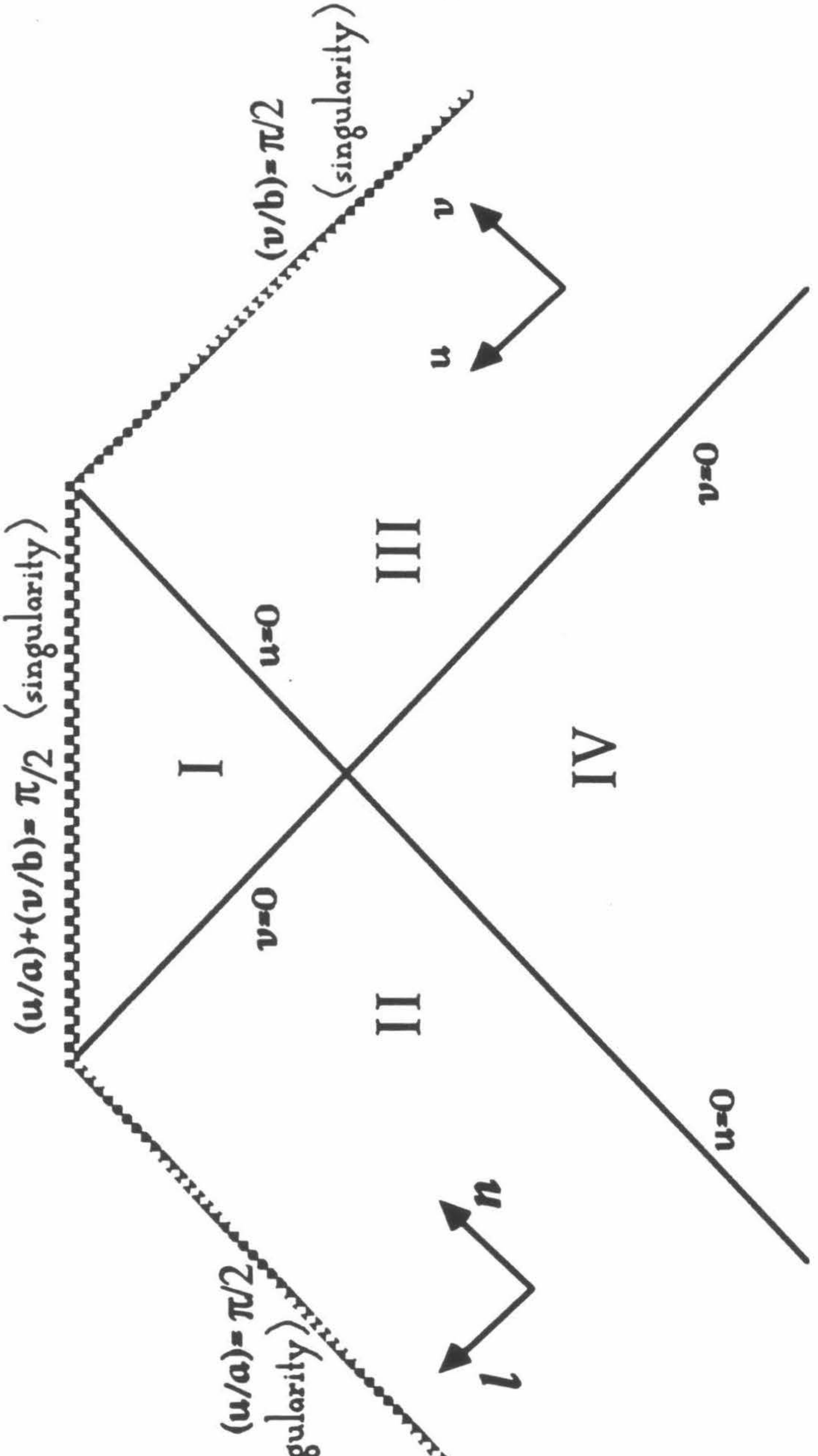
FIG. 3. A two-dimensional projection of the geometry of the colliding plane-wave solution (2.4). The metric is analytic throughout each of the regions I, II, III and IV, but it suffers discontinuities in its derivatives across the boundaries between the adjacent regions. The interaction region I is locally isometric to region J (Fig. 2) of the interior Schwarzschild solution (2.1). The curvature singularity at $(u/a)+(v/b)=\pi/2$ corresponds, under this isometry, to the Schwarzschild singularity at $r=0$.

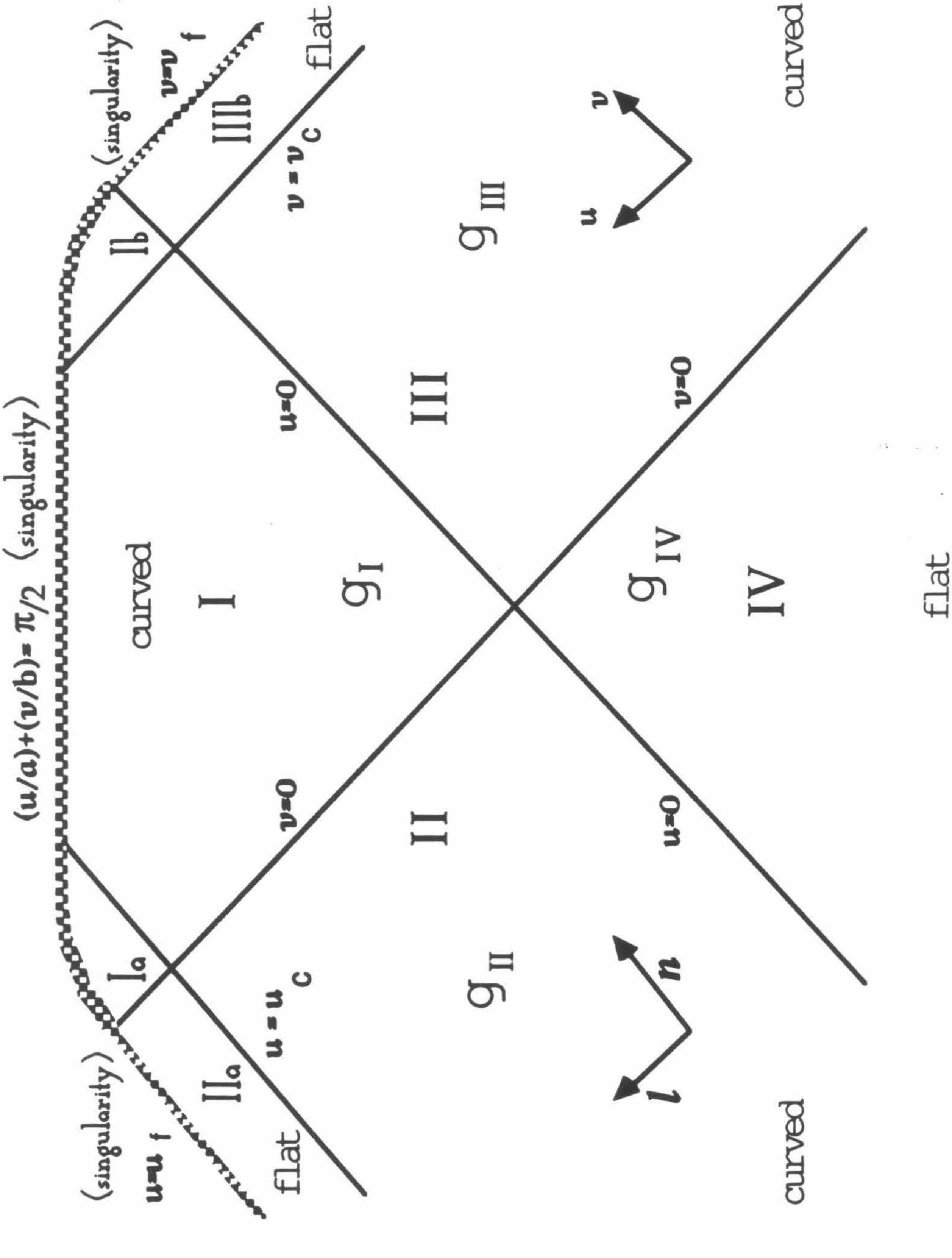
FIG. 4. Geometry of the colliding plane-wave solution that results from "cutting off" the incoming, colliding plane waves described by the solution (2.4). As explained in the text, the introduction of the secondary shocks along the surfaces $\{u=u_c\}$ and $\{v=v_c\}$ removes the curvature singularities on the focal planes $\{u=u_f\}$ and $\{v=v_f\}$. However, the geometry in the interaction region of this new solution is not everywhere described by the metric (2.4); the regions Ia and Ib are described by a different metric.



$\{v=0\} = \mathscr{N}_2$
$\mathscr{N}_1 = \{u=0\}$
u
v
I
II
III
IV
$\mathscr{N}_1$
$\mathscr{N}_2$
1
2


θ=0 (coordinate singularity)
r=M(1-cos θ)
r=M(1+cos θ)
J
r=0
r=0
r=2M, the horizon
(coordinate singularity)
r=M(1+cos θ)
r=M(1-cos θ)
θ=π (coordinate singularity)


(u/a)+(v/b)= π/2 (singularity)
(v/b)= π/2
(singularity)
(u/a)= π/2
(singularity)
u=0
v=0
I
II
III
IV
u
v
u
l
v=0
u=0


(u/a)+(v/b)= π/2 (singularity)
(singularity)
(singularity)
u=u f
v=v f
flat
flat
curved
curved
curved
flat
I
II
III
IV
Ia
Ib
IIa
IIIb
g I
g II
g III
g IV
u=0
v=0
u=0
v=0
u = u c
v = v c
u
v
l
n

CHAPTER 5

Structure of the Singularities Produced by Colliding Plane Waves

Published in Phys. Rev. D **38**, 1706 (1988).

ABSTRACT

When gravitational plane waves propagating and colliding in an otherwise flat background interact, they produce spacetime singularities. If the colliding waves have parallel (linear) polarizations, the mathematical analysis of the field equations in the interaction region is especially simple. Using the formulation of these field equations previously given by Szekeres, we analyze the asymptotic structure of a general colliding parallel-polarized plane-wave solution near the singularity. We show that the metric is asymptotic to an inhomogeneous Kasner solution as the singularity is approached. We give explicit expressions which relate the asymptotic Kasner exponents along the singularity to the initial data posed along the wave fronts of the incoming, colliding plane waves. It becomes clear from these expressions that for specific choices of initial data the curvature singularity created by the colliding waves degenerates to a coordinate singularity, and that a nonsingular Killing-Cauchy horizon is thereby obtained. Our equations prove that these horizons are unstable in the full nonlinear theory against small but generic perturbations of the initial data, and that in a very precise sense, "generic" initial data always produce all-embracing, spacelike curvature singularities without Killing-Cauchy horizons. We give several examples of exact solutions which illustrate some of the asymptotic singularity structures that are discussed in the paper. In particular, we construct a new family of exact colliding parallel-polarized plane-wave solutions, which create Killing-Cauchy horizons instead of a spacelike curvature singularity. The maximal analytic extension of one of these solutions across its Killing-Cauchy horizon results in a colliding plane-wave spacetime, in which a Schwarzschild black hole is created out of the collision between two plane-symmetric sandwich waves propagating in a cylindrical universe.

I. INTRODUCTION

Gravitational plane waves are among the simplest nontrivial exact solutions to the vacuum Einstein field equations that describe time-varying gravitational fields. Although the existence and the quantitative structure of these solutions have been known since the early days of general relativity,[1] the surprisingly rich qualitative features that they possess were not fully understood until the mid-1960s when Penrose[2] carried out his investigations on their global structure. (In fact, Penrose proposed the plane wave spacetimes as counter examples to a conjecture in global general relativity, which stated that any spacetime satisfying a sufficiently strong causality condition can be globally embedded in a high-dimensional Minkowski space.) The source of this rich global structure in plane-wave solutions is the focusing effect of gravitational plane waves, which is reviewed, for example, in Refs. 2 and 3, and in the introductory section of Ref. 4.

The presence of both spacelike and timelike nontrivial directions in exact (single) plane-wave solutions makes it possible to study interesting dynamical effects associated with the interaction of plane waves, without destroying the plane symmetry present in the original solutions. Thus, for example, it is not exceedingly difficult to write down solutions to the vacuum field equations that describe collisions of gravitational plane waves. The first such solution was discovered by Khan and Penrose[5] in their attempt to verify Penrose's earlier conjecture[2] that the focusing effect of single plane waves should cause the colliding waves to interact strongly and to eventually produce spacetime singularities. Several other solutions involving similar curvature singularities were obtained by Szekeres,[6] who formulated a general solution for the problem of colliding parallel-(linear)-polarized gravitational plane waves. Later Nutku and Halil[7] obtained a colliding plane-wave solution where the incoming waves had nonparallel linear polarizations; this solution too had a spacelike curvature singularity, similarly to the earlier solutions. The global structure of these early solutions is reviewed in Refs. 8 and 3.

The technique of generating colliding plane-wave solutions by the extension of suitable (but weakly restricted) plane-symmetric solutions to the field equations in the interaction region, pioneered by Khan and Penrose in Ref. 5, proved to be remarkably fertile in subsequent studies on colliding waves. Thus, using this technique, Chandrasekhar and Xanthopoulos[9] obtained many new solutions for both colliding purely gravitational plane waves and for colliding plane waves coupled with matter fields. Other solutions were obtained by the author in Ref. 10, where the Penrose-Khan prescription for generating colliding plane-wave solutions is reviewed, and compared with the direct method of solving the relevant initial-value problem, which, in the case of parallel-polarized waves, was worked out by Szekeres.[6]

A surprising result of the recent work on exact solutions for colliding plane waves was the discovery by Chandrasekhar and Xanthopoulos[11] of a solution, where the collision of the incoming waves (which are non-parallel-polarized) produces a nonsingular Killing-Cauchy horizon instead of a spacelike curvature singularity. The resulting metric can be analytically extended across this horizon to produce a maximal spacetime, whose singularities [which are timelike for the particular (i.e., maximal analytic) extension used by Chandrasekhar and Xanthopoulos[11]] could be avoided by observers traveling on timelike world lines, in striking contrast to the earlier solutions with their all-embracing, spacelike singularities which almost all observers are bound to encounter.[3] The structure, significance, and nongeneric nature of such Killing-Cauchy horizons in colliding plane-wave solutions (and in more general plane-symmetric spacetimes) are discussed extensively in Refs. 11, 4, and 3. The

Chandrasekhar-Xanthopoulos[11] solutions are relevant to the subject matter of the present paper, in that they point to a hitherto unsuspected richness in the structure of singularities produced by colliding plane waves. In fact, it was more or less widely believed[11] before the discovery of these solutions, that the singularity structure exhibited by the earlier exact solutions[5,6,7,8] was universal for colliding plane-wave spacetimes. And even after this remarkable discovery, one might be tempted to believe that the unusual structure of Chandrasekhar-Xanthopoulos[11] spacetimes is a result of the nonparallel configuration of the incoming polarizations, and that colliding plane waves with parallel polarizations will always produce singularities with the same global structure as the earlier exact solutions. One of the specific results of this paper is that this is not the case; in particular, in Sec. IV we present examples of exact solutions for colliding parallel-polarized plane waves, which possess nonsingular Killing-Cauchy horizons that are very similar in local structure to the horizons of the Chandrasekhar-Xanthopoulos[11] spacetimes.

The overall purpose of this paper is to explore in detail the structure of the spacetimes that result from the collisions of parallel-polarized plane waves, especially their singularity and Cauchy-horizon structures. The plan of the paper is as follows.

In Sec. II A, we give a very brief review of Szekeres's[6] formulation of the field equations and the characteristic initial-value problem for colliding parallel-linear-polarized plane waves, in the (u, v, x, y) coordinate system which we call "Rosen-type" and which is tuned to the plane-symmetry of the spacetime. Our presentation is necessarily brief, and the reader is referred to Ref. 6 for the full mathematical details, or to Ref. 10 for a short outline.

In Sec. II B, we perform a coordinate transformation to a new (α, β, x, y) coordinate system in which the mathematical analysis of the field equations simplifies considerably. Although this coordinate system and its properties were known to Szekeres,[6] he did not make extensive use of them since the coordinates (α,β) are badly behaved on the initial null surfaces where the initial data are posed. However, we will find this new coordinate system very useful both in discussing the general solution of the field equations (Section II B), and in discussing the asymptotic bahavior of the resulting spacetime (subsequent sections).

It will become clear in Sec. II B that some kind of singularity is associated with the "surface" $\alpha=0$ in a general colliding plane-wave spacetime. (Note: α is a timelike coordinate which monotonically decreases to zero along the world lines of all observers running into the singularity.) We show in Sec. III A that the spacetime metric asymptotically approaches an inhomogeneous Kasner[12] solution as α approaches zero, where the time coordinate t of the asymptotic Kasner spacetime is monotonically related to α, and the Kasner singularity at $t=0$ corresponds to the singularity at $\alpha=0$. We give explicit expressions which relate the spatially inhomogeneous asymptotic Kasner exponents along the singularity to the initial data posed along the wave fronts of the incoming, colliding plane waves. In general, these exponents depend on β, the spacelike coordinate running along the nontrivial spatial (z) direction in the spacetime.

Our discussion in Sec. III A indicates that for some specific choices of the initial data, the Kasner exponents (either locally, or globally for a finite interval in the spatial coordinate β) may take on the values associated with a degenerate Kasner solution. A degenerate Kasner spacetime is flat, and instead of a spacelike curvature singularity, it possesses a Killing-Cauchy horizon at $t=0$. It is then natural to expect that, when the asymptotic limit of the metric as $\alpha\rightarrow 0$ is a degenerate Kasner solution, our colliding plane-wave spacetime possesses a nonsingular Killing-Cauchy horizon at $\alpha=0$, across which the metric can be extended smoothly. However, to demonstrate this rigorously, we need to study the behavior of the spacetime curvature near α=0, and to show that the curvature is indeed well behaved when the metric approaches a degenerate Kasner limit at α=0. This would give us information about the asymptotic behavior of the *derivatives* of the metric as $\alpha \to 0$, complementing our analysis in Sec. III A of the asymptotic behavior of the metric itself. Thus, in Sec. III B, we derive expressions for the Newman-Penrose[13,3] curvature quantities (with respect to the null tetrad that we set up earlier in Sec. II A) in terms of the metric components in the (α,β,x,y) coordinate system. We then read out from these expressions the asymptotic structure of the curvature quantities as $\alpha \to 0$. This analysis indeed shows that when the metric is asymptotic to a degenerate Kasner solution, the curvature remains well behaved as $\alpha \to 0$.

We begin Sec. III C by recapitulating the principal conclusion of the analysis of Sec. III B: When the asymptotic limit of the solution is a degenerate Kasner metric, the colliding plane-wave spacetime possesses a Killing-Cauchy horizon (a coordinate singularity) at α=0 across which the curvature quantities are finite and well behaved. We note that spacetime can be extended through these horizons in infinitely many different ways; the geometry beyond the horizons cannot be determined from the initial data posed by the incoming, colliding plane waves. We then briefly recall our earlier work in Ref. 4, where we have proved general theorems stating the instability of such Killing-Cauchy horizons in any plane-symmetric spacetime against generic, plane-symmetric perturbations. In the specific case of the Killing-Cauchy horizons which occur at α=0 in our colliding plane-wave solutions, the existence of these instabilities is particularly clear: We discuss how our equations imply (i) that the horizons at α=0 are unstable in the full nonlinear theory against small but generic perturbations of the initial data (since such perturbations drive the asymptotic Kasner exponents away from the degenerate values), and (ii) that in a very precise sense, "generic" initial data always produce all-embracing, spacelike spacetime singularities at α=0 across which no extension of the metric is possible.

In Sec. IV we give several examples of exact solutions for colliding parallel-polarized plane waves, which illustrate some of the different asymptotic singularity structures that are discussed in the previous sections. Most of the examples we consider are new, and are discussed here for the first time. However, all of our examples have asymptotic Kasner exponents which are uniformly constant across the whole range of the spatial coordinate β. It seems particulary difficult to write down a full solution, expressible in closed form, which would exhibit a truly inhomogeneous asymptotic structure near the singularity α=0. By using the same line of reasoning that we have followed in Ref. 10, we construct exact colliding parallel-polarized plane-wave solutions, which produce Killing-Cauchy horizons at α=0 instead of a curvature singularity. The maximal analytic extension of one of these solutions across the horizon produces a colliding plane-wave spacetime with a surprising global structure.

In the concluding section, we briefly list the major results of the paper, and discuss some suggestions and plans for future research.

The notation and conventions of this paper are the same as in Refs. 3, 4, and 10. In particular, we adopt the metric signature (–,+,+,+), and we use the "rationalized" Newman-Penrose equations appropriate to this signature, which can be found, e.g., in Refs. 14 and 3.

II. THE FIELD EQUATIONS FOR COLLIDING PARALLEL-POLARIZED GRAVITATIONAL PLANE WAVES AND THEIR SOLUTION

A. Formulation of the problem in the Rosen-type (u,v,x,y) coordinate system

In any plane-symmetric spacetime (see Sec. III B of Ref. 3, or Sec. II of Ref. 4 for a careful definition of plane symmetry), there exists a canonical null tetrad[13] whose construction is described in Sec. III B of Ref. 3. In this null tetrad, which we call the standard tetrad, $\vec{l}$ and $\vec{n}$ are tangent to the two null geodesic congruences everywhere orthogonal to the plane-symmetry generating Killing vector fields $\vec{\xi}_1$ and $\vec{\xi}_2$, and $\vec{m}$ and its complex conjugate are linear combinations of the $\vec{\xi}_i$, $i=1, 2$. As is shown by Szekeres,[6] it follows from the presence of only two nontrivial dimensions in the spacetime, that we can find a local coordinate chart (u,v,x,y) in which $\vec{\xi}_i=\partial/\partial x^i$ $[(x^1,x^2)\equiv(x,y)]$, and in which the standard tetrad can be expressed as

$$\vec{l}=2e^{M(u,v)}\frac{\partial}{\partial u}+P^i(u,v)\frac{\partial}{\partial x^i} ,$$

$$\vec{n}=\frac{\partial}{\partial v}+Q^i(u,v)\frac{\partial}{\partial x^i} , \tag{2.1}$$

$$\vec{m}=\frac{1}{F(u,v)}\frac{\partial}{\partial x}+\frac{1}{G(u,v)}\frac{\partial}{\partial y} .$$

Here P^i,Q^i,M are real, and F,G are complex functions of (u,v), with $F^*G-G^*F\neq 0$ throughout the region on which strict plane symmetry[3,4] holds and on which the tetrad (2.1) and the coordinate chart (u,v,x,y) are well behaved. In the specific plane-symmetric spacetimes which describe gravitational plane waves propagating and colliding in an otherwise flat background, there will be a region, corresponding to the spacetime before the arrival of either plane wave, where the metric is flat. It is shown by Szekeres[6] (see also Sec. III B of Ref. 4), that the presence of such a flat region makes it possible to find a new coordinate system, which we still denote by (u, v, x, y), in which $P^i = Q^i = 0$ and the standard tetrad (2.1) takes the simpler form

$$\vec{l} = 2e^{M(u,v)} \frac{\partial}{\partial u}, \quad \vec{n} = \frac{\partial}{\partial v},$$

$$\vec{m} = \frac{1}{F(u,v)} \frac{\partial}{\partial x} + \frac{1}{G(u,v)} \frac{\partial}{\partial y}. \tag{2.2}$$

Finally, when the colliding plane waves have parallel linear polarizations, the tetrad components in Eq. (2.2) can be further restricted[6,10] to give

$$\vec{l} = 2e^{M} \frac{\partial}{\partial u}, \quad \vec{n} = \frac{\partial}{\partial v},$$

$$\vec{m} = N_1 \frac{\partial}{\partial x} + N_2 \frac{\partial}{\partial y}, \tag{2.3}$$

where

$$N_1 = \frac{1+i}{2} e^{(U-V)/2},$$

$$N_2 = \frac{1-i}{2} e^{(U+V)/2}, \tag{2.4}$$

with U and V real and with M, U, and V functions of u and v only. The presence of a difference between the linear polarizations of the incoming waves (or, the presence of a circular polarization component in any of the colliding waves) would manifest itself in the presence of a (u, v)-dependent relative phase factor between N_1 and N_2 in Eq. (2.4) above. The tetrad (2.3)–(2.4) gives rise to the metric

$$g = -e^{-M} du\, dv + e^{-U} (e^{V} dx^2 + e^{-V} dy^2). \tag{2.5}$$

Thus, the x-y part of the metric is in diagonal form uniformly at all points in the spacetime, and the Killing vector fields $\partial/\partial x$ and $\partial/\partial y$ are everywhere hypersurface orthogonal; each of these facts being equivalent to the assumption that the colliding plane waves have parallel (linear) polarizations.[6,10] The coordinate system (u,v,x,y) is uniquely determined, up to transformations of the form $u=f(u')$, $v=g(v')$, by demanding (i) that the metric in it has the above form (2.5) (hence, in particular, that the plane-symmetry generators are $\partial/\partial x^i$), and (ii) that in the flat region describing the spacetime before the arrival of either wave, (u,v,x,y) reduce to Minkowski coordinates. [Here, f and g are functions which are constrained to be of the form $f(u')=cu'$, $g(v')=v'/c$ in the *flat Minkowski region*, but which are completely arbitrary elsewhere. We will use this coordinate freedom below when we discuss the initial-value problem for the field equations.] Therefore, the coordinate system (u,v,x,y) is the direct analogue of the Rosen-type coordinates associated with each of the incoming, colliding plane waves. (For a discussion of different coordinate systems associated with plane-wave spacetimes, see Ref. 2, Sec. II of Ref. 3, and Sec. I of Ref. 4). We will thus call (u,v,x,y) the Rosen-type coordinates on the colliding plane-wave spacetime.

The vacuum field equations for the metric (2.5) are[6,15]

$$2(U_{,uu}+M_{,u}U_{,u})-U_{,u}{}^2-V_{,u}{}^2=0\,, \tag{2.6a}$$

$$2(U_{,vv}+M_{,v}U_{,v})-U_{,v}{}^2-V_{,v}{}^2=0\,, \tag{2.6b}$$

$$U_{,uv}-U_{,u}U_{,v}=0\,, \tag{2.6c}$$

$$V_{,uv}-\frac{1}{2}(U_{,u}V_{,v}+U_{,v}V_{,u})=0\,, \tag{2.6d}$$

where the integrability condition for the first two equations is satisfied by virtue of the last two, and yields the remaining field equation

$$M_{,uv}-\frac{1}{2}(V_{,u}V_{,v}-U_{,u}U_{,v})=0\,. \tag{2.7}$$

Therefore, it is sufficient to solve Eqs. (2.6c) and (2.6d) first and to obtain M by quadrature from the first two equations (2.6a) and (2.6b) afterward, since Eq. (2.7) as well as the integrability condition for Eqs. (2.6a) and (2.6b) are automatically satisfied as a result of Eqs. (2.6c) and (2.6d).

The initial-value problem associated with the field equations (2.6) and (2.7) is best formulated in terms of initial data posed on null (characteristic) surfaces. A natural choice for the initial characteristic surface is the surface made up of the two intersecting null hyperplanes which form the past wave fronts of the incoming plane waves, and which, by a readjustment of the null coordinates u and v if necesary, can be arranged to be the surfaces $\{u=0\}$ and $\{v=0\}$. The geometry of the resulting characteristic initial-value problem is depicted in Fig. 1. The initial data supplied by the plane wave propagating in the v direction (to the right in Fig. 1) is posed on the $u\geq 0$ portion of the surface $\{v=0\}$, and the initial data supplied by the plane wave propagating in the u direction (to the left in Fig. 1) is posed on the $v\geq 0$ portion of the surface $\{u=0\}$. In region IV, which represents the spacetime before the passage of either plane wave, the geometry is flat and all metric coefficients M, U, and V vanish identically. Now recall that there is a remaining coordinate freedom in the choice of the (u,v,x,y) coordinate system, given by the transformations of the form $u=f(u')$, and $v=g(v')$. This gauge freedom also manifests itself in the choice of initial data on the characteristic initial surface $\{u=0\}\cup\{v=0\}$: The choice of the initial data $\{M(u=0,v), M(u,v=0)\}$ for the metric function M is completely arbitrary, since, clearly, for a single plane wave [cf. Eq. (2.5)] $M(u)$ [$M(v)$] can be adjusted freely by coordinate transformations of the form $u=f(u')$ [$v=g(v')$]. [This arbitrariness (gauge freedom) in the choice of initial data in the (u,v) coordinates dissappears, when, as we will do in Sec. II B, one formulates the field equations in the (α,β) coordinate system. Then, any two different but *equivalent* choices of initial data for the functions V, U, and M in the (u,v) coordinates correspond, in the formalism of Sec. II B, to a *unique* choice of the functions $V(r,1)$ and $V(1,s)$ which determine the initial data. In fact, even the boost freedom (see below), which eventually remains in the choice of the (u,v) coordinates, is absent from the formalism based on the (α,β) coordinate system.] We will fix the above gauge freedom once and for all by posing our initial data so that

$$M(u=0,v)=M(u,v=0)\equiv 0\ . \tag{2.8}$$

Then the only remaining coordinate freedom in the problem is the scale (or boost) freedom given by the scaling (boost) transformations $u=cu'$, $v=v'/c$, where c is a positive constant. This remaining boost freedom is harmless however; in fact, it is even useful in carrying out computations involving colliding waves, when, for example, it is known from physical arguments that the results have to be scale invariant (see, e.g., the discussion in Refs. 15 and 16).

Our choice of gauge, Eq. (2.8), implies that the metric in region II (where $u\geq 0$, $v\leq 0$), describing the geometry of the incoming colliding wave that propagates in the v direction (to the right in Fig. 1), is given by

$$g_{\mathrm{II}}=-du\ dv+F_1^{\,2}(u)dx^2+G_1^{\,2}(u)dy^2\ , \tag{2.9}$$

and that the metric in region III (where $v\geq 0$, $u\leq 0$), describing the geometry of the incoming wave that propagates in the u direction (to the left in Fig. 1), is given by

$$g_{\mathrm{III}}=-du\,dv+F_2^2(v)dx^2+G_2^2(v)dy^2 . \tag{2.10}$$

Here, F_1, G_1 are C^1 (and piecewise C^2) functions of u (for $u \geq 0$), and F_2, G_2 are C^1 (and piecewise C^2) functions of v (for $v \geq 0$), which satisfy the initial conditions $F_1(u=0)=G_1(u=0)=F_2(v=0)=G_2(v=0)=1$ (coordinates Minkowski in IV), and satisfy the differential equations

$$\frac{F_1''(u)}{F_1(u)}+\frac{G_1''(u)}{G_1(u)}=0 , \qquad \frac{F_2''(v)}{F_2(v)}+\frac{G_2''(v)}{G_2(v)}=0 \tag{2.11}$$

for $u \geq 0$ and $v \geq 0$, respectively [these differential equations follow from the field equations (2.6)]. The initial data, induced on the characteristic initial surface $\{u=0\} \cup \{v=0\}$ by the colliding waves (2.9) and (2.10), are given by

$$U(u,v=0) \equiv U_1(u)=-\ln[F_1(u)G_1(u)] , \tag{2.12a}$$

$$V(u,v=0) \equiv V_1(u)=\ln\left[\frac{F_1(u)}{G_1(u)}\right] , \tag{2.12b}$$

$$U(u=0,v) \equiv U_2(v)=-\ln[F_2(v)G_2(v)] , \tag{2.12c}$$

$$V(u=0,v) \equiv V_2(v)=\ln\left[\frac{F_2(v)}{G_2(v)}\right] . \tag{2.12d}$$

If the colliding waves are sandwich plane waves (Sec. II of Ref. 3), we then have lengthscales f_1, f_2, a, f_1', f_2', and b such that

$$F_1(u)=\frac{F_1(a)}{a-f_1}(u-f_1) , \qquad G_1(u)=\frac{G_1(a)}{a-f_2}(u-f_2)$$

for $u \geq a$, (2.13)

and

$$F_2(v)=\frac{F_2(b)}{b-f_1'}(v-f_1') , \qquad G_2(v)=\frac{G_2(b)}{b-f_2'}(v-f_2')$$

for $v \geq b$. (2.14)

Although the initial data in the form of Eqs. (2.12) give the information about the incoming, colliding plane waves in an intuitively clear format [cf. Eqs. (2.9) and (2.10)], in the more precise mathematical description of the initial-value problem the initial data are completely determined by only the two freely specifiable functions $V_1(u)$, and $V_2(v)$. In other words, the initial data consist of

$$\{V_1(u), V_2(v)\} , \tag{2.15}$$

where $V_1(u)$ and $V_2(v)$ are C^1 (and piecewise C^2) functions for $u \geq 0$ and $v \geq 0$, respectively, which are freely specified except for the initial conditions $V_1(u=0)=V_2(v=0)=0$. In the linearized regime (when $V_1, V_2 \ll 1$), the functions V_1 and V_2 correspond to the time-dependent physical amplitudes of the incoming, colliding gravitational waves [cf. Eqs. (2.9) and (2.10)]. The remaining functions $U_1(u)$ and $U_2(v)$ are uniquely determined, by the initial data (2.15), through the constraint equations [cf. Eqs. (2.6a) and (2.6b)]

$$2U_{1,uu}-U_{1,u}{}^2=V_{1,u}{}^2 , \tag{2.16a}$$

$$2U_{2,vv}-U_{2,v}{}^2=V_{2,v}{}^2 , \tag{2.16b}$$

with the initial conditions $U_1(u=0)=U_2(v=0)=0$, $U_{1,u}(u=0)=U_{2,v}(v=0)=0$. Note that, if we define two new functions $f(u)$ and $g(v)$ by

$$f(u) \equiv e^{-U_1(u)/2}, \qquad g(v) \equiv e^{-U_2(v)/2}, \tag{2.17}$$

we can express Eqs. (2.16) in the form of "focusing" equations:

$$\frac{f_{,uu}}{f} = -\frac{1}{4} {V_{1,u}}^2, \tag{2.18a}$$

$$\frac{g_{,vv}}{g} = -\frac{1}{4} {V_{2,v}}^2, \tag{2.18b}$$

with the initial conditions $f(0)=g(0)=1$, $f'(0)=g'(0)=0$.

In Secs. III A and III B, when we discuss the asymptotic structure of the colliding plane-wave spacetime described by Eqs. (2.3)–(2.5), we will need the following equations which express the Newman-Penrose[13] curvature quantities in the null tetrad (2.3) and (2.4) in terms of the metric coefficients M, U, and V; the derivation of these equations can be found in Refs. 6 and 15:

$$\Psi_0 = 2ie^{2M}(M_{,u}V_{,u} + V_{,uu} - V_{,u}U_{,u}), \tag{2.19a}$$

$$\Psi_2 = -e^{M}M_{,uv}, \tag{2.19b}$$

$$\Psi_4 = \frac{i}{2}[(U_{,v} - M_{,v})V_{,v} - V_{,vv}], \tag{2.19c}$$

$$\Psi_1 = \Psi_3 = 0. \tag{2.19d}$$

B. The Field equations and their solution in the (α,β) coordinates

We now construct a new coordinate system (α,β,x,y), in which the field equations and the initial-value problem associated with them take simpler forms. This coordinate system is constructed as follows.

Consider the interaction region (region I in Fig. 1) where $u \geq 0$ and $v \geq 0$. This region is the domain of dependence[17] of the characteristic initial surface $\{u=0\} \cup \{v=0\}$, on which the initial-value problem defined by Eqs. (2.6), (2.8), (2.15), and (2.16) is to be solved. Consider the field equation (2.6c) in the interaction region. It follows from this equation that if we define

$$\alpha(u,v) \equiv e^{-U(u,v)} , \tag{2.20}$$

then, throughout the interaction region, $\alpha(u,v)$ satisfies

$$\alpha_{,uv}=0 , \tag{2.21}$$

the flat-space wave equation in two dimensions. Equation (2.21) suggests that we define another function, $\beta(u,v)$, such that

$$\beta_{,u}=-\alpha_{,u} , \qquad \beta_{,v}=\alpha_{,v} , \tag{2.22}$$

since, clearly, the integrability condition for Eqs. (2.22) is satisfied by virtue of Eq. (2.21). The general solution of Eq. (2.21) is

$$\alpha(u,v)=a(u)+b(v) , \tag{2.23}$$

where $a(u)$ and $b(v)$ are arbitrary functions. With this solution for α, Eqs. (2.22) yield

$$\beta(u,v)=-a(u)+b(v)+c , \tag{2.24}$$

where c is an arbitrary constant. Note that Eq. (2.20) defines α not only throughout the interior of the interaction region I where $u>0$, $v>0$, but also along the boundary $\{u=0\}\cup\{v=0\}$ of this region, which is the characteristic initial surface. Hence, the boundary values (2.12a), (2.12c) for the function $U(u,v)$ yield, through Eq. (2.20), the following boundary values for α:

$$\alpha(u,v=0)=e^{-U_1(u)}, \qquad \alpha(u=0,v)=e^{-U_2(v)}. \tag{2.25}$$

These initial values (2.25), when combined with the general solution (2.23) and the initial condition $U(u=0,v=0)=0$, immediately yield the unique solution

$$\alpha(u,v)=e^{-U_1(u)}+e^{-U_2(v)}-1 \tag{2.26}$$

for $\alpha(u,v)$, which holds throughout the interaction region. This solution, combining with Eq. (2.24) and setting the arbitrary constant c equal to zero, yields the solution

$$\beta(u,v)=e^{-U_2(v)}-e^{-U_1(u)} \tag{2.27}$$

for $\beta(u,v)$ and completes the construction of the new variables (α,β). To see that these variables actually define a new coordinate system, consider the two-form given by the exterior product $d\alpha\wedge d\beta$. When this two-form is nonzero throughout some region $\mathcal{U}$, it follows from the inverse function theorem[18] that the functions $\alpha(u,v)$ and $\beta(u,v)$ (together with the usual spatial coordinates x, y) constitute a regular coordinate system throughout $\mathcal{U}$. Now, Eqs. (2.26) and (2.27) give

$$d\alpha\wedge d\beta = 2U_1'(u)U_2'(v)e^{-[U_1(u)+U_2(v)]}\,du\wedge dv\ . \tag{2.28}$$

On the other hand, it immediately follows from Eqs. (2.16), or more clearly from the "focusing" equations (2.18), that as long as the initial data (2.15) are nontrivial for both incoming waves [i.e., as long as neither $V_1(u)$ nor $V_2(v)$ is identically zero], and as long as the initial surfaces $\{u=0\}$ and $\{v=0\}$ correspond to the true past wave fronts of the colliding waves [i.e., as long as $V_1(u)\neq 0$ and $V_2(v)\neq 0$ for *all* sufficiently small but positive u and v], we have

$$U_1'(u)>0\,,\, f'(u)<0 \quad \forall\, u>0\,,$$

$$U_2'(v)>0\,,\, g'(v)<0 \quad \forall\, v>0\,, \tag{2.29}$$

whereas $U_1'(u=0)=U_2'(v=0)=0$ because of the initial conditions [cf. Eqs. (2.16)]. Therefore we conclude [Eq. (2.28)], that as long as the initial data (2.15) are nontrivial for both colliding waves, and as long as the null surfaces $\{u=0\}$ and $\{v=0\}$ are the true past wave fronts, the functions (α,β,x,y) constitute a coordinate system which is regular wherever the coordinate system (u,v,x,y) is regular in the interior of the interaction region, $u>0$, $v>0$. On the other hand, the coordinates α,β are singular along the initial null surfaces $\{u=0\}$ and $\{v=0\}$. In other words, the singularities of the coordinate system (α,β,x,y) consist of the singularities of the (u,v,x,y) coordinates (when there are any), and the singularity along the initial characteristic surface $\{u=0\}\cup\{v=0\}$. Since the only place in the interaction region where the coordinates (u,v,x,y) can develop singularities is the "surface" $\{\alpha=0\}$ (see Sec. III A), it follows that the coordinate system (α,β,x,y) covers the domain of dependence of the initial surface $\{u=0\}\cup\{v=0\}$ regularly except for the singularities on $\{u=0\}$ and $\{v=0\}$.

The coordinates (α,β,x,y) enjoy a number of properties which make them useful in studying the field equations for colliding plane waves. First, the functions $\alpha(u,v)$ and $\beta(u,v)$ satisfy the wave equation in the two-dimensional Minkowski metric $-du\,dv$ (and, by conformal invariance, also in the two-dimensional metric $-e^{-M}du\,dv$). Hence, it follows that the $du\,dv$ part of the metric (2.5) will be in diagonal form [Eq. (2.43)] in the new coordinate system (α,β,x,y). Second, by performing the transformation (2.26) and (2.27) from the variables (u,v) to the new variables (α,β), we have eliminated one of the metric coefficients [namely the function $U(u,v)$] from Eq. (2.5), and absorbed it into the definition of the coordinate α. Therefore, the field equations in the new coordinate system [Eqs. (2.44)] will involve only two unknown variables, instead of the three functions M, V, and U involved in Eqs. (2.6). Finally, the Eqs. (2.26) and (2.27), which together with Eq. (2.20) yield the unique solution to the initial-value problem for $U(u,v)$ [Eq. (2.41)], also provide expressions for the new variables α and β purely in terms of the initial data on $\{u=0\}\cup\{v=0\}$. In other words, it is not necessary to solve any of the remaining field equations to perform the transformation from the (u,v,x,y) coordinates to the new (α,β,x,y) coordinate system.

We now proceed with the mathematical analysis of the initial-value problem defined by Eqs. (2.5), (2.6), (2.8), (2.15), and (2.16), in the new coordinate system (α,β,x,y). First we note the transformation rules

$$\partial_u=\alpha_{,u}(\partial_\alpha-\partial_\beta)\,, \tag{2.30a}$$

$$\partial_v=\alpha_{,v}(\partial_\alpha+\partial_\beta)\,, \tag{2.30b}$$

and their inverses

$$\partial_\alpha=\frac{1}{2}\left(\frac{1}{\alpha_{,u}}\partial_u+\frac{1}{\alpha_{,v}}\partial_v\right)\,, \tag{2.31a}$$

$$\partial_\beta=\frac{1}{2}\left(\frac{1}{\alpha_{,v}}\partial_v-\frac{1}{\alpha_{,u}}\partial_u\right)\,, \tag{2.31b}$$

which are derived using Eq. (2.22). Here, ∂_u, ∂_v, ∂_α, and ∂_β denote, respectively, the differential operators (≡vector fields) $\partial/\partial u$, $\partial/\partial v$, $\partial/\partial\alpha$, and $\partial/\partial\beta$. A short computation involving Eqs. (2.30) and (2.31) now gives

$$-du\,dv = \frac{1}{4\alpha_{,u}\alpha_{,v}}(-d\alpha^2 + d\beta^2)\,. \tag{2.32}$$

When inserted into the expression (2.5) for the metric and combined with Eq. (2.20), Eq. (2.32) yields the expression

$$g = \frac{e^{-M}}{4\alpha^2 U_{,u} U_{,v}}(-d\alpha^2 + d\beta^2) + \alpha(e^{V} dx^2 + e^{-V} dy^2) \tag{2.33}$$

for the spacetime metric, which is valid throughout the interaction region (region I in Fig. 1). Next, another short calculation using Eqs. (2.30) and (2.31) together with Eq. (2.21) gives

$$\partial_\alpha{}^2 - \partial_\beta{}^2 = \frac{1}{\alpha_{,u}\alpha_{,v}}\partial_u\partial_v\,, \tag{2.34}$$

where $\partial_\alpha{}^2$, $\partial_\beta{}^2$, and $\partial_u\partial_v$ denote the second-order differential operators $\partial^2/\partial\alpha^2$, $\partial^2/\partial\beta^2$, and $\partial^2/\partial u\,\partial v$, respectively. Combining Eq. (2.34) with the field equation (2.6d) and using Eq. (2.31a) yields

$$V_{,\alpha\alpha} + \frac{1}{\alpha}V_{,\alpha} - V_{,\beta\beta} = 0\,, \tag{2.35}$$

which is one of the field equations in the (α,β,x,y) coordinate system. To obtain the remaining field equations, we proceed as follows: First we note that we can rewrite the field equations (2.6a) and (2.6b) in the form

$$\frac{2}{e^{M} U_{,u} U_{,v}}(e^{M} U_{,u} U_{,v})_{,u} = U_{,u} + \frac{V_{,u}{}^2}{U_{,u}} + 2U_{,u}\,,$$

$$\frac{2}{e^M U_{,u} U_{,v}} (e^M U_{,u} U_{,v})_{,v} = U_{,v} + \frac{V_{,v}{}^2}{U_{,v}} + 2U_{,v} .$$

Thus, if we define a new function P by

$$e^P \equiv 4c \; e^M \; U_{,u} U_{,v} , \tag{2.36}$$

where c is an arbitrary constant having the dimensions of $(\text{length})^2$ [we will fix c later with our normalization condition Eq. (2.40)], then P satisfies

$$2P_{,u} = 3U_{,u} + \frac{V_{,u}{}^2}{U_{,u}} , \tag{2.37a}$$

$$2P_{,v} = 3U_{,v} + \frac{V_{,v}{}^2}{U_{,v}} . \tag{2.37b}$$

Combining Eqs. (2.37) with Eqs. (2.30) and using Eq. (2.20), we obtain

$$2\alpha_{,u} (P_{,\alpha} - P_{,\beta}) = -\frac{3\alpha_{,u}}{\alpha}$$

$$-\alpha\alpha_{,u} (V_{,\alpha}{}^2 + V_{,\beta}{}^2 - 2V_{,\alpha} V_{,\beta}) ,$$

$$2\alpha_{,v} (P_{,\alpha} + P_{,\beta}) = -\frac{3\alpha_{,v}}{\alpha}$$

$$-\alpha\alpha_{,v} (V_{,\alpha}{}^2 + V_{,\beta}{}^2 + 2V_{,\alpha} V_{,\beta}) ,$$

which, after some rearrangements, can be written in the form

$$(2P + 3\ln\alpha)_{,\alpha} = -\alpha (V_{,\alpha}{}^2 + V_{,\beta}{}^2) , \tag{2.38a}$$

$$(2P+3\ln\alpha)_{,\beta}=-2\alpha V_{,\alpha}V_{,\beta}\,. \tag{2.38b}$$

Equations (2.38) suggest that it will be convenient to define the combination $2P+3\ln\alpha$ as a new variable, which, together with the variable V, would uniquely determine the metric in the (α,β,x,y) coordinate system. Thus, after first introducing the two "normalization" length scales l_1 and l_2 by the equations

$$l_1\equiv\frac{1}{2U_{,u}(u_0,v_0)}\,,\qquad l_2\equiv\frac{1}{2U_{,v}(u_0,v_0)}\,, \tag{2.39a}$$

where (u_0,v_0), $u_0>0$, $v_0>0$ is an arbitrary, fixed point in the *interior* of the interaction region, we define a new function $Q(\alpha,\beta)$ by the relation

$$e^{Q/2}\equiv 4l_1l_2e^{M}\,U_{,u}U_{,v}\,\alpha^{3/2}\,. \tag{2.39b}$$

Using Eqs. (2.39a), we can now fix the constant c which occurs in Eq. (2.36):

$$c\equiv l_1l_2\,. \tag{2.40}$$

Note that the length scales l_1 and l_2 are determined by Eqs. (2.39a) in a well-defined manner, since by Eqs. (2.20) and (2.26)

$$\begin{aligned}U(u,v)&=-\ln\alpha(u,v)\\&=-\ln\left(e^{-U_1(u)}+e^{-U_2(v)}-1\right),\end{aligned} \tag{2.41}$$

so that

$$U_{,u}(u,v)=\frac{1}{\alpha(u,v)}U_1'(u)e^{-U_1(u)}\,,$$

$$U_{,v}(u,v)=\frac{1}{\alpha(u,v)}U_2'(v)e^{-U_2(v)}\,;$$

and therefore, by Eqs. (2.29), $U_{,u}(u,v)>0$, $U_{,v}(u,v)>0$ for any point (u,v) in the interior of the interaction region, where $u>0$, $v>0$, and where [as long as (u,v) is in the domain of dependence of the initial surface $\{u=0\}\cup\{v=0\}$ (cf. Secs. III A–III C)] $\alpha(u,v)>0$. It is now easy to obtain the remaining field equations, satisfied by the new variable $Q(\alpha,\beta)$: Combining Eq. (2.39b) with Eqs. (2.40) and (2.36), and then using Eqs. (2.38), we find

$$Q_{,\alpha}=-\alpha(V_{,\alpha}{}^2+V_{,\beta}{}^2)\,, \tag{2.42a}$$

$$Q_{,\beta}=-2\alpha V_{,\alpha}V_{,\beta}\,, \tag{2.42b}$$

where the integrability condition for Eqs. (2.42) is satisfied by virtue of the field equation (2.35) for $V(\alpha,\beta)$.

We are now in a position to write down the complete formulation, in the (α,β,x,y) coordinate system, of the metric and the field equations in the interaction region of a colliding parallel-polarized plane-wave spacetime. For this, we first combine Eq. (2.39b) with the expression (2.33) for the metric in the interaction region. This gives us the expression of the interaction region metric in terms of the two unknown variables V and Q. Then, we recall the field equation (2.35) for $V(\alpha,\beta)$, and combine it with the unique solution of the field equations (2.42) for $Q(\alpha,\beta)$, which we obtain by using the initial value of Q that follows from the normalization conditions Eqs. (2.39). As a result, we obtain the following expressions for the metric and the field equations in the interaction region of a colliding plane-wave spacetime:

$$g=e^{-Q(\alpha,\beta)/2}\,\frac{l_1l_2}{\sqrt{\alpha}}\,(-d\alpha^2+d\beta^2)$$

$$+\alpha(e^{V(\alpha,\beta)}dx^2+e^{-V(\alpha,\beta)}dy^2)\,, \tag{2.43}$$

where V and Q satisfy the following field equations:

$$V_{,\alpha\alpha}+\frac{1}{\alpha}V_{,\alpha}-V_{,\beta\beta}=0\ , \tag{2.44a}$$

$$Q(\alpha,\beta)=\int_{C:(\alpha_0,\beta_0)}^{(\alpha,\beta)}[-\alpha(V_{,\alpha}{}^2+V_{,\beta}{}^2)\,d\alpha$$

$$-2\alpha V_{,\alpha}V_{,\beta}\,d\beta\,]$$

$$+2M(\alpha_0,\beta_0)+3\ln\alpha_0\ . \tag{2.44b}$$

Here, $\alpha_0\equiv\alpha(u_0,v_0)$, $\beta_0\equiv\beta(u_0,v_0)$, $M(\alpha_0,\beta_0)\equiv M(u_0,v_0)$, and C is *any* (differentiable) curve in the (α,β) plane that starts at the initial point (α_0,β_0), and ends at the field point (α,β) at which Q is to be computed. The result of the integral in Eq. (2.44b) depends only on the end points of the curve C, since the integrability condition for Eqs. (2.42) is satisfied by virtue of the field equation (2.44a).

Equations (2.43) and (2.44) summarize the mathematical problem of colliding parallel-polarized plane waves in a remarkably compact form. In particular, the only unknown to be solved for is the function $V(\alpha,\beta)$ which satisfies the *linear* field equation (2.44a). Once $V(\alpha,\beta)$ is known, Q is determined by the explicit expression (2.44b) up to an unknown additive constant, which — by suitably choosing the initial point (u_0,v_0) [or (α_0,β_0)] — can be made arbitrarily small. The only disadvantage of this formalism based on the (α,β,x,y) coordinates is the coordinate singularity that the (α,β) chart develops on the characteristic initial surface $\{u=0\}\cup\{v=0\}$. This coordinate singularity causes, among other things, the function $Q(\alpha,\beta)$ to be logarithmically divergent (to $-\infty$) on the surfaces $\{u=0\}$ and $\{v=0\}$. Nevertheless, it is still possible to set up a well-defined initial-value problem for the function $V(\alpha,\beta)$, involving the initial data posed on the same characteristic surface $\{u=0\}\cup\{v=0\}$.

It becomes clear from Eqs. (2.43) and (2.44), that the "surface" $\{\alpha=0\}$ represents some kind of a singularity [either a spacetime singularity or (at least) a coordinate singularity] of the colliding plane-wave solution described by the metric (2.43). Since we are primarily interested in the behavior of the spacetime near this "surface" $\{\alpha=0\}$, which is bounded away from the coordinate singularity on the initial null surfaces, the formalism based on the new (α,β) variables is well suited to our objectives.

In the remaining two paragraphs of this section, we will describe the initial-value problem for the metric function $V(\alpha,\beta)$ and its solution. First, in the next paragraph, we explain how to pose the initial data given by Eq. (2.15), in the new formalism based on the (α,β) coordinates. Then, in the following paragraph, we give the explicit solution of this initial-value problem for $V(\alpha,\beta)$.

We begin by noting that [cf. Eqs. (2.26) and (2.27)] in the α,β coordinates the initial null surfaces $\{u=0\}$ and $\{v=0\}$ are expressed as (Fig. 1)

$$\{u=0\}\equiv\{\alpha-\beta=1\}\ , \qquad \{v=0\}\equiv\{\alpha+\beta=1\}\ . \tag{2.45}$$

Equations (2.45), together with Eq. (2.44a), suggest introducing the "characteristic" coordinates

$$r\equiv\alpha-\beta\ , \qquad s\equiv\alpha+\beta\ , \tag{2.46}$$

so that the initial null surfaces become

$$\{u=0\}\equiv\{r=1\}\ , \qquad \{v=0\}\equiv\{s=1\}\ . \tag{2.47}$$

In the new (r,s) coordinate system [Eqs. (2.46)], the field equation (2.44a) takes the form

$$V_{,rs}+\frac{1}{2(r+s)}(V_{,r}+V_{,s})=0\ , \tag{2.48}$$

which is a partial differential equation for the function $V(r,s)$. The initial-value problem for $V(r,s)$ consists of Eq. (2.48), and the initial data on the characteristic initial surface $\{r=1\}\cup\{s=1\}$ given by the freely specifiable functions $V(r,s=1)$ and $V(r=1,s)$. More precisely, the initial data consist of

$$\{V(r,1),V(1,s)\}\ , \tag{2.49}$$

where $V(r,1)$ and $V(1,s)$ are C^1 (and piecewise C^2) functions for $r\in(-1,1]$ and $s\in(-1,1]$, respectively, which are freely specified except for the initial conditions $V(r=1,1)=V(1,s=1)=0$. Once the initial-value problem (2.48) and (2.49) is solved for the function $V(r,s)$, the function $V(\alpha,\beta)$ is determined by the obvious expression

$$V(\alpha,\beta)\equiv V(r=\alpha-\beta,s=\alpha+\beta)\ . \tag{2.50}$$

There is a one-to-one correspondence between the initial data of the form (2.15), and initial data of the form (2.49), for the initial-value problem of colliding parallel-polarized plane waves. When initial data are given in the form of Eq. (2.15), i.e., when the functions $V_1(u)$ and $V_2(v)$ are specified, initial data in the form of Eq. (2.49) are uniquely determined in the following way: First, Eqs. (2.16) are solved with the given data $V_1(u)$ and $V_2(v)$, and the functions $U_1(u)$ and $U_2(v)$ are obtained as the unique solutions [cf. Eqs. (2.16) and the discussion following them]. Then, using the identities [cf. Eqs. (2.26) and (2.27) and Eq. (2.46)]

$$r=2a(u)=2e^{-U_1(u)}-1\ , \qquad s=2b(v)=2e^{-U_2(v)}-1$$

(2.51)

along the initial null surfaces $\{u=0\}$ and $\{v=0\}$, $u(r)$ and $v(s)$ are defined as the unique solutions to the implicit equations

$$r=2e^{-U_1[u(r)]}-1\ , \qquad s=2e^{-U_2[v(s)]}-1\ . \tag{2.52}$$

Finally, the initial data $\{V(r,1),V(1,s)\}$ in the form (2.49) are determined uniquely from the data $\{V_1(u),V_2(v)\}$ by

$$V(r,1)=V_1[u=u(r)]\ , \qquad V(1,s)=V_2[v=v(s)]\ . \tag{2.53}$$

Conversely, when initial data are given in the form of Eq. (2.49), i.e., when the functions $V(r,1)$ and $V(1,s)$ are specified, initial data in the form of (2.15) are uniquely determined in the following way: First, the differential equations

$$2U_{1,uu}-U_{1,u}{}^2$$

$$=4e^{-2U_1}U_{1,u}{}^2\,[V_{,r}(r=2e^{-U_1}-1,1)]^2\ , \tag{2.54a}$$

$$2U_{2,vv}-U_{2,v}{}^2$$

$$=4e^{-2U_2}U_{2,v}{}^2\,[V_{,s}(1,s=2e^{-U_2}-1)]^2\ , \tag{2.54b}$$

for the functions $U_1(u)$ and $U_2(v)$ are solved with the initial conditions $U_1(u=0)=U_2(v=0)=U_{1,u}(u=0)=U_{2,v}(v=0)=0$ [cf. Eqs. (2.16)]. Then, using Eqs. (2.52), the initial data $\{V_1(u),V_2(v)\}$ in the form (2.15) are determined uniquely from the data $\{V(r,1),V(1,s)\}$ by

$$V_1(u)=V(r=2e^{-U_1(u)}-1,1)\ ,$$

$$V_2(v)=V(1,s=2e^{-U_2(v)}-1)\ . \tag{2.55}$$

This completes the formulation of the initial-value problem for the function $V(\alpha,\beta)$, or, equivalently, for the function $V(r,s)$ [cf. Eq. (2.50)].

The solution to a two-dimensional linear hyperbolic initial-value problem of the form (2.48) and (2.49) is obtained by using the appropriate Riemann function (Sec. 4.4 of Ref. 18). Specifically, the Riemann function for Eq. (2.48) is a two-point function $A(r,s;\xi,\eta)$, which satisfies the adjoint[18] equation to Eq. (2.48);

$$A_{,rs}-\frac{1}{2(r+s)}(A_{,r}+A_{,s})+\frac{1}{(r+s)^2}A=0\ , \tag{2.56}$$

with the initial values

$$A(r,\eta;\xi,\eta)=\left(\frac{r+\eta}{\xi+\eta}\right)^{1/2}\ ,$$

$$A(\xi,s;\xi,\eta)=\left(\frac{\xi+s}{\xi+\eta}\right)^{1/2}\ . \tag{2.57}$$

Once the Riemann function A is known, the solution $V(r,s)$ of the initial-value problem (2.48) and (2.49) is given by (Sec. 4.4 of Ref. 18)

$$V(r,s)=A(1,1;r,s)V(1,1)+\int_1^s\left[V_{,s'}(1,s')+\frac{V(1,s')}{2(1+s')}\right]A(1,s';r,s)ds'$$

$$+\int_1^r\left[V_{,r'}(r',1)+\frac{V(r',1)}{2(1+r')}\right]A(r',1;r,s)dr'\ . \tag{2.58}$$

It is found by Szekeres in Ref. 6, that the unique Riemann function which solves Eq. (2.56) with the boundary values (2.57) is

$$A(r,s;\xi,\eta)=\left[\frac{r+s}{\xi+\eta}\right]^{1/2} P_{-1/2}\left[1+2\frac{(r-\xi)(s-\eta)}{(r+s)(\xi+\eta)}\right], \tag{2.59}$$

where $P_{-1/2}$ is the Legendre function P_ν for $\nu=-\frac{1}{2}$ [Ref. 19, Eqs. (8.820)–(8.222)]. Combining Eq. (2.59) with Eq. (2.58), and noting that $V(1,1)=0$ [Eq. (2.49)], we obtain the following explicit solution $V(r,s)$ of the initial-value problem (2.48) and (2.49):

$$V(r,s)=\int_1^s\left[V_{,s'}(1,s')+\frac{V(1,s')}{2(1+s')}\right]\left[\frac{1+s'}{r+s}\right]^{1/2} P_{-1/2}\left[1+2\frac{(1-r)(s'-s)}{(1+s')(r+s)}\right]ds'$$

$$+\int_1^r\left[V_{,r'}(r',1)+\frac{V(r',1)}{2(1+r')}\right]\left[\frac{1+r'}{r+s}\right]^{1/2} P_{-1/2}\left[1+2\frac{(1-s)(r'-r)}{(1+r')(r+s)}\right]dr'. \tag{2.60}$$

We have thus completed the full solution of the initial-value problem for colliding parallel-polarized gravitational plane waves, expressed in the (α,β,x,y) coordinate system that we constructed in the beginning of this section. We are now ready to study the asymptotic structure of the colliding plane-wave spacetime near the singularity $\alpha=0$.

III. THE ASYMPTOTIC STRUCTURE OF SPACETIME NEAR $\alpha=0$

A. The behavior of the metric near α=0 : An inhomogeneous Kasner singularity

Before embarking on a full mathematical analysis of the asymptotic structure of the metric (2.43) near α=0, we use the field equations (2.44) to make a few introductory observations about the asymptotic behavior of the functions $V(\alpha,\beta)$ and $Q(\alpha,\beta)$. These observations yield some preliminary insights into the asymptotic structure of the metric (2.43) which we will find useful both in this section and in the next one.

Our starting point is the solution of Eq. (2.44a) by the well-known method of separation of variables. Using this method, we easily find that the *formal* solution to Eq. (2.44a) can be written in the form

$$V(\alpha,\beta)=\int_{-\infty}^{+\infty}[(A_k \sin k\beta+B_k \cos k\beta)N_0(k\alpha)$$

$$+(C_k \sin k\beta+D_k \cos k\beta)J_0(k\alpha)]dk \ , \tag{3.1}$$

where J_0 and N_0 are the Bessel functions of the first and second kind, respectively. Using the series representations for J_0 and N_0 given by Eqs. (8.441) and (8.444) of Ref. 19, we find that Eq. (3.1) yields the expression

$$V(\alpha,\beta)=\varepsilon(\beta)\ln\alpha+\delta(\beta)+H(\alpha,\beta) \tag{3.2}$$

for $V(\alpha,\beta)$, where

$$\varepsilon(\beta)=\frac{2}{\pi}\int_{-\infty}^{+\infty}(A_k \sin k\beta+B_k \cos k\beta)dk \ , \tag{3.3a}$$

$$\delta(\beta)=\int_{-\infty}^{+\infty}(C_k \sin k\beta+D_k \cos k\beta)dk$$

$$+\frac{2}{\pi}\int_{-\infty}^{+\infty}[\gamma+\ln(\frac{1}{2}k)](A_k\sin k\beta+B_k\cos k\beta)dk \ , \tag{3.3b}$$

$$\lim_{\alpha\to 0} H(\alpha,\beta)\equiv 0 \ , \tag{3.3c}$$

and where γ is Euler's constant.[19] From Eq. (3.2) it immediately follows, using the field equation (2.44b), that the asymptotic structure of the function $Q(\alpha,\beta)$ near $\alpha=0$ is determined by

$$Q(\alpha,\beta)=-[\varepsilon(\beta)]^2\ln\alpha+\mu(\beta)+L(\alpha,\beta) \ , \tag{3.4}$$

where

$$\lim_{\alpha\to 0} L(\alpha,\beta)\equiv 0 \ , \tag{3.5}$$

and $\mu(\beta)$ is a (C^1) function of β determined by an expression similar to Eq. (3.3b). Note that the functions $H(\alpha,\beta)$ and $L(\alpha,\beta)$ [although they remain finite (in fact, vanish) as $\alpha\to 0$] are *not* smooth functions near $\alpha=0$. In fact, it follows from the series expansions for J_0 and N_0 [Eqs. (8.441)–(8.444) in Ref. 19], that, for example, the function $H(\alpha,\beta)$ has the behavior

$$\begin{aligned} H(\alpha,\beta)=&c_1(\beta)\alpha^2\ln\alpha+c_2(\beta)\alpha^4\ln\alpha+\ldots \\ &+c_k(\beta)\alpha^{2k}\ln\alpha+\ldots+d_1(\beta)\alpha^2 \\ &+d_2(\beta)\alpha^4+\ldots+d_k(\beta)\alpha^{2k}+\ldots \ , \end{aligned} \tag{3.6}$$

where $c_k(\beta)$ and $d_k(\beta)$ are functions of β determined by expressions similar to Eq. (3.3b). Equation (3.6), when combined with Eq. (3.2), yields a more detailed expression for the asymptotic structure of $V(\alpha,\beta)$ near $\alpha=0$:

$$V(\alpha,\beta)=\varepsilon(\beta)\ln\alpha+c_1(\beta)\alpha^2\ln\alpha+c_2(\beta)\alpha^4\ln\alpha+\ldots$$

$$+\delta(\beta)+d_1(\beta)\alpha^2+d_2(\beta)\alpha^4+\ldots . \tag{3.7}$$

Equations (3.7) and (3.4) summarize the asymptotic behavior of the metric functions V and Q near the singularity $\alpha=0$. But Eqs. (3.3) are not terribly useful for expressing the key functions $\varepsilon(\beta)$ and $\delta(\beta)$ in terms of the initial data (2.15) or (2.49) for the colliding plane waves. For this purpose, it is better to use the explicit solution (2.60) for the function $V(r,s)$ that we obtained in the last section. Thus, in the next paragraph, we will analyze the asymptotic behavior of the solution (2.60) near the singularity $\alpha=0$, and obtain explicit formulas for the functions $\varepsilon(\beta)$ and $\delta(\beta)$ expressed in terms of the initial data (2.49) for the incoming waves. Then, in the remainder of this section, we will use the analysis carried out so far to investigate the asymptotic structure of the spacetime metric (2.43) near the singular surface $\alpha=0$.

Note that in the (r,s) coordinate system of Sec. II B, the singularity $\alpha=0$ corresponds to $r+s=0$ [Eqs. (2.46)]. Combining Eq. (2.60) with Eq. (2.50), it is clear that the asymptotic structure of $V(\alpha,\beta)$ near $\alpha=0$ is determined by the asymptotic behavior of the function $P_{-1/2}(1+2z)$ as $z\rightarrow\infty$. To evaluate this asymptotic behavior, we first note that the integral representation [Eq. (8.822) in Ref. 19]

$$P_{-1/2}(z)=\frac{1}{\pi}\int_0^\pi (z+\sqrt{z^2-1}\cos\phi)^{-1/2}d\phi, \qquad \mathrm{Re}\, z>0 \tag{3.8}$$

can be rewritten in the form

$$P_{-1/2}(z)=\frac{2}{\pi}(z+\sqrt{z^2-1})^{-1/2}$$

$$\times K\left[\left(\frac{2\sqrt{z^2-1}}{z+\sqrt{z^2-1}}\right)^{1/2}\right], \tag{3.9}$$

where K is the complete elliptic integral of the second kind.[19] Subsequently, the asymptotic expression [Eq. (8.113) in Ref. 19]

$$K(k)=-\ln(\sqrt{1-k^2})+\ln 4+O(1-k^2)\ ,$$

when combined with Eq. (3.9), yields the asymptotic relation

$$P_{-\frac{1}{2}}(z)=\frac{\sqrt{2}}{\pi}z^{-\frac{1}{2}}\ln z+\frac{3\sqrt{2}\ln 2}{\pi}z^{-\frac{1}{2}}$$

$$+O\left[\frac{1}{z^2}\right]\ , \tag{3.10}$$

which in turn yields the desired asymptotic behavior

$$P_{-\frac{1}{2}}(1+2z)=\frac{1}{\pi}z^{-\frac{1}{2}}\ln z+\frac{3\ln 2}{\pi}z^{-\frac{1}{2}}$$

$$+O\left[\frac{1}{z^{3/2}}\right] \tag{3.11}$$

as $z\to\infty$. Now we combine Eq. (3.11) with Eq. (2.60) and Eq. (2.50), and then compare the resulting asymptotic form of $V(\alpha,\beta)$ with Eqs. (3.2) and (3.7) to read out the following explicit expressions for the functions $\varepsilon(\beta)$ and $\delta(\beta)$:

$$\varepsilon(\beta)=\frac{1}{\pi}\int_{\beta}^{1}\left[V_{,s}(1,s)+\frac{V(1,s)}{2(1+s)}\right]\frac{(1+s)}{\sqrt{(1+\beta)(s-\beta)}}ds$$

$$+\frac{1}{\pi}\int_{-\beta}^{1}\left[V_{,r}(r,1)+\frac{V(r,1)}{2(1+r)}\right]\frac{(1+r)}{\sqrt{(1-\beta)(r+\beta)}}dr\ , \tag{3.12a}$$

$$\delta(\beta)=-\int_{\beta}^{1}\left[V_{,s}(1,s)+\frac{V(1,s)}{2(1+s)}\right]\frac{(1+s)}{\sqrt{(1+\beta)(s-\beta)}}\left[\frac{2\ln 2}{\pi}+\frac{1}{\pi}\ln\left[\frac{(1+\beta)(s-\beta)}{1+s}\right]\right]ds$$

$$
-\int_{-\beta}^{1}\left[V_{,r}(r,1)+\frac{V(r,1)}{2(1+r)}\right]\frac{(1+r)}{\sqrt{(1-\beta)(r+\beta)}}\left[\frac{2\ln 2}{\pi}+\frac{1}{\pi}\ln\left(\frac{(1-\beta)(r+\beta)}{1+r}\right)\right]dr\ .
$$

(3.12b)

We note that Eq. (3.12a) can be rewritten in the simpler form

$$
\varepsilon(\beta)=\frac{1}{\pi}\frac{1}{\sqrt{1+\beta}}\int_{\beta}^{1}[(1+s)^{1/2}V(1,s)]_{,s}\left(\frac{s+1}{s-\beta}\right)^{1/2}ds
$$

$$
+\frac{1}{\pi}\frac{1}{\sqrt{1-\beta}}\int_{-\beta}^{1}[(1+r)^{1/2}V(r,1)]_{,r}\left(\frac{r+1}{r+\beta}\right)^{1/2}dr\ . \tag{3.13}
$$

The timelike coordinate α is a parameter which monotonically decreases to zero along the world line of any observer approaching the singularity. Consider the space-time metric (2.43) in the vicinity of such a world line as the observer approaches the singularity $\alpha=0$ at a fixed spatial coordinate β. According to Eqs. (3.4) and (3.7), the asymptotic behavior of the metric along the observer's world line as $\alpha\rightarrow 0$ can be expressed as

$$
g(\beta)\sim e^{-\mu(\beta)/2}l_1 l_2\,\alpha^{q_1(\beta)}(-d\alpha^2+d\beta^2)
$$

$$
+e^{\delta(\beta)}\,\alpha^{q_2(\beta)}dx^2+e^{-\delta(\beta)}\,\alpha^{q_3(\beta)}dy^2\ ,
$$

(3.14)

where

$$
q_1(\beta)=\frac{1}{2}[\varepsilon^2(\beta)-1]\ ,\qquad q_2(\beta)=1+\varepsilon(\beta)\ ,
$$

$$q_3(\beta)=1-\varepsilon(\beta)\ . \tag{3.15}$$

On the right-hand side of Eq. (3.14), all quantities that depend on β are to be regarded as constants when interpreting the metric $g(\beta)$ as the asymptotic limit of the metric (2.43); this asymptotic metric describes a region of spacetime which is arbitrarily large in the Killing x, y directions, but which extends (in general) very little [over a range in β small enough for the variation in $\varepsilon(\beta)$ to be negligible] in the β direction, and which covers a range $(0,\eta)$ in the coordinate α where η is arbitrarily small ($\eta\to 0$). Now, notice that the quantity $q_1(\beta)$ is always greater than -2 [$q_1\geq -\frac{1}{2}$ by Eqs. (3.15)]. Thus, we can introduce a new timelike coordinate t

$$t\equiv\alpha^{(q_1+2)/2}\ , \qquad \alpha\equiv t^{2/(q_1+2)}\ , \tag{3.16}$$

which is monotonically related to α, and in which the singularity $\alpha=0$ is located at $t=0$. In terms of this new timelike coordinate t, the asymptotic metric $g(\beta)$ of Eq. (3.14) takes the form

$$\begin{aligned} g(\beta) \sim & -\frac{4l_1 l_2 e^{-\mu(\beta)/2}}{[q_1(\beta)+2]^2}\, dt^2 + l_1 l_2 e^{-\mu(\beta)/2}\, t^{2p_3}\, d\beta^2 \\ & + e^{\delta(\beta)}\, t^{2p_1}\, dx^2 + e^{-\delta(\beta)}\, t^{2p_2}\, dy^2\ , \end{aligned} \tag{3.17}$$

where

$$p_3(\beta)=\frac{\varepsilon^2(\beta)-1}{\varepsilon^2(\beta)+3}\ , \tag{3.18a}$$

$$p_1(\beta)=\frac{2[1+\varepsilon(\beta)]}{\varepsilon^2(\beta)+3}\ , \tag{3.18b}$$

$$p_2(\beta)=\frac{2[1-\varepsilon(\beta)]}{\varepsilon^2(\beta)+3} \ . \tag{3.18c}$$

It is easily seen from Eqs. (3.18) that the exponents $p_1(\beta)$, $p_2(\beta)$, and $p_3(\beta)$ satisfy the Kasner relations[12]

$$\begin{aligned} p_1(\beta)+p_2(\beta)+p_3(\beta)&=p_1^2(\beta)+p_2^2(\beta)+p_3^2(\beta) \\ &=1 \ , \end{aligned} \tag{3.19}$$

for all values of $\varepsilon(\beta)$. Therefore, the asymptotic limit of the metric (2.43) as $\alpha\to 0$ at a fixed spatial position β is a vacuum Kasner solution, which, after absorbing the constant terms on the right-hand side of Eq. (3.17) into the definition of the coordinates, and for simplicity using units in which lengths are dimensionless, can be represented in the form

$$g(\beta)=-dt^2+t^{2p_1}dx^2+t^{2p_2}dy^2+t^{2p_3}dz^2 \ , \tag{3.20}$$

where the asymptotic Kasner exponents p_k, $k=1,2,3$, are given by Eqs. (3.18) and satisfy the relations (3.19). Equations (3.18), when combined with Eq. (3.13), provide the explicit formulas which express the asymptotic Kasner exponents $p_k(\beta)$ along the singularity in terms of the initial data (2.49) for the colliding waves.

The Kasner solution[12] defined by the global spacetime metric (3.20) has the following curvature tensor:

$$\begin{aligned} \vec{R} &= \sum_{j=1}^{3} \frac{p_j(p_j-1)}{t^2}(X_0\otimes\omega^j+X_j\otimes\omega^0)\otimes\omega^0\wedge\omega^j \\ &+\sum_{j<k}\frac{p_jp_k}{t^2}(X_j\otimes\omega^k-X_k\otimes\omega^j)\otimes\omega^j\wedge\omega^k \ , \end{aligned} \tag{3.21}$$

where the orthonormal basis $\{X_\mu\}$ and its dual $\{\omega^\nu\}$ are given by

$$X_0=\frac{\partial}{\partial t}\,, \qquad X_1=t^{-p_1}\frac{\partial}{\partial x}\,,$$

$$X_2=t^{-p_2}\frac{\partial}{\partial y}\,, \qquad X_3=t^{-p_3}\frac{\partial}{\partial z}\,, \tag{3.22a}$$

$$\omega^0=dt\,, \qquad \omega^1=t^{p_1}dx\,,$$

$$\omega^2=t^{p_2}dy\,, \qquad \omega^3=t^{p_3}dz\,. \tag{3.22b}$$

A number of fundamental properties of the Kasner spacetime can easily be deduced from the expression (3.21) of the curvature tensor. First, it becomes clear that the vacuum field equations are equivalent to the algebraic relations (3.19) for the exponents p_k. Next, a short computation using Eq. (3.21) gives

$$R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}=-\frac{16}{t^4}p_1p_2p_3 \tag{3.23}$$

provided that the vacuum conditions (3.19) are satisfied. Assuming that (3.19) hold, the following main conclusions are then obtained. (i) The "surface" $t=0$ is a curvature singularity of the vacuum Kasner solution *unless* one of the exponents is zero. This can only happen if (p_1,p_2,p_3) is equal to a permutation of (1,0,0), in which case we assume, without loss of generality, that $p_1=1, p_2=p_3=0$. For these values of the exponents (a degenerate Kasner solution) the metric (3.20) is flat [Eq. (3.21) gives $\boldsymbol{R}\equiv 0$]; the surface $\{t=0\}$ represents a Killing-Cauchy horizon[4] (a coordinate singularity) across which spacetime can be extended, e.g., to yield the maximal Minkowski space. The spacelike Killing vector $\partial/\partial x$ becomes null on this Killing-Cauchy horizon $\{t=0\}$. (ii) If all exponents are nonzero (the nondegenerate case), then one and only one of the exponents is strictly negative, while the other two are strictly positive. And finally, a straightforward application of the geodesic deviation equation with the curvature tensor (3.21) reveals that (iii) in a nondegenerate Kasner solution, timelike geodesic congruences which run into the singularity converge together in those spatial directions for which $p_k > 0$, and diverge apart in the direction for which $p_k < 0$. In other words, physical three-volumes get squashed in the two spatial directions with positive exponents, while they get infinitely stretched in the remaining direction with the negative exponent as the singularity is approached.[12]

After this brief interlude on the Kasner solution, we now return to the discussion of the asymptotic limit (3.17) and (3.18) of the colliding plane-wave metric (2.43). For much more detailed expositions on the Kasner solution (including its generalizations and their application to cosmology), the reader is referred to the literature listed as Ref. 12.

The following conclusions are easily obtained from Eqs. (3.17)–(3.20) combined with the results of our brief review of the Kasner solution: (i) If $|\varepsilon(\beta)| < 1$, then $p_1(\beta)$ and $p_2(\beta)$ are both positive and $p_3(\beta)$ is negative. This corresponds to an *anastigmatic*[2,3] singularity structure at $(\alpha=0,\beta)$; that is, focusing takes place in both the x and y directions. In particular, if the incoming plane waves are sandwich waves and either purely anastigmatic[2,3] or very nearly anastigmatic (i.e., if they have focal lengths f_1, f_2 [cf. Eqs. (2.13)–(2.14)] which are either equal ($f_1=f_2$) or satisfy $|f_2-f_1|/f_1 \ll 1$), and if both incoming waves are sufficiently weak [i.e., if $V_1, V_2 \ll 1$, cf. Eq. (2.15)], then Eq. (3.13) implies that at least throughout a large subinterval of the range $(-1,1)$ of β, $|\varepsilon(\beta)|$ will be much smaller than 1. Thus, under these circumstances, the structure of the singularity will be mostly anastigmatic. (ii) If, on the contrary, $|\varepsilon(\beta)| > 1$, then $p_3(\beta)$ is positive and one of $p_1(\beta), p_2(\beta)$ is negative. This corresponds to an astigmatic singularity structure at (α=0,β); that is, focusing occurs in only one of the two transverse directions x, y, whereas in the other direction an infinite defocusing takes place. In particular, if the incoming plane waves are highly astigmatic ($|f_2-f_1|/f_1 \gg 1$), or if they are sufficiently strong ($V_1, V_2 \sim 1$), then it is possible to have an interval in β with $|\varepsilon(\beta)|>1$, that is, an interval in β with an astigmatic singularity structure at α=0. (See, however, our second example in Sec. IV in which colliding highly astigmatic plane waves create a purely anastigmatic singularity.) (iii) Finally, if $|\varepsilon(\beta)|=1$, then $p_3(\beta)=0$ and one of $p_1(\beta)$, $p_2(\beta)$ is 1 whereas the other is zero. In this case the asymptotic metric $g(\beta)$ near α=0 is a degenerate Kasner solution (3.20) with either $(p_1,p_2,p_3)=(1,0,0)$ or $(p_1,p_2,p_3)=(0,1,0)$.

It seems evident that if the quantity $\varepsilon(\beta)$ is different from ± 1 (across an interval in β or at any point $\beta=\beta_0$), then the colliding plane-wave solution (2.43) has a curvature singularity at (α=0,β). On the other hand, in view of our conclusion (iii) in the above paragraph, it is also quite natural to expect that if $\varepsilon(\beta)\equiv\pm 1$ throughout an interval (β_1,β_2) in β, then the portion $\{\alpha=0,\beta_1<\beta<\beta_2\}$ of the surface $\{\alpha=0\}$ is *not* a curvature singularity, but instead it represents a nonsingular Killing-Cauchy horizon of the colliding plane-wave spacetime on which one of the spacelike Killing vector fields $\partial/\partial x$, $\partial/\partial y$ becomes null, and across which the metric can be smoothly extended.[4] However, our analysis so far is not sufficient to reach these conclusions rigorously. The reason is that although we now know the asymptotic limit of the metric (2.43) explicitly [Eq. (3.17)], we do not yet have full control on the asymptotic behavior of the spacetime curvature near the singularity α=0. In other words, in view of the presence of a whole series of logarithmic terms in the expansion (3.7) of $V(\alpha,\beta)$ [and similarly in the expansion (3.4) of $Q(\alpha,\beta)$], it is not clear *a priori* that the asymptotically Kasner nature of the metric as $\alpha\rightarrow 0$ implies the corresponding asymptotically

Kasner (as $t \to 0$) behavior of the spacetime curvature (which involves the derivatives of the metric). Thus, in the following section, we are going to study the behavior of the curvature associated with the metric (2.43) near the singularity $\alpha=0$.

B. The behavior of curvature near $\alpha=0$

Since in most of the literature on colliding plane waves[5,6,7,9,11] the spacetime curvature is studied in terms of the Newman-Penrose curvature quantities, we will also find it convenient to carry out our analysis of curvature using the curvature quantities (2.19) with respect to the standard tetrad (2.3) and (2.4). Equations (2.19) express these curvature quantites in terms of the tetrad coefficients M, U, and V, and in the Rosen-type coordinate system (u,v,x,y) for the colliding plane-wave spacetime. In this section, we will first obtain the corresponding formulas expressing the *same functions* Ψ_0, Ψ_2, and Ψ_4 in terms of the metric functions $V(\alpha,\beta)$ and $Q(\alpha,\beta)$, and in our favorite (α,β,x,y) coordinate system. Then, using these expressions, we will read out the asymptotic behavior of the curvature quantities as $\alpha \to 0$.

Consider first Eq. (2.19b) for the quantity Ψ_2. Combining this equation with Eq. (2.34) and Eq. (2.39b), and using Eq. (2.20), we obtain

$$\Psi_2=-\frac{e^{Q/2}}{4l_1 l_2}\alpha^{1/2}(\partial_\alpha{}^2-\partial_\beta{}^2)M \ . \tag{3.24}$$

Now note the following identities

$$\ln(U_{,u}U_{,v})=\ln(\alpha_{,u})+\ln(\alpha_{,v})-2\ln\alpha \ , \tag{3.25a}$$

$$\partial_u\partial_v\ln(\alpha_{,u})=0 \ , \qquad \partial_u\partial_v\ln(\alpha_{,v})=0 \ , \tag{3.25b}$$

which are derived by using Eqs. (2.20) and (2.21), respectively. If we take the logarithm of both sides in Eq. (2.39b) and apply the operator $\partial_\alpha^2-\partial_\beta^2$ on both sides of the result, and if we then use Eqs. (2.34) and (3.25) to simplify, we obtain the identity

$$(\partial_\alpha^2-\partial_\beta^2)M=\frac{1}{2}(Q_{,\alpha\alpha}-Q_{,\beta\beta}-\alpha^{-2})\,, \tag{3.26}$$

which, when combined with Eq. (3.24), yields the desired expression for the quantity Ψ_2:

$$\Psi_2=-\frac{e^{Q(\alpha,\beta)/2}}{8l_1l_2}\,\alpha^{1/2}\left[Q_{,\alpha\alpha}-Q_{,\beta\beta}-\frac{1}{\alpha^2}\right]. \tag{3.27}$$

The calculation of the corresponding expressions for the remaining curvature quantities Ψ_0 and Ψ_4 proceeds along similar lines. Substituting Eq. (2.30a) in the expression (2.19a) for Ψ_0, and then making use of the identity

$$(\partial_\alpha-\partial_\beta)M=\frac{1}{2}\left[Q_{,\alpha}-Q_{,\beta}+\frac{1}{\alpha}\right]-\frac{\alpha_{,uu}}{\alpha_{,u}{}^2}\,,$$

together with the Eqs. (2.20) and (2.21), we obtain

$$\Psi_0=\frac{i\,\alpha}{8l_1{}^2l_2{}^2}\,\frac{e^{Q(\alpha,\beta)}}{\alpha_{,v}{}^2}\Big[\frac{1}{2}(V_{,\alpha}-V_{,\beta})(Q_{,\alpha}-Q_{,\beta}+\frac{3}{\alpha})$$

$$+V_{,\alpha\alpha}+V_{,\beta\beta}-2V_{,\alpha\beta}\Big]\,.$$

(3.28)

And, substituting Eq. (2.30b) in the expression (2.19c) for Ψ_4, and then making use of the identity

$$(\partial_\alpha+\partial_\beta)M=\frac{1}{2}\left[Q_{,\alpha}+Q_{,\beta}+\frac{1}{\alpha}\right]-\frac{\alpha_{,vv}}{\alpha_{,v}{}^2}\,,$$

together with the Eqs. (2.20) and (2.21), we obtain

$$\Psi_4 = -\frac{i}{2}\alpha_{,v}{}^2\left[\frac{1}{2}(V_{,\alpha}+V_{,\beta})(Q_{,\alpha}+Q_{,\beta}+\frac{3}{\alpha})\right.$$

$$\left. +V_{,\alpha\alpha}+V_{,\beta\beta}+2V_{,\alpha\beta}\right] . \tag{3.29}$$

Note that the quantity $\alpha_{,v}$ that occurs in the expressions (3.28) and (3.29) is not fully expressed in terms of the metric functions $Q(\alpha,\beta)$ and $V(\alpha,\beta)$. However, by Eq. (2.26), $\alpha_{,v}=-U_{2,v}e^{-U_2(v)}$, and thus by the Eqs. (2.29) and the discussion preceding them, $\alpha_{,v}$ is nonzero and finite for all $-1<\beta<1$ in the limit $\alpha\to0$. Thererefore, the multiplicative factors involving $\alpha_{,v}$ in Eqs. (3.28) and (3.29) do not contribute to the qualitative asymptotic behavior of the curvature quantities near $\alpha=0$. Hence, we will not attempt to further express the quantity $\alpha_{,v}$ in terms of the metric functions Q and V; instead, we will regard it as a (nonzero) constant, multiplying the asymptotic limits of Ψ_0 and Ψ_4 as $\alpha\to0$ at a *fixed* spatial location β.

Before proceeding with the analysis of the asymptotic behaviors of Ψ_2, Ψ_0, and Ψ_4 near $\alpha=0$, we rewrite Eqs. (3.27)–(3.29) in terms of the metric function $V(\alpha,\beta)$. After eliminating the terms that involve the derivatives of $Q(\alpha,\beta)$ by making use of Eqs. (2.42), we obtain

$$\Psi_2 = -\frac{e^{Q(\alpha,\beta)/2}}{8l_1l_2}\,\alpha^{1/2}\left[-V_{,\alpha}{}^2-V_{,\beta}{}^2+2\alpha V_{,\alpha}(V_{,\beta\beta}-V_{,\alpha\alpha})-\frac{1}{\alpha^2}\right] , \tag{3.30}$$

$$\Psi_0 = \frac{i}{8l_1{}^2l_2{}^2\alpha_{,v}{}^2}\,e^{Q(\alpha,\beta)}\,\alpha$$

$$\times\left[\frac{1}{2}(V_{,\alpha}-V_{,\beta})\left[2\alpha V_{,\alpha}V_{,\beta}-\alpha V_{,\alpha}{}^2-\alpha V_{,\beta}{}^2+\frac{3}{\alpha}\right]+V_{,\alpha\alpha}+V_{,\beta\beta}-2V_{,\alpha\beta}\right] , \tag{3.31}$$

$$\Psi_4=-\frac{i}{2}\alpha_{,v}{}^2\left[\frac{1}{2}(V_{,\alpha}+V_{,\beta})\left(-2\alpha V_{,\alpha}V_{,\beta}-\alpha V_{,\alpha}{}^2-\alpha V_{,\beta}{}^2+\frac{3}{\alpha}\right)+V_{,\alpha\alpha}+V_{,\beta\beta}+2V_{,\alpha\beta}\right] . \tag{3.32}$$

Now, substituting the asymptotic expressions (3.7) and (3.4) for $V(\alpha,\beta)$ and $Q(\alpha,\beta)$ into Eqs. (3.30)–(3.32), we obtain, after some lengthy algebra, the following equations revealing the asymptotic behavior of the curvature quantities Ψ_2, Ψ_0, and Ψ_4, as $\alpha \to 0$ at a fixed spatial coordinate β:

$$\Psi_2(\beta) \sim -\frac{e^{\mu(\beta)/2}}{8l_1l_2}\,\alpha^{[1-\varepsilon^2(\beta)]/2}\left(\frac{\varepsilon^2(\beta)-1}{\alpha^2}+\frac{1}{\alpha}O(\alpha)\right) . \tag{3.33}$$

$$\Psi_0(\beta) \sim \frac{ie^{\mu(\beta)}}{8l_1{}^2l_2{}^2\alpha_{,v}{}^2}\,\alpha^{[1-\varepsilon^2(\beta)]}$$

$$\times\left(\frac{\varepsilon(\beta)[1-\varepsilon^2(\beta)]}{2\alpha^2}+\frac{3[\varepsilon^2(\beta)-1][\varepsilon'(\beta)\ln\alpha+\delta'(\beta)]}{2\alpha}-\frac{2\varepsilon'(\beta)}{\alpha}+\frac{1}{\alpha}O(\alpha)\right) . \tag{3.34}$$

$$\Psi_4(\beta) \sim -\frac{i}{2}\alpha_{,v}{}^2$$

$$\times\left(\frac{\varepsilon(\beta)[1-\varepsilon^2(\beta)]}{2\alpha^2}-\frac{3[\varepsilon^2(\beta)-1][\varepsilon'(\beta)\ln\alpha+\delta'(\beta)]}{2\alpha}+\frac{2\varepsilon'(\beta)}{\alpha}+\frac{1}{\alpha}O(\alpha)\right) . \tag{3.35}$$

In Eqs. (3.33)–(3.35), $O(\alpha)$ denote the remaining terms which are always of the form

$$O(\alpha)\equiv(\ )\alpha\ln\alpha+(\ )\alpha^2\ln\alpha+\ldots+(\ )\alpha(\ln\alpha)^2$$

$$+(\ )\alpha^2(\ln\alpha)^2+\ldots+(\ )\alpha+(\ )\alpha^2+\ldots+ , \tag{3.36}$$

where the "( )" denote well-behaved quantities that depend only on β.

Equations (3.33)–(3.35) provide a clear demonstration of our earlier statement (Sec. III A) that whenever $|\varepsilon(\beta)|\neq 1$ (across an interval in β or at an isolated point $\beta=\beta_0$), the colliding plane-wave spacetime possesses a curvature singularity at $(\alpha=0,\beta)$. [The asymptotic form of the curvature invariant (3.23) can be computed using Eqs. (3.33)–(3.35) along with the identities $\Psi_1=\Psi_3\equiv 0$; it is easily seen that as $\alpha\rightarrow 0$ this invariant diverges in accordance with Eqs. (3.23) and (3.16), i.e., as $\sim\alpha^{-[\varepsilon^2(\beta)+3]}$, whenever $|\varepsilon(\beta)|\neq 1$.] In order to prove our second statement (Sec. III A), that when $|\varepsilon(\beta)|\equiv 1$ throughout an interval (β_1,β_2) the surface $\{\alpha=0,\beta_1<\beta<\beta_2\}$ is a nonsingular Killing-Cauchy horizon,[4] we will need to perform a somewhat more detailed analysis of the asymptotic behavior of $V(\alpha,\beta)$ near $\alpha=0$. Thus, in the few remaining paragraphs of this section, we will present such an analysis and see that our conclusions indeed provide a proof for this second statement. Then, in the next section (Sec. III C), we will discuss the physical significance and the instabilities of these Killing-Cauchy horizons which occur at $\alpha=0$.

Before proceeding with our discussion, we note that when $|\varepsilon(\beta)|\equiv 1$ across an interval in β, all divergent terms in the expressions (3.33)–(3.35) vanish except (possibly) for the logarithmically divergent terms which could be introduced by the remainders $O(\alpha)/\alpha$ [Eq. (3.36)]. To learn more about these logarithmic terms, consider the expansion (3.7) for $V(\alpha,\beta)$. Equations (3.12) and (3.13) give expressions for the two most important coefficients $\varepsilon(\beta)$ and $\delta(\beta)$ which occur in this expansion, and the other coefficients $c_k(\beta)$ and $d_k(\beta)$ can be computed from the original field equation (2.44a) for V: The following expressions for the derivatives of V that occur in Eq. (2.44a) are obtained straightforwardly by using Eq. (3.7):

$$\frac{1}{\alpha}V_{,\alpha}=\frac{\varepsilon(\beta)}{\alpha^2}+2c_1(\beta)\ln\alpha+4c_2(\beta)\alpha^2\ln\alpha$$

$$+[c_1(\beta)+2d_1(\beta)]+[c_2(\beta)+4d_2(\beta)]\alpha^2$$

$$+O(\alpha^4)\,, \tag{3.37a}$$

$$V_{,\alpha\alpha}=-\frac{\varepsilon(\beta)}{\alpha^2}+2c_1(\beta)\ln\alpha+12c_2(\beta)\alpha^2\ln\alpha$$

$$+[3c_1(\beta)+2d_1(\beta)]$$

$$+[7c_2(\beta)+12d_2(\beta)]\alpha^2+O(\alpha^4)\,, \tag{3.37b}$$

$$V_{,\beta\beta}=\varepsilon''(\beta)\ln\alpha+c_1''(\beta)\alpha^2\ln\alpha+c_2''(\beta)\alpha^4\ln\alpha$$

$$+\delta''(\beta)+d_1''(\beta)\alpha^2+d_2''(\beta)\alpha^4+O(\alpha^6)\,. \tag{3.37c}$$

Inserting Eqs. (3.37) in the field equation (2.44a) and collecting together the coefficients of identical terms in α, we obtain the identities

$$c_1(\beta)=\frac{1}{4}\varepsilon''(\beta)\,, \qquad c_2(\beta)=\frac{1}{64}\varepsilon''''(\beta)\,,\ \ldots.\,, \tag{3.38a}$$

$$d_1(\beta)=\frac{1}{4}[\delta''(\beta)-\varepsilon''(\beta)]\,,$$

$$d_2(\beta)=\frac{1}{128}[2\delta''''(\beta)-3\varepsilon''''(\beta)]\,,\ \ldots.\,, \tag{3.38b}$$

which express all of the coefficients $c_k(\beta)$ and $d_k(\beta)$ in Eq. (3.7) in terms of the coefficients $\varepsilon(\beta)$ and $\delta(\beta)$.

It now becomes clear that when $|\varepsilon(\beta)|\equiv 1$ throughout an interval (β_1,β_2) [in fact, whenever $\varepsilon(\beta)$ is constant across such an interval], all the coefficients $c_k(\beta)$, $k\geq 1$ vanish for $\beta\in(\beta_1,\beta_2)$. In that case, the expansion (3.7) of $V(\alpha,\beta)$ *does not contain any logarithmic terms except for the leading term* $\varepsilon(\beta)\ln\alpha$. In particular, the derivatives

$V_{,\alpha}, V_{,\beta}, V_{,\alpha\alpha}, V_{,\beta\beta}, V_{,\alpha\beta}$ *contain no logarithmic terms whatsoever in* α, for $\beta\in(\beta_1,\beta_2)$. Therefore, by Eqs. (3.30)–(3.32), the remainder terms $O(\alpha)$ in Eqs. (3.33)—(3.35) also do not involve *any* logarithmic terms in α across the same interval; in other words

$$O(\alpha)=(\ )\alpha+(\ )\alpha^2+\ldots \qquad \forall\ \beta\in(\beta_1,\beta_2)\,. \tag{3.39}$$

Combining Eq. (3.39) with Eqs. (3.33)—(3.35), we find that we have proved the following result.

If $|\varepsilon(\beta)|\equiv1$ *throughout an interval* (β_1,β_2) *in* β, *then the curvature quantities* Ψ_0, Ψ_2, *and* Ψ_4 *are all bounded (≡finite, but in general nonzero) as* $\alpha\to0$, *whenever* β *belongs to this interval* (β_1,β_2); *i.e., all curvature quantities are perfectly well behaved across the surface* $\{\alpha=0,\beta_1<\beta<\beta_2\}$.

Clearly, if $|\varepsilon(\beta_0)|=1$ at an isolated point $\beta=\beta_0$, and furthermore if $\varepsilon'(\beta_0)=\varepsilon''(\beta_0)=0$, then by the Eqs. (3.38a), $c_1(\beta_0)=0$, and consequently

$$O(\alpha)=(\ )\alpha^2\ln\alpha+(\ )\alpha^3\ln\alpha+\ldots+(\ )\alpha^2(\ln\alpha)^2$$

$$+\ldots++(\ )\alpha+(\ )\alpha^2+\ldots \qquad \text{at } \beta=\beta_0\,. \tag{3.40}$$

Therefore, combining Eq. (3.40) with the Eqs. (3.33)–(3.35) as above, we obtain the following similar result.

If $|\varepsilon(\beta)|=1$ *at an isolated point* $\beta=\beta_0$, *and if, in addition, the first two derivatives of* $\varepsilon(\beta)$ *at the point* β_0 *vanish, then the curvature quantities* Ψ_0, Ψ_2, *and* Ψ_4 *are bounded (≡finite, but in general nonzero) as* $\alpha\to0$ *at the point* $\beta=\beta_0$.

C. Instability of the Killing-Cauchy horizons that occur at $\alpha=0$

We begin this section by rephrasing, in a somewhat more precise format, the three fundamental conclusions of the preceding section (Sec. III B).

(i) When $|\varepsilon(\beta_0)|\neq 1$, the two-surface $\{\alpha=0, \beta=\beta_0, -\infty<x<+\infty, -\infty<y<+\infty\}$ is a curvature singularity of the colliding plane-wave metric (2.43). This singularity is of asymptotically (nondegenerate) Kasner type, and it is in general inhomogeneous in the spatial β direction.

(ii) When $|\varepsilon(\beta)|=1$ at an isolated point $\beta=\beta_0$, i.e., if the C^1 function $|\varepsilon(\beta)|$ maps a small interval containing β_0 into a real interval containing 1 *in such a way that the inverse image of 1 is a single point (namely, β_0)*, then there are two possibilities: If $\varepsilon'(\beta_0)=\varepsilon''(\beta_0)=0$, then the two-surface $\mathcal{P}=\{\alpha=0, \beta=\beta_0, -\infty<x<+\infty, -\infty<y<+\infty\}$ is *not* a curvature singularity, but it still represents a spacetime singularity since no extension of the metric is possible across $\mathcal{P}$. Such an extension does not exist because in any spacetime neighborhood of $\mathcal{P}$ there are boundary points corresponding to true curvature singularities; consequently, any extension of the metric beyond the two-surface $\mathcal{P}$ would be incompatible with the topological manifold structure of the spacetime. [Note that, by Eqs. (3.14) and (3.15), β is a regular coordinate near β_0 when $|\varepsilon(\beta_0)|=1$. Also note that similar topological singularities frequently occur in the exact solutions for colliding plane waves, see, e.g., Refs. 8, 5, and 3.] If, on the other hand, either one (or both) of $\varepsilon'(\beta_0)$, $\varepsilon''(\beta_0)$ are nonzero, then $\mathcal{P}$ is a genuine curvature singularity of the colliding plane wave metric (2.43).

(iii) Finally, when $|\varepsilon(\beta)|\equiv 1$ throughout an interval (β_1,β_2), the three-surface $\mathcal{S}=\{\alpha=0, \beta_1<\beta<\beta_2, -\infty<x<+\infty, -\infty<y<+\infty\}$ is a nonsingular Killing-Cauchy horizon for the colliding plane-wave spacetime. The asymptotic Kasner exponents $[p_1(\beta), p_2(\beta), p_3(\beta)]$ take one of the degenerate values (1,0,0) [if $\varepsilon(\beta)=+1$] or (0,1,0) [if

$\varepsilon(\beta)=-1$] for $\beta\in(\beta_1,\beta_2)$ [Eqs. (3.18)], and correspondingly one of the spacelike Killing vectors $\partial/\partial x$ or $\partial/\partial y$ becomes a null vector on $\mathcal{S}$. As can be seen easily by inspection of the metric (2.43), $\mathcal{S}$ is a null hypersurface in spacetime, and the Killing vector that becomes null on $\mathcal{S}$ is tangent to the null geodesic generators of $\mathcal{S}$. In fact, $\mathcal{S}$ is a "Killing-Cauchy horizon of type II" in the terminology of Ref. 4, where the reader can find a much more detailed description of such horizons. The spacetime curvature is perfectly well behaved across $\mathcal{S}$, and consequently $\mathcal{S}$ represents only a coordinate singularity of the (α,β,x,y) [or equivalently the (u,v,x,y)] coordinate system; it is possible to extend the metric and the spacetime beyond $\mathcal{S}$ after constructing a new admissible coordinate chart that covers $\mathcal{S}$ and its spacetime neighborhood regularly.

Now suppose $|\varepsilon(\beta)|\equiv 1$ across some subinterval I of the range $(-1,1)$ of β. (Note that I might not be a connected interval.) Thus, the metric (2.43) can be extended beyond the null surface $\mathcal{S}=\{\alpha=0,\ \beta\in I,\ x,y\}$ in a perfectly smooth manner. This extension is not unique, however; the initial data posed by the incoming colliding plane waves do not uniquely single out a specific extension among the infinitely many possibilities. Therefore, the Killing-Cauchy horizon $\mathcal{S}$ is a future Cauchy horizon[17] for the initial characteristic surface $\{u=0\}\cup\{v=0\}$, i.e., $\mathcal{S}$ represents a future boundary for the domain of dependence $D^+[\{u=0\}\cup\{v=0\}]$ of this initial surface. Since this means a breakdown, beyond the surface $\mathcal{S}$, of the predictability of the spacetime geometry from the initial data posed on $\{u=0\}\cup\{v=0\}$ (or, equivalently, a breakdown of global hyperbolicity[17]), the occurrence of these Killing-Cauchy horizons in colliding plane-wave spacetimes may seem to contradict the cosmic censorship hypothesis,[20,21] or at least a version of this hypothesis suitably formulated for plane-symmetric spacetimes.[4] But recall that a careful formulation of cosmic censorship[20] always insists that the hypothesis holds only for "generic" spacetimes, where the notion of "genericity" is conveniently left unspecified so that it can be interpreted appropriately for specific examples. In fact, there are many "counter-examples" to cosmic censorship, which, in one way or another, fail to satisfy the criterion of "genericity."[20,21] Perhaps the best-known such examples are the maximal Reissner-Nordstrom and Kerr solutions; the inner horizons of these solutions constitute Cauchy horizons for all partial Cauchy surfaces located in the asymptotically flat region, and therefore cause the breakdown of global hyperbolicity in the corresponding maximal spacetimes. However, it is now well known[22] that these inner horizons are unstable against a large class of linearized perturbations (such as gravitational waves, electromagnetic radiation, ...). It is therefore expected (but not yet fully proved), that in the interior of any rotating or charged black hole which is formed via "generic" gravitational collapse, the growth of these linear instabilities would destroy the inner horizon, turn it into a (spacelike) curvature singularity, and thereby restore the global hyperbolicity of the resulting spacetime.

Now, physically, though not in a formal mathematical way, the Killing-Cauchy horizon $\mathcal{S}$ of the colliding plane-wave solution (2.43) is similar to the inner Cauchy horizons of the Kerr and Reissner-Nordstrom solutions (which are also Killing-Cauchy horizons). To better understand the physical significance of the issue of the stability of the horizon $\mathcal{S}$, consider the geometry of the colliding plane-wave spacetime depicted in Fig. 2. For enhanced dramatical effect, we have assumed in this figure that the interval I [across which $|\varepsilon(\beta)|\equiv 1$] is a disconnected interval made up of several connected pieces $I_1, I_2, \ldots, I_n$. Hence the Killing-Cauchy horizon $\mathcal{S}$ is also disconnected; it consists of several distinct horizons $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$. The spacetime is extended beyond each of the horizons $\mathcal{S}_i$ in a different way; and there is also a large amount of freedom in the choice of each individual extension. In particular, one can choose the extensions in such a way that the n horizons $\mathcal{S}_i$ act as doorways (through the otherwise singular surface $\alpha=0$), which can be used by the observers living in the interaction region of the colliding plane-wave solution as "tunnels" into n different spacetimes, each *causally* disconnected from all the others. Or, any two distinct horizons $\mathcal{S}_i$, $\mathcal{S}_j$ may be joined, through suitable extensions, to the same spacetime but at different locations in time and space, thus giving the interaction region observers the possibility to influence the extended spacetime "simultaneously" at two different timelike-separated points (breakdown of global hyperbolicity in the extended spacetime). Although these possibilities are intriguing, clearly they cannot be realized (or at least they cannot be made physically plausible) unless the Killing-Cauchy horizon $\mathcal{S}$ is stable — unlike the unstable inner horizons of the Kerr and Reissner-Nordstrom solutions — against small perturbations of the colliding plane-wave spacetime. [In fact, similar speculations (such as using the interior regions as wormholes for spacetime travel) were made on the global structure of the maximal Kerr and Reissner-Nordstrom solutions;[17] these speculations were later rendered implausible by the instability results[22] we mentioned above (however, see Reference 23 in this context, where the wormhole concept is revisited and resurrected in an unexpected direction).] Thus, for example, any "realistic" attempt to "build" a spacetime tunnel between two different universes by means of generating and colliding two gravitational plane waves would fail, unless the Killing-Cauchy horizons $\mathcal{S}_i$ at $\alpha=0$ are stable; in other words, unless the set of all initial data from which such horizons evolve constitutes an open subset (with respect to an appropriate topology[4]), or a subset with nonvanishing volume (with respect to an appropriate measure) in the set of all plane-symmetric initial data.

Now the horizons $\mathcal{S}_i$ are not the first examples of Killing-Cauchy horizons produced by colliding plane waves. As we have discussed in the Introduction, the occurrence of Killing-Cauchy horizons in colliding plane-wave spacetimes was first discovered by Chandrasekhar and Xanthopoulos[11] when they produced several exact solutions which contained such horizons. Shortly after this work, Chandrasekhar and Xanthopoulos[24] discovered that the presence of a perfect fluid with (energy density)=pressure, or the presence of null dust, in their spacetime destroys the horizon in the full nonlinear Einstein theory. Independently of and simultaneously with this discovery, the author[4] formulated and proved general theorems which established the instability of Killing-Cauchy horizons in *any* plane-symmetric spacetime against generic, linearized plane-symmetric perturbations. (In addition, there already exists a considerable amount of literature[22,21] on the instabilities of several particular examples of Killing-Cauchy horizons, and of general compact Cauchy horizons.[25]) However, except for the above-mentioned example of Chandrasekhar and Xanthopoulos[24] involving null fluids, the nonlinear growth of these linear instabilities and the subsequent transformation of the horizon into a singularity have remained only as plausible conjectures. Note that this situation is quite similar to the state of knowledge on the instability of the inner horizons of the Kerr and Reissner-Nordstrom spacetimes; there most of the convincing instability results are valid only for linearized perturbations, and we do not even have a Chandrasekhar-Xanthopoulos-type[24] analysis for special kinds of nonlinear perturbations.[22] (See, however, Ref. 26 where a qualitative argument is given for the full nonlinear instability of the Reissner-Nordstrom Killing-Cauchy horizon.)

It is therefore remarkable that the formalism which we have described thus far provides concise and rigorous proofs that (i) the Killing-Cauchy horizons $\mathcal{S}_i$ at $\alpha=0$

are unstable in the *full nonlinear theory* against small but generic (plane-symmetric) perturbations of the initial data for the colliding plane waves, and (ii) that in a very specific sense, "generic" initial data in the form (2.15) or (2.49) always produce all-embracing, spacelike spacetime singularities without Killing-Cauchy horizons at $\alpha=0$. Moreover, the proofs of these statements are almost trivial. We shall demonstrate them in the following remaining few paragraphs of this section:

It is clear from our analysis thus far, that the structure of the singularities at $\alpha=0$ is completely determined by a single quantity: namely, the C^1 function $\varepsilon(\beta)$. This function $\varepsilon(\beta)$ is C^1 on the interval $(-1,1)$ because of Eq. (3.13) and because of our insistence [Eqs. (2.15) and (2.49)] that the initial data $V(r,1)$, $V(1,s)$ be C^1 functions. Now consider the space of all such C^1 functions on $(-1,1)$, which we will denote by F. This space F can be made into a Banach space[27] after endowing it with a suitable norm (such as the sup- or L^p norms[27]) and constructing its completion; but the precise choice of the norm is immaterial for the discussion that follows. Now the functional $\mathcal{E}$, which assigns a unique function $\varepsilon(\beta)$ to each choice of initial data in the form (2.49) or (2.15), can be regarded as a mapping from the space of all possible initial data to the Banach space F of all possible functions $\varepsilon(\beta)$. This mapping $\mathcal{E}$ is known to us in explicit form; using Eq. (3.13), we can write

$$\mathcal{E}: \{V(r,1), V(1,s)\} \longrightarrow \varepsilon(\beta)\,,$$

$$\varepsilon(\beta) = \frac{1}{\pi}\,\frac{1}{\sqrt{1+\beta}}\int_{\beta}^{1}[(1+s)^{1/2}V(1,s)]_{,s}\left(\frac{s+1}{s-\beta}\right)^{1/2} ds$$

$$+\frac{1}{\pi}\,\frac{1}{\sqrt{1-\beta}}\int_{-\beta}^{1}[(1+r)^{1/2}V(r,1)]_{,r}\left(\frac{r+1}{r+\beta}\right)^{1/2} dr\,. \tag{3.41}$$

Or, using the Eqs. (2.51), we can write equivalently

$$\mathcal{E}:\ \{V_1(u),V_2(v)\}\ \rightarrow\ \varepsilon(\beta)\ ,$$

$$\varepsilon(\beta)=-\frac{1}{\pi}\left[\frac{2}{1+\beta}\right]^{1/2}\int\limits_0^{v(\beta)}[e^{-U_2(v)/2}V_2(v)]_{,v}\left[\frac{2e^{-U_2(v)}}{2e^{-U_2(v)}-1-\beta}\right]^{1/2}dv$$

$$-\frac{1}{\pi}\left[\frac{2}{1-\beta}\right]^{1/2}\int\limits_0^{u(-\beta)}[e^{-U_1(u)/2}V_1(u)]_{,u}\left[\frac{2e^{-U_1(u)}}{2e^{-U_1(u)}-1+\beta}\right]^{1/2}du\ , \qquad (3.42a)$$

where $U_1(u)$, $U_2(v)$ are determined by Eqs. (2.16), and $v(\beta)$, $u(-\beta)$ are defined by

$$\beta=2e^{-U_2[v(\beta)]}-1\ , \qquad -\beta=2e^{-U_1[u(-\beta)]}-1\ . \qquad (3.42b)$$

Since it is obviously more transparent, we will use the representation of the mapping $\mathcal{E}$ given by Eq. (3.41) [instead of Eqs. (3.42)] throughout the present discussion. We will first prove the assertion (ii) that we have made in the last paragraph above, namely that for generic initial data in the form (2.49), the surface $\alpha=0$ is an all-embracing spacetime singularity which does not involve any Killing-Cauchy horizons. We will then discuss the assertion (i) that the Killing-Cauchy horizons $\mathcal{S}_i$ are non-linearly unstable against generic perturbations in the initial data, and will see that its proof follows very easily from the proof of assertion (ii).

Now Eq. (3.41) tells us that we can represent the map $\mathcal{E}$ symbolically as

$$\mathcal{E}:\ D\equiv F\oplus F\ \rightarrow\ F\ , \qquad (3.43)$$

where D, the space of all initial data in the form $\{V(r,1),V(1,s)\}$, has been identified with the direct sum of Banach spaces $F\oplus F$ [cf. Eq. (2.49) and the discussion following it]. Consider, for each $\delta>0$, the subset H_δ of F given by

$$H_\delta \equiv \{\varepsilon(\beta) \in F \mid \text{there exists a connected subinterval in } (-1,1) \text{ of length} \geq \delta$$

$$\text{across which } |\varepsilon(\beta)| \equiv 1\} \,. \tag{3.44}$$

It is clear that H_δ is a closed subset of F with the property that its complement, H_δ', is dense in F with respect to the Banach space (norm) topology. We shall define a closed subset of a Banach space with this property as a *nongeneric subset;* i.e., a nongeneric subset in a Banach space B is a closed subset whose complement is dense in B. [This notion of a "nongeneric" subset intuitively corresponds to a physicist's notion of genericity. However, our notion does not necessarily coincide with the more frequently used notion of a "subset with measure zero." In fact, even in finite-dimensional Banach spaces there exist nongeneric subsets with nonzero Lebesgue measure (e.g., the fat Cantor set in the unit interval as a subset of R^1).[28] It is not yet clear whether our topological notion of genericity can be replaced with a measure theoretical alternative so as to leave the conclusions of this section intact. (See also the remarks at the end of the next paragraph in this connection.)] Thus, H_δ is a nongeneric subset of F for any $\delta>0$. Note also that $H_{\delta_1} \supset H_{\delta_2}$ whenever $\delta_1 \leq \delta_2$.

It is now clear from the conclusions (i), (ii), and (iii) which we have listed in the beginning of this section, that if $\varepsilon(\beta)$ is an element of F that does not belong to H_δ for any δ, then the corresponding colliding plane-wave solution possesses an all-embracing spacetime singularity at $\alpha=0$; this singularity is in general a curvature singularity, possibly crisscrossed with isolated (with respect to β) topological noncurvature singularities. Therefore, in order to prove our assertion (ii), we need only to prove that for all $\delta>0$, the inverse image $\mathcal{E}^{-1}(H_\delta)$ of H_δ under the map $\mathcal{E}$ is a nongeneric subset of the space of all initial data D. [The reader might be puzzled at this point as to why the subset $\mathcal{E}^{-1}(\bigcup_{\delta>0} H_\delta)$ of D is not what needs to be proved nongeneric. The answer lies in physics: From a physically realistic standpoint, there is always an absolute short-distance cutoff (a lower bound) on the length of a connected interval in β; this lower bound δ_c on δ is given by the Planck length l_P (or more precisely by $\delta_c = l_P / \sqrt{l_1 l_2}$). A Killing-Cauchy horizon that extends less than a Planck length δ_c in the β direction will almost certainly be indistinguishable, in its semiclassical manifestations, from a spacetime singularity. Furthermore, the subset $\bigcup_{\delta>0} H_\delta \subset F$ in question is *not* a nongeneric subset; in fact, it is easy to show that $\bigcup_{\delta>0} H_\delta$ is dense in F and hence is neither closed nor nongeneric. Thus, $\mathcal{E}^{-1}(\bigcup_{\delta>0} H_\delta)$ also is *not* nongeneric, and possibly it is dense in D. With our present somewhat naive (but physically satisfactory) notion of genericity, $\mathcal{E}^{-1}(\bigcup_{\delta>0} H_\delta)$ cannot be properly shown to be "nongeneric." In fact, it may be helpful to note that our genericity concept is similarly unable to identify the subset of all rational numbers in the unit interval as a "nongeneric" subset; the notion of the Lebesgue measure,[29] and not just a topological notion like ours, is needed to implement a formulation of genericity powerful enough to handle such questions effectively. It is conceivable that in our case too, a suitable extension of the notion of measure to infinite dimensional Banach spaces could yield both a more appropriate formulation and a proof for the "nongenericity" of the subset $\mathcal{E}^{-1}(\bigcup_{\delta>0} H_\delta)$.]

Turn now to the proof of the assertion that $\mathcal{E}^{-1}(H_\delta) \subset D$ is nongeneric for any $\delta>0$. For this proof, we need to consider some basic properties of the mapping $\mathcal{E}: D \to F$. First of all, it is easy to see that $\mathcal{E}$ is an onto map; that is, for any element $\varepsilon(\beta)$ in F, there exists a choice (in fact infinitely many choices) of initial data in D which would yield, under the map $\mathcal{E}$, precisely the element $\varepsilon(\beta)$. [In fact, the inverse image $\mathcal{E}^{-1}(q)$ of any point $q \equiv \varepsilon(\beta) \in F$ is an infinite set in D; to find just one element in this set, take $V(1,s)$ to be any function and solve the resulting integral equation

(3.41) for $V(r,1)$.] The second basic property of $\mathcal{E}$ is that $\mathcal{E}$ is defined on the whole Banach space D; that is, the domain of $\mathcal{E}$ is D. [To see this, apply a formal partial integration on both of the integrals in Eq. (3.41); the result can be written in a form which does not involve any differentiations of the functions $V(r,1)$ and $V(1,s)$.] And finally, $\mathcal{E}$ is a continuous linear mapping from D onto F. This follows (i) by first noting that $\mathcal{E}$ is a closed linear operator[27] [linearity of $\mathcal{E}$ is obvious from Eq. (3.41); closedness of $\mathcal{E}$ follows since $\mathcal{E}$ is essentially the composition of a differentiation operator (which is closed) and an integral operator (which is continuous)], and (ii) then using the closed graph theorem (Sec. II. 6 of Ref. 27) which says that a closed, onto linear mapping $\mathcal{E}: D \rightarrow F$ with domain $=D$ is continuous. Now we are ready to prove that $\mathcal{E}^{-1}(H_\delta) \subset D$ is nongeneric: Since $\mathcal{E}$ is continuous, $\mathcal{E}^{-1}(H_\delta)$ is a closed subset of D. To see that the complement of $\mathcal{E}^{-1}(H_\delta)$ in D is dense, use the open mapping theorem (Sec. II. 5 of Ref. 27) to conclude that $\mathcal{E}$ is an open map. If the complement of $\mathcal{E}^{-1}(H_\delta)$ were not dense in D, $\mathcal{E}^{-1}(H_\delta)$ would contain an open subset, and the open map $\mathcal{E}$ would send this open set onto an open subset of H_δ in F. This is impossible, since the subset H_δ is nongeneric and hence cannot contain an open set. This contradiction demonstrates that $\mathcal{E}^{-1}(H_\delta)$ is a nongeneric subset of D for any $\delta>0$.

It is now very easy to prove our remaining assertion: namely, that the Killing-Cauchy horizons $\mathcal{S}_i$ are nonlinearly unstable against generic perturbations of the initial data. Consider a given choice of initial data represented by a point p in the Banach space D. If the colliding plane-wave spacetime which evolves from these initial data p possesses Killing-Cauchy horizons $\mathcal{S}_i$ at $\alpha=0$, then there is a $\delta_0>0$ such that $p \in \mathcal{E}^{-1}(H_{\delta_0})$ (just take δ_0 as the size of the smallest horizon $\mathcal{S}_j$). Consequently, $p \in \mathcal{E}^{-1}(H_\delta)$ for each $\delta \leq \delta_0$. Now for each such $\delta \leq \delta_0$, no matter how small, the set

$\mathcal{E}^{-1}(H_\delta)$ to which p belongs is a nongeneric subset of D. Therefore, for each $\delta \leq \delta_0$, no matter how small, a generic perturbation of the point p representing the initial data will push p outside the subset $\mathcal{E}^{-1}(H_\delta)$; in other words, for any fixed but arbitrarily small $\delta > 0$, a generic perturbation of the initial data p would destroy all horizons $\mathcal{S}_i$ at $\alpha=0$ that are of length $\geq \delta$, and turn them into spacetime singularities. In fact, since from a physically realistic standpoint all Killing-Cauchy horizons have to be of a size large compared to the Planck size δ_c, even the nongenericity of just the set $\mathcal{E}^{-1}(H_{\delta_c})$ is sufficient to conclude that the horizons $\mathcal{S}_i$ are nonlinearly unstable against plane-symmetric perturbations.

To get an intuitive feeling about these instabilities, it might be useful to think of Eq. (3.41) as describing some kind of a superposition, at $\alpha=0$, of the two wave forms described by the functions $V(r,1)$ and $V(1,s)$ which constitute the initial data (see Fig. 2). Killing-Cauchy horizons form at $\alpha=0$ only when this superposition results in a "perfectly destructive interference" ($|\varepsilon(\beta)| \equiv 1$) across some interval in the spatial coordinate β. Any generic perturbation in the wave forms $V(r,1)$ and $V(1,s)$ causes small imperfections in the precision of this destructive interference [$|\varepsilon(\beta)|$ slightly deviates from 1]; and any small deviation from perfect destructive interference is sufficient to turn the Killing-Cauchy horizons into spacetime singularities (Fig. 2).

IV. EXAMPLES OF EXACT SOLUTIONS WHICH EXHIBIT SOME OF THE ABOVE-DISCUSSED ASYMPTOTIC SINGULARITY STRUCTURES

Our first example is the well-known Khan-Penrose[5] solution for colliding impulsive plane waves. The reader is referred to the original references[5,6,8] for comprehensive descriptions of the Khan-Penrose solution; here we will only discuss it from the point of view of our analysis in Sec. III above. The initial data for the Khan-Penrose solution written in the form of Eq. (2.15) are

$$V_1(u)=\ln\left(\frac{1+(u/a)}{1-(u/a)}\right) ,$$

$$V_2(v)=\ln\left(\frac{1+(v/b)}{1-(v/b)}\right) , \tag{4.1}$$

which give, by Eqs. (2.16),

$$U_1(u)=-\ln\left(1-\frac{u^2}{a^2}\right) ,$$

$$U_2(v)=-\ln\left(1-\frac{v^2}{b^2}\right) . \tag{4.2}$$

From Eqs. (4.2), (2.26), (2.27), and (2.51), we obtain the explicit forms of the various coordinate transformations we have discussed in Sec. II B,

$$\alpha=1-\frac{u^2}{a^2}-\frac{v^2}{b^2} , \qquad \beta=\frac{u^2}{a^2}-\frac{v^2}{b^2} ,$$

$$r=1-2\frac{u^2}{a^2} , \qquad s=1-2\frac{v^2}{b^2} , \tag{4.3}$$

which, when combined with the Eq. (4.1), yield the Khan-Penrose initial data in the form (2.49):

$$V_{\rm KP}(r,1)=\ln\left(\frac{1+\sqrt{(1-r)/2}}{1-\sqrt{(1-r)/2}}\right) ,$$

$$V_{\rm KP}(1,s)=\ln\left(\frac{1+\sqrt{(1-s)/2}}{1-\sqrt{(1-s)/2}}\right) . \tag{4.4}$$

When computed with the initial data (4.4), the explicit solution (2.60) gives an expression for $V_{KP}(\alpha,\beta)$ in closed form. If we take the normalization point (u_0,v_0) as $[\frac{1}{2}(\sqrt{3}-1)a\,,\,\frac{1}{2}(\sqrt{3}-1)b\,]$, and insert $V_{KP}(\alpha,\beta)$ into Eq. (2.44b), the function $Q_{KP}(\alpha,\beta)$ can also be evaluated in closed form. Finally, by combining these results with Eq. (2.43), the Khan-Penrose metric is found to have the following expression in the $(\alpha,\beta,x\,,y\,)$ coordinate system:

$$g_{KP}=\frac{ab}{\sqrt{\alpha}}\,\frac{\alpha^2}{[(1-\alpha)^2-\beta^2]^{\frac{1}{2}}[(1+\alpha)^2-\beta^2]^{\frac{1}{2}}[\sqrt{(1-\alpha)^2-\beta^2}+\sqrt{(1+\alpha)^2-\beta^2}]^2}(-d\,\alpha^2+d\,\beta^2)$$

$$+\alpha\left(\frac{\sqrt{1+\alpha-\beta}+\sqrt{1-\alpha-\beta}}{\sqrt{1+\alpha-\beta}-\sqrt{1-\alpha-\beta}}\right)\left(\frac{\sqrt{1+\alpha+\beta}+\sqrt{1-\alpha+\beta}}{\sqrt{1+\alpha+\beta}-\sqrt{1-\alpha+\beta}}\right)dx^2$$

$$+\alpha\left(\frac{\sqrt{1+\alpha-\beta}-\sqrt{1-\alpha-\beta}}{\sqrt{1+\alpha-\beta}+\sqrt{1-\alpha-\beta}}\right)\left(\frac{\sqrt{1+\alpha+\beta}-\sqrt{1-\alpha+\beta}}{\sqrt{1+\alpha+\beta}+\sqrt{1-\alpha+\beta}}\right)dy^2\,. \tag{4.5}$$

After expressing α and β in terms of u and v as in Eq. (4.3), Eq. (4.5) reduces to the standard expression[5] of the Khan-Penrose metric in the Rosen-type $(u\,,v\,,x\,,y\,)$ coordinate system. Inspection of Eq. (4.5) shows that $q_1(\beta)$, $q_2(\beta)$, and $q_3(\beta)$ [cf. Eqs. (3.14) and (3.15)] for the Khan-Penrose solution are equal to the *constant* values $\frac{3}{2}$, -1, and 3, respectively. This implies that $[p_1(\beta),p_2(\beta),p_3(\beta)]$ [Eqs. (3.16) and (3.17)] are equal to the constant values $(-\frac{2}{7}\,,\frac{6}{7}\,,\frac{3}{7})$, and using the inverse relations to Eqs. (3.18) given by

$$\varepsilon(\beta)=-1-2\frac{p_3(\beta)}{p_2(\beta)}\quad \text{if } p_2(\beta)\neq 0\,,$$

$$\varepsilon(\beta)=1+2\frac{p_3(\beta)}{p_1(\beta)}\quad \text{if } p_1(\beta)\neq 0\,, \tag{4.6}$$

these equations in turn imply that $\varepsilon(\beta)\equiv-2$. Thus, for the Khan-Penrose solution

$$\varepsilon_{\mathrm{KP}}(\beta)\equiv-2\,,\quad p_1(\beta)\equiv-\frac{2}{7}\,,\quad p_2(\beta)\equiv\frac{6}{7}\,,$$

$$p_3(\beta)\equiv\frac{3}{7}\quad \forall\,\beta\in(-1,1)\,. \tag{4.7}$$

Numerical computation of the integrals in Eq. (3.13) with the Khan-Penrose initial data (4.4) indicates that both of the two terms on the right-hand side of Eq. (3.13) [involving the integrals of $V_{\mathrm{KP}}(r,1)$ and $V_{\mathrm{KP}}(1,s)$] are separately constant, and equal to -1 for all β. Note that since $|\varepsilon(\beta)|>1$, the Khan-Penrose singularity is an astigmatic one [cf. the discussion following Eq. (3.23) in Sec. III A]. This is not surprising, since the incoming plane waves of the Khan-Penrose solution [which are described by the initial data (4.1) and Eqs. (4.2)] are (i) highly astigmatic (one of the focal lengths is infinite whereas the other is a or b), and (ii) very strong (both V_1 and V_2 are of order unity).

Turn now to our second example; a colliding plane-wave spacetime described by the initial data

$$V^{\times}(r,1)=-V_{\mathrm{KP}}(r,1)\,,\qquad V^{\times}(1,s)=V_{\mathrm{KP}}(1,s)\,. \tag{4.8}$$

Since the integrals in Eq. (3.13) are both linear in their respective arguments $V(r,1)$ and $V(1,s)$, and since for the Khan-Penrose initial data (4.4) these integrals both take the constant value -1, it follows that, for the colliding plane-wave solution which evolves from the initial data (4.8),

$$\varepsilon^{\times}(\beta)\equiv0\,,\quad p_1(\beta)\equiv\frac{2}{3}\,,\quad p_2(\beta)\equiv\frac{2}{3}\,,$$

$$p_3(\beta)\equiv-\frac{1}{3} \quad \forall\,\beta\in(-1,1)\,. \tag{4.9}$$

Therefore, as $|\varepsilon^{\times}(\beta)|<1$, the solution developing from (4.8) has a purely anastigmatic singularity structure at $\alpha=0$. This is interesting, because the incoming waves described by the data (4.8) are highly astigmatic. In fact, both incoming plane waves are impulsive waves identical in structure to the incoming waves of the Khan-Penrose solution [it is easy to see that $U_1(u)$, $U_2(v)$, α, β, r, and s for the initial data (4.8) have exactly the same forms as in the Khan-Penrose solution where they are given by Eqs. (4.2) and (4.3)]. The only exception to this identical structure is that one of the waves (namely, the wave that propagates in the v direction, see Fig. 1) has its direction of astigmatism "twisted" with respect to the other; in other words, one of the waves focuses in the x direction and defocuses in the y direction, whereas focusing by the other wave occurs with the roles of the x and y directions interchanged. Now, the solution $V^{\times}(\alpha,\beta)$ of the initial-value problem given by the field equation (2.44a) and the initial data (4.8) is easy to find: It is immediately seen after a short calculation that $V_{KP}(\alpha,\beta)$ is the sum of two pieces that separately satisfy Eq. (2.44a); and therefore, by the linearity of Eq. (2.44a), taking the difference of these pieces instead of their sum produces the unique solution of Eq. (2.44a) which satisfies the initial conditions (4.8). However, with this solution for $V^{\times}(\alpha,\beta)$, the integral in Eq. (2.44b) cannot be computed analytically to yield an expression for the function $Q^{\times}(\alpha,\beta)$ in closed form. Nevertheless, since the coordinate transformations between the (α,β) and (u,v) coordinates are known explicitly [Eqs. (4.3)], we can still write down the interaction-region metric for our solution in the following semiclosed form:

$$g^{\times}=-8\frac{u'v'e^{-Q^{\times}[\alpha(u,v),\beta(u,v)]/2}}{\sqrt{1-u'^2-v'^2}}du\,dv$$

$$+(1-u'^2-v'^2)\left(\frac{\sqrt{1-u'^2}+v'}{\sqrt{1-u'^2}-v'}\right)\left(\frac{\sqrt{1-v'^2}-u'}{\sqrt{1-v'^2}+u'}\right)dx^2$$

$$+(1-u'^2-v'^2)\left(\frac{\sqrt{1-u'^2}-v'}{\sqrt{1-u'^2}+v'}\right)\left(\frac{\sqrt{1-v'^2}+u'}{\sqrt{1-v'^2}-u'}\right)dy^2 ,$$

(4.10)

where $u'\equiv u/a$, $v'\equiv v/b$, $Q^{\times}$ is given by Eq. (2.44b), and $\alpha(u,v)$, $\beta(u,v)$ are determined by Eqs. (4.3). The metric in the remaining regions II, III, and IV (Fig. 1) of the solution (4.10) is found by extending (4.10) via the Penrose prescription.[5,9,10] Inspection of Eq. (4.10) makes it apparent that in the vicinity of the singularity $\alpha=0$ ($u'^2+v'^2=1$ in the u,v coordinates), the asymptotic behavior of the metric is characterized by $\varepsilon(\beta)\equiv 0$ [cf. Eqs. (3.17) and (3.18)].

Our third example is the colliding plane-wave spacetime which develops from the initial data

$$V_{½}(r,1)=\frac{1}{2}V_{\mathrm{KP}}(r,1) , \qquad V_{½}(1,s)=\frac{1}{2}V_{\mathrm{KP}}(1,s) ,$$

(4.11)

or, equivalently

$$V_1(u)=\frac{1}{2}\ln\left(\frac{1+(u/a)}{1-(u/a)}\right) ,$$

$$V_2(v)=\frac{1}{2}\ln\left(\frac{1+(v/b)}{1-(v/b)}\right) . \qquad (4.12)$$

Unlike with the Khan-Penrose solution, Eqs. (2.16) cannot be solved analytically with the initial data (4.12); and consequently, $U_1(u)$, $U_2(v)$, $\alpha(u,v)$ and $\beta(u,v)$ cannot be expressed in closed form for the solution (4.11). On the other hand, from the linear dependence [Eq. (3.13)] of $\varepsilon(\beta)$ on the initial data $\{V(r,1),V(1,s)\}$, it is very easy to see that the asymptotic structure of the solution (4.11) near $\alpha=0$ is characterized by

$$\varepsilon_{1/2}(\beta)\equiv-1\,,\quad p_1(\beta)\equiv0\,,\quad p_2(\beta)\equiv1\,,$$

$$p_3(\beta)\equiv0\quad\forall\,\beta\in(-1,1)\,. \tag{4.13}$$

Therefore (Sec. III. C), the solution (4.11) possesses a nonsingular Killing-Cauchy horizon at $\alpha=0$ across which the spacetime can be smoothly extended. Although the metric for this solution cannot be expressed in closed form in the Rosen-type u,v coordinate system [since the transformation to (α,β) coordinates is not available in analytic form], it can be easily computed in the (α,β) coordinates: By the linearity of the field equation (2.44a), it is clear that $V_{1/2}(\alpha,\beta)=\tfrac{1}{2}V_{KP}(\alpha,\beta)$; this implies, by Eq. (2.44b), that up to an additive constant, $Q_{1/2}(\alpha,\beta)=\dfrac{1}{4}Q_{\mathrm{KP}}(\alpha,\beta)$. Therefore, combining Eq. (2.43) with Eq. (4.5), we obtain

$$g_{1/2}=\frac{c_0\,ab}{\{[(1-\alpha)^2-\beta^2]^{1/2}[(1+\alpha)^2-\beta^2]^{1/2}[\sqrt{(1-\alpha)^2-\beta^2}+\sqrt{(1+\alpha)^2-\beta^2}]^2\}^{1/4}}(-d\alpha^2+d\beta^2)$$

$$+\alpha\left(\frac{\sqrt{1+\alpha-\beta}+\sqrt{1-\alpha-\beta}}{\sqrt{1+\alpha-\beta}-\sqrt{1-\alpha-\beta}}\right)^{1/2}\left(\frac{\sqrt{1+\alpha+\beta}+\sqrt{1-\alpha+\beta}}{\sqrt{1+\alpha+\beta}-\sqrt{1-\alpha+\beta}}\right)^{1/2}dx^2$$

$$+\alpha\left(\frac{\sqrt{1+\alpha-\beta}-\sqrt{1-\alpha-\beta}}{\sqrt{1+\alpha-\beta}+\sqrt{1-\alpha-\beta}}\right)^{1/2}\left(\frac{\sqrt{1+\alpha+\beta}-\sqrt{1-\alpha+\beta}}{\sqrt{1+\alpha+\beta}+\sqrt{1-\alpha+\beta}}\right)^{1/2}dy^2\,, \tag{4.14}$$

where c_0 is a numerical constant. Although the solution (4.14) is the first example of an exact colliding *parallel-polarized* plane-wave solution producing Killing-Cauchy horizons at $\alpha=0$, it has one undesirable feature: The incoming plane waves described by the data (4.12) are not sandwich waves; that is, the focal plane[10,3] of each single incoming wave represents a curvature singularity of the single plane-wave spacetime instead of just a coordinate singularity. (Readers can convince themselves of this fact by inspecting the behavior of the curvature [Eqs. (2.19)] in the single plane-wave spacetimes defined by Eqs. (4.12), (2.16), and (2.8). For a more detailed discussion of these issues, see Sec. II of Ref. 10.) As a result, it seems exceedingly difficult to carry out and analyze a maximal extension (of which we know there are infinitely many) of the spacetime (4.14) beyond the Killing-Cauchy horizon $\{\alpha=0\}$. In our final example below, we will discuss another exact colliding plane-wave solution which similarly produces a Killing-Cauchy horizon at $\alpha=0$, and we will see that the above-mentioned difficulty with singular focal planes does not arise in this solution. In fact, the maximal analytic extension of this solution across the horizon is readily available and produces a maximal colliding plane-wave spacetime with a surprising global structure.

We now turn to this final example: a family of colliding *parallel-polarized* plane wave solutions producing nonsingular Killing-Cauchy horizons at $\alpha=0$. These solutions are derived by a procedure that is almost identical to the procedure by which we have constructed the infinite-parameter family of exact solutions discussed in Ref. 10. To follow the *details* of our presentation, the reader *must* refer to Ref. 10; however, the qualitative features of our example can be understood from the discussion here, and especially from the Figs. 3, 4, and 5. The equation numbers that refer to equations of Ref. 10 will be denoted by a prefix "10"; for example, Eq. (10.3.4) refers to Eq. (3.4) of Ref. 10.

Consider the colliding parallel-polarized plane-wave solution described by Eq. (10.2.4). In Ref. 10, this solution was constructed from the interior Schwarzschild metric in the following four steps (Fig. 3).

(i) The coordinate transformation (10.2.2) was carried out to define a new set of coordinates (u',v',x,y) in terms of the Schwarzschild coordinates (r,θ,ϕ,t), and the interior Schwarzschild metric was expressed [Eq. (10.2.3)] in terms of these new coordinates.

(ii) Two length scales a and b were introduced by rescaling the null coordinates u' and v' through the relations $u'=u/a$, $v'=v/b$, where $ab=4M^2$.

(iii) The resulting interaction-region metric (10.2.3) was then extended beyond the null surfaces $\{u=0\}$ and $\{v=0\}$ by the Penrose prescription:[5,9] $u/a \rightarrow (u/a)H(u/a)$, $v/b \rightarrow (v/b)H(v/b)$, where H is the Heaviside step function.

(iv) The global topology of the resulting spacetime was changed from $S^2 \times R^2$ to R^4 by means of the coordinate transformation (10.2.2) and the nonanalytic extension (iii) across the null surfaces $\{u=0\}$ and $\{v=0\}$. In Fig. 3, we have indicated these null surfaces by their expressions in terms of the Schwarzschild coordinates; these expressions are $\{r=M(1+\cos\theta)\}$ and $\{r=M(1-\cos\theta)\}$ for $\{u=0\}$ and $\{v=0\}$, respectively.

The interaction region of the resulting colliding plane-wave solution (10.2.4) is locally isometric to the region denoted by IV in Fig. 3. In particular, the Schwarzschild singularity at $r=0$ corresponds, under this isometry, to the singularity at $\alpha=0$ [at $(u/a)+(v/b)=\pi/2$ in the Rosen-type coordinates of Ref. 10] created by the colliding waves. In Ref. 10, the above steps (i)–(iv) were repeated almost identically for an infinite-parameter family of regular interior Weyl solutions generalizing the interior Schwarzschild solution. As a result, the infinite-parameter family (10.3.18)–(10.3.22) of colliding plane-wave solutions was obtained. Use of Eqs. (4.6) and inspection of the solution (10.2.4) reveal that for this solution (which corresponds to all parameters d_k being zero), and for all the other solutions (10.3.18)–(10.3.22) (as long as all but finitely many d_k are zero), the asymptotic behavior of the metric near the singularity is characterized by

$$\varepsilon(\beta)\equiv-3\,, \quad p_1(\beta)\equiv-\frac{1}{3}\,, \quad p_2(\beta)\equiv\frac{2}{3}\,,$$

$$p_3(\beta)\equiv\frac{2}{3} \quad \forall\,\beta\in(-1,1)\,. \tag{4.15}$$

Note that by combining Eqs. (4.15) with Eqs. (3.33)–(3.35), we can reproduce the results of Ref. 10 dealing with the asymptotic behavior of the curvature quantities near the singularity $(u/a)+(v/b)=\pi/2$.

Now, in order to obtain colliding plane-wave solutions which produce Killing-Cauchy horizons at $\alpha=0$, we simply reverse the roles of regions I and IV in Fig. 3; that is, we take region I to be our interaction region, and apply the steps (i)–(iv) above to this new interaction-region metric. This results in a new colliding plane-wave solution whose metric is easily seen to be given by Eq. (10.2.3), but this time for $u<0$, $v<0$ (which describe region I) instead of $u>0$, $v>0$ (which describe region IV). Therefore, redefining u and v as $-u$ and $-v$, respectively, the interaction-region metric of the new solution can be written in the form

$$g_{\mathrm{I}}=-\left[1+\sin\left(\frac{u}{a}+\frac{v}{b}\right)\right]^2 du\,dv$$

$$+\left(\frac{1-\sin\left(\frac{u}{a}+\frac{v}{b}\right)}{1+\sin\left(\frac{u}{a}+\frac{v}{b}\right)}\right)dx^2$$

$$+\left[1+\sin\left(\frac{u}{a}+\frac{v}{b}\right)\right]^2\cos^2\left(\frac{u}{a}-\frac{v}{b}\right)dy^2\,, \tag{4.16}$$

where the interaction region on which the metric (4.16) is defined is given by $\{u>0, v>0\}$. Note that Eq. (4.16) can be obtained by applying the simple transformations $u \to -u$, $v \to -v$ to Eq. (10.2.3). The metric on the rest of the solution (4.16) (i.e., in regions II, III, and IV) is obtained by extending the interaction-region metric (4.16) via the Penrose prescription;[5,9] $(u/a) \to (u/a)H(u/a)$, $(v/b) \to (v/b)H(v/b)$. It is clear from Eq. (4.16) that the solution thus obtained has an asymptotic structure near $\alpha=0$ characterized by

$$\varepsilon(\beta)\equiv 1\,, \quad p_1(\beta)\equiv 1\,, \quad p_2(\beta)\equiv 0\,,$$

$$p_3(\beta)\equiv 0 \quad \forall\, \beta \in (-1,1)\,. \tag{4.17}$$

Therefore, $\alpha=0$ is a nonsingular Killing-Cauchy horizon produced by the colliding plane-wave solution (4.16). By applying exactly the same reasoning as above to the infinite-parameter family (10.3.18)–(10.3.22) of colliding plane-wave solutions, we obtain an infinite-parameter family of generalizations of the solution (4.16). These generalized solutions can be found by simply applying the transformations $u \to -u$, $v \to -v$ *throughout* Eqs. (10.3.18)–(10.3.22). As long as all but finitely many of the parameters d_k are nonzero, the generalized solutions all have the same asymptotic structure near $\alpha=0$ characterized by the exponents (4.17); i.e., all generalized solutions create Killing-Cauchy horizons at $\alpha=0$. The proof that the colliding plane-wave spacetime (4.16) and its generalizations described above are genuine solutions (in the sense of distributions) to the vacuum Einstein equations is provided by exactly the same arguments with which we showed the solutions (10.3.18)–(10.3.22) of Ref. 10 to be genuine vacuum solutions.

The interaction region of the solution (4.16) is locally isometric to region I (Fig. 3) of the interior Schwarzschild solution. In particular, the Killing-Cauchy horizon at

$\alpha=0$ [at $(u/a)+(v/b)=\pi/2$ in the Rosen-type coordinates] corresponds, under this isometry, to the horizon $\{r=2M\}$ of the Schwarzschild spacetime. This interaction region I of the solution (4.16) is formed by the collision of single plane waves whose forms in the precollision regions II and III (cf. Fig. 1) are

$$g_{\mathrm{II}}=-[1+\sin(u/a)]^2 du\, dv + \left[\frac{1-\sin(u/a)}{1+\sin(u/a)}\right] dx^2 + [1+\sin(u/a)]^2\cos^2(u/a)dy^2\ , \tag{4.18a}$$

$$g_{\mathrm{III}}=-[1+\sin(v/b)]^2 du\, dv + \left[\frac{1-\sin(v/b)}{1+\sin(v/b)}\right] dx^2 + [1+\sin(v/b)]^2\cos^2(v/b)dy^2\ . \tag{4.18b}$$

In contrast to Eq. (10.2.5), the incoming plane waves (4.18) are true sandwich waves; that is, the focal planes $u=\pi a/2$ and $v=\pi b/2$ of the incoming waves (4.18) represent nonsingular Killing-Cauchy horizons in the respective single plane-wave spacetimes. Similarly, it is easy to see (i) that the infinite-parameter family of generalizations of the solution (4.16) all have interaction regions locally isometric to an analogous region I in the interiors of the corresponding Weyl solutions, and (ii) that the Killing-Cauchy horizons created by these generalized solutions correspond to the horizons of the Weyl solutions from which they are derived. For each of these generalized solutions (as long as all but finitely many d_k are zero), the incoming single plane waves have a similar structure to the plane waves (4.18), and hence are also true sandwich waves.

In the remaining paragraphs of this section, we will concentrate on the colliding plane-wave solution (4.16) and discuss its properties in detail. In particular, we will show that the maximal analytic extension of (4.16) across the Killing-Cauchy horizon $\{\alpha=0\}$ is easy to find and yields a (weakly) asymptotically flat extended spacetime. The infinite-parameter family of generalizations of the solution (4.16) have qualitatively identical features with the solution (4.16), *provided* that all but finitely many parameters d_k are zero. In particular, each of these generalized solutions can be extended analytically across the horizon in a similar fashion; however, the extended spacetimes are not asymptotically flat since the generalized solutions are derived from Weyl solutions which violate asymptotic flatness.[10]

We first consider in some detail the structure of the colliding plane-wave spacetime (4.16) near the Killing-Cauchy horizon $\{\alpha=0\}$. Note that when we apply the coordinate transformation (10.2.2) to the interior Schwarzschild solution, and later extend the metric nonanalytically into the precollision regions as described in the steps (i)–(iv) above, we change the topology of the resulting spacetime from $S^2 \times R^2$ to R^4 (cf. Ref. 10). Thus (Fig. 4), the topology of the Schwarzschild horizon is also changed from $S^2 \times R^1$ to R^3 (when we regard the Schwarzschild horizon as the progenitor of the Killing-Cauchy horizon $\{\alpha=0\}$ which has topology R^3). The spacelike plane-symmetry-generating Killing vector $\partial/\partial x$ which becomes null on the horizon $\{\alpha=0\}$ corresponds, under the transformations (10.2.2), to the Killing vector $\partial/\partial t$ of the Schwarzschild spacetime which becomes null on the Schwarzschild horizon. The bifurcation two-sphere $\mathcal{S}^2$ of the Schwarzschild horizon, on which $\partial/\partial t$ vanishes, corresponds in our solution (4.16) to the crease singularity (≡bifurcation set[4]) of the Killing-Cauchy horizon $\{\alpha=0\} \equiv \{(u/a)+(v/b)=\pi/2\}$, on which the Killing vector $\partial/\partial x$

(which is tangent to the null generators of the horizon) vanishes (Fig. 4). Of course, when the solution (4.16) is *not* extended beyond the Killing-Cauchy horizon $\{\alpha=0\}$, the remaining Killing vector $\partial/\partial y$ is *not* cyclic, in contrast to the corresponding Killing vector $\partial/\partial\phi$ of the Schwarzschild spacetime which is cyclic. Therefore, the bifurcation set of the Killing-Cauchy horizon $\{\alpha=0\}$ has the topology R^2 in the unextended spacetime, in contrast to the bifurcation *sphere* $\mathcal{S}^2$ of the Schwarzschild horizon. [Note that, strictly speaking the Killing-Cauchy horizon $\{\alpha=0\}$ is *not* part of the spacetime manifold for the *unextended* colliding plane wave solution (4.16); in other words, the unextended solution (4.16) is represented by those points in Fig. 4 which lie *strictly* to the past of the horizon $\{\alpha=0\}$.] It is easy to see that the curvature quantities Ψ_0, Ψ_2, and Ψ_4 [Eqs. (2.19)] for the solution (4.16) are all finite and well behaved near and on the Killing-Cauchy horizon $\{(u/a)+(v/b)=\pi/2\}$. For example, the quantity Ψ_2 is given by

$$\Psi_2=-\frac{2}{ab}\,\{1+\sin[(u/a)+(v/b)]\}^{-3} \quad ; \tag{4.19}$$

and the other curvature quantities also exhibit similar smooth behavior at $(u/a)+(v/b)=\pi/2$. Hence, clearly, the spacetime can be extended smoothly beyond the horizon $\{\alpha=0\}$ to obtain a maximal colliding plane-wave solution. In fact, since the metric in the interaction region I (Fig. 3) is everywhere locally isometric to interior Schwarzschild, and the maximal analytic extension of the Schwarzschild metric across the Schwarzschild horizon is well known, the maximal analytic extension of the colliding plane-wave solution (4.16) across the Killing-Cauchy horizon $\{\alpha=0\}$ is very easy to describe. In the following final paragraph we will discuss the local description and the global structure of this maximal analytic extension.

Before proceeding with the detailed description of the extension, note that the two boundary points of the unextended solution (4.16) denoted by P and Q in Fig. 4 are spacetime singularities. The reason is that the amplitudes of the delta-function contributions to the curvature quantities along the null surfaces $\{u=0\}$ and $\{v=0\}$ diverge as these two points are approached. However, aside from these two singular points P and Q located on the bifurcation set of the Killing-Cauchy horizon $\{\alpha=0\}$, the geometry at all the other boundary points of the unextended spacetime is perfectly smooth and well behaved. Now, the local description of the maximal analytic extension of (4.16) is particularly clear: Near the Killing-Cauchy horizon $\{(u/a)+(v/b)=\pi/2\}$, the metric is locally isometric to the Schwarzschild solution near the horizon $\{r=2M\}$. It is clear that, because of the specific time orientation that we are using on the unextended colliding plane-wave spacetime, this Schwarzschild horizon to which our Killing-Cauchy horizon $\{\alpha=0\}$ corresponds is the *past* horizon of the Schwarzschild spacetime, rather than the future one. Construct, then, the usual Kruskal-type regular coordinate system on the (past) horizon $\{\alpha=0\}$, and simply extend the solution as *the maximal analytic extension* of the metric in such a coordinate system. Clearly, this would give us *precisely* the usual Schwarzschild solution outside the past horizon. However, just as the maximal analytic extension of a metric like $dx^2+\sin^2 x\, dy^2$ *forces* on us the fact that the coordinate y is periodic, and forces on us the fact that the metric represents a two-sphere; so also here, the maximal analytic extension in the above-described manner *forces* the coordinate y to be 2π periodic; i.e., forces us to identify any two points $(u,v,x,y+2\pi n)$ and $(u,v,x,y+2\pi m)$ *throughout* the spacetime, including, of course, the regions I, II, III, and IV lying *before* the Killing-Cauchy horizon. Thus, the maximal analytic extension of (4.16) yields us an exact solution, which describes the collision of two plane-symmetric sandwich gravitational waves propagating in a *cylindrical universe* with topology

$R^3{\times}S^1$. When these waves collide, they produce a Killing-Cauchy horizon, which, (i) when added to the unextended spacetime region $I^-[\{(u/a)+(v/b)=\pi/2\}]$ (with topology $R^3{\times}S^1$) causes the topology of the extended spacetime to become $R^2{\times}S^2$ because of the above identifications, and (ii) encloses a spacetime singularity (the future Schwarzschild singularity) that is spacelike. In fact, the collision produces a Schwarzschild black hole, complete with its future horizon and the two asymptotically flat regions. In Fig. 5, we have tried to depict symbolically the global structure of this maximal colliding plane-wave spacetime. Note that, after the maximal analytic extension is carried out, the singular points P and Q of the solution (4.16) are contained in the bifurcation sphere $\mathcal{S}^2$ of the extended Schwarzschild horizon. Also note that it might be helpful to visualize, as we have done in Fig. 5, the cylindrical spacetime with topology $R^3{\times}S^1$ (representing the history of the colliding waves "before" the Killing-Cauchy horizon forms) as the direct product of R^1 (representing the time direction) with a finite-sized but open-ended cylinder $R^2{\times}S^1$ (representing a slice of constant time). As we also explain in Fig. 5, the cylinder $R^1{\times}S^1$ is topologically equivalent (homeomorphic) to a twice-punctured two-sphere; therefore, the slices of constant time $R^2{\times}S^1$ are homeomorphic to the direct product of R^1 with a twice-punctured two-sphere. When the Killing-Cauchy horizon $\{\alpha=0\}$ forms and the spacetime is extended beyond it in the above-described manner, the missing pairs of points of these "twice-punctured" spheres are supplied by points from the horizon, and thereby the extended spacetime acquires the topological structure of $R^2{\times}S^2$, instead of the topology $R^3{\times}S^1$ of the original cylindrical background on which the colliding plane waves propagate.

V. CONCLUSIONS

We can summarize the main results of this paper as follows.

(i) In a suitable coordinate system, the structure of the singularities produced by colliding parallel-polarized gravitational plane waves can be analyzed in full generality and detail. This analysis (a) reveals that the asymptotic structure of these singularities are of inhomogeneous Kasner type, and (b) provides explicit expressions for the asymptotic Kasner exponents in terms of the initial data posed by the incoming, colliding plane waves.

(ii) For specific choices of initial data for the colliding waves, the asymptotically Kasner form that the spacetime metric takes near the singularity can be that of a degenerate Kasner solution. In this case, the curvature singularities created by the colliding waves degenerate to coordinate singularities, and nonsingular Killing-Cauchy horizons are thereby obtained. The mathematical formalism that is built in this paper proves (a) that these horizons are unstable in the full nonlinear theory against small but generic perturbations of the initial data, and (b) that in a very precise sense, "generic" initial data always produce all-embracing, spacelike curvature singularities without Killing-Cauchy horizons.

(iii) An abundance of exact colliding parallel-polarized plane-wave solutions can be constructed, which exemplify some of the asymptotic singularity structures discussed in general terms in this paper. In particular, an infinite-parameter family of such solutions are found which create Killing-Cauchy horizons instead of curvature singularities. The analytic extension of one of these solutions across its Killing-Cauchy horizon results in a maximal spacetime, in which a Schwarzschild black hole is created out of the collision between two plane-symmetric sandwich waves propagating in a cylindrical universe.

There are a few specific directions for further research along the lines discussed in this paper that are worth listing. These are the following.

(i) A similar study of the more general problem of colliding plane waves with arbitrarily oriented polarizations. It will be interesting to find out whether the fundamental aspects of our results (i) and (ii) above remain intact after the new degree of freedom associated with a discrepancy in the incoming polarizations enters the problem. (In fact, recent work by the author[30] shows that this is indeed the case.)

(ii) A similar analysis of the problem of colliding plane waves coupled with matter fields.[9] Again the most interesting targets for such an inquiry will be understanding the validity of the results (i) and (ii) above under the presence of a nonzero stress-energy tensor.

(iii) Finally, an analysis of the structure of singularities produced by colliding *almost*-plane waves (see Refs. 3, 15, and 31 in this connection). Although such an analysis may well be beyond the capabilities of current analytical techniques, the question of whether the relaxation of strict plane symmetry in the initial data will cause the asymptotic singularity structure to deviate significantly from inhomogeneous Kasner (and, e.g., to become inhomogeneous mixmaster[12] or some more general structure) is an extremely interesting one.

Note added in proof. After this paper had been accepted for publication, the author learned that solutions similar to the solution derived from the Schwarzschild metric and studied in Sec. IV here have been discovered and studied from another viewpoint previously and independently by Ferrari, Ibanez, and Bruni.[32]

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grant No. AST 85-14911.

REFERENCES FOR CHAPTER 5

[1] J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).

[2] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[3] U. Yurtsever, Phys. Rev. D **37**, 2803 (1988) (Chapter 3).

[4] U. Yurtsever, Phys. Rev. D **36**, 1662 (1987) (Chapter 2).

[5] K. Khan and R. Penrose, Nature (London) **229**, 185 (1971).

[6] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

[7] Y. Nutku and M. Halil, Phys. Rev. Lett. **39**, 1379 (1977).

[8] R. A. Matzner and F. J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[9] S. Chandrasekhar and B. C. Xanthopoulos, Proc. R. Soc. London **A398**, 223 (1985); **A402**, 37 (1985); **A402**, 205 (1985); **A403**, 189 (1986).

[10] U. Yurtsever, Phys. Rev. D. **37**, 2790 (1988) (Chapter 4).

[11] S. Chandrasekhar and B. C. Xanthopoulos, Proc. R. Soc. London **A** (to be published).

[12] E. Kasner, Am. J. Math. **43**, 217 (1921); E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz. Nauk **80**, 391 (1964) [Sov. Phys. Uspekhi **6**, 495 (1963)]; C. W. Misner, Phys. Rev. Lett. **22**, 1071 (1969); V. A. Belinsky and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **57**, 2163 (1969) [Sov. Phys. JETP 30, 1174 (1970)]; V. A. Belinsky, I. M. Khalatnikov, and E. M. Lifshitz, Adv. Phys. **19**, 525 (1970).

[13] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

[14] R. Price and K. S. Thorne, Phys. Rev. D **33**, 915 (1986) (Appendix A).

[15] U. Yurtsever, (in preparation).

[16] U. Yurtsever, (in preparation).

[17] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).

[18] P. R. Garabedian, *Partial Differential Equations* (Wiley, New York, 1967).

[19] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1980).

[20] R. Penrose in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[21] F. J. Tipler, C. J. S. Clarke, and G. F. R. Ellis, in *Einstein Centenary Volume*, edited by A. Held and P. Bergmann (Plenum, NewYork, 1979).

[22] M. Simpson and R. Penrose, Int. J. Theor. Phys. **7**, 183 (1973); J. M. McNamara, Proc. R. Soc. London **A358**, 499 (1978); **A364**, 121 (1978); Y. Gursel, I. D. Novikov, V. D. Sandberg, and A. A. Starobinsky, Phys. Rev. D **19**, 413 (1979); **20**, 1260 (1979); R. A. Matzner, N. Zamorano, and V. D. Sandberg, ibid. **19**, 2821 (1979); S. Chandrasekhar and J. B. Hartle, Proc. R. Soc. London **A384**, 301 (1982).

[23] M. Morris and K. S. Thorne, Am. J. Phys. **56**, 395 (1988).

[24] S. Chandrasekhar and B. C. Xanthopoulos, Proc. R. Soc. London **A** (to be published).

[25] V. Moncrief and J. Isenberg, Commun. Math. Phys. **89**, 387 (1983); J. Math. Phys. **26**, 1024 (1985).

[26] F. J. Tipler, Phys. Rev. D **16**, 3359 (1977).

[27] K. Yosida, *Functional Analysis* (Springer, Berlin, 1985).

[28] I am indebted to R. Geroch for pointing this out to me.

[29] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam, 1982).

[30] *U.* Yurtsever, Physical Review D, submitted (also available as Caltech preprint GRP-184) (Chapter 7).

[31] U. Yurtsever, Phys. Rev. D **38**, 1731 (1988) (Chapter 6).

[32] V. Ferrari and J. Ibanez, Gen. Relativ. Gravit. **19**, 405 (1987); V. Ferrari, J. Ibanez, and M. Bruni, Phys. Rev. D **36**, 1053 (1987); V. Ferrari and J. Ibanez, Proc. R. Soc. London **A** (to be published). I am grateful to V. Ferrari for bringing these references to my attention.

FIGURE CAPTIONS FOR CHAPTER 5

FIG. 1. The two-dimensional geometry of the characteristic initial-value problem for colliding plane waves. The null surfaces $\{u=0\}$ and $\{v=0\}$ are the past wave fronts of the incoming plane waves 1 and 2. Initial data corresponding to waves 1 and 2 are posed, respectively, on the upper portions of the surfaces $\{v=0\}$ and $\{u=0\}$ that are adjacent to the interaction region I. The geometry in the region IV is flat, and the geometry in regions II and III is given by the metric describing the incoming waves 1 and 2, respectively. The geometry of the interaction region I is uniquely determined by the solution of the above initial-value problem. The directions in which the various lines of constant coordinates u, v, α, β, r, and s run are also indicated, along with the descriptions of the initial null surfaces in these different coordinate systems.

FIG. 2. The geometry of a colliding plane-wave solution (2.43) for which $|\varepsilon(\beta)|\equiv 1$ throughout an interval I in β. The interval I is disconnected and is made up of several connected pieces $I_1, I_2, ..., I_n$. Since the surface $\alpha=0$ corresponds to a Killing-Cauchy horizon whenever $\beta \in I$, the singularity at $\alpha=0$ is interrupted by the n Killing-Cauchy horizons $\mathcal{S}_1$, $\mathcal{S}_2$, ..., $\mathcal{S}_n$ which are located along the intervals $I_1, I_2, ..., I_n$, respectively. Through each of the horizons $\mathcal{S}_i$ the spacetime curvature is finite and well behaved. Consequently, across each horizon $\mathcal{S}_i$ the metric can be extended smoothly to a maximal spacetime (possibly a different one for each i) whose choice is essentially arbitrary. In particular, the observers living in the interaction region I can use these horizons $\mathcal{S}_i$ as tunnels along which they can travel (e.g., following the timelike world lines γ in the figure) into different universes. Equation (3.13) in the text, which relates the function $\varepsilon(\beta)$ to the initial data for the colliding waves, can be regarded as describing a kind of superposition, at $\alpha=0$, of the two wave forms constituting the

initial data for the incoming colliding plane waves. For example, to compute $\varepsilon(\beta)$ at the point p in the figure, the initial data located in the cross-hatched portions of the two initial surfaces are superposed through Eq. (3.13). In order to have $|\varepsilon(\beta)|$ constant across a connected interval like the interval I_1, it is necessary to adjust the initial data so as to cancel precisely the two separate contributions to $\varepsilon(\beta)$ which would arise as the point p is moved across the interval I_1.

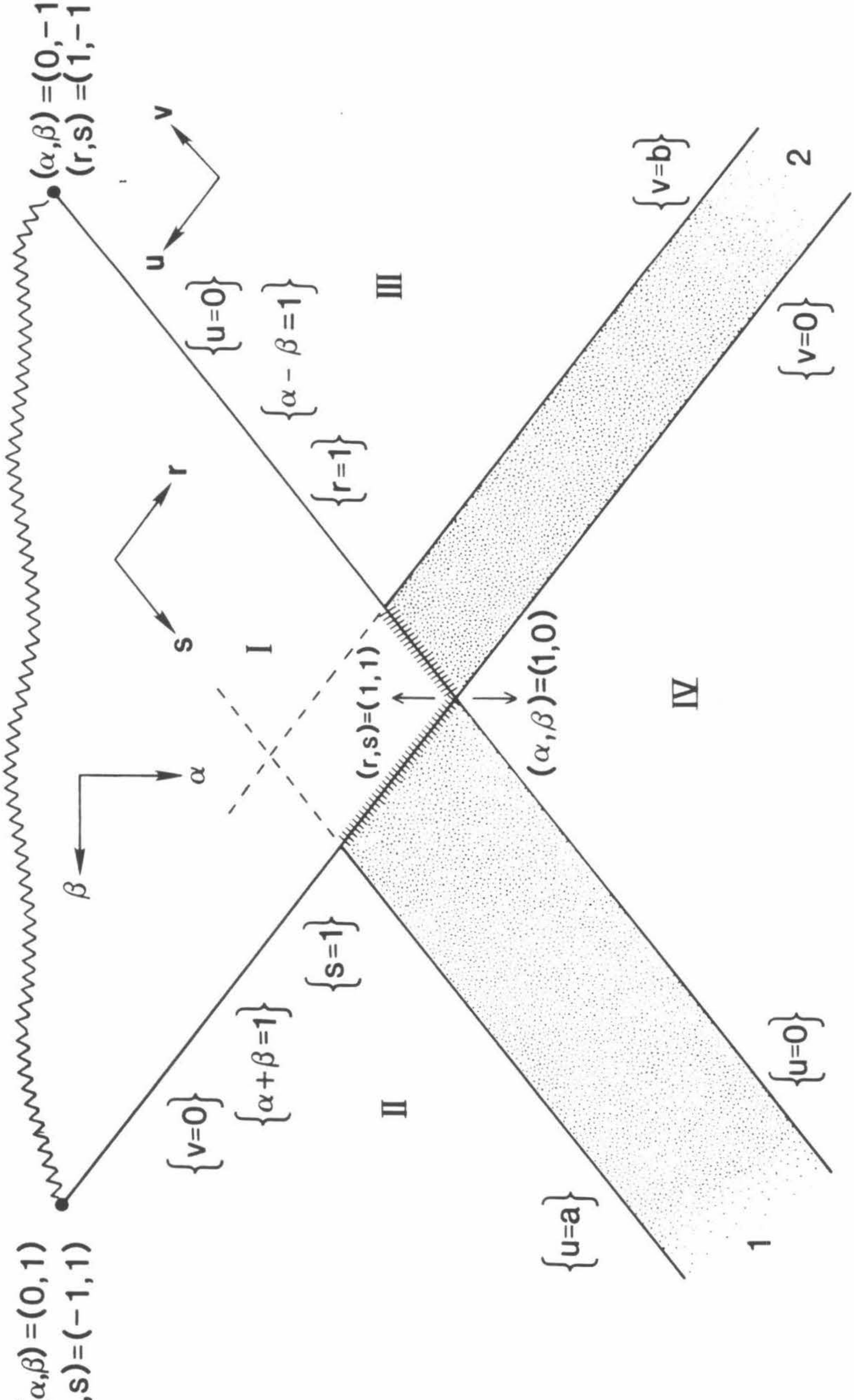
FIG. 3. The region I in Schwarzschild spacetime to which the interaction region of the colliding plane-wave solution (4.16) is locally isometric. This region I is shown shaded in this figure, which is drawn in a $\{t=\text{const}\}$, $\{\phi=0,\pi\}$ plane. As explained in the text, the geometry in region I is extended nonanalytically beyond the null surfaces $r=M(1+\cos\theta)$ and $r=M(1-\cos\theta)$, which correspond to the wave fronts $\{u=0\}$ and $\{v=0\}$, respectively. After this extension, the geometry in regions II and III represents incoming single plane sandwich waves [Eqs. (4.18)]; and region IV is flat. The interaction region I is bounded by a Killing-Cauchy horizon which corresponds to the event horizon of the Schwarzschild spacetime at $\{r=2M\}$. In Ref. 10, we have used the region IV of the interior Schwarzschild spacetime as the interaction region of the colliding plane-wave solution (10.2.4); the solution of Ref. 10 was obtained by exactly the same procedure as the solution (4.16) which we outline in Sec. IV here.

FIG. 4. The global structure of the colliding plane-wave solution (4.16). One of the spacelike Killing directions, namely, the y direction, is suppressed. The remaining spacelike Killing vector $\partial/\partial x$ becomes null on the Killing-Cauchy horizon as depicted; in fact, it is tangent to the null generators of the horizon. This Killing vector $\partial/\partial x$ corresponds to the Killing vector $\partial/\partial t$ of the Schwarzschild spacetime, which

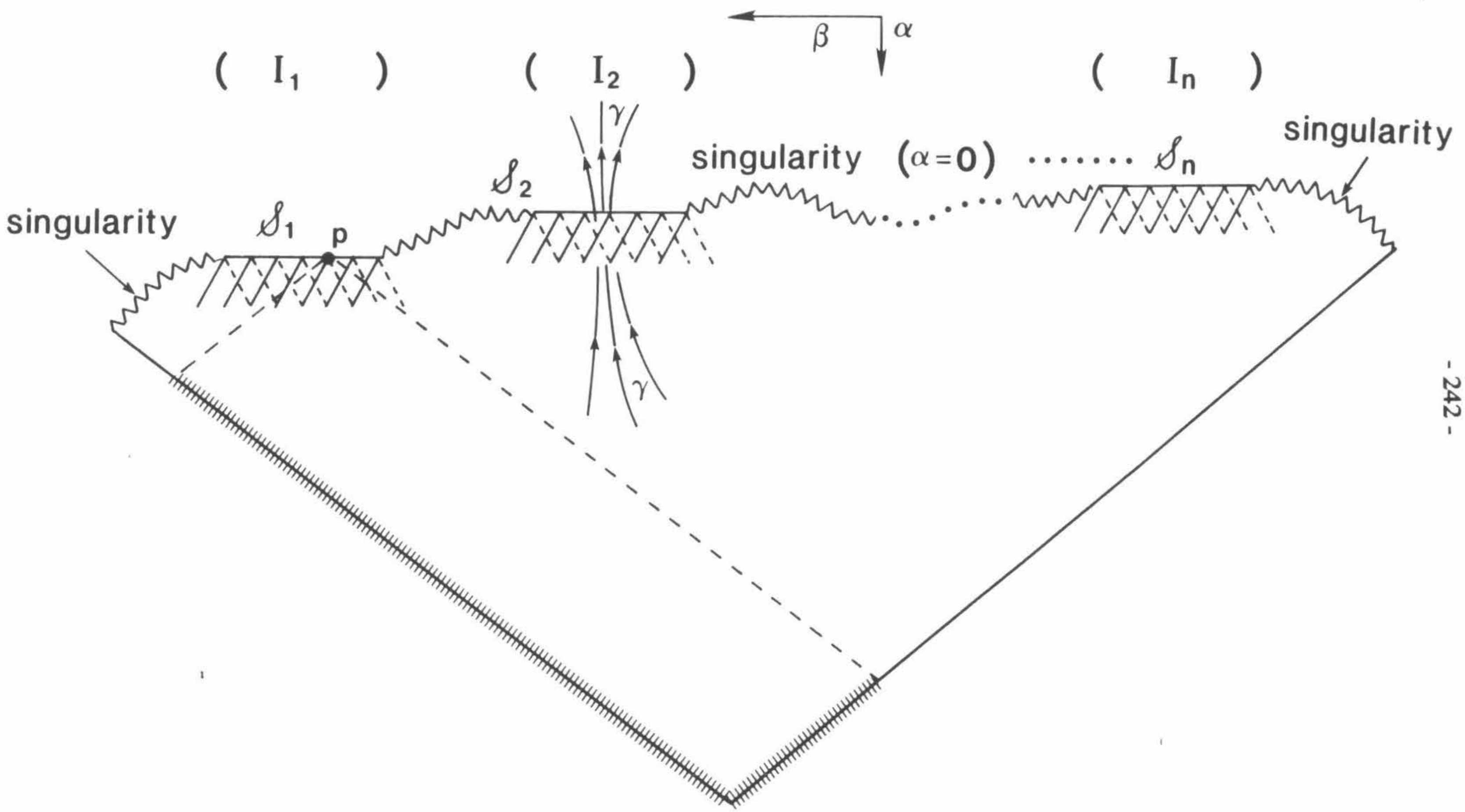
similarly becomes null on the event horizon. The suppressed Killing vector $\partial/\partial y$ corresponds, under the local isometry with the Schwarzschild spacetime, to the cyclic Killing vector $\partial/\partial\phi$. However, in constructing the solution (4.16), we have changed the topology of the spacetime from the topology $R^2 \times S^2$ of Schwarzschild to R^4. Therefore, the Killing vector $\partial/\partial y$ is no longer cyclic. In particular, the bifurcation two-sphere $\mathcal{S}^2$ of the Schwarzschild horizon is "torn" open in our solution to a bifurcation set which has topology R^2. The Killing vector field $\partial/\partial x$, which becomes null on the horizon, vanishes on this bifurcation set. Although the focal planes for each of the incoming plane waves in the solution (4.16) are nonsingular, the points P and Q where these focal planes intersect the bifurcation set of the Killing-Cauchy horizon represent spacetime singularities (see the discussion in Sec. IV).

FIG. 5. The global structure of the maximal colliding plane-wave spacetime obtained by analytically extending the solution (4.16) across its Killing-Cauchy horizon using Kruskal-type global coordinates. The description in the figure is only symbolic, and is intended to help the reader in visualizing the true geometry of this maximal extension. As explained in the text (Sec. IV), the maximal analytic extension of the metric (4.16) (leading to the Schwarzschild spacetime outside the past horizon) causes the (Killing-) coordinate y to become cyclic, and thereby causes the topology of the extended spacetime to become $S^2 \times R^2$ instead of R^4 [see Fig. 4 for a description of the global structure of the unextended solution (4.16)]. Similarly, this change in the topological nature of the coordinate y implies that the region of the maximal spacetime which lies to the past of the Killing-Cauchy horizon has topology $R^3 \times S^1$ instead of R^4. Since this region describes the history of the colliding plane waves before they create the Killing-Cauchy horizon, it follows that the incoming waves propagate and collide in a

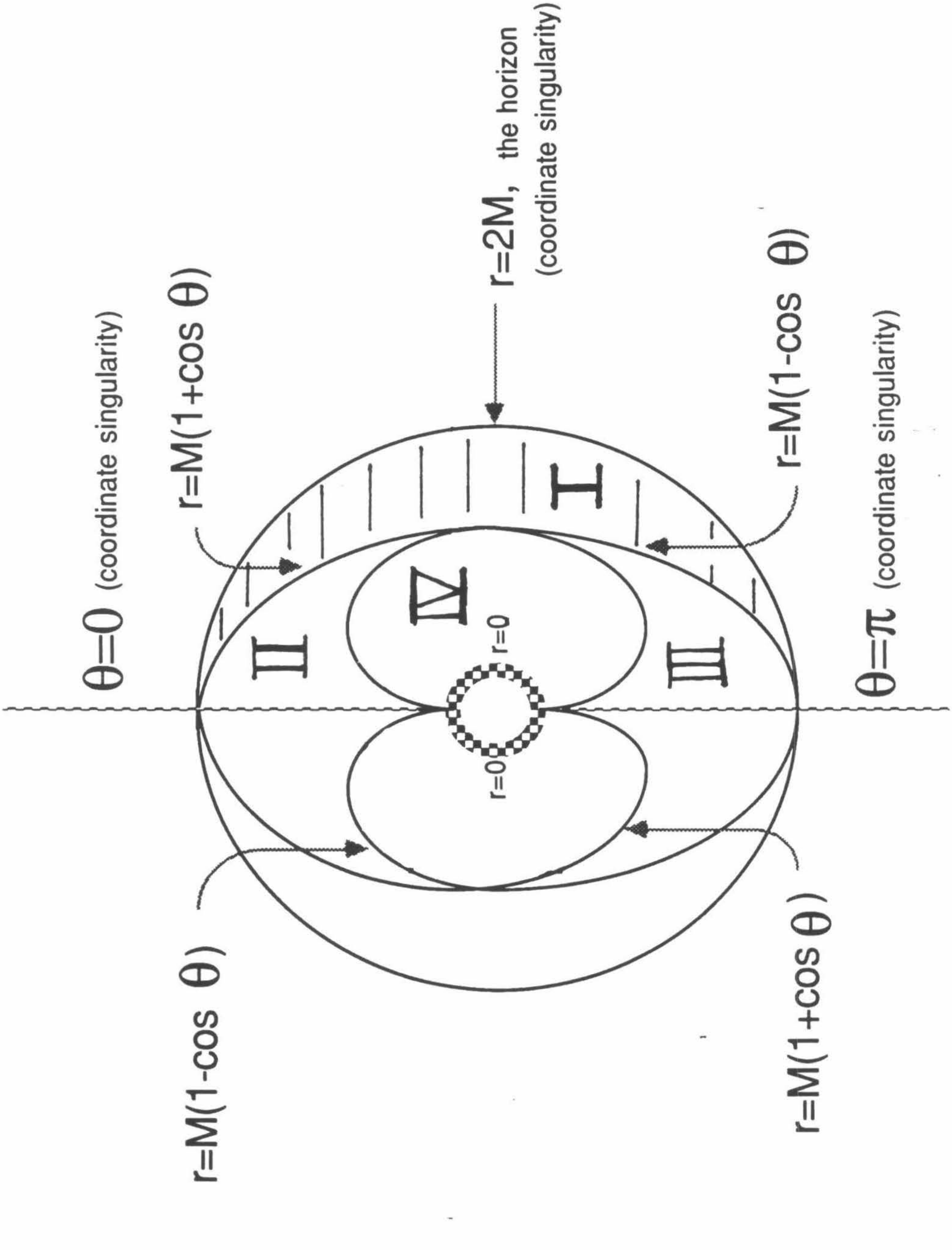
cylindrical universe with topology $R^3{\times}S^1$. Note that, after the maximal analytic extension is carried out, the singular points P and Q of the solution (4.16) are contained in the bifurcation sphere $\mathcal{S}^2$ of the extended Schwarzschild horizon. Also note that it might be helpful to visualize, as is done in the figure here, the cylindrical spacetime with topology $R^3{\times}S^1$ (representing the history of the colliding waves "before" the Killing-Cauchy horizon forms) as the direct product of R^1 (representing the time direction) with a finite-sized but open-ended cylinder $R^2{\times}S^1$ (representing a slice of constant time). The cylinder $R^1{\times}S^1$ is topologically equivalent (homeomorphic) to a twice-punctured two-sphere; therefore, the slices of constant time $R^2{\times}S^1$ are homeomorphic to the direct product of R^1 with a twice-punctured two-sphere. When the Killing-Cauchy horizon $\{\alpha{=}0\}$ forms and the spacetime is extended beyond it, the missing pairs of points of these "twice-punctured" spheres are supplied by points from the horizon, and thereby the extended spacetime acquires the topological structure of $R^2{\times}S^2$, instead of the topology $R^3{\times}S^1$ of the original cylindrical background on which the colliding plane waves propagate. Thus, the analytic extension of the solution (4.16) gives a maximal spacetime, in which a Schwarzschild black hole is created out of the collision between two plane-symmetric sandwich waves propagating in a cylindrical universe.

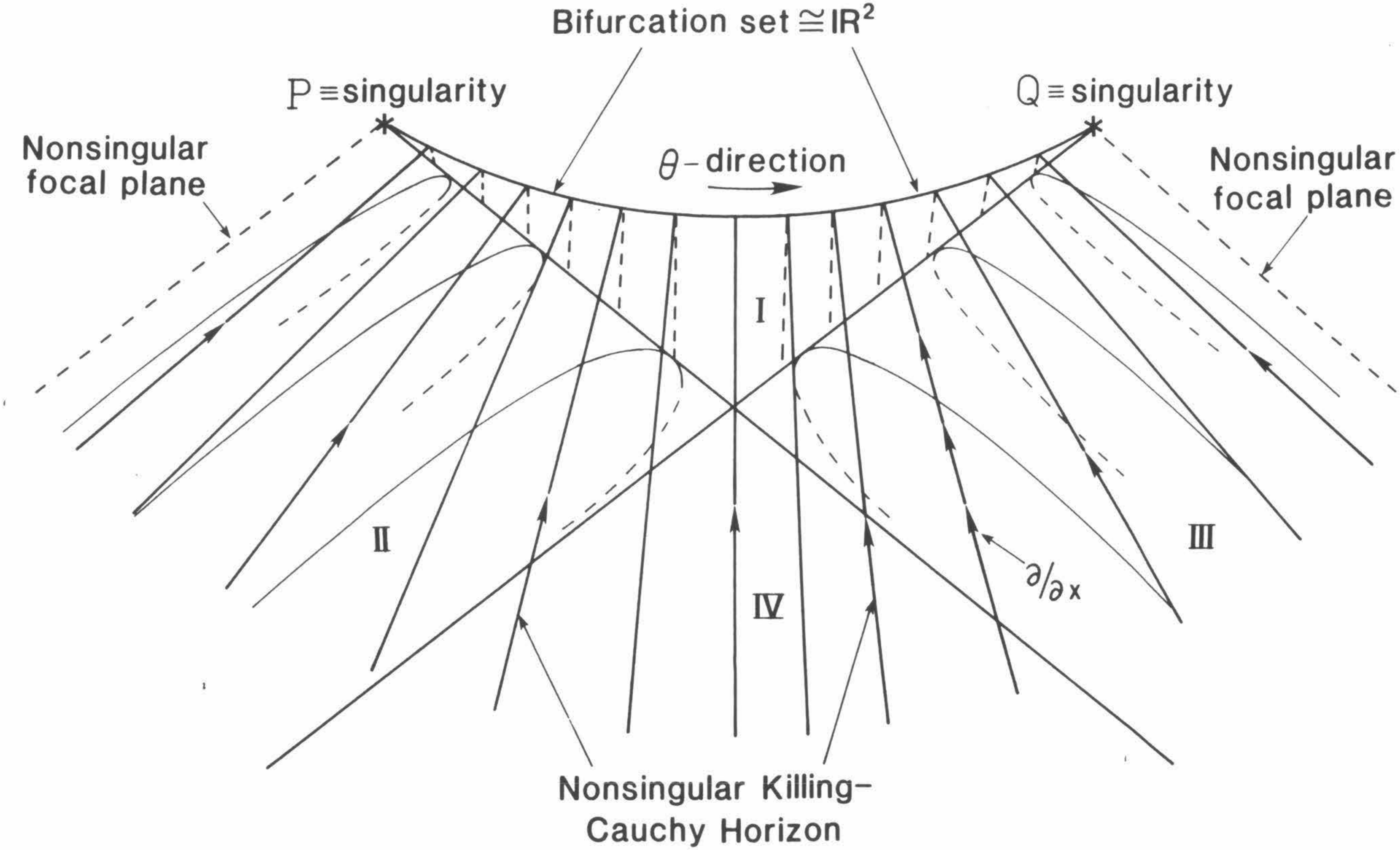


{α=0} ≡ {r+s=1}
(α,β)=(0,−1)
(r,s)=(1,−1)
(α,β)=(0,1)
(r,s)=(−1,1)
v
u
r
s
α
β
{u=0}
{α − β=1}
{r=1}
III
I
II
IV
{v=0}
{α+β=1}
{s=1}
(r,s)=(1,1)
(α,β)=(1,0)
{v=b}
{v=0}
{u=0}
{u=a}
1
2



β
α
(I_1)
(I_2)
(I_n)
γ
γ
singularity
singularity
singularity
singularity ($\alpha = 0$)
$\mathcal{S}_1$
p
$\mathcal{S}_2$
$\mathcal{S}_n$


r=2M, the horizon
(coordinate singularity)
r=M(1+cos θ)
r=M(1-cos θ)
θ=0 (coordinate singularity)
θ=π (coordinate singularity)
r=M(1-cos θ)
r=M(1+cos θ)
I
II
III
IV
r=0
r=0


Bifurcation set ≅ IR²
P ≡ singularity
Q ≡ singularity
Nonsingular
focal plane
θ - direction
Nonsingular
focal plane
I
II
III
IV
∂/∂x
Nonsingular Killing-
Cauchy Horizon

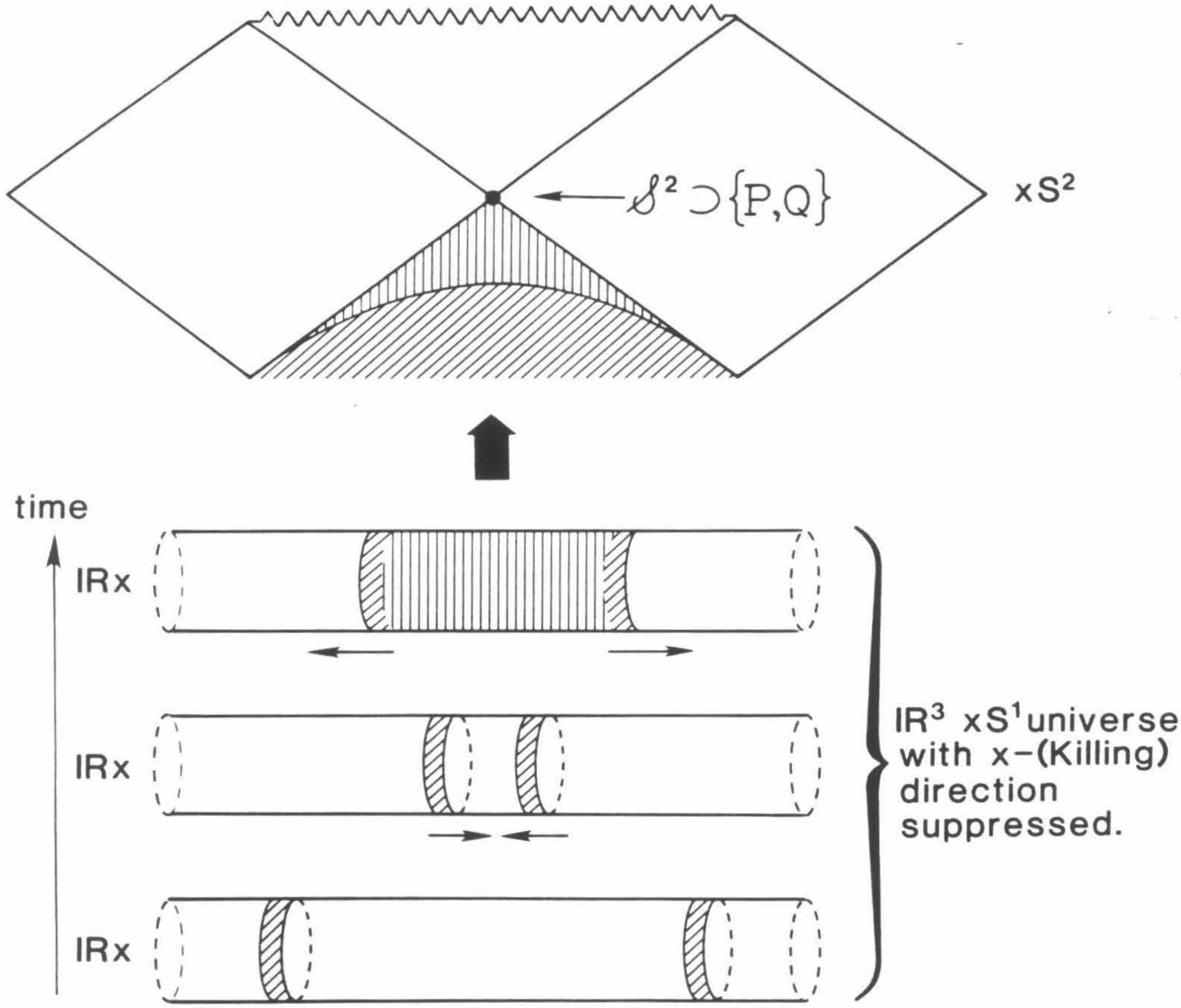

$\mathcal{S}^2 \supset \{P,Q\}$
$\times S^2$
time
IRx
IRx
IRx
IR³ xS¹universe
with x-(Killing)
direction
suppressed.

FIG. 5

CHAPTER 6

Singularities in the Collisions of Almost-Plane Gravitational Waves

Published in Physical Review D **38**, 1731 (1988).

ABSTRACT

It is well known that when gravitational plane waves propagating on an otherwise flat background collide, they produce spacetime singularities. In this paper we consider the problem of whether (or under what conditions) singularities can be produced by the collision of gravitational waves with finite but very large transverse sizes. On the basis of (nonrigorous) order-of-magnitude considerations, we discuss the outcome of the collision in two fundamentally different regimes for the parameters of the colliding waves; these parameters are the transverse sizes $(L_T)_i$, typical amplitudes h_i, typical reduced wavelengths $\lambda\!\!\!^{-}_i \equiv \lambda_i/2\pi$, thicknesses a_i, and focal lengths $f_i \sim \lambda\!\!\!^{-}_i{}^2/a_i h_i{}^2$ $(i=1,2)$ of the waves 1 and 2. For the first parameter regime where $(L_T)_1 \gg (L_T)_2$ and $h_1 \gg h_2$, we conjecture the following. (i) If $(L_T)_2 \ll \sqrt{\lambda\!\!\!^{-}_2 f_1}(\frac{h_1}{h_2})^{1/4}$, the almost-plane wave 2 will be focused by wave 1 down to a finite, minimum size, then diffract and disperse [Fig. 1(a)]. (ii) If $(L_T)_2 \gg \sqrt{\lambda\!\!\!^{-}_2 f_1}(\frac{h_1}{h_2})^{1/4}$ (and if wave 1 is sufficiently anastigmatic), wave 2 will be focussed by wave 1 so strongly that it forms a singularity surrounded by a horizon, and the end result is a black hole flying away from wave 1 [Fig. 1(b)]. For the second parameter regime where $(L_T)_1 \sim (L_T)_2 \equiv L_T$ and $h_1 \sim h_2$, we conjecture that if $L_T \gg \sqrt{f_1 f_2} \equiv f$, a horizon forms around the two colliding waves shortly before their collision, and the collision produces a black hole that is at rest with respect to the reference frame in which $f_1 \sim f_2 \sim f$ (Fig. 2). As a first step in proving this conjecture, we give a rigorous analysis of the second regime in the case $L_T \gg f$, for the special situation of colliding parallel-polarized (almost-plane) gravitational waves which are exactly plane-symmetric across a region of transverse size $\gg f$, but which fall off in an arbitrary way at larger transverse distances. Our rigorous analysis shows that this collision is guaranteed to produce a spacetime singularity with

the same local structure as in an exact plane-wave collision, but it does not prove that the singularity is surrounded by a horizon.

I. INTRODUCTION AND OVERVIEW

In a previous paper,[1] we have discussed and reviewed the general structure of plane wave and colliding plane-wave spacetimes in general relativity, and we have argued on the basis of previous work,[2–8] both by the author and largely by others, that (because of the focusing effect of gravitational plane waves), collisions between plane waves generically produce spacelike, all-embracing spacetime singularities. We have also posed, in the introductory section of the same paper,[1] the problem of colliding *almost-plane* gravitational waves; that is, the problem of whether (or under what conditions) spacetime singularities can be produced by the collision of gravitational waves with finite but very large transverse "spatial" sizes. General gravitational-wave spacetimes are defined and analyzed in Secs. IV A–IV E of Ref. 1. The detailed definition of almost-plane waves (which are a special case of gravitational wave spacetimes), and a perturbative analysis of the spacetimes produced by their collisions are the subjects of a future paper.[9] In the present paper, we formulate and prove a very specific theorem (based on the full nonlinear Einstein equations), which demonstrates that provided the colliding almost-plane waves have sufficiently "large" transverse size and satisfy a certain fairly strong restriction, their collision is guaranteed to produce spacetime singularities.

It may be useful to make clear right at the outset, that the greatest significance of this singularity result (and of the almost-plane-wave collision problem in general) does not really lie in showing that the creation of singularities by colliding exact plane waves is a "generic" property of colliding gravitational waves; that is, in showing that the singularities are *not* artifacts of the infinite amount of "energy" carried by the colliding, transversely infinite plane waves. Rather, the greatest significance of the almost-plane-wave collision problem lies with its potential to provide substantial insights into some fundamental issues in general relativity (such as cosmic censorship, the structure of generic singularities, ...), which can only be explored by studying the dynamics of fully nonlinear gravitational fields in the absence of symmetries. From this point of view, the most interesting aspect of colliding almost-plane waves is not that they represent gravitational waves with finite energy, but rather that they represent gravitational waves propagating and colliding in a background spacetime which is asymptotically flat. [In fact, even the presently available exact solutions for colliding plane waves demonstrate that the creation of singularities from plane-wave collisions has nothing to do with the infinite energy present in the incoming waves: Just take any one of these solutions and identify the points of the spacetime in such a way to make the Killing x,y coordinates cyclic; i.e., identify any two points of the form $(u,v,x+2\pi n,y+2\pi m)$ and $(u,v,x+2\pi k,y+2\pi l)$ throughout the spacetime (see below for a description of the Rosen-type coordinates u,v,x,y for a colliding plane-wave solution). The result is a colliding plane-wave spacetime, that represents the creation of spacetime singularities from the collision between two plane-symmetric gravitational waves of finite "size" (thus of finite energy), which propagate and collide in a toroidal universe with topology $S^1 \times S^1 \times R^2$.]

Before giving a detailed description of the plan of this paper, we will first present an informal overview of the physically important aspects of the almost-plane-wave collision problem.

For our purposes in this paper, an almost-plane wave is a gravitational wave spacetime[1] on which there exist (i) a local coordinate system (u,v,x,y), and (ii) a length scale L_T that characterizes the variation in the x, y directions of the components of geometric quantities. We assume (iii) that throughout the intersection of a suitable partial Cauchy surface Σ with the wave's *central region* (which has the form

$\mathcal{C}=\{\, |x| < L_T,\ |y| < L_T,\ u,\ v\,\}$), the metric components and other quantities are very nearly equal to the corresponding quantities for an exact plane-wave spacetime, and (iv) that the curvature components fall off arbitrarily (but in a manner consistent with the constraint equations) as $x^2+y^2 \rightarrow \infty$ at constant u and v. (See Ref. 9 for a more precise and detailed definition of almost-plane waves.) Hence, a local coordinate system (u, v, x, y) of the above type for an almost-plane wave is directly analogous to the Rosen-type coordinate systems[1] defined on exact plane-wave spacetimes. In Ref. 9 we also assume, in order for our perturbation formalism to make sense, that the dimensionless amplitude h of an almost-plane wave is everywhere small compared to unity. However, this restriction on the typical magnitude of h for an almost-plane wave is not necessary for the singularity theorem that we prove in this paper.

Now consider two almost-plane gravitational waves propagating and colliding on an otherwise flat background. If the central regions of the two waves collide with each other (which we will always assume to be the case), then [at least in some neighborhood of the characteristic initial surface[3,1] $\mathcal{N}=\mathcal{N}_1 \cup \mathcal{N}_2$ formed by the initial wave fronts $\mathcal{N}_1$, $\mathcal{N}_2$ of the colliding waves (Fig. 3)], it is possible[3,1] to set up a local coordinate system in which the conditions (ii)—(iv) above are satisfied for *both* colliding waves simultaneously; but possibly with different transverse length scales $(L_T)_1$ and $(L_T)_2$. In this coordinate system, the initial data supplied by the almost-plane wave 1 and posed on the initial null surface $\mathcal{N}_2$ are very nearly equal, throughout $\mathcal{C}_1 \cap \mathcal{N}_2$, to the initial data posed by a corresponding exact plane wave 1; and the initial data supplied by the almost-plane wave 2 and posed on the initial null surface $\mathcal{N}_1$ are very nearly equal, throughout $\mathcal{C}_2 \cap \mathcal{N}_1$, to the initial data supplied by a corresponding exact plane wave 2. (See Refs. 3 and 10, and Fig. 3 below for a detailed description of the characteristic initial-value problem for colliding exact plane waves.) The fundamental problem of colliding almost-plane gravitational waves is then to determine whether (or under what conditions on the initial data) the evolution of these data produces spacetime singularities. There are two qualitatively different parameter regimes for this problem on which some partial results are available at the time of this writing. We will now discuss these two regimes separately.

One regime is characterized by $(L_T)_1 \gg (L_T)_2$ and $h_1 \gg h_2$; that is, one of the almost-plane waves (namely, wave 1) is much larger in its transverse size and much stronger in its amplitude than the other (wave 2). In this case, we expect that the problem can be treated successfully by regarding the weaker wave 2 as a small disturbance propagating on a background spacetime predominantly determined by the stronger wave 1. Because of the focusing effect[11,1,4] of gravitational waves, the background wave 1 will then act like a lense, focusing the weaker almost-plane wave into a smaller and smaller transverse size and a larger and larger amplitude as it propagates farther towards the focal plane[1,4] of the background almost-plane wave (Fig. 1). This focusing will continue until it is terminated in one of the following two ways: Either the amplitude h_2 of the almost-plane wave 2 reaches a value large compared to unity *before* diffraction effects have a chance to counterbalance the focusing effect of the background; or diffraction wins over focusing before wave 2 gets to be nonlinearly strong. In the first case [Fig. 1(b)], a large amount of mass energy is focused into such a small spacetime region that a closed-trapped surface forms, followed by a spacetime singularity; this singularity would be hidden inside a black hole (event horizon) if the cosmic censorship hypothesis is valid under these circumstances [Fig. 1(b)]. In the second case [Fig. 1(a)], the beam of gravitational radiation corresponding to the almost-plane wave 2 passes through a narrow waist near the focal plane, but is dispersed afterwards by diffraction, without getting sufficiently nonlinearly strong to form a black hole [Fig. 1(a)]. It is easy to estimate, in rough order of magnitude, the quantitative criterion for the first alternative to be the case: The length scale at which diffraction effects would tend to dominate the propagation of wave 2 is $(L_T)_2^{\,2}/\lambda\!\!\!^{-}_2$, where $\lambda\!\!\!^{-}_2$ is a typical wavelength for wave 2. If the first focal length[1] of wave 1 is f_1 (and if wave 1 is sufficiently anastigmatic[1] so that both focal lengths are approximately the same), then singularities will form when $(L_T)_2^{\,2}/\lambda\!\!\!^{-}_2 \gg f_1$. A more careful analysis indeed shows that singularities form when

$$(L_T)_2 \gg (\lambda\!\!\!^{-}_2 f_1)^{1/2} \left(\frac{h_1}{h_2} \right)^{1/4} . \tag{1.1}$$

Though they are all physically plausible, the specific conjectures on which the above scenario is based [and in particular the singularity criterion Eq. (1.1)] have not yet been proved rigorously. In Ref.9, which is devoted to a perturbative analysis of the collision problem in the regime we have just discussed, we will present some partial results in the direction of the above conjectures, which are obtained under a number of idealizations. The *tentative* results of this analysis indicate (i) that if the background wave is sufficiently anastigmatic[1] (i.e., if the difference between the focal lengths of the background is sufficiently small), then singularities will form (presumably via the collapse of the almost-plane wave 2 into a black hole after it has been focused by the background) when the inequality (1.1) is satisfied, and (ii) that if the background wave is highly astigmatic,[11,1] then singularities will form only when there is sufficient mass-energy in the colliding waves to form a closed-trapped surface *already on the initial characteristic surface*, before the evolution of the almost-plane wave 2 begins and focusing takes place. The detailed discussion and derivation of these results will be found in Ref. 9.

Although our tentative results[9] so far in support of the singularity criterion (1.1) have been obtained in the first parameter regime, where $(L_T)_1 \gg (L_T)_2$ and $h_1 \gg h_2$, it seems evident that the criterion (1.1) is valid regardless of the relative strengths and transverse sizes of the colliding waves. In other words, we find it natural to generalize our conjectures on the singularity criterion (1.1) to include the second parameter regime (see below) where $(L_T)_1 \sim (L_T)_2$ and $h_1 \sim h_2$, as well as the first regime where $(L_T)_1 \gg (L_T)_2$ and $h_1 \gg h_2$. More precisely, we conjecture the following: (i) If $(L_T)_2 \ll \sqrt{\lambda_2 f_1}(h_1/h_2)^{1/4}$ and $(L_T)_1 \ll \sqrt{\lambda_1 f_2}(h_2/h_1)^{1/4}$ (where f_2 denotes the first focal length of the second wave), then the colliding waves will be focused by each other down to a finite, minimum size, then diffract and disperse. (ii) If $(L_T)_2 \gg \sqrt{\lambda_2 f_1}(h_1/h_2)^{1/4}$ and $(L_T)_1 \gg \sqrt{\lambda_1 f_2}(h_2/h_1)^{1/4}$ (and waves 1 and 2 are sufficiently anastigmatic), then the waves will be focused by each other so strongly that they form two separate singularities surrounded by event horizons, and the end result is two black holes flying away from each other.

The second parameter regime for the almost-plane-wave collision problem can be characterized as the complement of the first regime; that is, as the regime in which there are no *a priori* restrictions on the relative orders of magnitude of the parameters $(L_T)_1$, $(L_T)_2$, h_1, and h_2. However, in going through the discussion below, the readers might find it useful to assume that $(L_T)_1 \sim (L_T)_2 \equiv L_T$ and $h_1 \sim h_2$, as would be true for a typical problem that belongs to the second parameter regime. This is by far the most difficult parameter regime for the collision problem, since in this case it is not as easy to build a formalism based on the decoupling of the spacetime geometry to a background part and a small-disturbance part, as it is in the first regime above. Nevertheless, it is still possible, using some general causality arguments, to make a useful guess of a sufficient condition for the formation of singularities in this parameter regime. To do this, assume for simplicity that the colliding almost-plane waves have amplitudes small compared to unity: $h_1 \ll 1, h_2 \ll 1$. Then, the metric in the interaction region will be approximately flat throughout a large neighborhood of the initial characteristic surface $\mathcal{N}$. [In fact, if the colliding waves were *exactly plane symmetric* and had amplitudes h_1, h_2 of the same typical magnitude (much smaller than unity), then the geometry of the interaction region would be everywhere nearly flat except in a small spacetime "neighborhood" of the curvature singularity created by the colliding plane waves. In this neighborhood, the spacetime curvature generated by the nonlinear interaction and focusing of the plane waves becomes important; in fact, curvature diverges towards the singularity.] Now, returning to the original collision problem, where the initial data posed on $\mathcal{N}$ are "almost-plane symmetric" with a transverse length scale L_T, the readers can easily convince themselves (e.g., by drawing a simple spacetime diagram) that a plausible guess for a (sufficient) singularity criterion in this case would be

$$L_T \gg (f_1 f_2)^{1/2} . \qquad (1.2)$$

The plausability argument which leads to Eq. (1.2) is the following: Since the geometry is nearly flat throughout most of the interaction region, to a good approximation the causal structure of the interaction region is that of Minkowski spacetime. This means that during the time interval $\sim \tau$ of evolution, where τ measures the proper timelike distance between the singularity (which would be produced if the waves were exactly plane) and the initial surface, the information about the finite size of the colliding waves (which is contained in the initial data) can only propagate across a distance $\sim \tau$ along the transverse x, y directions. It is not hard to see that τ in this case is of the order of $(f_1 f_2)^{1/2} \equiv f$. Therefore, when the inequality (1.2) is satisfied, i.e., when

$L_T \gg f \sim \tau$, the information about the non-plane-symmetric, transversely finite nature of the colliding waves cannot reach the spatial locus of the singularity quickly enough to abort its formation.

Now, in view of our conjectures on the first singularity criterion, Eq. (1.1), it is natural to expect that for a typical collision problem the singularity criterion (1.2) will be an "overkill," i.e., that the collision would actually produce a spacetime singularity even for values of L_T much smaller than the values demanded by Eq. (1.2). Nevertheless, the result (1.2), or more precisely its more restricted version which we prove in this paper, is the only available singularity result which is valid for the second parameter regime. On intuitive grounds, we expect that the regime defined by Eq. (1.2) describes a situation where there is so much mass energy in the colliding almost-plane waves that a black-hole horizon forms slightly before their collision, and engulfs the central regions of both colliding waves, which subsequently collapse into a singularity inside the hole (Fig. 2). More quantitatively, we note that in the center-of-mass frame of the colliding almost-plane waves the total mass energy associated with the waves will be $M \sim \sqrt{M_1 M_2}$, where M_1 and M_2 are the total mass energy $M_i \sim (L_T)_i{}^2 a_i h_i{}^2/\lambda\!\!\!^{-}_i{}^2 \sim (L_T)^2/f_i$ in the two colliding waves in the original reference frame. Notice that $M \sim (L_T)^2/\sqrt{f_1 f_2} = (L_T)^2/f$; and thus our criterion for singularity formation (1.2) is equivalent to $M \gg L_T$, which says that the total spatial region occupied by the waves at the moment of their collision in their center-of-mass frame is smaller than their Schwarzschild radius. This suggests that a horizon has already formed around them by the time they collide. (This interpretation, as we attempt to illustrate in Fig. 2, is also in accordance with the local structure of the singularity whose existence is rigorously proved in Sec. III below: In general that singularity stretches continuously from the focal region of one wave into that of the other.

However, our rigorous analysis does not prove the existence of an event horizon enclosing the singularity.) This is in contrast with the singularity formation under the first singularity criterion [cf. Eq. (1.1)], where according to our conjectures [at least in the regime where $(L_T)_1$ and $(L_T)_2$ are comparable] the two colliding almost-plane waves collapse into and form two distinct black holes, which subsequently fly apart from each other [cf. Fig. 1(b)].

The singularity theorem which we prove in this paper is based on the following two idealizations. (i) We assume that the colliding almost-plane waves — though they belong to the second parameter regime above — are in fact *exactly plane-symmetric throughout their respective central regions* $\mathcal{C}_1$, $\mathcal{C}_2$; more precisely, the initial data supplied by the almost-plane wave 1 and posed on the initial null surface $\mathcal{N}_2$ are *precisely equal*, throughout $\mathcal{C}_1 \cap \mathcal{N}$, to the initial data posed by a corresponding exact plane wave 1; and the initial data supplied by the almost-plane wave 2 and posed on the initial null surface $\mathcal{N}_1$ are *precisely equal*, throughout $\mathcal{C}_2 \cap \mathcal{N}_1$, to the initial data supplied by a corresponding exact plane wave 2. Outside the central regions, these initial data fall off at large transverse distances (in a manner consistent with the constraint equations on the initial surface). However, the exact way in which this falloff occurs is inconsequential to the statement and proof of our singularity result. (ii) We assume that the colliding waves have parallel and constant linear polarizations. Under these two assumptions, we will find a lower bound for the transverse length scale L_T [Eq. (2.6)], above which singularities are guaranteed to be produced by the collision. In order of magnitude, this lower bound reduces to $(f_1 f_2)^{1/2}$ when the amplitudes of the colliding waves are small compared to unity.

The following is a brief description of the contents of the remaining two sections.

The rest of our analysis below is based entirely on our earlier paper;[12] consequently, readers *must* refer to Ref. 12 for the details underlying the discussions in Sec. II and in the Appendix. In fact, the remaining sections of this paper constitute a natural and specific application of the general formalism which we have developed in Ref. 12.

In Sec. II, we will use the results obtained in Ref. 12 to derive the asymptotic causal structure of a generic, colliding exact plane-wave spacetime near the singularity. We will then use a completely general lemma in general relativity (Lemma 2) to argue that, if the initial data for the colliding exact plane waves are cut off outside a sufficiently large transverse region (thereby making the waves almost-plane), then that cutoff will never influence the central region of the colliding plane-wave spacetime; and, in particular, the central region will evolve a singularity identical in local structure to that for the exact plane-wave case. This argument directly leads to the singularity result of the paper, which we state in the form of a theorem towards the end of Sec. II.

In the Appendix, we present a specific example for a colliding plane-wave solution, which evolves from physically reasonable initial data in the sense that the colliding plane waves (i) have amplitudes small compared to unity, and (ii) have well-defined wavelengths small compared to their focal lengths. This example illustrates the argument used in Sec. II to derive the singularity criterion (1.2) from the more general singularity theorem.

Our notation and other conventions throughout this paper are the same as in Ref. 12. Equation numbers that refer to equations of Ref. 12 will be denoted by a prefix "12"; for example, Eq. (12.2.37) refers to Eq. (2.37) of Ref. 12.

II. THE ASYMPTOTIC CAUSAL STRUCTURE OF A GENERIC COLLIDING EXACT PLANE-WAVE SPACETIME NEAR ITS SINGULARITY; AND A SINGULARITY RESULT FOR COLLIDING ALMOST-PLANE WAVES

In Sec. III A of Ref. 12, we have found that the asymptotic limit as $\alpha \to 0$ of the spacetime metric (12.2.43) for colliding parallel-polarized plane-waves is an inhomogeneous Kasner[13] metric given by Eq. (12.3.17), where the Kasner exponents $p_i(\beta)$ are expressed in terms of initial data for the colliding waves through Eqs. (12.3.18) and (12.3.13). A brief description of the coordinate system (α,β,x,y) and of the relevant initial-value problem is given in Fig. 3.

Consider now a Kasner solution defined by a global spacetime metric of the general form (12.3.17):

$$g = -a\, dt^2 + bt^{2p_3} d\beta^2 + ct^{2p_1} dx^2 + dt^{2p_2} dy^2 \,, \tag{2.1}$$

where a, b are constants having the dimensions of (length)^2, c, d are dimensionless constants, t, β are dimensionless coordinates, and the exponents p_k, k=1,2,3 satisfy the Kasner relations Eqs. (12.3.19). Let p be a point *arbitrarily close* to the singularity t=0. We will investigate the behavior of the past null cone of this point p; in particular, we are interested in evaluating the transverse dimensions of the domain which this past null cone circumscribes in a spatial slice of the form $\{t=\tau\}$, $\tau>0$, or, more precisely, the transverse dimensions of the domain $C_\tau \equiv J^-(p) \cap \{t=\tau\}$, $\tau>0$. Let the integrals of motion $g(\gamma_*, \partial/\partial x)$, $g(\gamma_*, \partial/\partial y)$, and $g(\gamma_*, \partial/\partial\beta)$ (associated with the Killing vector fields $\partial/\partial x$, $\partial/\partial y$, and $\partial/\partial\beta$) along a past-directed null geodesic γ from p be denoted by

$$g(\gamma_*, \partial/\partial x) \equiv C_x \,, \qquad g(\gamma_*, \partial/\partial y) \equiv C_y \,,$$

$$g(\gamma_*, \partial/\partial\beta) \equiv C_\beta . \tag{2.2}$$

Then, a short computation shows that the total (coordinate) distances in the x, y, and β directions, traveled by the null geodesic γ as it reaches from the point p to the spacelike slice $C_\tau = \{t = \tau\}$, are given by

$$(\Delta x)_\tau = \int_0^\tau \frac{C_x}{c\, t^{2p_1}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, t^{2p_3}} + \dfrac{C_x^2}{c\, t^{2p_1}} + \dfrac{C_y^2}{d\, t^{2p_2}} \right]^{1/2}} dt \ , \tag{2.3a}$$

$$(\Delta y)_\tau = \int_0^\tau \frac{C_y}{d\, t^{2p_2}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, t^{2p_3}} + \dfrac{C_x^2}{c\, t^{2p_1}} + \dfrac{C_y^2}{d\, t^{2p_2}} \right]^{1/2}} dt \ , \tag{2.3b}$$

$$(\Delta \beta)_\tau = \int_0^\tau \frac{C_\beta}{b\, t^{2p_3}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, t^{2p_3}} + \dfrac{C_x^2}{c\, t^{2p_1}} + \dfrac{C_y^2}{d\, t^{2p_2}} \right]^{1/2}} dt \ . \tag{2.3c}$$

It easily follows from the above equations that the maximum transverse (coordinate) dimensions of the domain C_τ in the x, y, and β directions are

$$\max(\Delta x)_\tau = 2(\Delta x)_\tau (C_\beta = C_y = 0)$$

$$= 2\left(\frac{a}{c} \right)^{1/2} \frac{1}{1-p_1} \tau^{1-p_1} , \tag{2.4a}$$

$$\max(\Delta y)_\tau = 2(\Delta y)_\tau (C_\beta = C_x = 0)$$

$$= 2\left[\frac{a}{d}\right]^{1/2} \frac{1}{1-p_2}\tau^{1-p_2}, \tag{2.4b}$$

$$\max(\Delta\beta)_\tau = 2(\Delta\beta)_\tau (C_x = C_y = 0)$$

$$= 2\left[\frac{a}{b}\right]^{1/2} \frac{1}{1-p_3}\tau^{1-p_3}. \tag{2.4c}$$

Clearly, if the Kasner solution (2.1) is nondegenerate (Sec. III A of Ref. 12), i.e., if the exponents p_k are all different from 1 [or equivalently by Eqs. (12.3.19), are all different from zero], then the causal past $J^-(p)$ of the point p intersects any spatial slice $\{t=\tau\}$ in a *compact* region $\mathcal{C}_\tau$, whose coordinate dimensions approach finite limits given by Eqs. (2.4) as the point p approaches the singularity at $t=0$. Note [cf. Eq. (2.1)] that the proper physical size of $\mathcal{C}_\tau$ (in all directions) depends only on τ, a, and the exponents p_k; in other words, the *proper* dimensions of $\mathcal{C}_\tau$ are independent of b, c, and d (as they should be since the constants b, c, d can be absorbed into a redefinition of the coordinates β, x, y).

Now, we have proved in Sec. III C of Ref. 12 that any interval in β, across which the exponents $[p_1(\beta), p_2(\beta), p_3(\beta)]$ take the degenerate values (1,0,0) or (0,1,0), represents a Killing-Cauchy horizon at $\alpha=0$ which is unstable in the full nonlinear theory against small but generic perturbations of the initial data (12.2.49). We have also proved that according to the specific notion of genericity which we have introduced in Sec. III C of Ref. 12, generic initial data in the form (12.2.49) always produce all-embracing, spacelike spacetime singularities at $\alpha=0$; such singularities are always associated with Kasner exponent values $p_k(\beta)$ that are each different from 1 at all $\beta \in (-1,1)$ except for isolated, discrete points (Sec. III B of Ref. 12). Now consider a point p in a generic colliding plane-wave spacetime, arbitrarily close to the singularity $\alpha=0$ at a fixed spatial point β, and, for simplicity, at $x=y=0$. Consider the past-directed null geodesics generating the past null cone of p, and compute, along each such null geodesic, the integrals which represent the total transverse (coordinate) distances along the x and y directions traveled by the geodesic as it reaches from the point p to the initial surface $\mathcal{N}$. It is clear from Eq. (12.2.43), that the contributions to these integrals from t values that are bounded away from the singularity $t=0$ ($\equiv\alpha=0$) are finite. Moreover, the contributions from arbitrarily near $t=0$ are arbitrarily well approximated by the values (2.4) obtained by using the asymptotic form (12.3.17) of the spacetime metric. Since for a generic solution and a generic value of β (i.e., a value different from the isolated discrete values for which the Kasner exponents are degenerate) the exponents $p_k(\beta)$ are each different from 1, it follows from Eqs. (2.4) that the contributions to these integrals from near the singularity are also finite for all (generic) values of β in a generic colliding plane-wave spacetime. Therefore, we have proved the following result.

Lemma 1: The intersection $J^-(p)\cap\mathcal{N}$ between the initial surface $\mathcal{N}$ and the causal past $J^-(p)$ of any (generic) point p in the interaction region of a *generic*, colliding (parallel-polarized) plane-wave spacetime is a compact set, whose transverse ($\equiv x,y$) dimensions approach finite limits (i.e., remain bounded from above) as the point p approaches the singularity at $\alpha=0$.

In fact, when the point p has a β value sufficiently far away from the edge points $\beta=+1$ and $\beta=-1$ (e.g., for $-\frac{1}{2}<\beta<\frac{1}{2}$), β remains approximately constant along the null geodesics from p which extend farthest in the x, y directions; hence, the asymptotic limit (12.3.17) of the metric remains a good approximation along these geodesics. Therefore, for such a point p approaching $\alpha=0$ at, say, $-cae\frac{1}{2}<\beta<\frac{1}{2}$, we can estimate the limits of the maximum transverse (coordinate) dimensions of $J^{-}(p)\cap\mathcal{N}$ by the quantities

$$L_x(\beta)=\frac{4e^{-[\delta(\beta)+\mu(\beta)/2]/2}}{[1-\varepsilon(\beta)]^2}\sqrt{l_1 l_2}\,, \tag{2.5a}$$

and

$$L_y(\beta)=\frac{4e^{[\delta(\beta)-\mu(\beta)/2]/2}}{[1+\varepsilon(\beta)]^2}\sqrt{l_1 l_2}\,. \tag{2.5b}$$

These expressions are obtained from Eqs. (2.4) (i) by substituting the values for the constants a, c, and d that are found upon comparing Eq. (2.1) with Eq. (12.3.17), and (ii) by putting $\tau=1$ [since $\alpha=t=1$ on the collision plane $\{u=v=0\}$ in $\mathcal{N}$ (Fig. 3)].

The readers who are familiar with global methods will recognize that Lemma 1 above is closely related to a well-known more general result in relativity. This result states that for any point p contained in (the interior of) the domain of dependence[14] $D^{+}(S)$ of a partial Cauchy surface S, the subset $J^{-}(p)\cap S$ in S is compact (see, for example, the proposition 6.6.6 in Ref. 14). In view of this result, the above lemma can be rephrased in the following equivalent form, which states that as a (generic) point p in the interaction region falls off the "edge" of the colliding plane-wave spacetime (by approaching the singularity $\alpha=0$), it does not fall off the "edge" ($\equiv$the boundary) of the domain of dependence $D^{+}(\mathcal{N})$ of the initial surface $\mathcal{N}$.

Lemma 1 (second version): In a generic colliding (parallel-polarized) plane-wave spacetime, the singularity $\{\alpha=0\}$ represents a future c boundary,[14] whose (generic) "points" [which are "terminal indecomposable past sets" (Sec. 6.8 of Ref. 14)] intersect the initial surface $\mathcal{N}$ in subsets with compact closure. In other words, *unless* the colliding plane-wave solution possesses Killing-Cauchy horizons at $\{\alpha=0\}$ destroying its global hyperbolicity [which can only occur for "nongeneric" initial data (see Sec. III C of Ref. 12)], the (generic) points of the singularity $\{\alpha=0\}$ (when they are considered as points on the future causal boundary of the spacetime) can be regarded as part of the domain of dependence $D^+(\mathcal{N})$ of the initial surface $\mathcal{N}$.

Turn now to the discussion of the following general problem (Fig. 4): Let $(\mathcal{M},g)$ be a spacetime, and Σ be a partial Cauchy surface in $(\mathcal{M},g)$ on which gravitational initial data [whose maximal development gives the metric on $D^+(\Sigma)$] are posed. Replace these initial data on Σ with a new set of initial data, which coincide with the old data on some closed subset S of Σ. What kind of general statement can then be made about the spacetime which evolves from these new initial data? Note that, if the gravitational field equations were a *linear* hyperbolic system, then the subset S would have a *fixed* domain of dependence, bounded by a characteristic surface which is uniquely determined by the boundary ∂S of S, but which is independent of the solution and which therefore is indifferent to any change in the initial data on Σ. In that case, it is obvious that the new solution would coincide with the old solution throughout this domain of dependence $D^+(S)$ of S. However, for a *nonlinear* hyperbolic system such as the Einstein field equations, the characteristics *do* depend on the particular solution considered; consequently, the simple "causality" argument valid for a *linear* hyperbolic system *cannot* be applied to the Einstein equations in the same obvious way, and thus *cannot* guarantee that the new and old solutions coincide throughout $D^+(S)$.

Now, return to the specific situation depicted in Fig. 4, where the initial data on Σ are modified in the cross-hatched portions of the initial surface but are left intact throughout an open subset $\mathcal{U}$ of Σ which contains the closed subset S. It is not difficult to see, after thinking about Fig. 4 for a while, that the only way in which the solution on $D^+(S)$ can change is through the formation of a shock wave, which propagates on a *noncharacteristic* (spacelike) surface and penetrates into $D^+(S)$ from the outside, thereby changing the solution on its wake from the old to the new. In fact, such shock waves *do* develop from the solutions of many nonlinear hyperbolic systems (e.g., the Burgers's equation[15]) and are frequently encountered in fluid mechanics where they correspond to real physical phenomena. In general relativity however, shock waves of this kind *cannot exist* in vacuum gravitational fields. The reason is that any surface of discontinuity (shock wave) for a solution to the *vacuum* Einstein equations *must* be a null (characteristic) surface, provided that the metric is continuous (C^0) across the surface of discontinuity. (See any review on the Cauchy problem such as Refs. 16 and 17. In the nonvacuum case, the surfaces of discontinuity may be timelike; they can never be spacelike unless matter fields which violate the dominant energy condition[14] are present.)

One might worry about the demand that the metric be C^0. Suppose there were a surface, spacelike on the front and timelike or null on the back, across which the metric is not C^0. By a slight abuse of terminology, we will still call such a surface of discontinuity a "gravitational shock." We will argue now that such a shock wave must be regarded as a spacetime singularity. First of all, a shock wave of this form cannot be a pure coordinate artifact (i.e., cannot be eliminated by a coordinate transformation), since if it were, then the new coordinate system in which the metric is continuous would define the physically appropriate admissible local atlas, and the original coordinates would simply be discarded as inadmissible. On the other hand, we can assume that the spacetime curvature is bounded across and near the shock front, since otherwise the shock would obviously be a curvature singularity and our claim would be proved. Therefore, our shock front is a surface of discontinuity representing a jump in the spacetime metric (and corresponding to a distribution solution of the vacuum field equations); and there is either a topological (as with a conical singularity) or a geometrical (as with a delta-function singularity in the Riemann curvature) *obstacle* against elimination of the jump by coordinate transformations. The readers can now easily see (e.g., by considering the behavior of null geodesics near the shock), that at least some causal geodesics must terminate somewhere on a shock front of this kind; therefore, our shock wave is a spacetime singularity at least in the sense of geodesical incompleteness. On the other hand, if we now take into account the result[18] that a nonsmooth metric can be uniformly approximated by a sequence of smooth metrics *only if* it is at least continuous (C^0), it then becomes apparent, that just as a conical singularity physically models an infinitesimally thin line distribution of unbounded curvature, so also our shock wave with a discontinuous metric must actually model an infinitesimally thin sheet of unbounded spacetime curvature. Consequently, the spacetime metric must be at least C^0 (in other words, the intrinsic geometry of the surface of the shock must be the same as measured from either side) across any genuine purely gravitational shock wave; and therefore all gravitational shock waves must propagate on null (characteristic) surfaces.

We can now summarize the conclusions of the last two paragraphs in the form of a lemma.

Lemma 2: Let $(\mathcal{M},g)$ be a spacetime and Σ be a partial Cauchy surface in $(\mathcal{M},g)$ on which gravitational initial data [whose development gives the metric on $D^+(\Sigma)$] are posed. Let $S\subset\Sigma$ be a closed subset, and $\mathcal{U}\subset\Sigma$ be an open subset containing S (Fig. 4). Suppose that the initial data on Σ are replaced with a new set of initial data which coincide with the original data throughout $\mathcal{U}$. Then, unless a spacetime singularity forms and penetrates into $D^+(S)$ from outside $D^+(S)$, the new solution coincides with the old solution throughout $D^+(S)$, where $D^+(S)$ denotes the domain of dependence of S with respect to the original metric and coincides with the domain of dependence of S with respect to the new metric.

Recall now Lemma 1 above, where we have demonstrated that for a generic colliding plane-wave spacetime any (generic) point p arbitrarily near the singularity $\alpha=0$ is contained in the domain of dependence of the initial surface $\mathcal{N}$; in fact, p is contained in the domain of dependence of a compact region in $\mathcal{N}$, whose transverse dimensions approach the finite limits estimated by Eqs. (2.5) as the point p approaches the singularity. Introducing the quantity L defined by

$$L \equiv \inf_{-\frac{1}{2}<\beta<+\frac{1}{2}} \max\,[L_x(\beta), L_y(\beta)]\ , \tag{2.6}$$

it is then clear that if $\mathcal{C}$ is a domain in the initial surface of the form $\mathcal{C}=\{\,|x|\leq L_T, |y|\leq L_T\,\}$, where L_T is much larger than L above, then there will be some points arbitrarily near the singularity $\alpha=0$ (at least in some neighborhood of the location $x=y=0$), which will be contained in the domain of dependence of the region $\mathcal{C}$. In other words, invoking the notions which we used in the second version of Lemma 1 above, the domain of dependence of such a (sufficiently large) subset $\mathcal{C}$ of $\mathcal{N}$ will contain some "points" on the singularity $\{\alpha=0\}$, at least in a neighborhood of the central plane $\{x=y=0\}$. Combining this result with Lemma 2 above, it is clear that we have proved the following singularity theorem.

Theorem: Let the initial data for two colliding almost-plane gravitational waves be identical to the initial data for two colliding (parallel-polarized) exact plane waves throughout a region $\mathcal{C}$ in the initial surface of the form $\mathcal{C}=\{\,|x|\leq L_T, |y|\leq L_T\,\}$. Let the corresponding initial data for this plane-symmetric portion be generic so that the maximal development of the complete plane-symmetric data produces all-embracing spacetime singularities at α=0. Let these plane-symmetric data be represented by the functions $V(r,1)$ and $V(1,s)$ as in Eq. (12.2.49). Construct, from these functions, the quantity $\varepsilon(\beta)$ defined by Eq. (12.3.13), and the quantity L defined by Eq. (2.6). Then, if $L_T \gg L$, the evolution of the almost-plane-symmetric data produces spacetime singularities; i.e., the colliding almost-plane waves create spacetime singularities.

Clearly, the local structure of the singularities which are guaranteed to exist by the above theorem will be precisely the same as the structure of the plane-symmetric singularities; i.e., these singularities will be locally of inhomogeneous Kasner type. Unfortunately however, the above theorem does not say anything about the *global* structure of the overall singularity created by the colliding almost-plane waves; this global structure (about which we have stated a number of conjectures in Sec. I and Fig. 2 of this paper) is the key to understanding the status of cosmic censorship in gravitational almost-plane-wave collisions.

Assume now that the colliding almost-plane waves both have amplitudes small compared to unity: $h_1 \ll 1$, $h_2 \ll 1$. In this paragraph, we will estimate the value of the lower bound L for L_T [Eq. (2.6)] in terms of the relevant parameters for these colliding waves. (For the demonstration in a specific example of some of the general arguments we use below, see the Appendix.) By Eqs. (12.3.12) and (12.3.13), the quantities $\varepsilon(\beta)$ and $\delta(\beta)$ are small compared to 1 in this case. Therefore, if we can choose the initial point (u_0,v_0) in such a way that the quantity $\mu(\beta)$ is also smaller than or of order unity, then by Eqs. (2.5) and (2.6) we could conclude that $L \sim \sqrt{l_1 l_2}$. In fact, such a choice is possible: if we fix u_0 and v_0 such that $\lambda_1 \ll u_0 \ll f_1$, and $\lambda_2 \ll v_0 \ll f_2$ (where f_1, f_2 are the first focal lengths and λ_1, λ_2 are the typical wavelengths of the colliding waves), then the point (u_0,v_0) belongs to a domain in the interaction region where (i) gravity is weak (since $u_0 \ll f_1$ and $v_0 \ll f_2$), so that U and the constant additive terms in Eq. (12.2.44b) are small compared to unity, and (ii) the integration path in Eq. (12.2.44b) is sufficiently far away (since $u_0 \gg \lambda_1$ and $v_0 \gg \lambda_2$) from the coordinate singularities on the initial null surfaces $\{u=0\}$ and $\{v=0\}$ (Sec. II B of Ref. 12), so that the contribution to $\mu(\beta)$ from the integrand in Eq. (12.2.44b) (which diverges towards the coordinate singularities on these initial null surfaces) is of order unity [Eq. (12.3.4)]. Moreover, with this choice for (u_0,v_0), Eqs. (12.2.39a) give

$$l_1 \sim f_1, \qquad l_2 \sim f_2, \tag{2.7}$$

and since $\mu(\beta)$ is of order 1 by the arguments above, Eqs. (2.7) yield, when combined with Eqs. (2.5) and (2.6), the following expression for the lower bound L, valid for colliding almost-plane waves with small amplitudes:

$$L \sim \sqrt{f_1 f_2}\,. \tag{2.8}$$

When combined with the theorem above, Eq. (2.8) constitutes a proof of the singularity criterion (1.2) in the special case where the colliding almost-plane waves are exactly plane-symmetric throughout their central regions.

Finally, we note that we have not given a proof of the criterion (1.2) for the more general case where the colliding waves are only "almost" (but not exactly) plane-symmetric throughout their central regions. As the readers might also have guessed in view of the theorem we have proved in this section, such a proof would immediately follow if we could prove that the solution on $D^+(\mathcal{N})$ depends *uniformly continuously* on the initial data on $\mathcal{N}$. We do not, however, expect this last statement to be valid, although we know, as a result of recent work by the author,[19] that the the former statement is true, namely that the singularity result of the theorem remains valid in the more general case where the colliding waves are not exactly plane symmetric over any region, but are sufficiently close to being plane symmetric across their central regions. (Note, in this context, that the well-known general theorems on the Cauchy problem which assert the continuous dependence of solutions on the initial data (such as the Cauchy stability theorem, see, e.g., Sec. 7.6 of Ref. 14) are valid with respect to the *compact-open* topology [i.e., the open topology based on convergence on *compact* subsets of $D^+(\mathcal{N})$], and not with respect to the *open* topology [i.e., the open topology based on (uniform) convergence on $D^+(\mathcal{N})$] on the spaces of all initial data on $\mathcal{N}$ and all four-metrics on $D^+(\mathcal{N})$.) More specifically, although the conclusion of the theorem on the *existence* of singularities remains intact under small perturbations of the initial data,[19] we expect the local structure of the singularities to change radically when the exact plane symmetry of the central region is broken; in fact, it is presumable that in this more general case the local inhomogeneous Kasner structure of the singularity would be replaced with a more general "generic" singularity structure (of the mixmaster type if one believes the Belinsky-Khalatnikov-Lifshitz[13] analysis of the structure of generic singularities, or of some more general type if one agrees with the critics of this analysis[20]), and therefore that with respect to the (uniform) *open topology* on the space of all metrics the solution would have changed *discontinuously* in response to the small change in the initial data.

ACKNOWLEDGMENTS

The author wishes to thank Kip S. Thorne for suggesting the specific example discussed in the Appendix. He also acknowledges partial financial support from the National Science Foundation under Grant No. AST 85-14911.

APPENDIX: A SPECIFIC EXAMPLE FOR A COLLIDING PLANE-WAVE SOLUTION WITH "PHYSICALLY REASONABLE" INITIAL DATA

Our example is described by the initial data [Eq. (12.2.15)]

$$V_1(u)=\frac{hu}{\lambda}\,,\qquad 0\le u\le 2\pi\lambda$$

$$=2\pi h\,,\qquad u>2\pi\lambda\,, \tag{A1a}$$

$$V_2(v)=\frac{hv}{\lambda}\,,\qquad 0\le v\le 2\pi\lambda$$

$$=2\pi h\,,\qquad v>2\pi\lambda\,, \tag{A1b}$$

where $h \ll 1$ is a constant representing the dimensionless amplitude of the colliding waves. [This is an approximation to the wave form that arises, in some observers' directions, from the near encounter of two stars, i.e., "gravitational bremsstrahlung" (see, e.g., Ref. 21, especially Fig. 2).] It follows from Eqs. (12.2.16)–(12.2.18) that

$$U_1(u)=-2\ln\left(1-\frac{h^2}{8\lambda^2}u^2\right)+O(h^4)\,,\qquad 0\le u\le 2\pi\lambda\,,$$

$$=-2\ln\left[\frac{\pi h^2}{2}\left(\frac{2}{\pi h^2}+\pi-\frac{u}{\lambda}\right)\right]+O(h^4)\,,\qquad u>2\pi\lambda\,, \tag{A2a}$$

$$U_2(v)=-2\ln\left(1-\frac{h^2}{8\lambda^2}v^2\right)+O(h^4)\,,\qquad 0\le v\le 2\pi\lambda\,,$$

$$=-2\ln\left[\frac{\pi h^2}{2}\left(\frac{2}{\pi h^2}+\pi-\frac{v}{\lambda}\right)\right]+O(h^4)\,,\qquad v>2\pi\lambda\,, \tag{A2b}$$

for the initial data (A1). It is easily seen from Eqs. (A2) that

$$f_1=f_2\equiv f=\frac{2\lambda}{\pi h^2}+\frac{4\pi}{3}\lambda+O(h^2)\approx\frac{2\lambda}{\pi h^2}\,. \tag{A3}$$

We choose the initial normalization point (u_0,v_0) as prescribed in the discussion at the end of Sec. II above; i.e., we take $\lambda\ll u_0\ll f$ and $\lambda\ll v_0\ll f$. Then, combining Eqs. (A2) with Eqs. (12.2.39a) gives

$$l_1\approx\frac{\lambda}{2\pi h^2}\sim f_1\,,\qquad l_2\approx\frac{\lambda}{2\pi h^2}\sim f_2\,, \tag{A4}$$

as we have argued in Sec. II. Consider the region of weak gravity for the colliding plane-wave spacetime determined by the data (A1). This region is described, e.g., by $\alpha>\frac{1}{2}$. We identify four distinct subsets of this weak gravity region as follows: $\{\alpha>\frac{1}{2},\, 0\le u, v\le 2\pi\lambda\}\equiv I$, $\{\alpha>\frac{1}{2},\, 0\le u\le 2\pi\lambda,\, v>2\pi\lambda\}\equiv II$, $\{\alpha>\frac{1}{2},\, u>2\pi\lambda,\, 0\le v\le 2\pi\lambda\}\equiv III$, and $\{\alpha>\frac{1}{2},\, u>2\pi\lambda,\, v>2\pi\lambda\}\equiv IV$. (These regions I, II, III, IV, written in italics, are not to be confused with the similarly named regions of Fig. 3.) Then, Eqs. (12.2.26) and (12.2.27) give, in the weak gravity region,

$$\begin{aligned}
\alpha&=1-\frac{h^2}{4\lambda^2}(u^2+v^2)+O(h^4) &&\text{in } I\,,\\
&=1-\frac{h^2}{4\lambda^2}u^2+\pi h^2\left[\pi-\frac{v}{\lambda}\right]+O(h^4) &&\text{in } II\,,\\
&=1-\frac{h^2}{4\lambda^2}v^2+\pi h^2\left[\pi-\frac{u}{\lambda}\right]+O(h^4) &&\text{in } III\,,\\
&=1+\pi h^2\left[2\pi-\frac{u+v}{\lambda}\right]+O(h^4) &&\text{in } IV\,,
\end{aligned} \tag{A5a}$$

and

$$\beta = \frac{h^2}{4\lambda^2}(u^2 - v^2) + O(h^4) \qquad \text{in } I\ ,$$

$$= h^2\left[\frac{u^2}{4\lambda^2} + \pi\left(\pi - \frac{v}{\lambda}\right)\right] + O(h^4) \qquad \text{in } II\ ,$$

$$= -h^2\left[\frac{v^2}{4\lambda^2} + \pi\left(\pi - \frac{u}{\lambda}\right)\right] + O(h^4) \qquad \text{in } III\ ,$$

$$= \frac{\pi h^2}{\lambda}(u - v) + O(h^4) \qquad \text{in } IV\ . \tag{A5b}$$

Combining Eqs. (A5) with the Eqs. (12.2.46), and using Eqs. (12.2.52) and (12.2.53) with the initial data (A1), we obtain the initial data in the form of Eq. (12.2.49):

$$V(r,1) = [2(1-r)]^{1/2} + O(h^4)\ , \qquad 1 \geq r \geq 1 - 2\pi^2 h^2$$

$$= 2\pi h\ , \qquad r < 1 - 2\pi^2 h^2\ , \tag{A6a}$$

$$V(1,s) = [2(1-s)]^{1/2} + O(h^4)\ , \qquad 1 \geq s \geq 1 - 2\pi^2 h^2$$

$$= 2\pi h\ , \qquad s < 1 - 2\pi^2 h^2\ . \tag{A6b}$$

It is easily seen from Eqs. (12.3.12) and (12.3.13) and Eqs. (A6) above, that provided $-\frac{1}{2} < \beta < \frac{1}{2}$, the quantities $\varepsilon(\beta)$ and $\delta(\beta)$ are both $\sim O(h)$. Therefore, our first assertion in Sec. II, namely, that with the above choice (A1) for the initial data we have $\varepsilon(\beta) \sim O(h)$ and $e^{\delta(\beta)} \sim 1$, is demonstrated for the example (A1). To demonstrate the remaining assertion, namely that with the above choice for the initial point (u_0, v_0) the quantity $\mu(\beta) \sim 1$, it is sufficient to show that the function $Q(\alpha,\beta)$ is ~ 1 at the point

(u_0, v_0) [since the subsequent contributions to $\mu(\beta)$ will be of order $O(h^2)$, see Eqs. (12.2.44b) and (12.3.4)]. Now note that in the weak gravity region the spacetime geometry is well approximated by $-du\,dv + dx^2 + dy^2$; i.e., by the flat Minkowski metric. By substituting the inverse relations for Eqs. (A5) above into this flat metric, and then comparing the resulting metric in the (α, β, x, y) coordinates with the general expression (12.2.43), we find

$$Q(\alpha,\beta) \approx -2\ln\left[2\pi^2 h^2 \left(\frac{\alpha}{(1-\alpha)^2-\beta^2}\right)^{1/2}\right] \qquad \text{in } I \ ,$$

$$\approx -2\ln\left[\sqrt{2}\pi h \left(\frac{\alpha}{1-\alpha+\beta}\right)^{1/2}\right] \qquad \text{in } II \ ,$$

$$\approx -2\ln\left[\sqrt{2}\pi h \left(\frac{\alpha}{1-\alpha-\beta}\right)^{1/2}\right] \qquad \text{in } III \ ,$$

$$\approx -\ln\alpha \qquad \text{in } IV \ . \tag{A7}$$

Now, since the normalization point (u_0, v_0) is chosen in such a way that $\lambda \ll u_0, v_0 \ll f$, it lies not only in the weak gravity region (as $u_0, v_0 \ll f$), but also in the subset IV of this region (as $u_0, v_0 \gg \lambda$) where α is neither very close to zero nor very close to unity; i.e., where we can assume $\alpha \sim \frac{1}{2}$. Thus, by Eqs. (A7) above, $Q(\alpha_0, \beta_0) \sim -\ln\alpha(u_0, v_0) \sim 1$, as was to be demonstrated.

REFERENCES FOR CHAPTER 6

[1] U. Yurtsever, Phys. Rev. D **37**, 2803 (1988) (Chapter 3).

[2] K. Khan and R. Penrose, Nature (London) **229**, 185 (1971).

[3] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

[4] Y. Nutku and M. Halil, Phys. Rev. Lett. **39**, 1379 (1977).

[5] F. J. Tipler, Phys. Rev. D **22**, 2929 (1980).

[6] R. A. Matzner and F. J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[7] S. Chandrasekhar and B. C. Xanthopoulos, Proc. R. Soc. London **A** (to be published).

[8] U. Yurtsever, Phys. Rev. D **36**, 1662 (1987) (Chapter 2).

[9] U. Yurtsever, (in preparation).

[10] U. Yurtsever, Phys. Rev. D **37**, 2790 (1988) (Chapter 4).

[11] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[12] U. Yurtsever, Phys. Rev. D **38**, 1706 (1988) (Chapter 5).

[13] E. Kasner, Am. J. Math. **43**, 217 (1921); E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz. Nauk **80**, 391 (1964) [Sov. Phys. Uspekhi **6**, 495 (1963)]; C. W. Misner, Phys. Rev. Lett. **22**, 1071 (1969); V. A. Belinsky and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **57**, 2163 (1969) [Sov. Phys. JETP 30, 1174 (1970)]; V. A. Belinsky, I. M. Khalatnikov, and E. M. Lifshitz, Adv. Phys. **19**, 525 (1970).

[14] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).

[15] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley Interscience, New York, 1974).

[16] Y. Choquet-Bruhat and J. York, in *Einstein Centenary Volume*, edited by A. Held and P. Bergmann (Plenum, New York, 1979).

[17] A. Fischer and J. E. Marsden, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[18] J. York (private communication).

[19] U. Yurtsever, Physical Review D, submitted (also available as Caltech preprint GRP-184) (Chapter 7).

[20] J. D. Barrow and F. J. Tipler, Phys. Rep. **56**, 371 (1979); J. D. Barrow, ibid. **85**, 1 (1982).

[21] S. J. Kovacks and K. S. Thorne, Astrophys. J. **229**, 62 (1978).

FIGURE CAPTIONS FOR CHAPTER 6

FIG. 1. Spacetime diagrams for the conjectured outcome of an almost-plane-wave collision in the first parameter regime. The parameters are the transverse sizes $(L_T)_i$, typical amplitudes h_i, typical reduced wavelengths $\lambda\!\!\!{}^{-}_i \equiv \lambda_i/2\pi$, thicknesses a_i, and focal lengths $f_i \sim \lambda\!\!\!{}^{-}_i{}^2/a_i h_i^2$ $(i=1,2)$ of the waves 1 and 2. The first parameter regime is given by $(L_T)_1 \gg (L_T)_2$ and $h_1 \gg h_2$; in this regime, the interaction of the colliding waves can be treated successfully by regarding the weaker wave 2 as a small disturbance propagating on a background spacetime predominantly determined by the stronger wave 1. (a) If $(L_T)_2 \ll \sqrt{\lambda\!\!\!{}^{-}_2 f_1}(h_1/h_2)^{1/4}$, the almost-plane wave 2 will be focused, because of the focusing effect of wave 1, into a waist near the focal plane of finite, minimum size; but it will then diffract and disperse, since linear diffraction effects [which become important at an affine distance $\sim (L_T)_2^2/\lambda\!\!\!{}^{-}_2$] ultimately dominate the nonlinear self-gravity effects (which become important near the focal plane, at an affine distance $\sim f_1$). (b) If $(L_T)_2 \gg \sqrt{\lambda\!\!\!{}^{-}_2 f_1}(h_1/h_2)^{1/4}$ (and if wave 1 is sufficiently anastigmatic, i.e., if its two focal lengths are approximately equal), then wave 2 will be focused by wave 1 so strongly that it forms a singularity surrounded by a horizon, and the end result is a black hole flying away from wave 1. Just before the final moments of its collapse into the black hole, near the spacetime region enclosed by a small square in the figure (b) above (where gravity is still weak), the total energy-momentum associated with the almost-plane wave 2 can be represented by a timelike vector $\vec{t}$. We expect that the invariant mass associated with $\vec{t}$ is a good estimate for the mass of the final black hole, and that the black hole forms at rest with respect to the rest frame defined by $\vec{t}$.

FIG. 2. Spacetime diagram for the conjectured outcome of an almost-plane-wave collision in the second parameter regime where $(L_T)_1 \sim (L_T)_2 \equiv L_T$ and $h_1 \sim h_2$. If

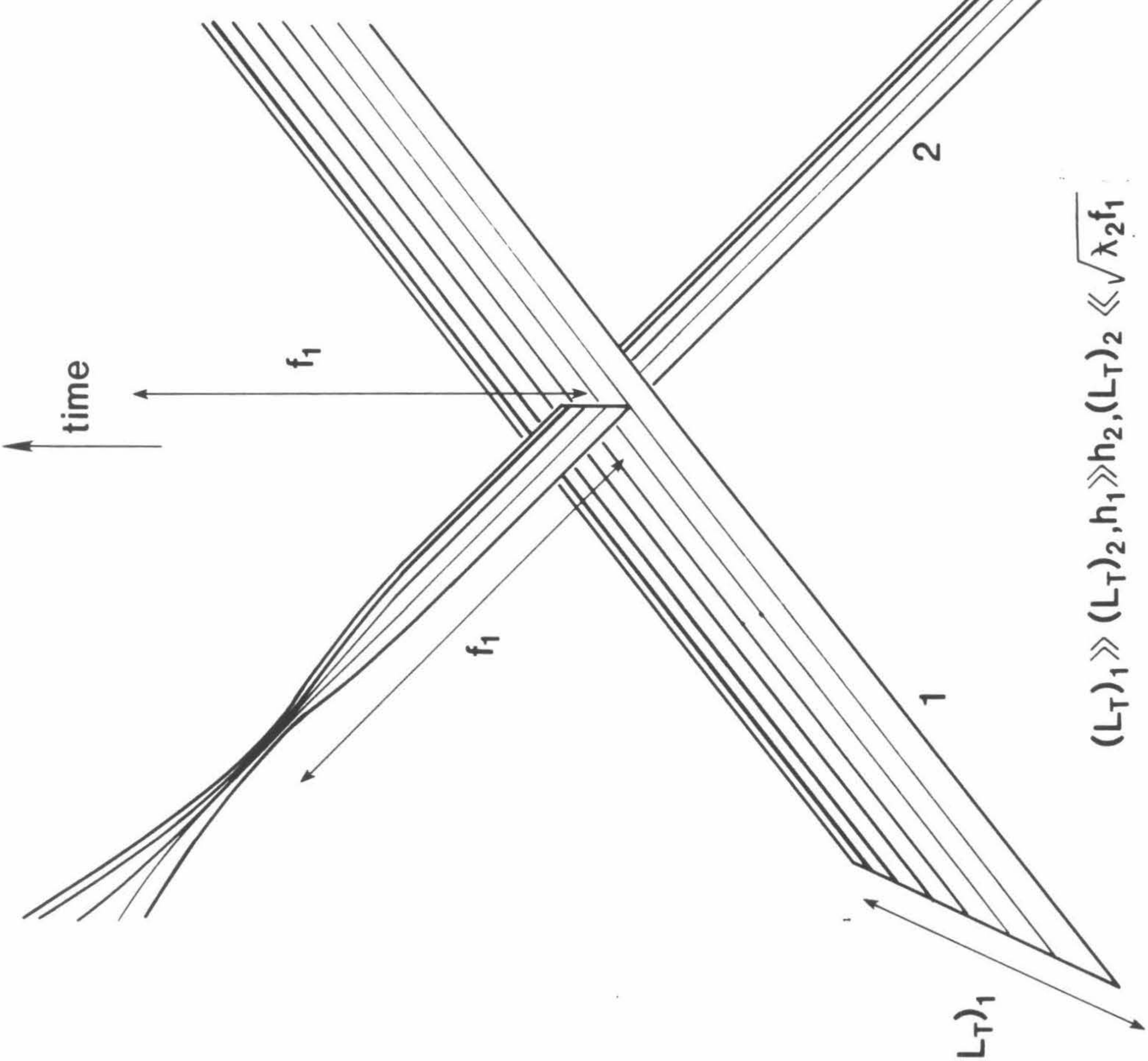
$L_T \gg \sqrt{f_1 f_2} \equiv f$, a horizon forms around the two colliding waves shortly before their collision, and the collision produces a black hole that is at rest with respect to the reference frame in which $f_1 \sim f_2 \sim f$. The mass of this black hole can be estimated as $M \sim \sqrt{M_1 M_2}$, where M_1 and M_2 are the total mass energy $M_i \sim (L_T)_i{}^2 a_i h_i{}^2/\lambda\!\!\!^-{}_i{}^2 \sim (L_T)^2/f_i$ in the two colliding waves. Note that the black hole forms when $M \sim (L_T)^2/f$ is much larger than L_T, i.e., when the (invariant) mass energy contained in the colliding waves is much larger than their transverse sizes.

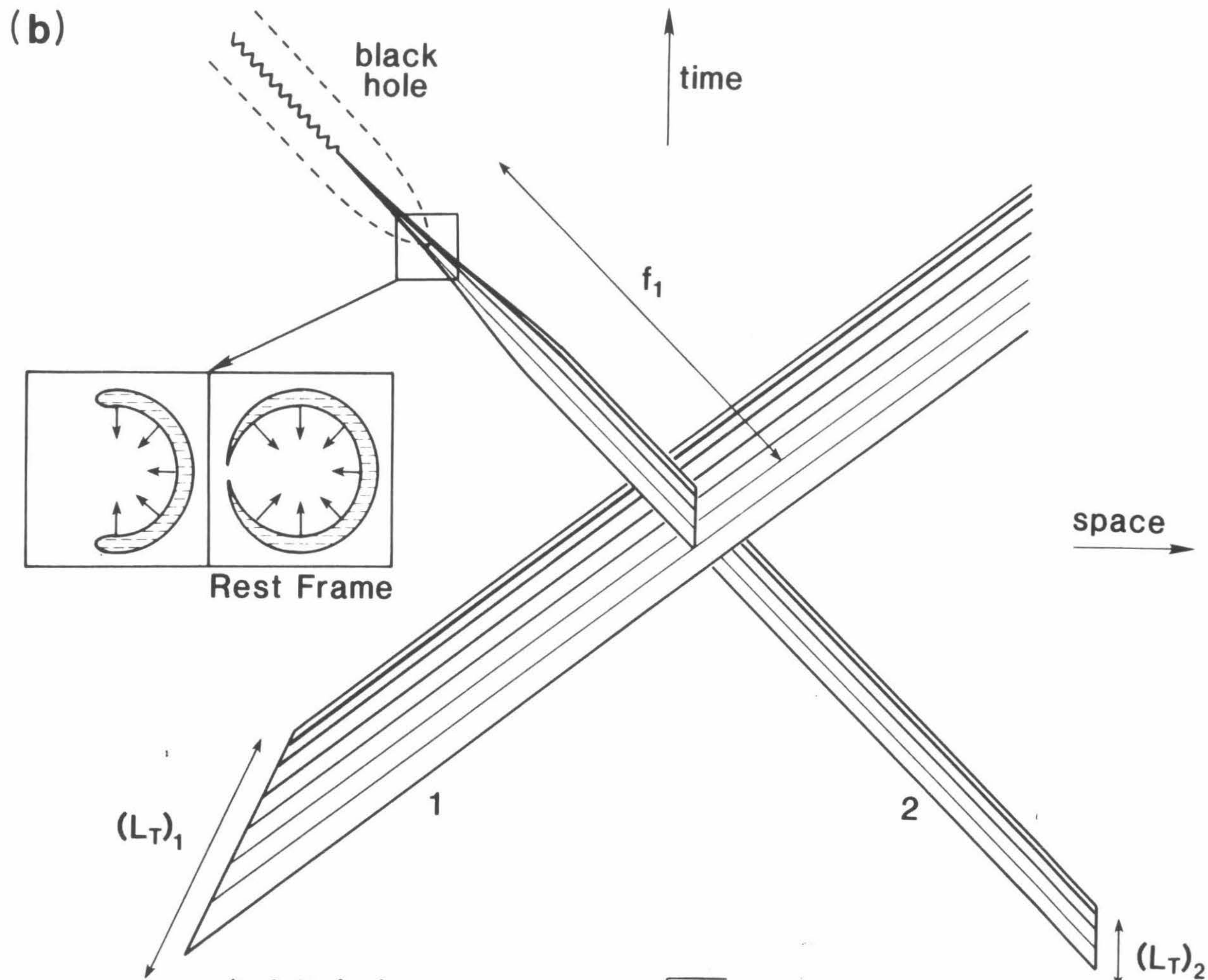
FIG. 3. The two-dimensional geometry of the characteristic initial-value problem for colliding plane waves. The null surfaces $\mathcal{N}_1=\{u=0\}$ and $\mathcal{N}_2=\{v=0\}$ are the past wave fronts of the incoming plane waves 1 and 2. Initial data corresponding to waves 1 and 2 are posed, respectively, on the upper portions of the surfaces $\mathcal{N}_1=\{v=0\}$ and $\mathcal{N}_2=\{u=0\}$ that are adjacent to the interaction region I. The geometry in the region IV is flat, and the geometry in regions II and III is given by the metric describing the incoming waves 1 and 2, respectively. The geometry of the interaction region I is uniquely determined by the solution of the above initial-value problem. The directions in which the various lines of constant coordinates u, v, α, β, r, and s run are also indicated, along with the descriptions of the initial null surfaces in these different coordinate systems.

FIG. 4. The geometry of the general initial-value problem which is the subject of Lemma 2 in the text. The original initial data on the partial Cauchy surface Σ are modified in the cross-hatched portions of the initial surface, but are left intact throughout an open subset $\mathcal{U}$ of Σ which contains a closed subset S. In the new solution which evolves from these modified initial data, the geometry on the domain of

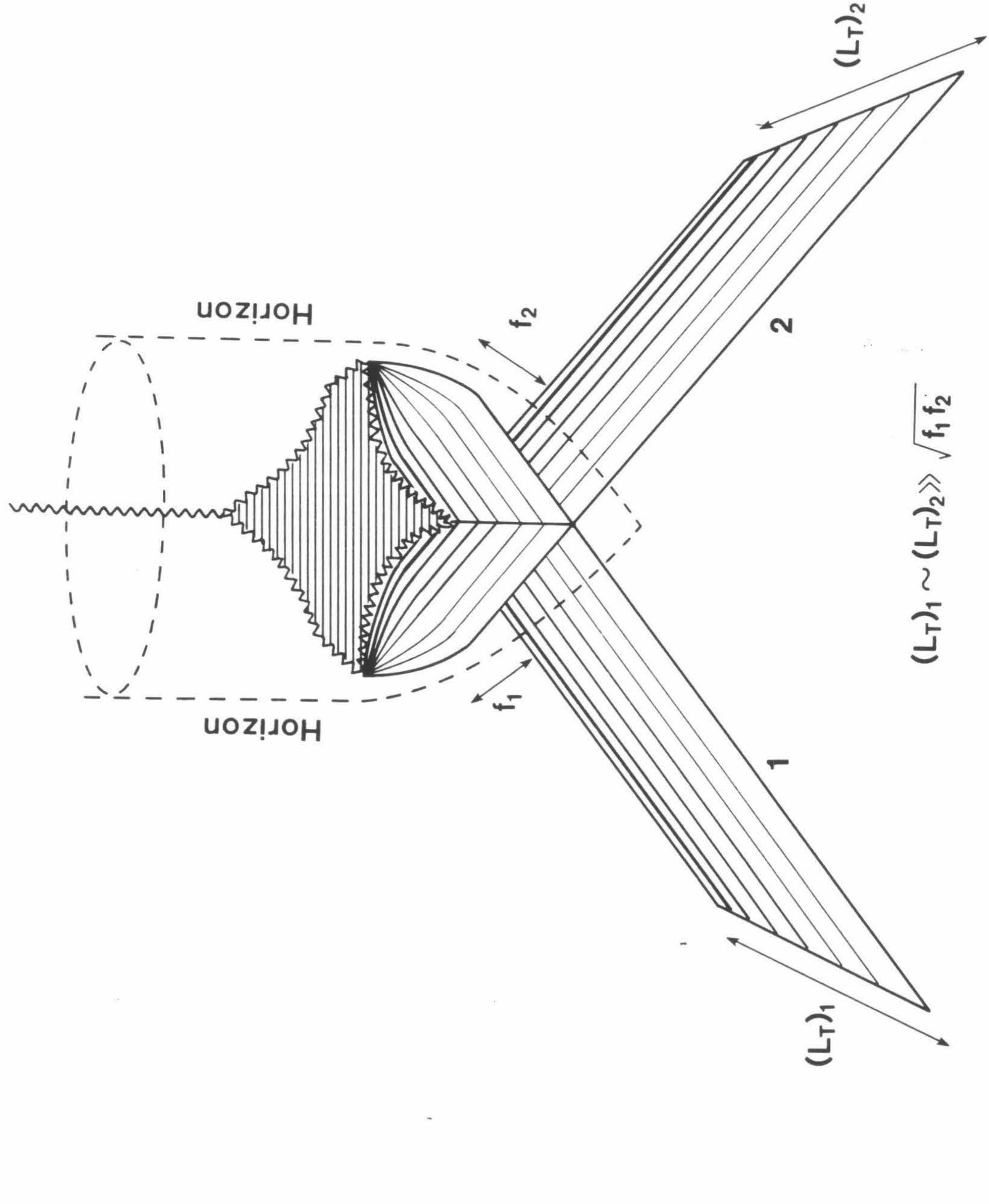
dependence $D^+(S)$ of S can change only if a gravitational shock wave forms outside and propagates into $D^+(S)$, transforming the solution on its wake from the original to the modified. However, gravitational shock waves always propagate on null (characteristic) surfaces; and the domain of dependence $D^+(S)$ (with respect to the original metric) is bounded by a null surface [the unique ingoing null surface $H^+(S)$ through ∂S]. Thus, no gravitational shocks can propagate into $D^+(S)$ from the outside, and consequently, the original geometry remains invariant throughout the domain of dependence $D^+(S)$ (which thereby coincides with the domain of dependence of S in the new solution).

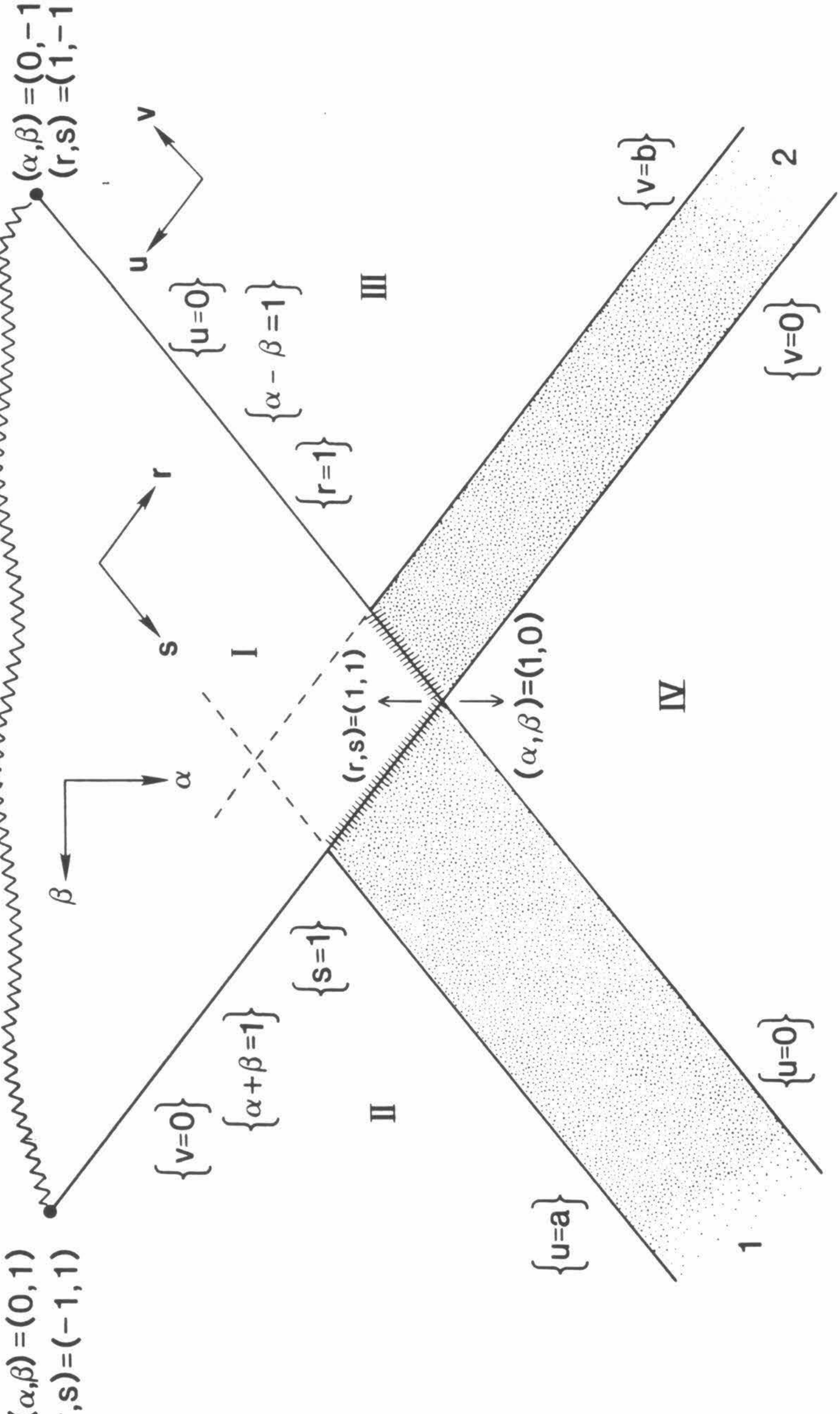
(a)

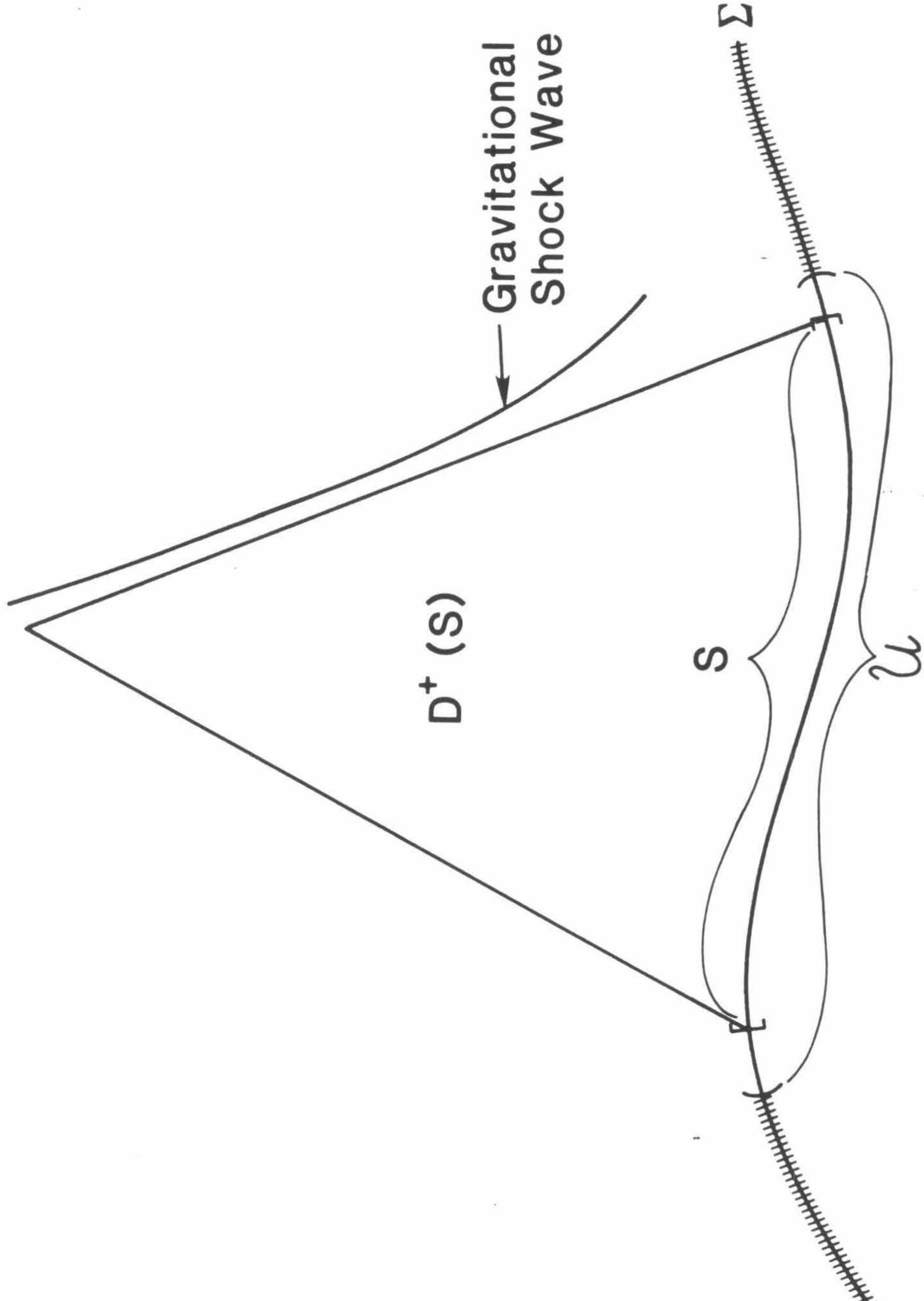

time
space
f_1
f_1
1
2
$(L_T)_1$
$(L_T)_2$
$(L_T)_1 \gg (L_T)_2, h_1 \gg h_2, (L_T)_2 \ll \sqrt{\lambda_2 f_1}$



$(L_T)_1 \gg (L_T)_2, h_1 \gg h_2, (L_T)_2 \gg \sqrt{\lambda_2 f_1}$, Wave 1 anastigmatic


Horizon
Horizon
f_1
f_2
1
2
$(L_T)_1$
$(L_T)_2$
$(L_T)_1 \sim (L_T)_2 \gg \sqrt{f_1 f_2}$


{α = 0} ≡ {r+s = 1}
(α,β) = (0,1)
(r,s) = (−1,1)
(α,β) = (0,−1)
(r,s) = (1,−1)
β
α
s
r
u
v
I
II
III
IV
{u=0}
{α − β = 1}
{r=1}
{s=1}
{α+β=1}
{v=0}
{u=a}
{v=b}
{v=0}
{u=0}
(r,s)=(1,1)
(α,β)=(1,0)
1
2


Gravitational Shock Wave
D⁺ (S)
S
𝒰
Σ

CHAPTER 7

Singularities and Horizons in the Collisions of Gravitational Waves

Submitted to Physical Review D.

ABSTRACT

It is well known that when gravitational plane waves propagating and colliding in an otherwise flat background interact, they produce singularities. In this paper we explore the structure of the singularities produced in the collisions of arbitrarily-polarized gravitational plane waves, and we consider the problem of whether (or under what conditions) singularities can be produced in the collisions of *almost-plane* gravitational waves with finite but very large transverse sizes. First we analyze the asymptotic structure of a general arbitrarily-polarized colliding plane-wave spacetime near its singularity. We show that the metric is asymptotic to a generalized inhomogeneous-Kasner solution as the singularity is approached. In general, the asymptotic Kasner axes as well as the asymptotic Kasner exponents along the singularity are functions of the spatial coordinate that runs tangentially to the singularity in the non-plane-symmetric direction. It becomes clear that for specific values of these asymptotic Kasner exponents and axes the curvature singularity created by the colliding waves degenerates to a coordinate singularity, and that a nonsingular Killing-Cauchy horizon is thereby obtained. Our analysis proves that these horizons are unstable in the full nonlinear theory against small but generic plane-symmetric perturbations of the initial data, and that in a very precise and rigorous sense, "generic" initial data for colliding arbitrarily-polarized plane waves always produce all-embracing, spacelike curvature singularities without Killing-Cauchy horizons. Next we turn to the problem of colliding almost-plane gravitational waves, and by combining the results that we obtain in this paper and in other previous papers with the Hawking-Penrose singularity theorem and the Cauchy stability theorem, we prove that if the initial data for two colliding almost-plane waves are sufficiently close to being exactly plane symmetric across a sufficiently large but bounded region of the initial surface, then

their collision must produce spacetime singularities. Although our analysis proves the existence of these singularities rigorously, it does not give any information about either their global structure (e.g. whether they are hidden behind an event horizon) or their local asymptotic behavior (e.g. whether they are of Belinsky-Khalatnikov-Lifshitz generic-mixmaster type).

I. INTRODUCTION AND OVERVIEW

With a short letter[1] published in Nature in 1971, K. Khan and R. Penrose announced their discovery of a new exact solution to the vacuum Einstein field equations; it described the interaction between two impulsive, plane-symmetric gravitational waves, propagating and colliding in an otherwise flat background spacetime. The collision was followed by a spacetime region in which the nonlinear interaction between the waves generated a gravitational field qualitatively different from the linear superposition of the two incoming fields. In fact, the spacetime curvature generated by the collision increased without bound along all timelike worldlines in the interaction region, and it ultimately diverged to form a spacetime singularity where the observers' worldlines reached and terminated in finite proper time. Despite its complicated local and global structure,[2] the physical interpretation of this solution was simple: Each of the two colliding plane waves generated a spacetime geometry in its wake which acted like an infinite, perfectly converging lens,[3] focusing any radiation field which passed through the plane wave while propagating in the opposite direction. When the two plane waves collided, each of them was thus perfectly focused by the other's background geometry; diffraction effects were prevented from counterbalancing this perfect focusing by the global exact-plane-symmetry of spacetime. As a result, while they propagated through the interaction region the amplitudes of the colliding waves grew without bound and ultimately diverged, creating a spacelike curvature singularity which bounded the interaction region in all future directions.

In the nearly two decades since the discovery of the Khan-Penrose[1] solution (and of the simultaneous discovery of other similar solutions by Szekeres[4]), the progress in the search for exact solutions describing colliding plane waves has been phenomenal, with significant contributions by many workers. Recent research in this field has particularly benefited from the carrying-over of the inverse-scattering techniques for generating stationary axisymmetric solutions (one spacelike and one timelike Killing vectors) of Einstein's equations to the problem of generating plane-symmetric solutions (two commuting spacelike Killing vectors). For a brief description of the history of these developments and a (necessarily incomplete) list of references, we refer readers to Refs. 5 and 6 (especially Sec. I of Ref. 5 and Sec. I of Ref. 6) and to the references cited therein.

In our view the greatest significance of the problem of colliding gravitational waves lies not with those aspects of it that are peculiar to specific exact solutions, but rather with its potential to provide insight into some of the broader issues in general relativity (such as cosmic censorship, structure of singularities, ...) which arise naturally in studying the dynamics of fully nonlinear gravitational fields. From this point of view, gravitational-wave collisions can be considered as the vacuum analogues of gravitational collapse, and as such they provide a framework in which issues like cosmic censorship can be discussed without the undue complications of a specifically chosen nonzero stress-energy tensor. In fact, we contend that among all the issues raised by the last two-decades of exact-solutions research on colliding plane waves the following two are the most important; and that owing to their inherent generality these issues are not likely to be completely resolved by work on exact solutions alone:

On the one hand, thanks to the work of Chandrasekhar and Xanthopoulos[7] who first discovered this phenomenon, we now know that colliding plane waves do not always create spacelike curvature singularities with a global structure similar to the singularity of the Khan-Penrose solution: for some choices of the incoming plane waves, their collision produces a nonsingular Killing-Cauchy horizon[8] at the points where ordinarily one would expect curvature singularities to form. The spacetime can then be extended smoothly across this horizon (in nonunique ways) to obtain several inequivalent, maximal solutions, which all evolve from the same initial data posed by the incoming, colliding plane waves (breakdown of predictability). It is therefore of fundamental importance to determine (i) under what conditions on the initial data (the incoming plane waves) the collision creates singularities and under what conditions it creates horizons, (ii) what are the local structures of the singularities and horizons thus created, and (iii) whether "generic" initial data (with respect to some appropriate notion of genericity) always produce "pure" spacetime singularities without Killing-Cauchy horizons, i.e., whether any breakdowns in global predictability can occur in "generic" gravitational plane-wave collisions. The issue here is of the *structure* of singularities produced by colliding plane waves.

On the other hand, it is natural to raise the issue of whether (or under what conditions) spacetime singularities can be produced by the collisions of gravitational waves which are not exactly plane-symmetric, but which have finite but very large transverse "spatial" sizes; i.e., by the collisions of *almost-plane* gravitational waves. This second issue is of the *existence* (and possibly also the structure) of singularities created in the collisions of almost-plane gravitational waves.

In a series of two papers published previously in this journal (Refs. 6 and 9), we attempted to resolve the above issues in the special case where the colliding waves had parallel constant-linear polarizations. Thus, in Ref. 6 we showed that the asymptotic structure of a colliding parallel-polarized plane-wave spacetime near its singularity can be completely and explicitly determined in terms of the initial data posed by the incoming waves. Our analysis proved that although Killing-Cauchy horizons can be produced in the collisions of parallel-polarized plane waves, these horizons are unstable in the full nonlinear theory against small but generic plane-symmetric perturbations of the initial data, and that in a very precise sense, "generic" initial data always produce all-embracing, spacelike curvature singularities without Killing-Cauchy horizons. In Ref. 9, we analyzed the collision between two almost-plane gravitational waves whose initial data across a bounded region of the initial surface were identical with the initial data posed by colliding parallel-polarized exactly plane waves, but fell off in an arbitrary way at larger transverse distances. We proved that if this bounded region of exact plane symmetry in the initial surface is sufficiently large, then the collision between the almost-plane waves is guaranteed to produce a spacetime singularity with the same local structure as in an exact plane-wave collision.

The work described in the present paper is a continuation of the work reported in Refs. 6 and 9. The main results of this paper are (i) the generalization of the results of Refs. 6 and 9 to the case where the polarizations of the colliding waves are entirely arbitrary (i.e. neither parallel nor constant-linear), and (ii) the proof of a much stronger version of the singularity theorem of Ref. 9; specifically, that if the initial data for two colliding almost-plane waves are *sufficiently close* to being exactly plane symmetric across a sufficiently large but bounded region of the initial surface, then their collision must produce spacetime singularities. Sections II–III and Sec. IV A below describe the above-mentioned generalization of the results of Ref. 6 and Ref. 9, respectively, whereas Sec. IV B is devoted to the new singularity theorem. The five appendices at the end of the paper deal with a number of issues of a more technical nature that are raised during the course of the analyses in Secs. II–IV. We note, however, that these Appendices (especially Appendices A, C, and D) contain a large amount of information, some of which might be useful in future research on questions that are left unresolved in this paper. We feel that any serious reading of the paper *must* include at least the three Appendices A, C, and D.

The more precise plan of this paper is as follows:

In Sec. II A, we give a very brief review of Szekeres's[4] formulation of the field equations and the characteristic initial-value problem for colliding arbitrarily-polarized plane waves, in the (u, v, x, y) coordinate system which we call "Rosen-type" and which is tuned to the plane symmetry of the spacetime. This formulation is entirely analogous to the corresponding formulation for the parallel-polarized case which we have discussed in Sec. II A of Ref. 6. Consequently, here we only present the essential facts and formulas that will be needed in later sections, and refer the reader to Sec. II A of Ref. 6 for the details of their derivation and meaning. In this section and throughout the paper, we try to maintain as much parallelism as possible between our presentation here and the presentation in Refs. 6 and 9. For this reason, readers may find it helpful to reference these two previous papers[6,9] while reading the present paper.

In Sec. II B, we perform a coordinate transformation to a new (α, β, x, y) coordinate system, in which the field equations and the initial-value problem associated with them take simpler forms. Again the construction and the properties of this new coordinate system are straightforward generalizations of the construction and properties of the (α, β) coordinates discussed in Sec. II B of Ref 6. However, while the field equations for colliding parallel-polarized plane waves (Sec. II B of Ref. 9) reduced in the (α, β) coordinates to a single *linear* hyperbolic equation for which an explicit Riemann function could be found,[4,6] in the general case the simplification achieved by this coordinate change, though substantial, is not as great: The field equations in the (α, β) coordinates reduce to a system of *nonlinear*, coupled hyperbolic partial differential equations (PDE) for two functions which represent the dimensionless amplitudes for the two independent modes of polarization. Although at present it seems unlikely (because of their high nonlinearity) that an explicit general solution (Riemann function) can be found for these equations, in Appendix C we discuss some interesting and suggestive aspects of this particular system of nonlinear PDE which might later prove useful in the search for such a general solution. A further disturbing consequence of this fundamental nonlinearity in the field equations for colliding nonparallel-polarized plane waves is that the global existence and uniqueness of their solutions may not be guaranteed. In the parallel-polarized case, it is guaranteed by the linearity of the single nontrivial field equation that there exists a unique, global solution defined throughout the domain of dependence of the initial surface, i.e., throughout the entire interaction region up to the "singularity" $\{\alpha=0\}$ at which either spacetime singularities or Killing-Cauchy horizons form (Secs. II B and III A of Ref. 6). In contrast, the field equations in the nonparallel-polarized case are nonlinear, and it is well known that solutions of nonlinear hyperbolic PDE do not in general exist globally. This raises the possibility that solutions of the field equations might break down at points which lie *within* the interaction region *before* the "singular" surface $\{\alpha=0\}$, and consequently the possibility that colliding nonparallel-polarized plane waves might create spacetime singularities in the region where $\alpha>0$; such singularities, if present, would not be treatable by analyzing the asymptotic structure of spacetime near $\alpha=0$. Fortunately, however, a careful analysis which we undertake in Appendix A shows that thanks to some very special properties possessed by the field equations, the global existence and uniqueness of their solutions can be proved despite the presence of strong nonlinearities. Therefore, the singularities and horizons created by colliding plane waves always lie on or beyond the surface $\{\alpha=0\}$.

Our discussions in Sec. II B and in Appendix A bring us to the analysis of the asymptotic structure of spacetime near $\alpha=0$. Relying on the results of Appendix B which show that as $\alpha\to 0$ the spatial-derivative terms in the field equations are asymptotically negligible compared to the α-derivative terms, we begin Sec. III A by studying the ordinary differential equations that are obtained by eliminating the spatial β-derivative terms from the field equations; this allows us to determine the asymptotic behavior of the metric functions near $\alpha=0$. We show that the spacetime metric asymptotically approaches a generalized inhomogeneous-Kasner[10] solution as α approaches zero, where the time coordinate t of the asymptotic Kasner spacetime is monotonically related to α, and the Kasner singularity at $t=0$ corresponds to the singularity at $\alpha=0$. We call this asymptotic inhomogeneous-Kasner structure "generalized" because unlike the parallel-polarized case in which the asymptotic Kasner exponents were associated with the fixed set of axes $\{x, y\}$ throughout the singularity (Sec. III A of Ref. 6), here in general the asymptotic Kasner axes are linear combinations of x, y and they vary across the singularity as functions of the spatial coordinate β. Since we do not have a general solution for the field equations in the nonparallel-polarized case, in contrast to Sec. III A of Ref. 6 we cannot in general relate the asymptotic Kasner exponents and/or axes along the singularity to the initial data posed along the wavefronts of the incoming, colliding plane waves. (See however Appendix C where one such relation is obtained in a special case.) As in Ref. 6, in general these asymptotic Kasner exponents as well as the asymptotic Kasner axes depend on β, the spacelike coordinate running along the nontrivial spatial (z) direction in the spacetime.

We begin Sec. III B with a discussion of Tipler's theorem,[11,12] which proves that in any vacuum, nonflat plane-symmetric spacetime there must exist either a spacetime singularity (where null geodesics terminate) or a Killing-Cauchy horizon (where the strict plane symmetry of spacetime breaks down). We note that the content of Tipler's theorem is made particularly transparent by our analysis of the asymptotic structure of colliding plane-wave spacetimes: On the one hand, it becomes clear from our discussion in Sec. III A that the asymptotic Kasner exponents and axes (throughout a connected interval in the spatial coordinate β) may take on the values associated with a degenerate Kasner solution. Since a degenerate Kasner spacetime is flat and possesses a Killing-Cauchy horizon at $t=0$ instead of a singularity, it follows that when the asymptotic Kasner exponents for the colliding plane-wave metric are degenerate a nonsingular Killing-Cauchy horizon forms at $\alpha=0$ across which spacetime can be extended smoothly. On the other hand, it is easily seen from the expressions of the Newman-Penrose curvature quantities in the (α,β) coordinates that if the asymptotic Kasner exponents are *nondegenerate*, then $\alpha=0$ is a curvature singularity. Next we observe that when a Killing-Cauchy horizon forms at $\alpha=0$, the spacetime can be extended through it in infinitely many different ways; the geometry beyond the horizon cannot be determined from the initial data posed by the incoming, colliding plane waves. We then briefly recall our earlier work in Ref. 8, where we proved general theorems stating the instability of such Killing-Cauchy horizons in any plane-symmetric spacetime against generic, plane-symmetric perturbations. For the special case of the Killing-Cauchy horizons which occur in collisions of parallel-polarized plane waves, our discussions in Sec. III C of Ref. 6 proved that in fact these instabilities render the set of "all" horizon-producing initial data "nongeneric" with respect to a very precise notion of nongenericity. More specifically, our analysis in Ref. 6 proved that the subset of all initial data which produce at least one connected Killing-Cauchy horizon larger than Planck size is nongeneric within the set of all colliding parallel-polarized plane-wave initial data. Correspondingly, by making use of the discussions in Appendices A and B, we prove in Sec. III B the generalization of this result (with the same notion of genericity as in Ref. 6) to the case of colliding arbitrarily-polarized plane waves. In addition, by introducing a more sophisticated notion of genericity which we describe in greater detail in Appendix D, we prove that the subset of *all* horizon-producing initial data (and not just the subset of those data which produce horizons larger than Planck size) is nongeneric within the set of all initial data for colliding plane waves. We also discuss why we believe that our topological notion of genericity (described in Appendix D) is more appropriate in general relativity than other possible "probabilistic" notions based on measure theory.

In Sec. IV A, using the conclusions we obtained in the previous sections, we prove the generalization of the singularity result that was proved for parallel-polarized colliding almost-plane waves in Sec. II of Ref. 9 to the case of colliding almost-plane waves with arbitrary polarizations. More specifically, we prove that if the initial data posed by two colliding almost-plane gravitational waves are (i) identical with the initial data posed by two colliding exactly plane waves (with arbitrary polarizations) across a bounded but sufficiently large region of the initial surface, and (ii) fall off in an arbitrary way (consistent with the constraint equations) at larger transverse distances, then the collision between the almost-plane waves is guaranteed to produce a spacetime singularity with the same local structure as in an exact plane-wave collision.

In Sec. IV B, we combine the Hawking-Penrose singularity theorem (Ref. 13 and Sec. 8.2 of Ref. 14), the Cauchy stability theorem,[15] and a lemma about the null cones in a nondegenerate Kasner spacetime which we discuss in Appendix E, to prove that the conclusion of the singularity theorem of Sec. IV A about the *existence* of singularities remains valid when the colliding almost-plane waves are not *exactly* plane-symmetric over any region, but are only *approximately* plane-symmetric across their central regions. In other words, we prove that if the initial data for two colliding almost-plane waves are *sufficiently close* to being exactly plane symmetric across a sufficiently large but bounded region of the initial surface, then their collision must produce spacetime singularities. Although our analysis proves the existence of these singularities rigorously, it does not give any information about either their global structure (e.g. whether they are hidden behind an event horizon) or their local asymptotic behavior (e.g. whether they are of Belinsky-Khalatnikov-Lifshitz[10] generic-mixmaster type).

Our notation and other conventions throughout this paper are the same as in Refs. 6 and 9. Equation numbers that refer to equations of Refs. 6 or 9 will be denoted by a prefix "6" or "9"; for example, Eq. (6.3.13) and Eq. (9.2.6) refer, respectively, to Eq. (3.13) of Ref. 6 and Eq. (2.6) of Ref. 9.

As in our previous papers,[6,9] here we are concerned exclusively with the collisions of purely gravitational (vacuum) waves. Whether the conclusions of Secs. II and III in this paper remain valid in the presence of matter fields coupled to the colliding plane waves is an interesting and unexplored question.

II. FIELD EQUATIONS FOR COLLIDING GRAVITATIONAL PLANE WAVES

A. Formulation of the problem in the Rosen-type (u, v, x, y) coordinate system

In any plane-symmetric spacetime (see Sec. III B of Ref. 12, or Sec. II of Ref. 8 for a careful definition of plane symmetry), there exists a canonical null tetrad[16] whose construction is described in Sec. III B of Ref. 12. In this null tetrad, which we call the standard tetrad, $\vec{l}$ and $\vec{n}$ are tangent to the two null geodesic congruences everywhere orthogonal to the plane-symmetry generating Killing vector fields $\vec{\xi}_1$ and

$\vec{\xi}_2$, and $\vec{m}$ and its complex conjugate are linear combinations of the $\vec{\xi}_i$, $i=1,2$. As is shown by Szekeres in Ref. 4 and discussed briefly in Sec. II A of Ref. 6, the special geometry of a colliding plane-wave spacetime allows us to find a local coordinate system (u,v,x,y) in which $\vec{\xi}_i = \partial/\partial x^i$ $[(x^1, x^2) \equiv (x,y)]$, and in which the standard tetrad can be brought into the form

$$\vec{l} = 2e^M \frac{\partial}{\partial u} , \quad \vec{n} = \frac{\partial}{\partial v} ,$$

$$\vec{m} = N_1 \frac{\partial}{\partial x} + N_2 \frac{\partial}{\partial y} , \tag{2.1}$$

with

$$N_1 = \frac{1}{\sqrt{2}} e^{(U-V)/2} \sqrt{\cosh W} \exp\{ \tfrac{1}{2} i \, [\sin^{-1}(\tanh W)] \} ,$$

$$N_2 = \frac{i}{\sqrt{2}} e^{(U+V)/2} \sqrt{\cosh W} \exp\{ -\tfrac{1}{2} i \, [\sin^{-1}(\tanh W)] \} , \tag{2.2}$$

where M, U, V, and W are real functions of u and v only. (Notice the slight phase difference between our choice for N_1 and N_2 here and that in Sec. II A of Ref.6 [Eqs. (6.2.4)]. The only equations in this paper that are affected by this discrepancy are the expressions for the Newman-Penrose curvature quantities [Eqs. (2.12) below] which differ from the corresponding expressions in Ref. 6 [Eqs. (6.2.19)] by factors of 2 or i.) The null tetrad given by Eqs. (2.1) and (2.2) gives rise to the metric

$$g = -e^{-M} du dv + e^{-U} [\cosh W (e^V dx^2 + e^{-V} dy^2) - 2 \sinh W \, dx dy] . \tag{2.3}$$

Thus, the functions $V(u,v)$ and $W(u,v)$ represent the dimensionless amplitudes of the two independent polarization modes in the gravitational radiation field (2.3).

The vacuum Einstein field equations for the metric (2.3) can be written in the form[4]

$$2(U_{,uu}+M_{,u}U_{,u})-U_{,u}{}^2-V_{,u}{}^2\cosh^2 W-W_{,u}{}^2=0\,, \tag{2.4a}$$

$$2(U_{,vv}+M_{,v}U_{,v})-U_{,v}{}^2-V_{,v}{}^2\cosh^2 W-W_{,v}{}^2=0\,, \tag{2.4b}$$

$$U_{,uv}-U_{,u}U_{,v}=0\,, \tag{2.4c}$$

$$V_{,uv}-\frac{1}{2}(U_{,u}V_{,v}+U_{,v}V_{,u})+(V_{,u}W_{,v}+V_{,v}W_{,u})\tanh W=0\,, \tag{2.4d}$$

$$W_{,uv}-\frac{1}{2}(U_{,u}W_{,v}+U_{,v}W_{,u})-V_{,u}V_{,v}\sinh W\cosh W=0\,, \tag{2.4e}$$

where the integrability condition for the first two equations is satisfied by virtue of the last three, and yields the remaining field equation

$$M_{,uv}-\frac{1}{2}(V_{,u}V_{,v}\cosh^2 W-U_{,u}U_{,v})-\frac{1}{2}W_{,u}W_{,v}=0\,. \tag{2.5}$$

It is sufficient to solve Eqs. (2.4c)–(2.4e) first and to obtain M by quadrature from the first two equations (2.4a) and (2.4b) afterward, since Eq. (2.5) as well as the integrability condition for Eqs. (2.4a) and (2.4b) are automatically satisfied as a result of Eqs. (2.4c)–(2.4e).

The initial-value problem associated with the field equations (2.4) and (2.5) is best formulated in terms of initial data posed on null (characteristic) surfaces. A natural choice for the initial characteristic surface is the surface made up of the two intersecting null hyperplanes which form the initial wave fronts of the incoming plane waves, and which, by a readjustment of the null coordinates u and v if necessary, can be arranged to be the surfaces $\{u=0\}$ and $\{v=0\}$. The geometry of the resulting characteristic initial-value problem is depicted in Fig. 1. The initial data supplied by the plane wave propagating in the v direction (to the right in Fig. 1) is posed on the $u \geq 0$ portion of the surface $\{v=0\}$, and the initial data supplied by the plane wave propagating in the u direction (to the left in Fig. 1) is posed on the $v \geq 0$ portion of the surface $\{u=0\}$. In region IV, which represents the spacetime before the passage of either plane wave, the geometry is flat and all metric coefficients M, U, V, and W vanish identically. Now recall our discussions in Sec. II A of Ref. 6 about the gauge freedom in the choice of the (u,v,x,y) coordinate system, and about how this freedom manifests itself in the choice of initial data on the characteristic initial surface $\{u=0\} \cup \{v=0\}$. For exactly the same reasons as described in those discussions, here as well as in Ref. 6 the choice of the initial data $\{M(u=0,v), M(u,v=0)\}$ for the metric function M is completely arbitrary. As we did in Ref. 6, we will fix this gauge freedom once and for all by posing our initial data so that

$$M(u=0,v) = M(u,v=0) \equiv 0 . \tag{2.6}$$

After making this gauge choice, it becomes clear from the field equations (2.4) that the initial data on $\{u=0\} \cup \{v=0\}$ are completely determined by only the four freely-specifiable functions $V_1(u) \equiv V(u,v=0)$, $W_1(u) \equiv W(u,v=0)$, $V_2(v) \equiv V(u=0,v)$, and $W_2(v) \equiv W(u=0,v)$. In other words, the initial data consist of

$$\{V_1(u), W_1(u), V_2(v), W_2(v)\} , \tag{2.7}$$

where $V_1(u)$, $W_1(u)$ and $V_2(v)$, $W_2(v)$ are C^1 (and piecewise C^2) functions for $u \geq 0$ and $v \geq 0$, respectively, which are freely specified except for the initial conditions $V_1(u=0) = W_1(u=0) = V_2(v=0) = W_2(v=0) = 0$. The remaining functions $U_1(u) \equiv U(u,v=0)$ and $U_2(v) \equiv U(u=0,v)$ which specify the initial values of the metric function $U(u,v)$ are uniquely determined, by the initial data (2.7), through the constraint equations [cf. Eqs. (2.4a) and (2.4b)]

$$2U_{1,uu} - U_{1,u}{}^2 = V_{1,u}{}^2 \cosh^2 W_1 + W_{1,u}{}^2 , \tag{2.8a}$$

$$2U_{2,vv} - U_{2,v}{}^2 = V_{2,v}{}^2 \cosh^2 W_2 + W_{2,v}{}^2 , \tag{2.8b}$$

with the initial conditions $U_1(u=0)=U_2(v=0)=0$, $U_{1,u}(u=0)=U_{2,v}(v=0)=0$. Note that, if we define two new functions $f(u)$ and $g(v)$ by

$$f(u) \equiv e^{-U_1(u)/2} , \qquad g(v) \equiv e^{-U_2(v)/2} , \tag{2.9}$$

we can express Eqs. (2.8) in the form of "focusing" equations:

$$\frac{f_{,uu}}{f} = -\frac{1}{4}(V_{1,u}{}^2 \cosh^2 W_1 + W_{1,u}{}^2) , \tag{2.10a}$$

$$\frac{g_{,vv}}{g} = -\frac{1}{4}(V_{2,v}{}^2 \cosh^2 W_2 + W_{2,v}{}^2) , \tag{2.10b}$$

with the initial conditions $f(0)=g(0)=1$, $f'(0)=g'(0)=0$. It immediately follows from Eqs. (2.10) and (2.9) that

$$f(u)<1 , \quad f'(u)<0 \quad \forall\, u>0 , \qquad g(v)<1 , \quad g'(v)<0 \quad \forall\, v>0 , \tag{2.11a}$$

$$U_1(u)>0 , \quad U'_1(u)>0 \quad \forall\, u>0 , \quad U_2(v)>0 , \quad U'_2(v)>0 \quad \forall\, v>0, \tag{2.11b}$$

as long as the initial data (2.7) are nontrivial for both incoming waves [i.e., as long as neither $V_1(u)$ and $W_1(u)$ nor $V_2(v)$ and $W_2(v)$ are identically zero], and as long as the initial surfaces $\{u=0\}$ and $\{v=0\}$ correspond to the true initial wave fronts of the colliding waves [i.e., as long as either $V_1(u)\neq0$ or $W_1(u)\neq0$ and either $V_2(v)\neq0$ or $W_2(v)\neq0$ for *all* sufficiently small but positive u and v], both of which conditions we will always assume throughout this paper.

In Secs. III A and III B below, when we discuss the asymptotic structure of the colliding plane-wave spacetime described by Eqs. (2.1)–(2.3), we will need the following equations that express the Newman-Penrose[16] curvature quantities in the null tetrad (2.1) and (2.2) in terms of the metric coefficients M, U, V, and W; the derivation of these equations can be found in Ref. 4:

$$\Psi_0 = -2e^{2M}\{[2V_{,u}W_{,u}\sinh W - V_{,u}(U_{,u}-M_{,u})\cosh W + V_{,uu}\cosh W]$$

$$-i[W_{,uu}-(U_{,u}-M_{,u})W_{,u}-V_{,u}{}^2\sinh W\cosh W]\}\ , \tag{2.12a}$$

$$\Psi_2 = e^{M}[M_{,uv}-i(V_{,v}W_{,u}-V_{,u}W_{,v})\cosh W]\ , \tag{2.12b}$$

$$\Psi_4 = -\frac{1}{2}\{[2V_{,v}W_{,v}\sinh W - V_{,v}(U_{,v}-M_{,v})\cosh W + V_{,vv}\cosh W]$$

$$+i[W_{,vv}-(U_{,v}-M_{,v})W_{,v}-V_{,v}{}^2\sinh W\cosh W]\}\ . \tag{2.12c}$$

$$\Psi_1 = \Psi_3 = 0\ . \tag{2.12d}$$

B. Field equations in the (α,β) coordinates

We now construct a new coordinate system in which the field equations and the initial-value problem associated with them take simpler forms. The construction and the properties of this new coordinate system are straightforward generalizations of the construction and properties of the (α,β) coordinates discussed in Sec. II B of Ref. 6. Consequently, here we will be somewhat concise in our presentation and refer the reader to Ref. 6 for details.

Consider the interaction region (region I in Fig. 1) where $u \geq 0$ and $v \geq 0$. This region is the *domain of dependence*[14] of the characteristic initial surface $\{u=0\} \cup \{v=0\}$, on which the initial-value problem defined by Eqs. (2.4) and (2.6)–(2.8) is to be solved. Consider the field equation (2.4c) in the interaction region. It follows from this equation that if we define

$$\alpha(u,v) \equiv e^{-U(u,v)} , \tag{2.13}$$

then, throughout the interaction region, $\alpha(u,v)$ satisfies

$$\alpha_{,uv} = 0 , \tag{2.14}$$

the flat-space wave equation in two dimensions. Equation (2.14) suggests that we define the complementary variable, $\beta(u,v)$, such that

$$\beta_{,u} = -\alpha_{,u} , \qquad \beta_{,v} = \alpha_{,v} . \tag{2.15}$$

Clearly, the integrability condition for Eqs. (2.15) is satisfied by virtue of Eq. (2.14). The initial-value problem for $\alpha(u,v)$ is easily solved, and when combined with Eq. (2.15), it yields the expressions[6]

$$\alpha(u,v) = e^{-U_1(u)} + e^{-U_2(v)} - 1 , \tag{2.16a}$$

$$\beta(u,v) = e^{-U_2(v)} - e^{-U_1(u)} , \tag{2.16b}$$

which complete the construction of the new variables (α,β). To see that these variables actually define a new coordinate system, note that by Eqs. (2.16)

$$d\alpha \wedge d\beta = 2U_1'(u)U_2'(v)e^{-[U_1(u)+U_2(v)]} du \wedge dv . \tag{2.17}$$

Therefore, from Eqs. (2.11b), Eq. (2.17), and the inverse function theorem,[17] it follows that the functions (α,β,x,y) constitute a regular coordinate system wherever the coordinate system (u,v,x,y) is regular in the interior of the interaction region, where $u>0$, $v>0$. On the other hand, by Eqs. (2.17) and the initial conditions for Eqs. (2.8), the coordinates α,β are singular along the initial null surfaces $\{u=0\}$ and $\{v=0\}$. In other words, the singularities of the coordinate system (α,β,x,y) consist of the singularities of the (u,v,x,y) coordinates (when there are any), and the singularity along the initial characteristic surface $\{u=0\}\cup\{v=0\}$. Since the only place in the interaction region where the coordinates (u,v,x,y) can develop singularities is the "surface" $\{\alpha=0\}$ (see Sec. III A), it follows that the coordinate system (α,β,x,y) covers the domain of dependence of the initial surface $\{u=0\}\cup\{v=0\}$ regularly except for the singularities on $\{u=0\}$ and $\{v=0\}$.

The coordinates (α,β,x,y) enjoy a number of properties which make them useful in studying the field equations for colliding plane waves. We will not list these properties here as they are discussed in detail in Sec. II B of Ref. 6; instead, we will proceed directly with the analysis of the initial-value problem (2.4) and (2.6)–(2.8) in the new coordinate system (α,β,x,y). First we note the transformation rules

$$\partial_u = \alpha_{,u}(\partial_\alpha - \partial_\beta)\,, \tag{2.18a}$$

$$\partial_v = \alpha_{,v}(\partial_\alpha + \partial_\beta)\,, \tag{2.18b}$$

and their inverses

$$\partial_\alpha = \frac{1}{2}\left(\frac{1}{\alpha_{,u}}\partial_u + \frac{1}{\alpha_{,v}}\partial_v\right)\,, \tag{2.19a}$$

$$\partial_\beta = \frac{1}{2}\left(\frac{1}{\alpha_{,v}}\partial_v - \frac{1}{\alpha_{,u}}\partial_u\right)\,, \tag{2.19b}$$

which are derived using Eq. (2.15). [For our notation, see the explanations following Eqs. (6.2.31) and (6.2.34) in Ref. 6.] A short computation involving Eqs. (2.18) and (2.19) now gives

$$-dudv = \frac{1}{4\,\alpha_{,u}\,\alpha_{,v}}(-d\,\alpha^2 + d\,\beta^2)\,. \tag{2.20}$$

When inserted into Eq. (2.3) and combined with Eq. (2.13), Eq. (2.20) yields the expression

$$g = \frac{e^{-M}}{4\alpha^2 U_{,u} U_{,v}}(-d\,\alpha^2 + d\,\beta^2)$$

$$+\,\alpha\,[\,\cosh W\,(e^{V} dx^2 + e^{-V} dy^2) - 2\sinh W\; dxdy\,] \tag{2.21}$$

for the spacetime metric, which is valid throughout the interaction region (region I in Fig. 1). Next, another short calculation using Eqs. (2.18) and (2.19) together with Eq. (2.14) gives

$$\partial_\alpha{}^2 - \partial_\beta{}^2 = \frac{1}{\alpha_{,u}\,\alpha_{,v}}\partial_u\,\partial_v\;. \tag{2.22}$$

Combining Eq. (2.22) with the field equations (2.4d) and (2.4e) and using Eqs. (2.18) and (2.19), we obtain the field equations satisfied by the amplitudes V and W in the $(\alpha,\beta,x\,,y\,)$ coordinate system:

$$V_{,\alpha\alpha} + \frac{1}{\alpha}V_{,\alpha} - V_{,\beta\beta} = 2\,(V_{,\beta}W_{,\beta} - V_{,\alpha}W_{,\alpha}\,)\tanh W\;, \tag{2.23a}$$

$$W_{,\alpha\alpha} + \frac{1}{\alpha}W_{,\alpha} - W_{,\beta\beta} = (V_{,\alpha}{}^2 - V_{,\beta}{}^2\,)\sinh W \cosh W\;. \tag{2.23b}$$

To obtain the remaining field equations, we proceed as follows: First we note that after defining a new function P by

$$e^{P} \equiv 4c\, e^{M}\, U_{,u} U_{,v} \,, \tag{2.24}$$

where c is an arbitrary constant having the dimensions of (length)^2 [we will fix c later with our normalization condition Eq. (2.28)], we can rewrite the field equations (2.4a) and (2.4b) in the form

$$2P_{,u} = 3U_{,u} + \frac{1}{U_{,u}}(V_{,u}{}^2 \cosh^2 W + W_{,u}{}^2) \,, \tag{2.25a}$$

$$2P_{,v} = 3U_{,v} + \frac{1}{U_{,v}}(V_{,v}{}^2 \cosh^2 W + W_{,v}{}^2) \,. \tag{2.25b}$$

Combining Eqs. (2.25) with Eqs. (2.18) and using Eq. (2.13) we obtain, after some rearrangements,

$$(2P + 3\ln\alpha)_{,\alpha} = -\alpha[(V_{,\alpha}{}^2 + V_{,\beta}{}^2)\cosh^2 W + W_{,\alpha}{}^2 + W_{,\beta}{}^2] \,, \tag{2.26a}$$

$$(2P + 3\ln\alpha)_{,\beta} = -2\alpha(V_{,\alpha} V_{,\beta} \cosh^2 W + W_{,\alpha} W_{,\beta}) \,. \tag{2.26b}$$

Equations (2.26) suggest that it will be convenient to define the combination $2P + 3\ln\alpha$ as a new variable, which, together with the variables V and W, would uniquely determine the metric in the (α,β,x,y) coordinate system. Thus, after first introducing the two "normalization" length scales l_1 and l_2 by the equations

$$l_1 \equiv \frac{1}{2U_{,u}(u_0,v_0)} \,, \qquad l_2 \equiv \frac{1}{2U_{,v}(u_0,v_0)} \,, \tag{2.27a}$$

where (u_0,v_0), $u_0>0$, $v_0>0$ is an arbitrary, fixed point in the *interior* of the interaction region, we define a new function $Q(\alpha,\beta)$ by the relation

$$e^{Q/2} \equiv 4 l_1 l_2 \, e^M \, U_{,u} U_{,v} \, \alpha^{3/2} \, . \tag{2.27b}$$

Using Eqs. (2.27a), we then fix the constant c which occurs in Eq. (2.24):

$$c \equiv l_1 l_2 \, . \tag{2.28}$$

Note that the length scales l_1 and l_2 are determined by Eqs. (2.27a) in a well-defined manner, since by Eqs. (2.13) and (2.16a)

$$U(u,v) = -\ln \alpha(u,v) = -\ln\left(e^{-U_1(u)} + e^{-U_2(v)} - 1\right) , \tag{2.29}$$

so that

$$U_{,u}(u,v) = \frac{1}{\alpha(u,v)} U_1'(u) e^{-U_1(u)} ,$$

$$U_{,v}(u,v) = \frac{1}{\alpha(u,v)} U_2'(v) e^{-U_2(v)} ;$$

and therefore, by Eqs. (2.11b), $U_{,u}(u,v)>0$, $U_{,v}(u,v)>0$ for any point (u,v) in the interior of the interaction region, where $u>0$, $v>0$, and where [as long as (u,v) is in the domain of dependence of the initial surface $\{u=0\} \cup \{v=0\}$] $\alpha(u,v)>0$. It is now easy to obtain the remaining field equations, satisfied by the new variable $Q(\alpha,\beta)$: Combining Eq. (2.27b) with Eqs. (2.28) and (2.24), and then using Eqs. (2.26), we find

$$Q_{,\alpha} = -\alpha\left[(V_{,\alpha}{}^2 + V_{,\beta}{}^2) \cosh^2 W + W_{,\alpha}{}^2 + W_{,\beta}{}^2 \right] , \tag{2.30a}$$

$$Q_{,\beta} = -2\alpha (V_{,\alpha} V_{,\beta} \cosh^2 W + W_{,\alpha} W_{,\beta}) , \tag{2.30b}$$

where the integrability condition for Eqs. (2.30) is satisfied by virtue of the field equations (2.23) for $V(\alpha,\beta)$ and $W(\alpha,\beta)$.

We now combine Eq. (2.27b) with the expression (2.21) for the metric in the interaction region. This gives us the expression of the interaction region metric in terms of the three unknown variables V, W, and Q. Then, by using the initial value of Q that follows from our normalization conditions Eqs. (2.27), we construct the unique solution $Q(\alpha,\beta)$ of the field equations (2.30) by quadrature. As a result, we obtain the following expressions for the metric and the field equations, valid in the interaction region of any arbitrarily-polarized colliding plane-wave spacetime:

$$g = e^{-Q(\alpha,\beta)/2} \frac{l_1 l_2}{\sqrt{\alpha}} (-d\alpha^2 + d\beta^2)$$

$$+ \alpha [\cosh W(\alpha,\beta) (e^{V(\alpha,\beta)} dx^2 + e^{-V(\alpha,\beta)} dy^2) - 2 \sinh W(\alpha,\beta)\, dx dy] , \qquad (2.31)$$

where V, W, and Q satisfy the following field equations:

$$V_{,\alpha\alpha} + \frac{1}{\alpha} V_{,\alpha} - V_{,\beta\beta} = 2 (V_{,\beta} W_{,\beta} - V_{,\alpha} W_{,\alpha}) \tanh W , \qquad (2.32a)$$

$$W_{,\alpha\alpha} + \frac{1}{\alpha} W_{,\alpha} - W_{,\beta\beta} = (V_{,\alpha}{}^2 - V_{,\beta}{}^2) \sinh W \cosh W . \qquad (2.32b)$$

$$Q(\alpha,\beta) = \int_{C:(\alpha_0,\beta_0)}^{(\alpha,\beta)} \{ -\alpha [(V_{,\alpha}{}^2 + V_{,\beta}{}^2) \cosh^2 W + W_{,\alpha}{}^2 + W_{,\beta}{}^2] \, d\alpha$$

$$- 2\alpha (V_{,\alpha} V_{,\beta} \cosh^2 W + W_{,\alpha} W_{,\beta}) \, d\beta \}$$

$$+ 2 M(\alpha_0,\beta_0) + 3 \ln \alpha_0 . \qquad (2.33)$$

Here, $\alpha_0 \equiv \alpha(u_0,v_0)$, $\beta_0 \equiv \beta(u_0,v_0)$, $M(\alpha_0,\beta_0) \equiv M(u_0,v_0)$, and C is *any* (differentiable) curve in the (α,β) plane that starts at the initial point (α_0,β_0), and ends at the field point (α,β) at which Q is to be computed. The result of the integral in Eq. (2.33) depends only on the end points of the curve C, since the integrability condition for Eqs. (2.30) is satisfied by virtue of the field equations (2.32).

Equations (2.31)–(2.33) summarize the initial-value problem for colliding plane waves in a conveniently compact form. The only unknowns that must be found by solving partial differential equations (PDE) are the functions $V(\alpha,\beta)$ and $W(\alpha,\beta)$ which satisfy the *nonlinear* system of coupled hyperbolic PDE (2.32). Once $V(\alpha,\beta)$ and $W(\alpha,\beta)$ are known, Q is determined by the explicit expression (2.33) up to an unknown additive constant, which – by suitably choosing the initial point (u_0,v_0) [or (α_0,β_0)] – can be made arbitrarily small. The only disadvantage of the formalism (2.31)–(2.33) is the coordinate singularity that the (α,β) chart develops on the characteristic initial surface $\{u=0\}\cup\{v=0\}$. This coordinate singularity causes, among other things, the function $Q(\alpha,\beta)$ to be logarithmically divergent (to $-\infty$) on the surfaces $\{u=0\}$ and $\{v=0\}$. However, it is still possible to set up a well-defined initial-value problem for the functions $V(\alpha,\beta)$ and $W(\alpha,\beta)$, using initial data posed on the same characteristic surface $\{u=0\}\cup\{v=0\}$. In addition, since we are primarily interested in the behavior of spacetime near the singular "surface" $\{\alpha=0\}$ well away from the coordinate singularity on the initial null surfaces, the above formalism based on (α,β) coordinates is well suited to our objectives.

To understand how to pose initial data for the field equations (2.32), first note that [cf. Eqs. (2.16)] in the α,β coordinates the initial null surfaces $\{u=0\}$ and $\{v=0\}$ are expressed as (Fig. 1)

$$\{u=0\}\equiv\{\alpha-\beta=1\}\,,\qquad \{v=0\}\equiv\{\alpha+\beta=1\}\,. \tag{2.34}$$

Equations (2.34) suggest introducing "characteristic" coordinates

$$r \equiv \alpha - \beta \,, \qquad s \equiv \alpha + \beta \,, \tag{2.35}$$

so that the initial null surfaces become (see Fig. 1)

$$\{u=0\} \equiv \{r=1\} \,, \qquad \{v=0\} \equiv \{s=1\} \,. \tag{2.36}$$

The initial-value problem for the functions V and W consists of the field equations (2.32), and the initial data on the characteristic initial surface $\{r=1\} \cup \{s=1\}$ given by the freely specifiable functions $V(r,s=1)$, $W(r,s=1)$ and $V(r=1,s)$, $W(r=1,s)$. More precisely, the initial data consist of

$$\{\ V(r,1), W(r,1), V(1,s), W(1,s)\ \} \,, \tag{2.37}$$

where $V(r,1)$, $W(r,1)$ and $V(1,s)$, $W(1,s)$ are C^1 (and piecewise C^2) functions for $r \in (-1,1]$ and $s \in (-1,1]$, respectively, which are freely specified except for the initial conditions $V(r=1,1) = W(r=1,1) = V(1,s=1) = W(1,s=1) = 0$.

There is a one-to-one correspondence between the initial data of the form (2.7), and initial data of the form (2.37) for the initial-value problem of colliding plane waves. When initial data are given in the form of Eq. (2.7), i.e., when the functions $V_1(u)$, $W_1(u)$ and $V_2(v)$, $W_2(v)$ are specified, initial data in the form of Eq. (2.37) are uniquely determined in the following way: First, Eqs. (2.8) are solved with the given data $V_1(u)$, $W_1(u)$ and $V_2(v)$, $W_2(v)$, and the functions $U_1(u)$ and $U_2(v)$ are obtained as the unique solutions. [Cf. the discussion following Eqs. (2.8)]. Then, using the identities [cf. Eqs. (2.16) and Eq. (2.35)]

$$r = 2e^{-U_1(u)} - 1 \,, \qquad s = 2e^{-U_2(v)} - 1 \tag{2.38}$$

along the initial null surfaces $\{u=0\}$ and $\{v=0\}$, $u(r)$ and $v(s)$ are defined as the unique solutions to the implicit equations

$$r=2e^{-U_1[u(r)]}-1\,,\qquad s=2e^{-U_2[v(s)]}-1\,. \tag{2.39}$$

Finally, the initial data $\{V(r,1), W(r,1), V(1,s), W(1,s)\}$ in the form (2.37) are determined uniquely from the data $\{V_1(u), W_1(u), V_2(v), W_2(v)\}$ by

$$V(r,1)=V_1[u=u(r)]\,,\qquad W(r,1)=W_1[u=u(r)]\,.$$

$$V(1,s)=V_2[v=v(s)]\,,\qquad W(1,s)=W_2[v=v(s)]\,. \tag{2.40}$$

Conversely, when initial data are given in the form of Eq. (2.37), i.e., when the functions $V(r,1)$, $W(r,1)$ and $V(1,s)$, $W(1,s)$ are specified, initial data in the form of (2.7) are uniquely determined in the following way: First, the differential equations

$$\begin{aligned}2U_{1,uu}-U_{1,u}{}^2\\ &=4e^{-2U_1}U_{1,u}{}^2\,\{[V_{,r}(r=2e^{-U_1}-1,1)]^2\cosh^2 W(r=2e^{-U_1}-1,1)\\ &+[W_{,r}(r=2e^{-U_1}-1,1)]^2\}\,,\end{aligned} \tag{2.41a}$$

$$\begin{aligned}2U_{2,vv}-U_{2,v}{}^2\\ &=4e^{-2U_2}U_{2,v}{}^2\,\{[V_{,s}(1,s=2e^{-U_2}-1)]^2\cosh^2 W(1,s=2e^{-U_2}-1)\\ &+[W_{,s}(1,s=2e^{-U_2}-1)]^2\}\,,\end{aligned} \tag{2.41b}$$

for the functions $U_1(u)$ and $U_2(v)$ are solved with the initial conditions $U_1(u=0)=U_2(v=0)=0$, $U_{1,u}(u=0)=U_{2,v}(v=0)=0$ [cf. Eqs. (2.8)]. Then, using Eqs. (2.39), the initial data $\{V_1(u), W_1(u), V_2(v), W_2(v)\}$ in the form (2.7) are determined uniquely from the data $\{V(r,1), W(r,1), V(1,s), W(1,s)\}$ by

$$V_1(u) = V(r = 2e^{-U_1(u)} - 1, 1) ,$$

$$W_1(u) = W(r = 2e^{-U_1(u)} - 1, 1) .$$

$$V_2(v) = V(1, s = 2e^{-U_2(v)} - 1) ,$$

$$W_2(v) = W(1, s = 2e^{-U_2(v)} - 1) . \tag{2.42}$$

This completes the formulation of the initial-value problem for the system of coupled nonlinear hyperbolic PDE (2.32).

III. ASYMPTOTIC STRUCTURE OF COLLIDING PLANE-WAVE SPACE-TIMES NEAR $\alpha=0$

A. Singularities and horizons at $\alpha=0$: A generalized inhomogeneous-Kasner asymptotic structure

It is clear from the expression (2.31) of the metric that the "surface" $\{\alpha=0\}$ represents some kind of singularity [either a spacetime singularity or (at least) a coordinate singularity] of the colliding plane-wave spacetime. In this section and the following Sec. III B, we will study the asymptotic behavior of the colliding plane-wave metric (2.31)–(2.33) near this singularity $\{\alpha=0\}$.

Before proceeding with the analysis of asymptotic structure, recall the conclusions of Sec. II B in Ref. 6, where the field equations for colliding *parallel-polarized* plane waves were studied in (α,β) coordinates. [Compare Eqs. (6.2.43) and (6.2.44) with Eqs. (2.31)–(2.33) above.] There the field equations reduced to a single *linear* hyperbolic PDE for $V(\alpha,\beta)$ [Eq. (6.2.44a)], followed by a quadrature for $Q(\alpha,\beta)$ [Eq. (6.2.44b)] similar to Eq. (2.33) above. [The readers can rederive these equations by simply putting $W \equiv 0$ in Eqs. (2.31)–(2.33).] It is well known that for linear hyperbolic PDE of the kind (6.2.44a) solutions with sufficiently smooth initial data exist globally (see, for example, Secs. 5.2 and 5.3 of Ref. 18 and p. 115 in Sec. 4.2 of Ref. 19). Therefore, it was guaranteed by the linearity of Eq. (6.2.44a) in Ref. 6 that the field equations for V and Q had unique global solutions defined throughout the domain of dependence of the initial surface, i.e., throughout the interaction region $\{\alpha>0\}$. In fact, a general solution (Riemann function[19]) for Eq. (6.2.44a) could be found in closed form [Eq. (6.2.59)], which yielded an explicit representation [Eq. (6.2.60)] of the global solution $V(\alpha,\beta)$ (for $\alpha>0$) in terms of initial data. This assured that the singularities [or Killing-Cauchy horizons (coordinate singularities)] created by colliding parallel-polarized plane waves always lie at or beyond the surface $\{\alpha=0\}$; this surface is in fact the boundary of the domain of dependence, and as Eq. (6.2.43) makes clear, some kind of singularity is always present there.

In contrast with the parallel-polarized case, the field equations (2.32) for arbitrarily-polarized colliding plane waves are nonlinear. It is a standard result (see e.g. Ref. 20 and Sec. VI.6 of Ref. 21) that quasilinear hyperbolic PDE of the form (2.32) always have unique, *local* solutions, defined in a neighborhood of the initial surface on which regular initial data are posed. On the other hand, it is also well known[22–32] that in general these local solutions do not exist *globally*; i.e., in general solutions of nonlinear hyperbolic PDE blow up or otherwise break down in finite time within the interior of their domain of dependence. [A particularly lucid example of this break-down-in-finite-time phenomenon for solutions of nonlinear hyperbolic PDE is discussed by S. Klainerman, following his Eq. (13) in Ref. 23.] We also note in this connection that thanks to the recent work of Klainerman,[23,28,31] Shatah, [25] Sideris, [29] Klainerman and Ponce, [27] and Christodoulou,[32] it is now known that for initial data which are sufficiently "small" in some appropriate sense, solutions of nonlinear hyperbolic PDE of the kind (2.32) *do* exist globally, i.e. throughout the domain of dependence of the initial surface. (See Appendix A for a somewhat more detailed discussion of this point.) In any case, as we have also discussed in the Introduction (Sec. I), if the global existence of solutions with arbitrary (not necessarily small) initial data were false for the field equations (2.32), then this would have the disturbing consequences that (i) colliding nonparallel-polarized plane waves might create singularities in the interior of the interaction region where $\alpha>0$, and (ii) these singularities, if present, would not be analyzable by studying the asymptotic spacetime structure near $\alpha=0$. Therefore, before the asymptotic-structure analysis of this section can be relied on to fully describe the singularity structure of colliding plane-wave spacetimes, it is necessary to have a proof that solutions of Eqs. (2.32) exist globally for *all* initial data.

Obviously, one way to prove this global existence result would be to obtain a general solution (Riemann function[19]) for Eqs. (2.32), in the same way as the Riemann function [Eq. (6.2.59)] of Ref. 6 yielded the explicit expression (6.2.60) of the solution V in terms of initial data, and thus provided a constructive proof for the global existence of V in the parallel-polarized case. It seems unlikely, however, that such a general solution can be found for the nonlinear system (2.32); hence the global existence of solutions for (2.32) must be proved using nonconstructive arguments. Indeed, such a nonconstructive proof *can* be provided, as we discuss in detail in Appendix A, thanks to some very special properties possessed by the field equations [especially the existence of the positive-definite conserved energy form Eq. (A18)]. Thus, our discussions in Appendix A prove that the singularities and Killing-Cauchy horizons (see below) created by colliding plane waves always lie at or beyond $\{\alpha=0\}$; no singularities ever occur in the interior of the interaction region where $\alpha>0$. [Incidentally, Appendix A also proves as a special case that the global solution (6.2.60) for $V(\alpha,\beta)$ coupled with $W\equiv 0$ is the unique solution of Eqs. (2.32) corresponding to initial data (2.37) with $W(r,1)=W(1,s)\equiv 0$; i.e., colliding plane waves which are initially parallel-polarized remain parallel-polarized everywhere after they scatter each other.] Furthermore, in Appendix B we use the results of Appendix A to prove that the spatial (β) derivative terms in the field equations (2.32) are asymptotically negligible compared to the timelike (α) derivative terms as the singularity $\{\alpha=0\}$ is approached. As we will heavily rely on these results in the discussions below, we suggest to those readers who desire greater logical completeness that they read Appendices A and B at this point, before proceeding with the rest of Secs. III A and B.

Since as $\alpha\to 0$ the β-derivative terms in Eqs. (2.32) are asymptotically negligible compared to the α-derivative terms (Appendix B), the asymptotic behaviors of V and W near $\alpha=0$ are identical with those of the solutions of the ordinary differential equations

$$V_{,\alpha\alpha}+\frac{1}{\alpha}V_{,\alpha}=-2V_{,\alpha}W_{,\alpha}\tanh W\ , \tag{3.1a}$$

$$W_{,\alpha\alpha}+\frac{1}{\alpha}W_{,\alpha}=V_{,\alpha}{}^2\sinh W\cosh W \tag{3.1b}$$

obtained from Eqs. (2.32) by ignoring all terms with β-derivatives.

Consider first Eq. (3.1a) for V. Dividing both sides by $V_{,\alpha}$ and integrating, we obtain

$$\ln|\alpha V_{,\alpha}| + 2\ln(\cosh W) = C \ , \tag{3.2}$$

which immediately yields

$$V_{,\alpha} = \frac{C}{\alpha \cosh^2 W} \ . \tag{3.3}$$

[Here and henceforth C will stand for an arbitrary (indefinite) constant.] Clearly, the constant C in Eq. (3.3) will in general depend on β. Thus, we rename the constant C of Eq. (3.3) as $\varepsilon_1(\beta)$, and then apply a further integration to obtain

$$V(\alpha,\beta) = \varepsilon_1(\beta) \int \frac{d\alpha}{\alpha \cosh^2 W} + \delta_1(\beta) + H_1(\alpha,\beta) \ , \tag{3.4a}$$

where

$$\lim_{\alpha \to 0} H_1(\alpha,\beta) \equiv 0 \ . \tag{3.4b}$$

Equations (3.4) determine the asymptotic behavior of $V(\alpha,\beta)$ once the asymptotic behavior of W is known.

To find the asymptotic behavior of $W(\alpha,\beta)$, consider Eq. (3.1b) for W and insert into it the expression for $V_{,\alpha}$ given by Eq. (3.3); this yields

$$\frac{1}{\alpha}(\alpha W_{,\alpha})_{,\alpha} = \frac{{\varepsilon_1}^2}{\alpha^2 \cosh^3 W} \sinh W \ . \tag{3.5}$$

Multiplying both sides of Eq. (3.5) by $2\alpha^2 W_{,\alpha}$ and integrating once after collecting all terms on the left-hand side, we obtain

$$(\alpha W_{,\alpha})^2 + \frac{{\varepsilon_1}^2}{\cosh^2 W} = C \equiv {\varepsilon_2}^2 \ , \tag{3.6}$$

where we have renamed the β-dependent constant C as $\varepsilon_2(\beta)$. We will always assume, without loss of generality, that by convention $\varepsilon_2 \geq 0$. Equation (3.6) can then be rewritten in the form

$$\pm\frac{d\alpha}{\alpha}=\frac{dW}{[\varepsilon_2{}^2-(\varepsilon_1{}^2/\cosh^2 W)]^{1/2}}$$

$$=\frac{1}{\varepsilon_2}\,\frac{\cosh W\; dW}{[\cosh^2 W-(\varepsilon_1{}^2/\varepsilon_2{}^2)]^{1/2}}\,. \tag{3.7}$$

The integration of Eq. (3.7) is elementary, and it yields the following two possibilities for the asymptotic behavior of $W(\alpha,\beta)$ near α=0:

$$W(\alpha,\beta)=\delta_2(\beta)\,\alpha^{\varepsilon_2(\beta)}+H_2(\alpha,\beta) \qquad \text{in which case } \left(\frac{\varepsilon_1(\beta)}{\varepsilon_2(\beta)}\right)^2 \text{ must equal 1 ,} \tag{3.8a}$$

$$W(\alpha,\beta)=\pm\varepsilon_2(\beta)\ln\alpha+\delta_2(\beta)+H_2(\alpha,\beta) \qquad \text{in which case } \left(\frac{\varepsilon_1(\beta)}{\varepsilon_2(\beta)}\right)^2 \text{ is arbitrary ,} \tag{3.8b}$$

where

$$\lim_{\alpha\to 0} H_2(\alpha,\beta)\equiv 0\,. \tag{3.8c}$$

Combining Eqs. (3.8) with Eqs. (3.4), we find that there are three and only three distinct possible asymptotic behaviors for V and W near α=0. We can express these three possible cases in the following final form:

Case (a). In this case $[\varepsilon_1(\beta)/\varepsilon_2(\beta)]^2$ must equal 1, and $\varepsilon_2(\beta)>0$:

$$V(\alpha,\beta)=\varepsilon_1(\beta)\ln\alpha+\delta_1(\beta)+H_1(\alpha,\beta)\,,$$

$$W(\alpha,\beta)=\delta_2(\beta)\,\alpha^{\varepsilon_2(\beta)}+H_2(\alpha,\beta)\,. \tag{3.9a}$$

Case (b). In this case $[\varepsilon_1(\beta)/\varepsilon_2(\beta)]^2$ is arbitrary, and $\varepsilon_2(\beta)>0$:

$$V(\alpha,\beta)=\frac{2\varepsilon_1(\beta)}{\varepsilon_2(\beta)}\,e^{\pm 2\delta_2(\beta)}\,\alpha^{2\varepsilon_2(\beta)}+\delta_1(\beta)+H_1(\alpha,\beta)\,,$$

$$W(\alpha,\beta)=\pm\varepsilon_2(\beta)\ln\alpha+\delta_2(\beta)+H_2(\alpha,\beta)\,. \tag{3.9b}$$

Case (c). In this case $\varepsilon_2(\beta)=\varepsilon_1(\beta)=0$:

$$V(\alpha,\beta)=\delta_1(\beta)+H_1(\alpha,\beta)\,,$$

$$W(\alpha,\beta)=\delta_2(\beta)+H_2(\alpha,\beta)\,. \tag{3.9c}$$

In all three cases (a)–(c) above the terms $H_i(\alpha,\beta)$ have the general form ($i\equiv 1, 2$)

$$H_i(\alpha,\beta)=\sum_{k=2}^{\infty} c^{(i)}{}_k(\beta)\,\alpha^k+\sum_{k=2,\,l=1}^{\infty} d^{(i)}{}_{kl}(\beta)\,\alpha^k\,(\ln\alpha)^l\,. \tag{3.10}$$

[Equation (3.10) follows from the expressions (3.3) and (3.6) for $V_{,\alpha}$ and $W_{,\alpha}$. In fact, Eqs. (3.3) and (3.6) constrain the form of $H_i(\alpha,\beta)$ even further than Eq. (3.10), and we will use these extra constraints below in deriving the asymptotic form of $Q(\alpha,\beta)$ near $\alpha=0$.]

The asymptotic behavior of the metric function $Q(\alpha,\beta)$ is obtained by combining Eqs. (3.9) and (3.10) with the field equation (2.33). The final result can be described as follows:

$$Case\ (a){:}\quad Q(\alpha,\beta) = -\varepsilon_1^{\,2}(\beta)\ln\alpha + \mu(\beta) + L(\alpha,\beta)\,, \tag{3.11a}$$

$$Case\ (b){:}\quad Q(\alpha,\beta) = -\varepsilon_2^{\,2}(\beta)\ln\alpha + \mu(\beta) + L(\alpha,\beta)\,, \tag{3.11b}$$

$$Case\ (c){:}\quad Q(\alpha,\beta) = \mu(\beta) + L(\alpha,\beta)\,, \tag{3.11c}$$

where

$$\lim_{\alpha\to 0} L(\alpha,\beta) \equiv 0\,, \tag{3.12}$$

but $L(\alpha,\beta)$ does not necessarily have the general form (3.10).

With Eqs. (3.9)–(3.12), we have completed our analysis of the asymptotic forms of the metric functions V, W, and Q near $\alpha=0$; at this point readers might find it useful to compare Eqs. (3.9)–(3.12) with the corresponding Eqs. (6.3.4)–(6.3.7) of Ref. 6 for the parallel-polarized case.

Now we are ready to analyze the asymptotic behavior of the arbitrarily-polarized colliding plane-wave metric (2.31) near the singular surface $\{\alpha=0\}$. We first note that the $x-y$ part of the metric (2.31), when considered as a two-dimensional symmetric tensor field on $\{u=\text{const},\ v=\text{const}\}$ sections, is positive-definite and nondegenerate,[33] i.e., it is a euclidean metric. (That this must be the case becomes clear when one recalls that by the definition of plane-symmetry[8,9] the Killing vectors $\partial/\partial x$ and $\partial/\partial y$ must span a spacelike two-dimensional plane in each tangent space. Only asymptotically, as $\alpha\to 0$, can this 2-plane become null.) Consequently, it is possible to diagonalize the $x-y$ part of the metric by using two spacelike, orthonormal 1-forms defined throughout the interaction region. When this is done, we find that the metric (2.31) can be brought into the diagonal form

$$g = e^{-Q(\alpha,\beta)/2}\frac{l_1 l_2}{\sqrt{\alpha}}(d\alpha^2 + d\beta^2) + \alpha(\omega^1\otimes\omega^1 + \omega^2\otimes\omega^2) \tag{3.13a}$$

with the orthogonal spacelike 1-forms

$$\omega^1 = \frac{e^{\hat{V}/2}}{(2\sinh\hat{V})^{1/2}} \left[P\, dx - \frac{\sinh W}{P}\, dy \right] , \tag{3.13b}$$

$$\omega^2 = \frac{e^{-\hat{V}/2}}{(2\sinh\hat{V})^{1/2}} \left[\frac{\sinh W}{P}\, dx + P\, dy \right] , \tag{3.13c}$$

where,

$$P \equiv (\sinh\hat{V} + \sinh V \cosh W)^{1/2} , \tag{3.13d}$$

$$\hat{V} \equiv \ln[\cosh V \cosh W + (\cosh^2 V \cosh^2 W - 1)^{1/2}] . \tag{3.13e}$$

A short computation using Eqs. (3.13) shows that when considered as functions of the variables V and W, the 1-forms ω^1 and ω^2 are discontinuous at $W=0$; the 1-form ω^1 (as well as ω^2) tends to two different limits as $W \to 0$ depending on whether $W \to +0$ or $W \to -0$. In contrast, the tensor field $\omega^1 \otimes \omega^1 + \omega^2 \otimes \omega^2$ depends on V and W smoothly; in fact

$$\omega^1 \otimes \omega^1 + \omega^2 \otimes \omega^2 \to e^V dx^2 + e^{-V} dy^2 \qquad \text{as } W \to \pm 0 .$$

Therefore, the discontinuities in the dependence of ω^i on V and W are unimportant when analyzing the asymptotic structure of the spacetime geometry (3.13a) near $\alpha=0$.

We now combine Eqs. (3.13) with Eqs. (3.9)–(3.12), and obtain the following final results for the asymptotic form of the metric (2.31) as $\alpha \to 0$:

Case (a). In this case $[\varepsilon_1(\beta)/\varepsilon_2(\beta)]^2$ must equal 1, and $\varepsilon_2(\beta)>0$:

$$g(\beta) \sim l_1 l_2 e^{-\mu(\beta)/2} \alpha^{[\varepsilon_1{}^2(\beta)-1]/2} (-d\alpha^2 + d\beta^2)$$

$$+e^{\delta_1(\beta)}\alpha^{1+\varepsilon_1(\beta)}dx^2+e^{-\delta_1(\beta)}\alpha^{1-\varepsilon_1(\beta)}dy^2 . \tag{3.14a}$$

Case (b). In this case $[\varepsilon_1(\beta)/\varepsilon_2(\beta)]^2$ is arbitrary, and $\varepsilon_2(\beta)>0$:

$$g(\beta)\sim l_1 l_2 e^{-\mu(\beta)/2}\alpha^{[\varepsilon_2{}^2(\beta)-1]/2}(-d\alpha^2+d\beta^2)+\alpha(\omega^1\otimes\omega^1+\omega^2\otimes\omega^2) ,$$

$$\omega^1(\beta)\sim e^{\mp\delta_2(\beta)/2}\alpha^{-\varepsilon_2(\beta)/2}(e^{\delta_1(\beta)/2}dx\pm e^{-\delta_1(\beta)/2}dy) ,$$

$$\omega^2(\beta)\sim\frac{e^{\pm\delta_2(\beta)/2}}{\cosh\delta_1(\beta)}\alpha^{\varepsilon_2(\beta)/2}(\mp e^{-\delta_1(\beta)/2}dx+e^{\delta_1(\beta)/2}dy) . \tag{3.14b}$$

Case (c). In this case $\varepsilon_2(\beta)=\varepsilon_1(\beta)=0$:

$$g(\beta)\sim l_1 l_2 e^{-\mu(\beta)/2}\alpha^{-1/2}(-d\alpha^2+d\beta^2)+\alpha(\omega^1\otimes\omega^1+\omega^2\otimes\omega^2) ,$$

$$\omega^1\sim\frac{s(\beta)}{[s^2(\beta)-1]^{1/2}}\left[q(\beta)\,dx-\frac{\sinh\delta_2(\beta)}{q(\beta)}dy\right] ,$$

$$\omega^2\sim\frac{1}{[s^2(\beta)-1]^{1/2}}\left[\frac{\sinh\delta_2(\beta)}{q(\beta)}dx+q(\beta)\,dy\right] ,$$

$$q(\beta)\equiv\left[\frac{s^2(\beta)-s(\beta)}{2s(\beta)}+\sinh\delta_1(\beta)\cosh\delta_2(\beta)\right]^{1/2} ,$$

$$s(\beta)\equiv\cosh\delta_1(\beta)\cosh\delta_2(\beta)+[\cosh^2\delta_1(\beta)\cosh^2\delta_2(\beta)-1]^{1/2} . \tag{3.14c}$$

In all three cases (3.14a–c), the asymptotic structure of the metric is generalized inhomogeneous-Kasner. The following equations are derived from Eqs. (3.14) in order to express this inhomogeneous-Kasner structure more precisely [compare also Eqs. (6.3.14)–(6.3.19) of Ref. 6]:

Case (a):

$$g(\beta) \sim -\frac{16\, l_1 l_2\, e^{-\mu(\beta)/2}}{[\varepsilon_1^{\,2}(\beta)+3]^2}\, dt^2 + l_1 l_2\, e^{-\mu(\beta)/2}\, t^{2p_3}\, d\beta^2$$

$$+\, e^{\delta_1(\beta)}\, t^{2p_1}\, dx^2 + e^{-\delta_1(\beta)}\, t^{2p_2}\, dy^2\,, \tag{3.15a}$$

where

$$t \equiv \alpha^{[\varepsilon_1^{\,2}(\beta)+3]/4}\,, \tag{3.15b}$$

and

$$p_3(\beta) = \frac{\varepsilon_1^{\,2}(\beta)-1}{\varepsilon_1^{\,2}(\beta)+3}\,, \quad p_1(\beta) = \frac{2[1+\varepsilon_1(\beta)]}{\varepsilon_1^{\,2}(\beta)+3}\,, \quad p_2(\beta) = \frac{2[1-\varepsilon_1(\beta)]}{\varepsilon_1^{\,2}(\beta)+3}\,. \tag{3.15c}$$

Case (b):

$$g(\beta) \sim -\frac{16\, l_1 l_2\, e^{-\mu(\beta)/2}}{[\varepsilon_2^{\,2}(\beta)+3]^2}\, dt^2 + l_1 l_2\, e^{-\mu(\beta)/2}\, t^{2p_3}\, d\beta^2$$

$$+\, e^{\mp\delta_2(\beta)}\, t^{2p_1}\, dX_{(\beta)}^{\;2} + \frac{e^{\pm\delta_2(\beta)}}{\cosh^2\delta_1(\beta)}\, t^{2p_2}\, dY_{(\beta)}^{\;2}\,, \tag{3.16a}$$

where

$$t \equiv \alpha^{[\varepsilon_2^{\,2}(\beta)+3]/4}\,, \tag{3.16b}$$

and

$$p_3(\beta) = \frac{\varepsilon_2^{\,2}(\beta)-1}{\varepsilon_2^{\,2}(\beta)+3}\,, \quad p_1(\beta) = \frac{2[1+\varepsilon_2(\beta)]}{\varepsilon_2^{\,2}(\beta)+3}\,, \quad p_2(\beta) = \frac{2[1-\varepsilon_2(\beta)]}{\varepsilon_2^{\,2}(\beta)+3}\,, \tag{3.16c}$$

$$X_{(\beta)} \equiv (e^{\delta_1(\beta)/2} x \pm e^{-\delta_1(\beta)/2} y) , \quad Y_{(\beta)} \equiv (\mp e^{-\delta_1(\beta)/2} x + e^{\delta_1(\beta)/2} y) . \tag{3.16d}$$

Case (c):

$$g(\beta) \sim -\frac{16}{9} l_1 l_2 e^{-\mu(\beta)/2} dt^2 + l_1 l_2 e^{-\mu(\beta)/2} t^{2p_3} d\beta^2$$

$$+ \frac{s^2(\beta)}{s^2(\beta)-1} t^{2p_1} dX_{(\beta)}{}^2 + \frac{1}{s^2(\beta)-1} t^{2p_2} dY_{(\beta)}{}^2 , \tag{3.17a}$$

where

$$t \equiv \alpha^{3/4} , \tag{3.17b}$$

and

$$p_3(\beta) = -\frac{1}{3} , \quad p_1(\beta) = \frac{2}{3} , \quad p_2(\beta) = \frac{2}{3} , \tag{3.17c}$$

$$X_{(\beta)} \equiv \left[q(\beta) x - \frac{\sinh\delta_2(\beta)}{q(\beta)} y \right] , \quad Y_{(\beta)} \equiv \left[\frac{\sinh\delta_2(\beta)}{q(\beta)} x + q(\beta) y \right] . \tag{3.17d}$$

Equations (3.15)–(3.17) demonstrate that at a fixed value of β the asymptotic limit of the spacetime metric (2.31) has the general form of a vacuum Kasner[10] solution:

$$g = -a\, dt^2 + b\, t^{2p_3} d\beta^2 + c\, t^{2p_1} dX^2 + d\, t^{2p_2} dY^2 , \tag{3.18}$$

where a, b are constants having the dimensions of (length)^2, c, d are dimensionless constants, t, β are dimensionless coordinates, and the exponents $p_k, k=1,2,3$ in all cases satisfy the Kasner relations [cf. Eqs. (3.15c), (3.16c), and (3.17c)]

$$p_1(\beta) + p_2(\beta) + p_3(\beta) = p_1{}^2(\beta) + p_2{}^2(\beta) + p_3{}^2(\beta) = 1 . \tag{3.19}$$

The coordinates X, Y are asymptotically-constant linear combinations [cf. Eqs. (3.16d) and (3.17d)] of the spacelike (Killing) coordinates x and y that determine the asymptotic Kasner axes along which the exponents p_1 and p_2 are defined (the exponent p_3 is always associated with the β axis). In fact, it becomes clear from Eqs. (3.15)–(3.17) that in general these asymptotic Kasner axes (defined by the coordinates $X_{(\beta)}$, $Y_{(\beta)}$), like the Kasner exponents $p_k(\beta)$, depend on the spatial coordinate β across the singularity: hence the rationale for our use of the term "generalized inhomogeneous-Kasner" to describe the asymptotic structures (3.15)–(3.17).[6]

If all of the exponents p_k are different from 1 [or equivalently by Eqs. (3.19) all are different from 0], then the Kasner spacetime (3.18) possesses a curvature singularity at $t=0$. (For a brief description of the geometry of the Kasner solution see Sec. III A of Ref. 6.) It follows that when $p_k(\beta)$ are similarly all different from 0 in any of the three cases (a)–(c) [Eqs. (3.15)–(3.17)], the colliding plane-wave spacetime (2.31) possesses a curvature singularity at $(\alpha=0,\beta)$. Conversely, when any of the p_k in Eq. (3.18) is equal to 1 (in which case both other exponents are zero), the metric (3.18) is flat (a degenerate Kasner solution[6]) and $\{t=0\}$ is a nonsingular Killing-Cauchy horizon[8] in the Kasner spacetime. Similarly, we claim that if any of the two exponents $p_1(\beta)$, $p_2(\beta)$ is identically equal to 1 across an interval (β_1,β_2) [the exponent $p_3(\beta)$ can never equal 1, see Eqs. (3.15c) and (3.16c)], then the surface $\{\alpha=0, \beta_1<\beta<\beta_2\}$ is a Killing-Cauchy horizon for the colliding plane-wave spacetime (2.31). More precisely, we claim the following:

(i) In case (a), the surface $\{\alpha=0, \beta_1<\beta<\beta_2\}$ is a Killing-Cauchy horizon if and only if

$$|\varepsilon_1(\beta)| \equiv 1 \qquad \forall\ \beta \in (\beta_1,\beta_2)\,. \tag{3.20a}$$

In this case, the spacelike Killing vector that becomes null on the horizon[8] is either $\partial/\partial x$ (when $\varepsilon_1 \equiv +1$) or $\partial/\partial y$ (when $\varepsilon_1(\beta) \equiv -1$).

(ii) In case (b), the surface $\{\alpha=0, \beta_1<\beta<\beta_2\}$ is a Killing-Cauchy horizon if and only if

$$\varepsilon_2(\beta) \equiv 1 \,, \qquad \delta_1(\beta) \equiv \text{const} \equiv \delta_1 \qquad \forall\ \beta \in (\beta_1,\beta_2)\,. \tag{3.20b}$$

In this case, the spacelike Killing vector that becomes null on the horizon is

$$\frac{\partial}{\partial Y_{(\beta)}} \equiv \frac{1}{2\cosh\delta_1} \left(\mp e^{-\delta_1/2} \frac{\partial}{\partial x} + e^{\delta_1/2} \frac{\partial}{\partial y} \right) . \tag{3.21}$$

In case (c), $(\alpha=0,\beta)$ is always a curvature singularity since the exponents $p_k(\beta)$ are all different from zero [Eqs. (3.17c)].

To prove the above claims (i) and (ii), we proceed exactly as we did in Ref. 6: First, we obtain the expressions of the Newman-Penrose curvature quantities (2.12) in the (α,β) coordinates. This can be done in precisely the same way as that explained in Sec. III B of Ref. 6; it gives (note that as in Ref. 6 the quantity $\alpha_{,v}$ in Eqs. (3.22) below is finite and nonvanishing as $\alpha\to 0$ [cf. Eqs. (2.16a) and (2.11b)]; consequently it can be regarded as a constant when analyzing the asymptotic behaviors of Ψ_2, Ψ_0, and Ψ_4 near the singularity)

$$\Psi_2 = \frac{e^{Q(\alpha,\beta)/2}}{8l_1l_2}\,\alpha^{1/2} \left(Q_{,\alpha\alpha} - Q_{,\beta\beta} - \frac{1}{\alpha^2} - 4i\,(V_{,\beta}W_{,\alpha} - V_{,\alpha}W_{,\beta})\cosh W \right) , \tag{3.22a}$$

$$\Psi_0 = -\frac{1}{8{l_1}^2{l_2}^2} \frac{e^{Q(\alpha,\beta)}\alpha}{{\alpha_{,v}}^2} \{ \cosh W\, [\tfrac{1}{2}(V_{,\alpha} - V_{,\beta})(Q_{,\alpha} - Q_{,\beta} + \frac{3}{\alpha})$$

$$+ V_{,\alpha\alpha} + V_{,\beta\beta} - 2V_{,\alpha\beta}\,]$$

$$+2\sinh W\,(V_{,\alpha}-V_{,\beta})(W_{,\alpha}-W_{,\beta})-i\,[\frac{1}{2}(W_{,\alpha}-W_{,\beta})(Q_{,\alpha}-Q_{,\beta}+\frac{3}{\alpha})$$

$$+W_{,\alpha\alpha}+W_{,\beta\beta}-2W_{,\alpha\beta}-(V_{,\alpha}{}^2+V_{,\beta}{}^2-2V_{,\alpha}V_{,\beta})\sinh W\cosh W\,]\,\}\,, \quad (3.22b)$$

$$\Psi_4=-\frac{1}{2}\,\alpha_{,v}{}^2\,\{\,\cosh W\,[\frac{1}{2}(V_{,\alpha}+V_{,\beta})(Q_{,\alpha}+Q_{,\beta}+\frac{3}{\alpha})$$

$$+V_{,\alpha\alpha}+V_{,\beta\beta}+2V_{,\alpha\beta}\,]$$

$$+2\sinh W\,(V_{,\alpha}+V_{,\beta})(W_{,\alpha}+W_{,\beta})+i\,[\frac{1}{2}(W_{,\alpha}+W_{,\beta})(Q_{,\alpha}+Q_{,\beta}+\frac{3}{\alpha})$$

$$+W_{,\alpha\alpha}+W_{,\beta\beta}+2W_{,\alpha\beta}-(V_{,\alpha}{}^2+V_{,\beta}{}^2+2V_{,\alpha}V_{,\beta})\sinh W\cosh W\,]\,\}\,. \quad (3.22c)$$

Next, we replace $Q_{,\alpha}$ and $Q_{,\beta}$ in Eqs. (3.22) with their values in terms of V and W given by Eqs. (2.30), and then substitute for V and W their asymptotic limits Eqs. (3.9) and (3.10) where the coefficients $c^{(i)}{}_k$ and $d^{(i)}{}_{kl}$ are obtained in terms of $\varepsilon_1, \varepsilon_2, \delta_1, \delta_2$ upon inserting Eqs. (3.9) into the field equations (2.32) [compare Eqs. (6.3.38) of Ref. 6]. Inspection of the resulting asymptotic expressions for the curvature quantities yields the following conclusions [compare Eqs. (6.3.33)–(6.3.35) of Ref. 6]:

(i) The surface $\{\alpha=0, \beta_1<\beta<\beta_2\}$ is a (connected) Killing-Cauchy horizon if and only if one of the two conditions (3.20a) or (3.20b) is satisfied throughout (β_1,β_2). When such a Killing-Cauchy horizon $\mathcal{S}$ forms, the curvature quantities Ψ_2, Ψ_0, and Ψ_4 are finite and well-behaved (but in general nonzero) through $\mathcal{S}$ as $\alpha\to 0$ at any $\beta\in(\beta_1,\beta_2)$.

(ii) Suppose the point $p\equiv(\alpha=0, \beta=\beta_0)$ does not belong to a Killing-Cauchy horizon, i.e., suppose there is no interval (β_1,β_2) containing β_0 throughout which one of the conditions (3.20a) or (3.20b) is satisfied. Then p corresponds to a curvature singularity of the colliding plane-wave spacetime except when one of the following is true at p:

$$\text{In case (a)}, \quad \varepsilon_1(\beta_0)=\pm 1\,, \quad \varepsilon_1'(\beta_0)=\varepsilon_1''(\beta_0)=0\,.$$

$$\text{In case (b)}, \quad \varepsilon_2(\beta_0)=1\,, \quad \varepsilon_2'(\beta_0)=\varepsilon_2''(\beta_0)=\delta_1'(\beta_0)=\delta_1''(\beta_0)=0\,.$$

Although under any one of the above circumstances p is not a curvature singularity (Ψ_2, Ψ_0, and Ψ_4 are finite as $\alpha\to 0$ *at* $\beta=\beta_0$), it still corresponds to a spacetime singularity since there is no topological neighborhood around p which is completely free of neighboring curvature singularities (cf. the assumption that p does not belong to a Killing-Cauchy horizon).

It has become clear in this section that the asymptotic behavior of a general colliding plane-wave spacetime near its singularity is completely characterized by the four functions $\varepsilon_1(\beta)$, $\varepsilon_2(\beta)$, $\delta_1(\beta)$, and $\delta_2(\beta)$. In contrast with Ref. 6 where the corresponding functions $\varepsilon(\beta)$ and $\delta(\beta)$ in the parallel-polarized case could be expressed explicitly in terms of initial data [Eqs. (6.3.13) and (6.3.12b)], here such expressions cannot be found in general due to the absence of a Riemann function for Eqs. (2.32). Consequently it is not in general possible to relate the asymptotic Kasner axes and exponents along the singularity $\alpha=0$ to the initial data (2.37) posed along the wavefronts of the incoming plane waves. In Appendix C, when we discuss some intriguing aspects of the field equations (2.32) which might some day prove useful in the search for a Riemann function, we also indicate an interesting special case in which one of the asymptotic structure-functions can be expressed explicitly in terms of the initial data posed by the colliding waves [Eq. (C6)].

B. Instability and nongenericity of the Killing-Cauchy horizons that occur at α=0

Our analysis in the previous section proved that whenever the "surface" $\{\alpha=0\}$ is free of Killing-Cauchy horizons, it represents a curvature singularity of the colliding plane-wave spacetime (2.31). In fact, that this must be true in general in *any* plane-symmetric spacetime is the content of a singularity theorem due to Tipler.[11] (A discussion of this theorem emphasizing its relevance to Killing-Cauchy horizons as well as to singularities can be found in Sec. III B of Ref. 12.) More precisely, Tipler's theorem proves that any nonflat, plane-symmetric spacetime in which the null convergence condition[14] holds is either null-geodesically incomplete or possesses a region where its strict plane symmetry[8,12] breaks down; i.e., the spacetime either contains singularities (where null geodesics terminate) or Killing horizons (where at least one of the plane-symmetry-generating spacelike Killing vectors becomes null).

The horizons $\mathcal{S}$ that occur in colliding plane-wave spacetimes are Killing horizons since as discussed in Sec. III A [Eqs. (3.20) and (3.21)] there exists a spacelike, constant (hence Killing) linear combination of the Killing vectors $\partial/\partial x$ and $\partial/\partial y$ which becomes null on $\mathcal{S}$. As a consequence, on $\mathcal{S}$ the Rosen-type coordinates (u,v,x,y) [and also the coordinates (α,β,x,y)] break down, developing coordinate singularities similar to those developed by (t,β,X,Y) at the surface $\{t=0\}$ of the *degenerate* (flat) Kasner solution (3.18). As another consequence of this breakdown of strict plane symmetry, the past-directed null generators of $\mathcal{S}$ (which are tangent to the Killing direction that becomes null on $\mathcal{S}$) fail to intersect the initial characteristic surface $\mathcal{N}\equiv\{u=0\}\cup\{v=0\}$; i.e., $\mathcal{S}$ is outside the domain of dependence[14] $D^+(\mathcal{N})$ of $\mathcal{N}$. In fact, it is easy to see that $\mathcal{S}$ constitutes precisely the future boundary of $D^+(\mathcal{N})$; more precisely, the Killing horizon $\mathcal{S}$ is at the same time a future *Cauchy horizon* for the initial characteristic surface $\mathcal{N}$.

It is well known that spacetime can be smoothly extended across the Killing-Cauchy horizon $\mathcal{S}$ in infinitely many different ways. The geometry of spacetime beyond $\mathcal{S}$ is not uniquely determinable by the initial data posed on $\mathcal{N}$; global predictability breaks down. Since these causal properties of the horizons $\mathcal{S}$ and their implications were discussed extensively in Sec. III C of Ref. 6 (see also Fig. 2 of Ref. 6), we will not repeat those discussions here. We will only note, as a particularly relevant implication of the breakdown of predictability, that the occurrence of horizons in the collisions of gravitational plane waves might appear to diminish the predictive power of Tipler's singularity theorem: If a horizon forms, existence of singularities cannot be proved; in fact when horizons are present the existence of singularities is false: there are examples[34] of exact solutions for nonvacuum colliding plane waves which have everywhere-nonsingular extensions beyond their Killing-Cauchy horizons.

We also recall our discussion in Ref. 6 of the strong cosmic censorship conjecture,[35,36] and of how, when suitably restricted to plane-symmetric spacetimes, it predicts the instability of the Killing-Cauchy horizons $\mathcal{S}$. These instabilities are also discussed extensively in the literature: On the one hand, there are examples of exact colliding plane-wave solutions whose horizons are destroyed and replaced by singularities when matter fields are introduced into the spacetime;[37] on the other hand, there are general theorems proving the linearized instability of arbitrary Killing-Cauchy horizons in plane-symmetric spacetimes,[8] and of compact Killing horizons in a general spacetime.[38] In fact, for the special case of the Killing-Cauchy horizons which occur in collisions of parallel-polarized plane waves, our discussions in Sec. III C of Ref. 6 prove that the instabilities render the set of horizon-producing initial data "nongeneric" with respect to a very precise notion of nongenericity. More specifically, our analysis in Ref. 6 proves that the subset of all initial data which produce at least one connected Killing-Cauchy horizon larger than Planck size is nongeneric within the set of all colliding parallel-polarized plane-wave initial data. Correspondingly, by making use of the results of Appendices A and B and of Sec. III A, we will prove below the generalization of this result (with the same notion of genericity as in Ref. 6) to the case of colliding arbitrarily-polarized plane waves. In addition, by using a more sophisticated notion of genericity described in detail in Appendix D, we will prove that the subset of *all* horizon-producing initial data (and not just the subset of those data which produce horizons larger than Planck size) is nongeneric within the set of all initial data for colliding plane waves. We will also discuss why we believe that our topological notion of genericity (described in Appendix D) is more appropriate in general relativity than other possible "probabilistic" notions based on measure theory. Note that these results (i) fully restore the predictive power of Tipler's singularity theorem: *generic* gravitational plane-wave collisions always produce "pure" spacetime singularities without Killing-Cauchy horizons, and (ii) similarly yield a proof of "plane-symmetric" strong cosmic censorship:[35,36] generic plane-symmetric gravitational initial data always evolve into *inextendible* globally-hyperbolic maximal developments. [To be more precise, our analysis proves (ii) only within the class of plane-symmetric metrics which can be brought into the form (2.3); this class includes (but is larger than) all metrics that are flat in some open set somewhere in spacetime.[4,8]]

To prove our results on the nongenericity of plane-symmetric Killing-Cauchy horizons, we proceed as follows. We first make the space D of all initial data in the form (2.37),

$$D \equiv \{ p \mid p \equiv [V(r,1), W(r,1), V(1,s), W(1,s)] \} \ , \tag{3.23}$$

a Banach space[39] completed under the norm (say)

$$\|p\| \equiv \left[\int_{-1}^{1} (|V(r,1)|^2 + |W(r,1)|^2)\, dr + \int_{-1}^{1} (|V(1,s)|^2 + |W(1,s)|^2)\, ds \right]^{1/2}$$

(the precise choice of the norm is immaterial). Similarly, the space F of all asymptotic structure-functions,

$$F \equiv \{ f \mid f \equiv [\varepsilon_1(\beta), \delta_1(\beta), \varepsilon_2(\beta), \delta_2(\beta)] \} \ , \tag{3.24}$$

can be made a Banach space after completion with respect to the norm

$$\|f\| \equiv \left[\int_{-1}^{1} (|\varepsilon_1(\beta)|^2 + |\delta_1(\beta)|^2 + |\varepsilon_2(\beta)|^2 + |\delta_2(\beta)|^2)\, d\beta \right]^{1/2}$$

(again the precise choice of the norm is unimportant). The vector space structures in both D and F are defined pointwise; thus, under the above norms both D and F are isomorphic to standard L^2 spaces.[39] We also construct the space $A \equiv \{ q \mid q \equiv [f, \sigma(\beta)] \}$ of all possible asymptotic behaviors. Here $f \in F$, and $\sigma(\beta)$ is a function with values in the (discrete) flag set $\{a, +b, -b, c\}$; the flag $\sigma(\beta)$ determines which of the four possible asymptotic behaviors described by the structure functions f and Eqs. (3.9a–c) is actually assumed by (V, W) near $\alpha=0$ and at β. Obviously, the function $\sigma(\beta)$ is not continuous in general; however it can be assumed to be Lebesgue measurable[40] on $(-1,1)$. Also, in order to have each point of A correspond to a distinct asymptotic behavior, we impose the restrictions that $\sigma(\beta) = c$ if and only if $\varepsilon_1(\beta) = \varepsilon_2(\beta) = 0$, and that $\sigma(\beta) = a$ only if $|\varepsilon_1(\beta)| = |\varepsilon_2(\beta)|$ *or* $\delta_2(\beta) = 0$, for all $q \in A$, $q = [f, \sigma(\beta)]$. We make A a complete metric space by introducing the distance function

$$d(q,q') \equiv \|f - f'\| + \int_{-1}^{1} [1 - \delta_{\sigma(\beta)\,\sigma'(\beta)}]$$

$$\times \{ \delta_{\sigma(\beta)\,a}\, \delta_{\sigma'(\beta)\,c}\, [\,|\varepsilon_1(\beta)| + |\delta'_2(\beta)|\,] + \delta_{\sigma(\beta)\,c}\, \delta_{\sigma'(\beta)\,a}\, [\,|\varepsilon'_1(\beta)| + |\delta_2(\beta)|\,]$$

$$+ \delta_{\sigma(\beta)\,\pm b}\, \delta_{\sigma'(\beta)\,c}\, [\,|\varepsilon_2(\beta)| + |\varepsilon_1(\beta)|\,] + \delta_{\sigma(\beta)\,c}\, \delta_{\sigma'(\beta)\,\pm b}\, [\,|\varepsilon'_2(\beta)| + |\varepsilon'_1(\beta)|\,]$$

$$+ \delta_{\sigma(\beta)\,a}\, \delta_{\sigma'(\beta)\,\pm b}\, [\,|\varepsilon'_2(\beta)| + |\delta'_2(\beta)|\,] + \delta_{\sigma(\beta)\,\pm b}\, \delta_{\sigma'(\beta)\,a}\, [\,|\varepsilon_2(\beta)| + |\delta_2(\beta)|\,]\,\}\; d\beta,$$

where $q \equiv [f, \sigma(\beta)]$, $q' \equiv [f', \sigma'(\beta)]$, $\delta_{\sigma \pm b} \equiv \delta_{\sigma + b} + \delta_{\sigma - b}$, $\delta_{\sigma\sigma'}$ denotes the Kronecker delta symbol, and the integral over β is the Lebesgue integral with respect to the standard Lebesgue measure on $(-1,1)$.[40] This elaborate structure of the distance function d is introduced in order to make sure that q approaches q' $[d(q,q') \to 0]$ if and only if the asymptotic behavior described by q approaches that described by q' [cf. Eqs. (3.9)].

By the global existence and uniqueness of solutions of the field equations (2.32) (Sec. III A and Appendices A and B), there exists a well-defined map

$$\mathcal{E} : D \to A \ . \tag{3.25}$$

To every $p \in D$, the map $\mathcal{E}$ assigns the unique $q \in A$ that determines the asymptotic behavior near α=0 of the global solution which evolves from p. Moreover, it follows from the global well-posedness[18,19] of the initial-value problem for (V,W) that $\mathcal{E}$ is a continuous map. [By "global well-posedness," we mean the property that solutions of the initial-value problem carry the initial data on a hypersurface Σ_1 onto the data induced on a future hypersurface Σ_2 in a continuous way; i.e. the property that solutions on compact subsets of $D^+(\Sigma_1)$ depend continuously on their initial values on Σ_1. Once global existence and uniqueness of solutions are proved (Appendix A), global well-posedness follows from standard arguments; see Ref. 18, Ref. 21, and Sec. 4.2 of Ref. 19.] We claim that the map $\mathcal{E}$ has an inverse

$$\mathcal{E}^{-1} : A \to D$$

which is also continuous. To see this, note that given $q \in A$, $q \equiv [f, \sigma(\beta)]$, we can determine a unique solution (V, W) in the following way: Using the structure functions $\varepsilon_1(\beta)$, $\delta_1(\beta)$, $\varepsilon_2(\beta)$, $\delta_2(\beta)$ provided by f, we determine the asymptotic limit (3.9) for V and W. [The ambiguity as to which Eq. (3.9) to use will be resolved by the flag $\sigma(\beta)$.] Inserting these expressions (3.9) and (3.10) of V and W into the field equations (2.32), we can compute all the coefficients $c^{(i)}{}_k$ and $d^{(i)}{}_{kl}$ of Eq. (3.10) in terms of f; this yields an asymptotic solution for (V, W). On a spacelike surface in the vicinity of α=0, this asymptotic solution induces well-posed initial data, and by global existence and uniqueness (Appendix A) these data can be evolved back onto the initial surface where they induce the desired initial data $p = \mathcal{E}^{-1}(q) \in D$. Clearly, by this construction $\mathcal{E}(p)$=q and $\mathcal{E}^{-1}[\mathcal{E}(p)]$=$p$, thus, $\mathcal{E}^{-1}$ is a genuine inverse for $\mathcal{E}$. Again by arguments based on global well-posedness of the initial-value problem for (V, W), $\mathcal{E}^{-1} : A \to D$ is a continuous map. Thus, $\mathcal{E} : D \to A$ is a homeomorphism.

In the parallel-polarized ($W \equiv 0$) case of Ref. 6, the homeomorphism $\mathcal{E}$ is known in explicit form. There, D is the Banach space of all data of the form $[V(r,1), V(1,s)]$, A is the Banach space of all pairs $[\varepsilon(\beta), \delta(\beta)]$ [which in the general case correspond to $\varepsilon_1(\beta)$ and $\delta_1(\beta)$], and $\mathcal{E}$ is the *linear* map $D \to A$ given by the integral equations (6.3.13) and (6.3.12b). (Note that in this case $A \equiv F$; i.e. no flags σ are necessary to distinguish between different cases of asymptotic behavior [in other words, in this case $\sigma(\beta) \equiv a$ and $\varepsilon_2(\beta) = \delta_2(\beta) \equiv 0$, cf. Eqs. (6.3.4)–(6.3.7)].) The inverse of $\mathcal{E}$, $\mathcal{E}^{-1}$, is defined by solving these integral equations for $V(r,1)$ and $V(1,s)$ given q=$[\varepsilon(\beta), \delta(\beta)]$.

Both $\mathcal{E}$ and $\mathcal{E}^{-1}$ are linear continuous (bounded) maps. Therefore $\mathcal{E}: D \rightarrow A$ is a continuous Banach space isomorphism. Note that the construction of an asymptotic solution V from given q is explicitly carried-out in Ref. 6 via Eqs. (6.3.7) and (6.3.38).

Now we return to the general (arbitrarily-polarized) case, and for each $\delta>0$ we define a subset H_δ of A as follows:

$$
\begin{aligned}
H_\delta \equiv \{\, [f, \sigma(\beta)] \in A \mid & \text{ there exists a connected subinterval of length } \geq \delta \\
& \text{ in } (-1,1) \text{ across which } \varepsilon_1(\beta) \equiv \pm 1 \,\} \\
\bigcup \ \{\, [f, \sigma(\beta)] \in A \mid & \text{ there exists a connected subinterval of length } \geq \delta \\
& \text{ in } (-1,1) \text{ across which } \varepsilon_2(\beta) \equiv 1 \text{ and } \delta_1(\beta) \equiv \text{const} \,\} \ . \qquad (3.26)
\end{aligned}
$$

By Eqs. (3.20), if $p \in D$ is such that the evolution of p creates at least one connected Killing-Cauchy horizon of β-length $\geq \delta$, then p must belong to $\mathcal{E}^{-1}(H_\delta)$. (See Fig. 2 of Ref. 6.) Clearly, H_δ is a nongeneric subset in the sense of Ref. 6: H_δ is closed and its complement is dense in A. Since $\mathcal{E}$ is a homeomorphism, this implies that $\mathcal{E}^{-1}(H_\delta)$ is nongeneric in D for each $\delta>0$. Taking $\delta = \delta_P \equiv l_P / \sqrt{l_1 l_2}$ where l_P is the Planck length, this proves that the set of all initial data in D which create at least one connected Killing-Cauchy horizon of larger than Planck size is a nongeneric subset [since it is contained in the nongeneric subset $\mathcal{E}^{-1}(H_{\delta_P})$]. By the same arguments as in Sec. III C of Ref. 6, this is equivalent to proving the full nonlinear instability of the Killing-Cauchy horizons at $\alpha=0$ against generic, plane-symmetric perturbations of the initial data.

Now, assuming that the reader has read through Appendix D, we consider the nongenericity of the set of *all* horizon-producing initial data. We introduce the subset

$$W \equiv \mathcal{E}^{-1}\Big(\bigcup_{\delta>0} H_\delta\Big) = \bigcup_{\delta>0} \mathcal{E}^{-1}(H_\delta)$$

of D. The set of *all* horizon-producing initial data, $W_H \subset D$, is contained in W : $W_H \subset W$ [cf. Eqs. (3.20) and (3.26)]. Since $\mathcal{E}: D \to A$ is a homeomorphism, we have: (i) each $\mathcal{E}^{-1}(H_\delta)$ is a closed set with empty interior, (ii) $\mathcal{E}^{-1}(H_{\delta_2}) \supset \mathcal{E}^{-1}(H_{\delta_1})$ [since $H_{\delta_2} \supset H_{\delta_1}$ by Eq. (3.26)] whenever $\delta_2 < \delta_1$, and (iii) $\bigcup_{\delta>0} \mathcal{E}^{-1}(H_\delta) = W$. As the Banach space D, being a complete metric space, is a Baire space, (i)–(iii) imply that the subset $W \subset D$ is thin in the sense of Appendix D. Therefore, by the definition of nongeneric subsets given in Appendix D, the subset $W_H \subset W \subset D$ of *all* horizon-producing initial data is nongeneric within the space of all plane-symmetric initial data D.

Finally, we make a few remarks on the use of the intuitive notion of genericity in theoretical physics. When physicists use the adjective "generic" they may be referring to any one of two fundamentally different intuitive notions, although the distinction is often not stated explicitly. One of these notions has an essentially *probabilistic* nature: Suppose a system (or a person/observer) chooses a set of parameters (initial conditions, integration constants, model parameters,...) out of a continuum of possibilities, and suppose there is evidence that in general the choice is made *at random*. Then the physicists' notions of "nongeneric choice" or "nongeneric outcome" would nicely correspond to the mathematical notion of "measure zero;[40]" i.e., a nongeneric choice ($\equiv$ a choice with zero probability) would be one that belongs to a subset of measure zero within the set of all choices. The second notion, on the other hand, has a *constructive* nature: Suppose there is a system or a theoretical model that is to be constructed out of a continuum of possibilities; an initial-value problem is a nice example of such a model. Here "genericity" is the issue of whether the model continues to "behave" in the same way when it is perturbed slightly away from its original form, i.e., the issue of whether the model is constructible in practice (compare the concept of "structural stability" in the theory of dynamical systems[41]). Consequently, genericity in this case is best formulated mathematically as a *topological* condition since the fundamental notions involved in its intuitive description are notions of "neighborhood," e.g. notions like "slightly perturbed," "nearby," and "stable." [In fact, the probabilistic and topological concepts of genericity are *not* compatible with each other mathematically; for example, (as has been pointed out to us by R. Geroch[42]) the unit interval admits topological homeomorphisms under which closed nowhere-dense subsets with zero Lebesgue-measure are carried onto closed nowhere-dense subsets with positive measure!] It is our view that the notion of genericity that is appropriate in general relativity, and in any other similar dynamical-evolution context, is the second *topological* notion as opposed to the more common probabilistic one. We hope that the specific topological concept of genericity discussed in Appendix D will find other useful applications in relativity besides the application that we have described in this section.

IV. SINGULARITIES IN THE COLLISIONS OF ALMOST-PLANE GRAVITATIONAL WAVES

A. A singularity result for colliding almost-plane waves whose initial data are exactly plane symmetric across a sufficiently large region of the initial surface

The content and derivation of the results of this section are so much in parallel with those of Sec. II in Ref. 9 that here we will give only the precise statements of the main conclusions, and brief comments about their derivation. To put the material of this section in proper context, we recommend that readers consult the detailed discussions in Secs. I and II of Ref. 9.

In this paper, as in Ref. 9, we will define an almost-plane wave as a gravitational wave spacetime[12] on which there exist (i) a local coordinate system (u, v, x, y), and (ii) a length scale L_T that characterizes the variation in the x, y directions of the components of geometric quantities, such that: (iii) throughout the intersection of a suitable partial Cauchy surface Σ with the wave's *central region* [which has the form $\mathcal{C}=\{\, |x|<L_T,\ |y|<L_T,\ u, v \,\}$], the metric components and other quantities are very nearly equal to the corresponding quantities for an exact plane-wave spacetime; and (iv) the curvature components fall off to zero arbitrarily (but in a manner consistent with the constraint equations on Σ) as $x^2+y^2 \to \infty$ at constant u and v. When we consider two almost-plane gravitational waves colliding on an otherwise flat background, we will always assume that the central regions of the two waves collide with each other. Then [at least in some neighborhood of the characteristic initial surface $\mathcal{N}=\mathcal{N}_1 \cup \mathcal{N}_2$ formed by the initial wave fronts $\mathcal{N}_1$, $\mathcal{N}_2$ of the colliding waves (Fig. 1)], it is possible[9,12] to set up a local coordinate system in which the conditions (ii)—(iv) above are satisfied for *both* colliding waves simultaneously; but possibly with different transverse length scales $(L_T)_1$ and $(L_T)_2$. In this coordinate system, the initial data supplied by the almost-plane wave 1 and posed on the initial null surface $\mathcal{N}_2$ are very nearly equal, throughout $\mathcal{C}_1 \cap \mathcal{N}_2$, to the initial data posed by a corresponding

exact plane wave 1; and the initial data supplied by the almost-plane wave 2 and posed on the initial null surface $\mathcal{N}_1$ are very nearly equal, throughout $\mathcal{C}_2 \cap \mathcal{N}_1$, to the initial data supplied by a corresponding exact plane wave 2. The fundamental problem of colliding almost-plane gravitational waves is then to determine whether (or under what conditions on the initial data) the evolution of these data produces spacetime singularities.

The following Lemma is proved in exactly the same way as Lemma 1 of Ref. 9; its derivation uses only the result (Secs. III A and B) that the asymptotic limit of a *generic* colliding plane-wave metric is an inhomogeneous-(nondegenerate)-Kasner solution. Restricted to the parallel-polarized case, this fact was also the only ingredient in the proof of Lemma 1 of Ref. 9.

Lemma 1: The intersection $J^-(q) \cap \mathcal{N}$ between the initial surface $\mathcal{N} \equiv \mathcal{N}_1 \cup \mathcal{N}_2$ (Fig. 1), and the causal past $J^-(q)$ of any (generic[9]) point q in the interaction region of a *generic*, arbitrarily-polarized colliding plane-wave spacetime is a compact set, whose transverse ($\equiv x, y$) dimensions approach finite limits (i.e., remain bounded from above) as the point q approaches the singularity at α=0.

In fact, when the point q has a β value sufficiently far away from the edge points β=+1 and β=−1 (e.g., for $-\frac{1}{2}<\beta<\frac{1}{2}$), β remains approximately constant along the past-directed null geodesics from q which extend farthest in the x, y directions; hence, the asymptotic limit (3.18) of the metric (with β-dependent coefficients a, b, c, and d) remains a good approximation along these geodesics. Furthermore, the coordinates x, y are constant linear combinations of $X_{(\beta)}$, $Y_{(\beta)}$, and in general at least one of the coefficients in each combination is of order 1 whereas the other may be small compared to 1 [cf. Eqs. (3.16d) and (3.17d)]. Therefore, for such a point q approaching α=0 at, say, $-\frac{1}{2}<\beta<\frac{1}{2}$, we can estimate the limits of the maximum transverse (coordinate) dimensions of $J^{-}(q)\cap\mathcal{N}$ by the quantities [compare Eqs. (9.2.4) of Ref. 9]

$$L_x(\beta)=2\left[\frac{a(\beta)}{c(\beta)}\right]^{1/2}\frac{1}{1-p_1(\beta)}\,, \tag{4.1a}$$

and

$$L_y(\beta)=2\left[\frac{a(\beta)}{d(\beta)}\right]^{1/2}\frac{1}{1-p_2(\beta)}\,, \tag{4.1b}$$

where the constants $a(\beta)$, $c(\beta)$, $d(\beta)$, and the exponents $p_1(\beta)$, $p_2(\beta)$ are found upon comparing Eq. (3.18) with either Eq. (3.15a), Eq. (3.16a), or Eq. (3.17a), depending on whether the asymptotic behavior of the metric is described by Case (a), Case (b), or Case (c), respectively. [Compare Eqs. (9.2.5) of Ref. 9.]

As in Ref. 9, Lemma 1 can be rephrased in the following equivalent form:

Lemma 1 (second version): In a generic colliding (arbitrarily-polarized) plane-wave spacetime, the singularity $\{\alpha=0\}$ represents a future c boundary,[14] whose (generic) "points" [which are "terminal indecomposable past sets" (Sec. 6.8 of Ref. 14)] intersect the initial surface $\mathcal{N}$ in subsets with compact closure. In other words, *unless* the colliding plane-wave solution possesses Killing-Cauchy horizons at $\{\alpha=0\}$ destroying its global hyperbolicity [which can only occur for "nongeneric" initial data (Sec. III B)], the (generic) points of the singularity $\{\alpha=0\}$ (when they are considered as points on the future causal boundary of spacetime) can be regarded as part of the domain of dependence $D^{+}(\mathcal{N})$ of the initial surface $\mathcal{N}$.

The following result was discussed and proved in Sec. II of Ref. 9 (see Lemma 2 and Fig. 4 of Ref. 9).

Lemma 2: Let $(\mathcal{M},g)$ be a spacetime and Σ be a partial Cauchy surface in $(\mathcal{M},g)$ on which gravitational initial data [whose development gives the metric on $D^+(\Sigma)$] are posed. Let $S\subset\Sigma$ be a closed subset, and $\mathcal{U}\subset\Sigma$ be an open subset containing S (Fig. 4 of Ref. 9). Suppose that the initial data on Σ are replaced with a new set of initial data which coincide with the original data throughout $\mathcal{U}$. Then, unless a spacetime singularity forms and penetrates into $D^+(S)$ from outside $D^+(S)$, the new solution coincides with the old solution throughout $D^+(S)$. Here $D^+(S)$ denotes the domain of dependence of S with respect to the original metric and coincides with the domain of dependence of S with respect to the new metric.

Now, introducing the quantity L defined by

$$L \equiv \inf_{-\frac{1}{2}<\beta<+\frac{1}{2}} \max\,[L_x(\beta), L_y(\beta)]\,, \tag{4.2}$$

and combining Lemma 2 with the second version of Lemma 1, it becomes clear that we have obtained a proof for the following singularity theorem.

Theorem 1: Let the initial data for two colliding almost-plane gravitational waves be identical to the initial data for two colliding arbitrarily-polarized exact plane waves throughout a region $\mathcal{C}$ in the initial surface of the form $\mathcal{C}=\{\,|x|\leq L_T, |y|\leq L_T\,\}$. Let the corresponding initial data for this plane-symmetric portion be generic so that the maximal development of the complete plane-symmetric data produces "pure" spacetime singularities at α=0 without Killing-Cauchy horizons (Sec. III B). Let these plane-symmetric initial data be represented by the point $p\equiv[V(r,1),W(r,1),V(1,s),W(1,s)]$ in the space D. Compute the image $[f,\sigma(\beta)]\equiv\mathcal{E}(p)\in A$ of p under the map $\mathcal{E}$ defined by Eq. (3.25) (see Sec. III B for notation). Using $[f,\sigma(\beta)]$, construct the quantities $L_x(\beta)$ and $L_y(\beta)$ defined by Eqs. (4.1), and the quantity L defined by Eq. (4.2). Then, if $L_T \gg L$, the evolution of the almost-plane-symmetric data produces spacetime singularities; i.e., the colliding almost-plane waves create spacetime singularities.

Clearly, singularities which are guaranteed to exist by the above theorem will have a local structure that is precisely the same as the structure of the plane-symmetric singularities; i.e., locally these singularities will be of generalized inhomogeneous-Kasner type.

Consider now the physically-interesting regime where the colliding almost-plane waves both have amplitudes small compared to unity: $h_1 \ll 1, h_2 \ll 1$. (This means that both $V(r,1)$, $V(1,s)$ and $W(r,1)$, $W(1,s)$ are small compared to 1; more precisely, the typical amplitude h for a general plane wave is defined by $h^2 \equiv h_+{}^2 + h_\times{}^2$, where h_+ and $h_\times$ are the typical magnitudes of V and W, respectively [cf. Eqs. (2.10)].) By Eqs. (6.3.12) and (6.3.13) of Ref. 6 and by the continuity of the map $\mathcal{E}$ [Eq. (3.25)], the quantities $\varepsilon_i(\beta)$ and $\delta_i(\beta)$ ($i \equiv 1, 2$) are small compared to 1 in this case. Therefore, if we can choose the initial point (u_0, v_0) [Eqs. (2.27a)] in such a way that the quantity $\mu(\beta)$ is also smaller than or of order unity, then by Eqs. (4.1), (4.2), and (3.15)–(3.17) we could conclude that $L \sim \sqrt{l_1 l_2}$. In fact, as demonstrated in the Appendix of Ref. 9, such a choice is possible: if we fix u_0 and v_0 such that $\lambda\!\!\!{}^-_1 \ll u_0 \ll f_1$, and $\lambda\!\!\!{}^-_2 \ll v_0 \ll f_2$ (where f_1, f_2 are the first focal lengths and $\lambda\!\!\!{}^-_1$, $\lambda\!\!\!{}^-_2$ are the typical wavelengths of the colliding waves), then the point (u_0, v_0) belongs to a domain in the interaction region where (i) gravity is weak (since $u_0 \ll f_1$ and $v_0 \ll f_2$), so that U and the constant additive terms in Eq. (2.33) are small compared to unity, and (ii) the integration path in Eq. (2.33) is sufficiently far away (since $u_0 \gg \lambda\!\!\!{}^-_1$ and $v_0 \gg \lambda\!\!\!{}^-_2$) from the coordinate singularities on the initial null surfaces $\{u=0\}$ and $\{v=0\}$ (Sec. II B), so that the contribution to $\mu(\beta)$ from the integrand in Eq. (2.33) (which diverges towards the coordinate singularities on these initial null surfaces) is of order unity [Eqs. (3.11)].

Moreover, with this choice for (u_0, v_0), Eqs. (2.27a) give

$$l_1 \sim f_1, \qquad l_2 \sim f_2. \tag{4.3}$$

Since by the above arguments $\mu(\beta)$ is of order 1, when combined with Eqs. (4.1), (4.2), and (3.15)–(3.17) Eqs. (4.3) finally yield the following order-of-magnitude estimate for L, valid for colliding almost-plane waves with small amplitudes:

$$L \sim \sqrt{f_1 f_2}. \tag{4.4}$$

Therefore, by Theorem 1, if the colliding almost-plane waves have small initial amplitudes and are exactly plane-symmetric across a region of size $L_T \gg \sqrt{f_1 f_2}$ over the initial surface, then their collision produces singularities. These singularities have the same (inhomogeneous-Kasner) local structure as the singularities produced by the exact-plane-wave collision.

B. Singularities produced by colliding almost-plane waves with arbitrary initial data: An existence theorem

In this section we will prove that the conclusions of Theorem 1 (Sec. IV A) about the *existence* of singularities in almost-plane-wave collisions remain valid when the colliding waves are only approximately (but not exactly) plane-symmetric throughout their central regions. More precisely, we will prove that if p is a choice of gravitational initial data on $\mathcal{N}$ that satisfies the conditions of Theorem 1 with $L_T \gg L$, then there exists a neighborhood $\mathcal{W}$ of p, open within the space of all gravitational initial data on $\mathcal{N}$, such that the Cauchy development of any data in $\mathcal{W}$ produces spacetime singularities. (For a still more precise statement see below.) Note that a proof of this statement would immediately follow if we could prove that the solutions on $D^+(\mathcal{N})$

depended *uniformly continuously* on the initial data on $\mathcal{N}$. In general this is false, because general theorems which assert the continuous dependence of solutions on initial data (such as the Cauchy stability theorem, see, e.g., Sec. 7.6 of Ref. 14) are valid with respect to the *compact-open* topology [i.e., the open topology based on convergence on *compact* subsets of $D^+(\mathcal{N})$], and not with respect to the *open* topology [i.e., the open topology based on (uniform) convergence on $D^+(\mathcal{N})$] on the spaces of all initial data on $\mathcal{N}$ and all four-metrics on $D^+(\mathcal{N})$. [We will denote by $\mathcal{D}$ and $\mathcal{G}$ these spaces of all (vacuum) initial data on $\mathcal{N}$ and all Lorentz metrics on $D^+(\mathcal{N})$, respectively, both topologized with the compact-open topology. The space $\mathcal{D}$ should not be confused with the Banach space D of all *plane-symmetric* vacuum data on $\mathcal{N}$ (Sec. III B).] To see more intuitively why uniform-continuous dependence on initial data fails, recall (i) that singularities can be thought of as points "at infinity," and (ii) that when the initial data p are slightly perturbed their development cannot remain uniformly close to the original solution all the way to infinity (i.e. all the way up to the singularities). The main content of the singularity theorem of this section lies in showing how to get around this failure of uniform-continuous dependence in the specific case of colliding almost-plane gravitational waves.

We first list three Lemmas whose corollaries will directly lead to the proof of our singularity theorem:

Lemma 3: In a nondegenerate Kasner spacetime [Eq. (3.18)], the future null cone $\dot{J}^+(q)$ of any point q starts to reconverge near the singularity $\{t=0\}$, i.e., on each future-directed null geodesic from q the convergence $\hat{\theta}$ (Sec. 4.2 of Ref. 14) of the null generators of $\dot{J}^+(q)$ becomes negative near $t=0$.

The proof of Lemma 3 is given in Appendix E.

Corollary 1: Let $p \in \mathcal{D}$ denote a choice of vacuum initial data on $\mathcal{N}$ that describes colliding almost-plane gravitational waves, and let p satisfy the conditions of Theorem 1 with $L_T \gg L$. Then, for every point q in the Cauchy development of p that lies sufficiently close to the singularity whose existence is guaranteed by Theorem 1, the future null cone $\dot{J}^+(q)$ of q starts to reconverge near $\{\alpha=0\}$; i.e., on each future null geodesic from q the convergence $\hat{\theta}$ of null generators of $\dot{J}^+(q)$ becomes negative near $\alpha=0$.

This corollary follows immediately from Lemma 3, Theorem 1, and the result (Secs. III A and B) that the asymptotic singularity structure of a generic colliding plane-wave spacetime is inhomogeneous-nondegenerate-Kasner.

Lemma 4: Let $p \in \mathcal{D}$ be vacuum initial data which satisfy the conditions of Theorem 1 with $L_T \gg L$. Then p has an open neighborhood $\mathcal{W}$ in $\mathcal{D}$ such that for any $d \in \mathcal{W}$ the maximal Cauchy development of d contains points q whose future null cones $\dot{J}^+(q)$ start to reconverge.

Proof: In the maximal development of p we can find a compact region $\mathcal{K}$ containing at least some of the points q whose null cones reconverge according to Corollary 1. Furthermore, for at least one such point q, we can obviously also arrange (without destroying the compactness of $\mathcal{K}$) that $\mathcal{K}$ contains a spherical section through the null cone $\dot{J}^+(q)$ of q at which the convergence $\hat{\theta}$ of each null generator of $\dot{J}^+(q)$ is negative. Clearly (since the topology on $\mathcal{G}$ is the compact-open topology), the maximal development of p has an open neighborhood $\mathcal{U}$ in the space of all metrics $\mathcal{G}$, such that these properties of the compact region $\mathcal{K}$ and the point $q \in \mathcal{K}$ continue to hold under any metric on $\mathcal{K}$ that comes from $\mathcal{U}$. The Einstein map, which assigns to every initial data in $\mathcal{D}$ its maximal Cauchy development in $\mathcal{G}$, is continuous by the Cauchy stability theorem.[15] Therefore, the inverse image of $\mathcal{U}$

under the Einstein map is an open subset $\mathcal{W}$ of $\mathcal{D}$, and it is easy to see that this $\mathcal{W}$ satisfies the properties required by the Lemma.

Lemma 5 [Hawking-Penrose (1970) singularity theorem[13,14]]: Spacetime is causal-geodesically incomplete if:

(i) $R_{\mu\nu} K^{\mu} K^{\nu} \geq 0$ for every non-spacelike vector $\vec{K}$.

(ii) The causal genericity condition (condition 4.4.5 of Ref. 14) is satisfied.

(iii) The chronology condition holds (there are no closed timelike curves).

(iv) There exists a point q such that on every future directed null geodesic from q the convergence $\hat{\theta}$ of the generators of $\dot{J}^{+}(q)$ becomes negative.

Lemma 5 is stated and proved as Theorem 8.2.2 in Ref. 14.

The following singularity theorem is now obtained as a direct corollary of Lemma 5:

Theorem 2: Let $p \in \mathcal{D}$ be vacuum initial data which satisfy the conditions of Theorem 1 with $L_T \gg L$. Let $\mathcal{W} \subset \mathcal{D}$ be that open neighborhood of p in $\mathcal{D}$ whose existence and properties are demonstrated in Lemma 4. Then, for any $d \in \mathcal{W}$ one of the following is true:

(a) The maximal Cauchy development of d is a maximal (inextendible[14]) spacetime. In this case, this unique inextendible spacetime satisfies conditions (i) (since the maximal development is vacuum), (ii) [cf. Eqs. (3.22)], (iii) (since the maximal development is globally hyperbolic), and (iv) (since Lemma 4 holds for the neighborhood $\mathcal{W}$) of Lemma 5, and therefore it is causal-geodesically incomplete (singular).

(b) The maximal (W^4–) Cauchy development [14,43] of d is bounded by shock waves through which spacetime is extendible but not in a smooth (W^4) way [here W^k denotes the space of metrics which belong to the Sobolev spaces $W^k(\mathcal{V})$ for all spacetime-regions $V \subset M$ with smooth boundary and compact closure; for details see Secs. 7.4 and 7.6 of Ref. 14]. It is generally believed[14,43] (although not yet proved) that in this case there will be an extension of the maximal development *through* the shock waves, which is uniquely determined by the initial data d and for which conditions (i) and (iii) of Lemma 5 are satisfied. If this is the case, then by the Cauchy stability theorem and the choice of W conditions (ii) and (iv) will also hold. Thus, if the extension is maximal (i.e. if no Cauchy horizons are encountered), then it will be an inextendible causal-geodesially incomplete (singular) spacetime by Lemma 5.

(c) The maximal (Cauchy) development of d obtained as in (a) [or the maximal development-extension obtained by maximally applying (b)] is bounded by Cauchy horizons; thereby it is extendible. [Note that these Cauchy horizons (if they occur) have nothing to do with the *Killing-Cauchy* horizons (Secs. III A and B) which are excluded a priori by the assumption (Theorem 1) that the central plane-symmetric portion of the initial data p are generic.] In this case, those extensions beyond the Cauchy horizon(s) for which conditions (i) and (iii) of Lemma 5 are everywhere satisfied [note that conditions (ii) and (iv) are always satisfied for any extension] will give maximal spacetimes which are causal-geodesically incomplete (singular) by Lemma 5. For those extensions beyond the Cauchy horizon(s) which violate conditions (i) or (iii) of Lemma 5, the incompleteness of the extended (maximal) spacetime cannot be proved.

On the other hand, if the strong cosmic censorship hypothesis[35,36] holds, then the outcome (c) above is "nongeneric;" hence

Corollary: If the strong cosmic censorship conjecture[35,36] holds (at least in vacuum) and $W \subset D$ is chosen as in Theorem 2, then the unique maximal (inextendible) spacetime obtained from the maximal Cauchy development of any "generic" initial data $d \in W$ is causal-geodesically incomplete (singular).

Combined with Eq. (4.4), Theorem 2 can be rephrased (roughly) as saying that if two colliding almost-plane waves with small initial amplitudes are sufficiently close to being exactly plane symmetric across a region of size $L_T \gg \sqrt{f_1 f_2}$ on the initial surface, then their collision produces spacetime singularities. Note that the theorem does not give any quantitative information about the "size" of the open neighborhood $\mathcal{W}$ (cf. Lemma 4); i.e., it does not indicate with what degree of accuracy the initial data of the colliding waves must approximate exact plane symmetry in order to produce singularities. Likewise, although the theorem proves the existence of the singularities rigorously, it does not give any information about either their global structure (e.g., whether they are hidden behind an event horizon[9]) or their local asymptotic behavior (e.g., whether they are of Belinsky-Khalatnikov-Lifshitz[10] generic-mixmaster type).

ACKNOWLEDGMENTS

I would like to thank Valeria Ferrari and Basilis Xanthopoulos for informative discussions. This work was partially supported by the National Science Foundation under Grant No. AST 85-14911.

APPENDIX A: PROOF OF GLOBAL EXISTENCE AND UNIQUENESS FOR SOLUTIONS OF THE FIELD EQUATIONS FOR COLLIDING PLANE WAVES

In this appendix we will study the field equations

$$V_{,\alpha\alpha}+\frac{1}{\alpha}V_{,\alpha}-V_{,\beta\beta}=2\,(V_{,\beta}W_{,\beta}-V_{,\alpha}W_{,\alpha})\tanh W\ , \tag{2.32a}$$

$$W_{,\alpha\alpha}+\frac{1}{\alpha}W_{,\alpha}-W_{,\beta\beta}=(V_{,\alpha}{}^2-V_{,\beta}{}^2)\sinh W\cosh W \tag{2.32b}$$

for colliding arbitrarily-polarized plane waves. We will prove that for *any* smooth initial data

$$\{ V(r,1), W(r,1), V(1,s), W(1,s) \} , \tag{2.37}$$

the solution (V,W) of the initial-value problem (2.32) and (2.37) exists globally and is unique throughout the domain of dependence $D^+(\mathcal{N}) = \{\alpha-\beta \leq 1, \alpha+\beta \leq 1\} \cap \{\alpha > 0\}$ of the characteristic initial surface

$$\mathcal{N} \equiv \{r=1, -1 < s \leq 1\} \cup \{s=1, -1 < r \leq 1\} , \quad r \equiv \alpha-\beta , \quad s \equiv \alpha+\beta \tag{A1}$$

on which the smooth initial data (2.37) are posed. Notice that here we regard the problem (2.32) and (2.37) as a hyperbolic initial-value problem defined on an *ordinary Euclidean space* R^2, rather than as a problem defined on the interaction region of a four-dimensional Lorentzian colliding plane-wave spacetime (2.31). In this formulation, the Euclidean R^2 on which (2.32) and (2.37) are to be solved is determined simply by the Euclidean coordinates (α,β) [or (r,s)]; the geometry of this Euclidean space and of the characteristic initial-value problem (2.32), (2.37), and (A1) are described in Fig. 2 (cf. also Fig. 1).

Before we actually prove the global existence and uniqueness of solutions for the initial-value problem (2.32), (2.37), and (A1), we will first describe how this problem can be transformed into an equivalent problem in ordinary four-dimensional Minkowski spacetime. It will turn out that the results of this appendix and also of Appendix B below are much easier to obtain for this Minkowski-space problem than the original initial-value problem described above. To explain how this equivalent problem arises, we first introduce a "fiducial" four-dimensional spacetime with the metric

$$g_M \equiv -d\alpha^2 + d\beta^2 + \alpha^2 d\xi^2 + d\eta^2 , \tag{A2}$$

and we consider the invariant wave equations

$$\Box V \equiv V^{;\mu}{}_{;\mu} = -2\, g_M(\nabla V, \nabla W)\tanh W \equiv -2\, V^{;\mu} W_{;\mu} \tanh W \ , \tag{A3a}$$

$$\Box W \equiv W^{;\mu}{}_{;\mu} = -2\, g_M(\nabla V, \nabla V)\sinh W \cosh W \equiv V^{;\mu} V_{;\mu} \sinh W \cosh W \ , \tag{A3b}$$

defined on this fiducial background (A2). When written explicitly in terms of the (α,β,ξ,η) coordinates, Eqs. (A3) take the form

$$-V_{,\alpha\alpha} - \frac{1}{\alpha}V_{,\alpha} + V_{,\beta\beta} + \frac{1}{\alpha^2}V_{,\xi\xi} + V_{,\eta\eta} = -2\,(V_{,\beta}W_{,\beta} + \frac{1}{\alpha^2}V_{,\xi}W_{,\xi}$$

$$+ V_{,\eta}W_{,\eta} - V_{,\alpha}W_{,\alpha})\tanh W, \tag{A4a}$$

$$-W_{,\alpha\alpha} - \frac{1}{\alpha}W_{,\alpha} + W_{,\beta\beta} + \frac{1}{\alpha^2}W_{,\xi\xi} + W_{,\eta\eta} = (V_{,\beta}{}^2 + \frac{1}{\alpha^2}V_{,\xi}{}^2$$

$$+ V_{,\eta}{}^2 - V_{,\alpha}{}^2)\sinh W \cosh W, \tag{A4b}$$

and when compared with Eqs. (2.32) they immediately show that the solutions $V(\alpha,\beta)$, $W(\alpha,\beta)$ of the field equations (2.32) correspond *precisely* to the (ξ,η)–independent solutions (V, W) of the invariant wave equations (A3). The advantage of introducing the fiducial spacetime (A2) now becomes clear after one realizes that the metric (A2) is in fact *flat*: By introducing the new coordinates

$$T = -\alpha\cosh\xi\ , \quad X = -\alpha\sinh\xi\ , \quad Y = \eta\ , \quad Z = \beta \tag{A5}$$

in terms of which

$$\alpha = (T^2 - X^2)^{1/2}\ , \quad \beta = Z\ , \quad \xi = \tanh^{-1}(X/T)\ , \quad \eta = Y\ , \tag{A6}$$

we find that

$$g_M \equiv -d\alpha^2 + d\beta^2 + \alpha^2 d\xi^2 + d\eta^2 = -d\,T^2 + dX^2 + dY^2 + dZ^2 \,. \tag{A7}$$

In fact, a short computation using Eqs. (A5) and (A6) gives

$$\frac{\partial}{\partial\alpha} = \frac{1}{\alpha}\left[T\frac{\partial}{\partial T} + X\frac{\partial}{\partial X}\right], \quad \frac{\partial}{\partial\beta} = \frac{\partial}{\partial Z},$$

$$\frac{\partial}{\partial\xi} = X\frac{\partial}{\partial T} + T\frac{\partial}{\partial X}, \quad \frac{\partial}{\partial\eta} = \frac{\partial}{\partial Y}\,. \tag{A8}$$

Therefore, the spacetime (A2) is precisely the wedge $\{\,|T| > |X|\,, T < 0\}$ in Minkowski space, and (α,β,ξ,η) are the usual wedge coordinates, tuned to the plane-symmetric structure on the wedge that arises due to the presence of the Killing vectors $\partial/\partial Y = \partial/\partial\eta$ (which generates translations) and $X\,\partial/\partial T + T\,\partial/\partial X = \partial/\partial\xi$ [which generates (spacelike) Lorentz boosts] (see Sec. I of Ref. 8 for a more detailed discussion of the geometry of this wedge region). The invariant wave equations (A3) can now be rewritten in the form

$$\Box V = -2\,\nabla V \cdot \nabla W \,\tanh W\,, \tag{A9a}$$

$$\Box W = (\nabla V)^2 \sinh W \cosh W\,, \tag{A9b}$$

where $(\nabla V)^2 \equiv \nabla V \cdot \nabla V$, and "$\Box$" and "$\cdot$" denote the usual wave operator and the usual Lorentzian inner product on Minkowski spacetime, respectively. The term "invariant wave equations" for Eqs. (A9) or (A3) expresses the fact that if (V,W) is any solution to Eqs. (A9) then $(V\circ\phi\,, W\circ\phi)$ is also a solution, where ϕ is any isometry (i.e., any Poincare transformation) on the (flat) spacetime (A7); that is, isometries of the spacetime leave the solutions invariant. This in particular implies that solutions of Eqs. (A9) are mapped onto solutions under translations along ξ and η [$\equiv$ boosts along X and translations along Y; see Eqs. (A8)].

Notice that we have now obtained a complete reformulation of the initial-value problem Eqs. (2.32), (2.37), and (A1): (i) Instead of the solutions $V(\alpha,\beta)$, $W(\alpha,\beta)$ of Eqs. (2.32), we deal with the plane-symmetric [$\equiv$ (ξ,η)–independent] solutions of the invariant wave equations (A9) on the Minkowski wedge $\{\, |T| > |X|\,, T<0\}$. We write these nonlinear wave equations (A9) in the form

$$V_{,kk} - V_{,TT} = -2\,(V_{,k}\,W_{,k} - V_{,T}\,W_{,T}\,)\tanh W\ , \tag{A10a}$$

$$W_{,kk} - W_{,TT} = (V_{,k}\,V_{,k} - V_{,T}\,V_{,T}\,)\sinh W\,\cosh W\ , \tag{A10b}$$

where $x^k \equiv x^1, x^2, x^3 \equiv X\,, Y\,, Z$, and we adopt the summation convention that repeated spacelike orthonormal indices $k\,, l\,, m\,,...$ are summed over regardless of whether they are contracted or not. (ii) Instead of posing the initial data for (V, W) in the form (2.37) and (A1), we pose plane-symmetric [$\equiv$ (ξ,η)–independent] initial data for Eqs. (A9) [or equivalently for Eqs. (A10)] on the characteristic initial surface

$$\begin{aligned} \mathcal{C} \equiv & \{\, (T^2 - X^2)^{1/2} - Z = 1, T<0,\ 0<(T^2-X^2) \leq 1\,\} \\ & \cup \{\, (T^2 - X^2)^{1/2} + Z = 1, T<0,\ 0<(T^2-X^2) \leq 1\,\}\ . \end{aligned} \tag{A11}$$

The surface $\mathcal{C}$ is a null hypersurface in the fiducial Minkowski spacetime (A7); in fact $\mathcal{C}$ is generated by null geodesics that are orthogonal to the spacelike two-surface $\mathcal{Z} \equiv \{\alpha = (T^2 - X^2)^{1/2} = 1,\ \beta = Z = 0\}$ inside the Minkowski wedge, i.e., by those null generators of $\dot{J}^+(\mathcal{Z})$ that have their past endpoints on $\mathcal{Z}$. [The readers can see without much difficulty that in the three-dimensional Minkowski space where the Y dimension is absent, $\mathcal{C}$ (where it is a two-dimensional hypersurface $\equiv {}^{(3)}\mathcal{C}$) would be made up of two symmetrically-configured half-null-cones intersecting each other at $\mathcal{Z}$; the apex of each half-null-cone would lie on the crease $\{X = T = 0\}$ of the horizon $\{\, |T| = |X|, T \leq 0\}$. The surface $\mathcal{C}$ in the four-dimensional case (a three-dimensional null hypersurface) is obtained by just sweeping this two-dimensional ${}^{(3)}\mathcal{C}$ through spacetime parallel to the Y direction.] The two-dimensional (with Z and Y directions suppressed) geometry of this initial-value problem is depicted in Fig. 3. From the invariant character of the nonlinear wave equations (A9), it immediately follows that once we prove the global existence and uniqueness of solutions for Eqs. (A9) with arbitrary initial data posed on an arbitrary initial surface in Minkowski spacetime, this would automatically prove the global existence and uniqueness of solutions for the initial-value problem (2.32), (2.37), and (A1). In particular, when plane-symmetric [$\equiv$ (ξ,η)–independent] initial data for (V,W) are posed on $\mathcal{C}$, the unique global solution (V,W) of the above initial-value problem (A9)–(A11) would be everywhere independent of (ξ,η) (i.e., it would be everywhere plane-symmetric); these functions $V(\alpha,\beta)$ and $W(\alpha,\beta)$ would therefore constitute the unique global solution of Eqs. (2.32) corresponding to initial data (2.37) that have the same functional form as the data posed on $\mathcal{C}$ [expressed in (α,β) or (r,s) coordinates].

The introduction of the fiducial four-dimensional Minkowski space (A7) has transformed the problem (2.32), (2.37), and (A1) into a problem in ordinary flat spacetime. [Note that this fiducial flat space (A7) is entirely "fictitious," i.e., there is no geometric relationship between the spacetime (A7) and the colliding plane-wave spacetime (2.31).] More specifically, by embedding the two-dimensional hyperbolic initial-value problem (2.32), (2.37), and (A1) in a higher-dimensional flat space (from where it is recovered under the restriction of plane symmetry), we have eliminated the singular terms involving $1/\alpha$ from Eqs. (2.32) [compare Eqs. (2.32) with Eqs. (A9)].

The focusing effect described by these singular terms of Eqs. (2.32) has been transformed, in the new formulation (A9)–(A11), into the geometric effect of the exact plane symmetry imposed on the initial data. More precisely, the domain of dependence of the new initial surface $\mathcal{C}$ [Eq. (A11)] is (cf. Fig. 3)

$$D^+(\mathcal{C}) = \{\ |T| > |X|, T < 0\ \} \cap J^+(\mathcal{C})\,. \tag{A12}$$

In particular, the horizon $\{|T| = |X|, T \leq 0\} \equiv \{\alpha = 0\}$ of the Minkowski wedge is the future Cauchy horizon $H^+(\mathcal{C})$ of $\mathcal{C}$ [more precisely, $H^+(\mathcal{C}) = \{|T| = |X|, T \leq 0\} \cap J^+(\mathcal{C})$]; in fact the region $J^-(q) \cap \mathcal{C}$ becomes unboundedly large in the ξ direction as any arbitrary point q of the wedge approaches the horizon (Fig. 3). As a result, when the initial data posed on $\mathcal{C}$ have a plane-symmetric [(ξ,η)–independent] structure, the data "seen" by any field point q become infinitely extended in the ξ direction as q approaches the horizon $\{\alpha=0\}$. This effect in the formalism (A9)–(A11) is the geometric counterpart of the focusing effect caused by the singular $1/\alpha$ terms in Eqs. (2.32). In particular, it now becomes very clear why the solutions (V,W) of Eqs. (2.32) in general develop singularities at $\alpha=0$ (Sec. III A): The global existence of solutions for the initial-value problem (A9)–(A11) (which we will prove below) guarantees that (V,W) are smooth throughout the domain of dependence $D^+(\mathcal{C})$ of the initial surface $\mathcal{C}$, but not necessarily on $\mathcal{C}$'s Cauchy horizon $\{|T| = |X|, T \geq 0\} \equiv \{\alpha = 0\}$ where the field points are influenced by an infinitely large sector of the initial data (Fig. 3).

In the remaining paragraphs of this appendix we will explain how the global existence and uniqueness of solutions for the system (A9)–(A11) are proved. We remark that the above-discussed specific technique of "resolving" the singularities (i.e., the $1/\alpha$ terms) of the system (2.32), (2.37), and (A1) by embedding it into a higher dimensional problem [Eqs. (A9)–(A11)] might prove useful more generally, i.e., in studying other PDE with similar singular coefficients. (Note also that this technique is quite similar to the well-known method of "resolution of singularities" frequently used in the qualitative theory of ordinary differential equations; see, for example, Refs. 41 and 44.)

We now turn to the proof of global existence for Eqs. (A9). The proof of *local existence and uniqueness* for any nonlinear hyperbolic system of the kind Eqs. (A9) is standard and can be found, among other places, in Sec. VI.6 of Ref. 21, and in Refs. 20, 22, 23, and 26. This local result can be stated as follows:

LE: Let Σ be any regular partial Cauchy surface (or a characteristic initial surface consisting of two intersecting null surfaces) in Minkowski space, and let $\{V_0, \dot{V}_0, W_0, \dot{W}_0\}$ ($\{V_0, W_0\}$) be regular initial data for Eqs. (A9) on Σ. Then, there exist a neighborhood $\mathcal{U}$ of Σ, and unique functions (V,W) defined on $\mathcal{U}$ which satisfy Eqs. (A9) on $\mathcal{U}$ and which induce the given initial data on Σ. If the data and Σ are C^∞, then (V,W) are C^∞ on $\mathcal{U}$.

In general, *global existence* for a *nonlinear* system of hyperbolic PDE of the kind Eqs. (A9) is false; see Refs. 22–32 and Secs. I and III A of this paper. Global existence means, in more precise terms, the following:

GE: Let Σ be any regular partial Cauchy surface in Minkowski space, and let $\{V_0, \dot{V}_0, W_0, \dot{W}_0\}$ be regular initial data for Eqs. (A9) on Σ. Then, there exist unique functions (V,W) defined *throughout the domain of dependence* $D^+(\Sigma)$ *of* Σ, which satisfy Eqs. (A9) on $D^+(\Sigma)$ and which induce the given initial data on Σ. If the data and Σ are C^∞, then (V,W) are C^∞ on $D^+(\Sigma)$.

From the recent work of Klainerman,[23,28,31] Shatah, [25] Sideris, [29] Klainerman and Ponce, [27] and Christodoulou,[32] we know that nonlinear wave equations of the type

Eqs. (A9) have global solutions for *small* initial data. More precisely:

GE for small initial data: Let Σ be any regular partial Cauchy surface and d be regular initial data for Eqs. (A9) on Σ. If d is small, ie., if the Sobolev norm $\| d \|$ of $\{V_0, \dot{V}_0, W_0, \dot{W}_0\}$ in some suitable Sobolev space[21,39] $W^{k,2}(\Sigma)$ is sufficiently small, then the conclusions of GE above are true for Σ and d.

Now, in order to prove GE for Eqs. (A9) for arbitrary Σ and arbitrary initial data, it is sufficient to prove the following reduced global existence result:

RGE: Let arbitrary regular initial data d for Eqs. (A9) be posed on $\Sigma \equiv \{T=0\}$, and let d be compact supported in an open ball $S_0 \subset \Sigma$ in Σ. [More precisely, $S_0 \equiv \{(X_i, 0) \mid (X_i - Y_i)(X_i - Y_i) < R^2\}$, for some fixed $(Y_i, 0)$ in $\Sigma \equiv \{T=0\}$, and $R > 0$.] Then, solutions (V, W) exist which are defined and satisfy Eqs. (A9) throughout the interior $D^+(S_0)$ of the null cone $H^+(\overline{S_0})$, and which coincide with the data d on Σ. These functions (V, W) are *unique*, and they are C^∞ in $D^+(S_0)$ if the initial data d are C^∞.

For Eqs. (A9), RGE implies the more general GE because the characteristics are independent of the specific solution (V, W): the characteristics of Eqs. (A9) are always fixed to be the null hypersurfaces of Minkowski spacetime. Thus, given an arbitrary partial Cauchy surface Σ and arbitrary data d on it, for any point $q \in D^+(\Sigma)$ we can apply the construction described in Fig. 4(a), and introduce a $\{T=0\}$ surface [with some suitable Lorentz coordinates (T, X, Y, Z)] in the vicinity of the compact region $J^-(q) \cap \Sigma$. This reduces the global existence problem for Σ to the problem of RGE, provided the data on Σ can be transferred onto $\{T=0\}$ by means of LE. If this fails, then we iteratively apply the construction described in Fig. 4(a) to the points of $\{T=0\}$ [Fig. 4(b)], and we continue this iteration until the new smaller $\{T=0\}$ surfaces fall into that small neighborhood of Σ on which local existence is guaranteed by

LE. Tracing our steps backwards by means of RGE after this last step is achieved, we see that the data on Σ can indeed be transferred to the *first* $\{T=0\}$ surface depicted in Fig. 4(a).

Remark 1: Once RGE and hence GE are proved as we will do below, then it follows that GE also holds when Σ is a characteristic initial surface consisting of two intersecting null hypersurfaces. This is because for a characteristic Σ and d we can apply the construction described in Fig. 5, and transfer the data d on Σ onto a space-like hypersurface Σ' which lies in that neighborhood of Σ where local existence is guaranteed by LE. Since global existence and uniqueness hold for Σ' and d', they consequently hold for Σ and d (see Fig. 5).

Remark 2: Here we will prove only the existence part of RGE; once existence is proved, global uniqueness follows from standard arguments as in Refs. 22 and 26.

Proof of RGE for Eqs. (A9): This proof uses three fundamental ingredients:

(i) *Conserved positive-definite energy form for Eqs. (A9):* One of the most intriguing and special properties of Eqs. (A9) is that they can be derived from a simple Lagrangian: Introducing the Lagrange density

$$\mathcal{L} \equiv -\frac{1}{2}\cosh^2 W\,(V^{,\mu}V_{,\mu}) - \frac{1}{2}W^{,\mu}W_{,\mu}\,, \tag{A13}$$

where $x^\mu \equiv x^0, x^1, x^2, x^3 \equiv T, X, Y, Z$ and Greek indices $\mu, \nu, \rho, \ldots$ take the values 0, 1, 2, 3, it is easily seen that the Euler-Lagrange equations

$$\frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}}{\partial V_{,\mu}} - \frac{\partial \mathcal{L}}{\partial V} = 0\,, \qquad \frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}}{\partial W_{,\mu}} - \frac{\partial \mathcal{L}}{\partial W} = 0\,, \tag{A14}$$

when applied to $\mathcal{L}$ of Eq. (A13), yield precisely the nonlinear invariant wave equations (A9a) and (A9b). Consequently, we can define a conserved stress-energy tensor

$$T_{\mu\nu} \equiv \eta_{\mu\nu} \mathcal{L} - \frac{\partial \mathcal{L}}{\partial V^{,\mu}} V_{,\nu} - \frac{\partial \mathcal{L}}{\partial W^{,\mu}} W_{,\nu} \ , \tag{A15}$$

which satisfies

$$T^{\mu\nu}{}_{,\nu} \equiv 0 \ . \tag{A16}$$

When combined with Eq. (A13), Eq. (A15) gives

$$T_{\mu\nu} = \cosh^2 W \, V_{,\mu} V_{,\nu} + W_{,\mu} W_{,\nu}$$

$$- \frac{1}{2} \eta_{\mu\nu} (\cosh^2 W \, V^{,\rho} V_{,\rho} + W^{,\rho} W_{,\rho}) \ . \tag{A17}$$

Therefore, the positive-definite energy form

$$T_{TT} = \frac{1}{2} [\cosh^2 W \, (V_{,k} V_{,k} + V_{,T}{}^2) + W_{,k} W_{,k} + W_{,T}{}^2] \tag{A18}$$

has the *conservation* property:

$$T_{TT,T} - T_{Tk,k} = 0 \ . \tag{A19}$$

Consequently, when compact-supported initial data for (V, W) are posed on the initial surface $\{T = \text{const} \equiv \tau\}$, the positive-definite conserved energy form Eq. (A18) satisfies

$$\frac{\partial}{\partial \tau} \int_{\{T = \tau\}} T_{TT} \, d^3 X = 0$$

$$= \frac{\partial}{\partial \tau} \int_{\{T = \tau\}} \frac{1}{2} [\cosh^2 W \, (V_{,T}{}^2 + V_{,k} V_{,k}) + W_{,T}{}^2 + W_{,k} W_{,k}] \, d^3 X \equiv 0 \tag{A20}$$

for all $\tau \geq T$.

(ii) *Energy inequality for Eqs. (A9)*: If the initial-value problem for Eqs. (A9) is posed as in the satetement of RGE (see above), then, combined with the positive-definiteness of T_{TT} , Eq. (A20) yields (consult Fig. 6 for a description of the relevant geometry)

$$\int_{S_T} \frac{1}{2} [\cosh^2 W \, (V_{,T}{}^2 + V_{,k} V_{,k}) + W_{,T}{}^2 + W_{,k} W_{,k}] d^3X$$

$$\leq \int_{S_0} \frac{1}{2} [\cosh^2 W \, (V_{,T}{}^2 + V_{,k} V_{,k}) + W_{,T}{}^2 + W_{,k} W_{,k}] d^3X \ , \tag{A21}$$

for all $T > 0$. [Here S_τ denotes the compact set $\{T = \tau\} \cap \overline{D^+(S_0)}$ (Fig. 6).]

(iii) *Independence of the characteristics of Eqs. (A9) from the solutions*: As we have noted before, for any solution (V, W) the characteristic surfaces of Eqs. (A9) are fixed to be the null hypersurfaces of Minkowski spacetime; i.e., they are independent of the solution.

Now, the proof of RGE follows from the following arguments:

The conservation property Eq. (A20) of the energy form Eq. (A18) implies that the $W^{1,2}$ Sobolev-norm of the initial data is conserved; hence the solution does not deteriorate in the L^2–sense. However, this fact by itself is not sufficient to prove RGE: the estimates for the "life-span" of solutions of nonlinear hyperbolic PDE in general depend on the norm of the data d in higher-order Sobolev spaces than $W^{1,2}$; e.g., they depend on the norm $\| d \|$ in $W^{k,2}(S_0)$ where $k \geq 5$. [See Refs. 22, 23, 26, and 27.] Nevertheless, the (standard) argument outlined in the following paragraph [which uses all three ingredients (i)–(iii) above] suffices to prove RGE:

When the initial data posed on S_0 are analytic, it follows from the Cauchy-Kovalewski theorem[18] that there exists a local analytic solution, determined by an explicit, convergent power series. As is shown in Ref. 20, the fact (iii) above and the energy inequality (A21) imply that in fact this unique analytic solution exists *globally* throughout $D^+(S_0)$. Now, for smooth but nonanalytic initial data d, one approximates d by a series of analytic data d_n; $d_n \to d$ as $n \to \infty$. When combined with (iii), the energy inequality (A21) then shows that the corresponding global analytic solutions (V_n, W_n) in $D^+(S_0)$ *converge* (in $W^{1,2}$) to a smooth global solution (V, W); these limits of V_n and W_n throughout $D^+(S_0)$ constitute the unique global solution of Eqs. (A9) with initial data d. The most crucial step of this proof lies in showing the convergence of the series of analytic solutions (V_n, W_n) throughout $D^+(S_0)$; the energy inequality (A21) is essential for doing so. For the details, consult Ref. 20, Ref. 30, and Sec. VI.5 of Ref. 21.

APPENDIX B: PROOF THAT THE SPATIAL-DERIVATIVE TERMS IN THE FIELD EQUATIONS FOR COLLIDING PLANE WAVES ARE ASYMPTOTICALLY NEGLIGIBLE NEAR $\alpha = 0$

In this appendix, we will prove that the (global) solutions $V(\alpha,\beta)$, $W(\alpha,\beta)$ of the field equations (2.32) have the same asymptotic behaviors near $\alpha = 0$ as the solutions of the ordinary differential equations

$$V_{,\alpha\alpha} + \frac{1}{\alpha} V_{,\alpha} + 2V_{,\alpha} W_{,\alpha} \tanh W = 0 , \tag{3.1a}$$

$$W_{,\alpha\alpha} + \frac{1}{\alpha} W_{,\alpha} - V_{,\alpha}{}^2 \sinh W \cosh W = 0 \tag{3.1b}$$

which are obtained from Eqs. (2.32) by ignoring all terms with β-derivatives. As in Appendix A, we will find that working exclusively with the standard problem (2.32), (2.37), and (A1) is not terribly useful, and we will work instead with the equivalent plane-symmetric [$\equiv(\xi,\eta)$–independent] Minkowski-space initial-value problem given by Eqs. (A9)–(A11).

We begin by introducing the following differential operators $\Lambda_\alpha^{(i)}[V,W]$ and $\Lambda_\beta^{(i)}[V,W]$ $(i\equiv 1,2)$, which are well-behaved *throughout* the Minkowski spacetime $\mathcal{M}$ of Eq. (A7) and which act on smooth functions (V,W) defined on $\mathcal{M}$:

$$\Lambda_\alpha^{(1)}[V,W]\equiv(T\partial_T+X\partial_X)^2V-(X\partial_T+T\partial_X)^2V$$

$$+2\,[\,(T\partial_T+X\partial_X)V\;(T\partial_T+X\partial_X)W$$

$$-(X\partial_T+T\partial_X)V\;(X\partial_T+T\partial_X)W\,]\tanh W\,, \tag{B1a}$$

$$\Lambda_\beta^{(1)}[V,W]\equiv-[\,V_{,YY}+V_{,ZZ}+2(V_{,Y}W_{,Y}+V_{,Z}W_{,Z})\tanh W\,]\,, \tag{B1b}$$

$$\Lambda_\alpha^{(2)}[V,W]\equiv(T\partial_T+X\partial_X)^2W-(X\partial_T+T\partial_X)^2W$$

$$+\{\,[(X\partial_T+T\partial_X)V]^2-[(T\partial_T+X\partial_X)V]^2\,\}\sinh W\cosh W\,, \tag{B2a}$$

$$\Lambda_\beta^{(2)}[V,W]\equiv-[\,W_{,YY}+W_{,ZZ}-(V_{,Y}{}^2+V_{,Z}{}^2)\sinh W\cosh W\,]\,, \tag{B2b}$$

where ∂_{x^μ} denotes the differential operator $\partial/\partial x^\mu$. Comparing Eqs. (B1) and (B2) with Eqs. (A4) and using Eqs. (A8), it is easy to see that *throughout the open wedge region* $\Lambda\equiv\{\,|T|>|X|,\,T<0\}$ in Minkowski space (where $\alpha>0$), the differential operators $\Lambda_\alpha^{(i)}[V,W]$ and $\Lambda_\beta^{(i)}[V,W]$ satisfy

$$\Lambda_\alpha^{(1)}[V,W]=\alpha^2\left[\,V_{,\alpha\alpha}+\frac{1}{\alpha}V_{,\alpha}-\frac{1}{\alpha^2}V_{,\xi\xi}+2\left(V_{,\alpha}W_{,\alpha}-\frac{1}{\alpha^2}V_{,\xi}W_{,\xi}\right)\tanh W\right]\,,$$

$$\Lambda_\beta^{(1)}[V,W]=-[\,V_{,\beta\beta}+V_{,\eta\eta}+2(V_{,\beta}W_{,\beta}+V_{,\eta}W_{,\eta})\tanh W\,]\,,$$

$$\Lambda_\alpha^{(2)}[V,W]=\alpha^2\left[\,W_{,\alpha\alpha}+\frac{1}{\alpha}W_{,\alpha}-\frac{1}{\alpha^2}W_{,\xi\xi}+\left(\frac{1}{\alpha^2}V_{,\xi}{}^2-V_{,\alpha}{}^2\right)\sinh W\cosh W\right]\,,$$

$$\Lambda_\beta^{(2)}[V,W] = -[\,W_{,\beta\beta} + W_{,\eta\eta} - (V_{,\beta}{}^2 + V_{,\eta}{}^2)\sinh W \cosh W\,]\,. \tag{B3}$$

It therefore becomes clear from Eqs. (A4) and (B3) that throughout the wedge Λ the invariant wave equations (A9) for V and W can be written in the form

$$\frac{1}{\alpha^2}\,\Lambda_\alpha^{(i)}[V,W] + \Lambda_\beta^{(i)}[V,W] = 0\,. \tag{B4}$$

On the other hand, if we introduce the differential operators

$$\angle_\alpha^{(1)}[V,W] \equiv V_{,\alpha\alpha} + \frac{1}{\alpha}V_{,\alpha} + 2V_{,\alpha}W_{,\alpha}\tanh W\;, \tag{B5a}$$

$$\angle_\beta^{(1)}[V,W] \equiv -2V_{,\beta}W_{,\beta}\tanh W - V_{,\beta\beta}\,, \tag{B5b}$$

$$\angle_\alpha^{(2)}[V,W] \equiv W_{,\alpha\alpha} + \frac{1}{\alpha}W_{,\alpha} - V_{,\alpha}{}^2\sinh W\cosh W\;, \tag{B6a}$$

$$\angle_\beta^{(2)}[V,W] \equiv V_{,\beta}{}^2\sinh W\cosh W - W_{,\beta\beta}\,, \tag{B6b}$$

which are well-behaved *throughout the open wedge* Λ but which are *singular* (in fact undefined) outside it, then we can rewrite the field equations (2.32) in the form

$$\angle_\alpha^{(i)}[V,W] + \angle_\beta^{(i)}[V,W] = 0\,, \tag{B7}$$

with the additional restriction that the solutions V and W must be plane symmetric, i.e. independent of (ξ,η).

Now consider given plane-symmetric initial data $\{V_0^{(\infty)}, W_0^{(\infty)}\}$ posed on the initial surface $\mathcal{C}$ of Eq. (A11) (see also Fig. 3). (The rationale for our notation will become clear in a moment.) For any $L > 0$, we construct a new set of initial data on $\mathcal{C}$ by the relations

$$V_0^{(L)} \equiv V_0^{(\infty)} f^{(L)}(\xi^2+\eta^2)\,, \qquad W_0^{(L)} \equiv W_0^{(\infty)} f^{(L)}(\xi^2+\eta^2)\,, \tag{B8a}$$

where $f^{(L)}(u)$ is a family of smooth functions in $C^\infty(R)$ satisfying (for each $L>0$)

$$f^{(L)}(u) = 1 \quad \text{for } u \le L^2\,, \qquad f^{(L)}(u) = 0 \quad \text{for } u \ge 4L^2\,,$$

$$\frac{d}{du} f^{(L)}(u) \le 0 \quad \forall\, u \in R\,. \tag{B8b}$$

In other words, the initial data $\{V_0^{(L)}, W_0^{(L)}\}$ are obtained by smoothly cutting off the plane-symmetric initial data $\{V_0^{(\infty)}, W_0^{(\infty)}\}$ at a distance $2L$ in the ξ and η directions. [The existence of smooth functions f with the properties (B8b) is a well-known result in elementary analysis; see, e.g., Lemma 1.10 of Ref. 17.] By Appendix A, for each $L>0$ there exists a global solution $(V^{(L)}, W^{(L)})$ of Eqs. (A9) [or equivalently of Eqs. (B4)] which is defined throughout the wedge Λ and which evolves from the initial data (B8) on $\mathcal{C}$. We claim that for any *finite* $L>0$, these solutions $(V^{(L)}, W^{(L)})$ are in fact smooth and well-behaved *on and across the Cauchy horizon* $H^+(\mathcal{C}) = \{\,|T| = |X|, T \le 0\} = \{\alpha = 0\}$ of $\mathcal{C}$. To see this, consider the construction depicted in Fig. 7: This figure describes how we build a new initial surface Π by (i) choosing an $R>0$ with $R>2L$, (ii) adjoining a smooth spacelike hypersurface Σ to the initial surface $\mathcal{C}$ at the cylindrical cross-section $\mathcal{C} \cap \{\xi^2+\eta^2 = R^2\}$ *through* $\mathcal{C}$, and finally (iii) discarding the portion of $\mathcal{C}$ that remains in the past of Σ (Fig. 7). [Note that the geometry described in Fig. 7 is fully symmetric in the ξ and η directions; consequently, the three-dimensional picture of Π with the β ($\equiv Z$) direction suppressed can be obtained by rotating Fig. 7 around the T axis.] On the new initial surface Π, we pose new initial data d for (V, W) by leaving the data as they are on $\mathcal{C}$ [i.e. $\{d \text{ on } \mathcal{C}\} \equiv \{V_0^{(L)}, W_0^{(L)}\}$] and by putting $d \equiv 0$ on Σ. Inspection of Fig. 7 makes it clear that throughout the subset of the wedge Λ that corresponds to the dotted region in Fig. 7, the global solution of the initial-value problem (Π, d) for Eqs. (A9) [or for Eqs. (B4)] is precisely equal to the solution $(V^{(L)}, W^{(L)})$. Moreover, it is also obvious from Fig. 7 that the domain of dependence of Π includes the horizon $H^+(\mathcal{C})$ as well as the region that lies beyond the horizon. Therefore, since by Appendix A the solution of the initial-value problem (Π, d) exists smoothly throughout $D^+(\Pi)$, we conclude that the solution $(V^{(L)}, W^{(L)})$ of the problem $(\mathcal{C}, d^{(L)})$ is also smooth at and across the horizon $H^+(\mathcal{C})$.

The following identities are now easily derived from Eqs. (B3), (B5), (B6), and (B8):

$$\lim_{L\to\infty} V^{(L)} = V^{(\infty)}, \qquad \lim_{L\to\infty} W^{(L)} = W^{(\infty)}. \tag{B9}$$

$$\angle_\alpha{}^{(i)}[V^{(\infty)}, W^{(\infty)}] = \lim_{L\to\infty} \frac{1}{\alpha^2} \Lambda_\alpha{}^{(i)}[V^{(L)}, W^{(L)}], \tag{B10a}$$

$$\angle_\beta{}^{(i)}[V^{(\infty)}, W^{(\infty)}] = \lim_{L\to\infty} \Lambda_\beta{}^{(i)}[V^{(L)}, W^{(L)}]. \tag{B10b}$$

By Eq. (B4), the nonlinear wave equations (A9) satisfied by $V^{(L)}$ and $W^{(L)}$ can be written in the form

$$\Lambda_\alpha{}^{(i)}[V^{(L)}, W^{(L)}] = -\alpha^2 \Lambda_\beta{}^{(i)}[V^{(L)}, W^{(L)}]. \tag{B11}$$

Since $\Lambda_\beta{}^{(i)}[V, W]$ are smooth differential operators well-behaved throughout Minkowski spacetime, and since by the above paragraph $V^{(L)}$ and $W^{(L)}$ are also well-behaved on and across the horizon $H^+(\mathcal{C}) = \{\alpha = 0\}$, Eq. (B11) proves that

$$\Lambda_\alpha{}^{(i)}[V^{(L)}, W^{(L)}] \to 0 \qquad \text{asymptotically as } \alpha \to 0\,. \tag{B12}$$

Moreover, it is clear from Eqs. (B1a), (B2a), and (B3) that the operators $\Lambda_\alpha{}^{(i)}[V, W]$

are *not* multiples of α^2, i.e., they *cannot* be written in the form $\alpha^2 \mathcal{P}^{(i)}[V,W]$ where $\mathcal{P}^{(i)}[V,W]$ are *smooth* operators throughout the Minkowski spacetime $\mathcal{M}$. Therefore, it follows from Eqs. (B11) and (B12) that the asymptotic behaviors of the solutions $(V^{(L)}, W^{(L)})$ as $\alpha \to 0$ are the same as those of the solutions $(V_{as}{}^{(L)}, W_{as}{}^{(L)})$ of

$$\Lambda_\alpha{}^{(i)}[V_{as}{}^{(L)}, W_{as}{}^{(L)}] \equiv 0 \,. \tag{B13}$$

On the other hand, in the wedge region Λ where $\alpha > 0$, we can rewrite Eq. (B13) (trivially) as

$$\frac{1}{\alpha^2}\, \Lambda_\alpha{}^{(i)}[V_{as}{}^{(L)}, W_{as}{}^{(L)}] \equiv 0 \,. \tag{B14}$$

Taking the limit of Eq. (B14) as $L \to \infty$ and using Eqs. (B9) and (B10), we obtain

$$0 \equiv \lim_{L \to \infty} \frac{1}{\alpha^2}\, \Lambda_\alpha{}^{(i)}[V_{as}{}^{(L)}, W_{as}{}^{(L)}] = \mathcal{L}_\alpha{}^{(i)}[V_{as}{}^{(\infty)}, W_{as}{}^{(\infty)}] \equiv 0 \,. \tag{B15}$$

When compared with Eqs. (B5a) and (B6a), Eq. (B15) proves our claim that the solutions $(V^{(\infty)}, W^{(\infty)})$ of the field equations (2.32) have the same asymptotic behaviors near $\alpha = 0$ as the solutions of the ordinary differential equations (3.1).

APPENDIX C: SOME REMARKS ON THE FIELD EQUATIONS FOR COLLIDING NONPARALLEL-POLARIZED PLANE WAVES

In this appendix, we will describe some interesting equivalent formulations of the field equations (2.32) for arbitrarily-polarized colliding plane waves; we hope that some of these alternative forms might eventually prove useful in the search for a general solution of Eqs. (2.32).

For the first reformulation, we introduce a 1-form $\Theta(\alpha,\beta)$ by the relation

$$\Theta \equiv \cosh^2 W \, dV \ . \tag{C1}$$

Denoting the α, β components of Θ by Θ_α and Θ_β, respectively (that is, putting $\Theta = \Theta_\alpha d\alpha + \Theta_\beta d\beta$), we can then express the field equations (2.32) purely in terms of Θ and the function $W(\alpha,\beta)$:

$$\frac{1}{\alpha}(\alpha\Theta_\alpha)_{,\alpha} - \Theta_{\beta,\beta} = 0 \ , \tag{C2a}$$

$$W_{,\alpha\alpha} + \frac{1}{\alpha} W_{,\alpha} - W_{,\beta\beta} = (\Theta_\alpha{}^2 - \Theta_\beta{}^2)\,\frac{\sinh W}{\cosh^3 W} \ ; \tag{C2b}$$

where Eqs. (C2a) and (C2b) are to be solved subject to the *auxilliary condition*

$$d\Theta = 2 \tanh W \, dW \wedge \Theta \ . \tag{C2c}$$

Now consider the special case determined by the *ansatz*

$$d\Theta \equiv 0 \ , \tag{C3}$$

which is equivalent to $dW \wedge dV \equiv 0$, and which is in turn equivalent to the existence of a functional relationship between V and W. The class of solutions that obey the condition (C3) includes all parallel-polarized ($W \equiv 0$) solutions, as well as solutions (V,W) that one obtains from parallel-polarized metrics (2.31) by effecting a constant linear transformation on the coordinates x and y, thereby introducing an artificial cross-polarization component W. However, the special class (C3) is clearly larger than the class of these essentially parallel-polarized solutions. In any case, if by utilizing the condition (C3) we introduce a new function $\tilde{V}(\alpha,\beta)$ that satisfies

$$d\tilde{V} \equiv \Theta = \cosh^2 W \, dV \ , \tag{C4}$$

then the field equations (C2) can be rewritten in terms of the two functions $\tilde{V}$ and W in the form

$$\tilde{V}_{,\alpha\alpha} + \frac{1}{\alpha}\tilde{V}_{,\alpha} - \tilde{V}_{,\beta\beta} = 0 \ , \tag{C5a}$$

$$W_{,\alpha\alpha} + \frac{1}{\alpha}W_{,\alpha} - W_{,\beta\beta} = (\tilde{V}_{,\alpha}{}^2 - \tilde{V}_{,\beta}{}^2)\,\frac{\sinh W}{\cosh^3 W} \ ; \tag{C5b}$$

where Eqs. (C5a) and (C5b) must be solved subject to the *auxilliary condition*

$$d\tilde{V} \wedge dW \equiv 0 \ . \tag{C5c}$$

The solution of the linear equation (C5a) can be found explicitly in terms of initial data; see Sec. II B of Ref. 6, especially Eqs. (6.2.44a) and (6.2.60). In fact, it becomes clear from Eq. (C5a) that in this special case [Eq. (C3)] we can express the asymptotic structure-function $\varepsilon_1(\beta)$ (Sec. III A) explicitly in terms of the initial data for $\tilde{V}(\alpha,\beta)$: Combining Eqs. (3.3) and (3.4a) with Eq. (C4), and comparing Eq. (C5a) with Eqs. (6.2.44a), (6.3.7), and (6.3.13) of Ref. 6, we obtain

$$\varepsilon_1(\beta) = \frac{1}{\pi}\,\frac{1}{\sqrt{1+\beta}}\int_{\beta}^{1}[(1+s)^{1/2}\,\tilde{V}(1,s)]_{,s}\left(\frac{s+1}{s-\beta}\right)^{1/2} ds$$

$$+ \frac{1}{\pi}\,\frac{1}{\sqrt{1-\beta}}\int_{-\beta}^{1}[(1+r)^{1/2}\,\tilde{V}(r,1)]_{,r}\left(\frac{r+1}{r+\beta}\right)^{1/2} dr \ . \tag{C6}$$

Returning now to the general case (C2), we note that the field equation (C2a) for Θ can be rewritten as

$$(\alpha\,\Theta_\alpha)_{,\alpha} = (\alpha\,\Theta_\beta)_{,\beta}\,. \tag{C7}$$

Equation (C7) implies that there exists a function $S(\alpha,\beta)$ that satisfies

$$S_{,\alpha} = \alpha\,\Theta_\beta\,, \qquad S_{,\beta} = \alpha\,\Theta_\alpha\,, \tag{C8}$$

and in turn, Eqs. (C8) can be expressed in the equivalent form

$$\Theta \equiv \frac{1}{\alpha} * dS\;, \tag{C9}$$

where the Hodge-star[17] operator "*" is defined with respect to the two-dimensional flat metric $(-d\alpha^2 + d\beta^2)$. In terms of the two functions $S(\alpha,\beta)$ and $W(\alpha,\beta)$, the field equations (2.32) [or equivalently Eqs. (C2)] can now be rewritten in the alternative form

$$d\,(\,\frac{1}{\alpha} * dS\,) = \frac{2\tanh W}{\alpha}\, dW \wedge * dS\;, \tag{C10a}$$

$$d\,(\,\alpha * dW\,) = -\frac{\sinh W}{\alpha\cosh^3 W}\, dS \wedge * dS\;, \tag{C10b}$$

with no auxilliary conditions.

APPENDIX D: A MORE SOPHISTICATED FORMULATION OF THE NOTION OF NONGENERICITY IN AN ARBITRARY BAIRE SPACE

Recall the simple definition that we introduced in Sec. III C of Ref. 6 to describe the nongenericity of a subset in an arbitrary Banach space. According to this definition, a subset is nongeneric if it is closed and has a dense complement, i.e., if it is a closed subset with empty interior. Although this notion of genericity is both intuitively plausible and broad enough to describe the nongenericity of larger-than-

Planck-size Killing-Cauchy horizons in colliding plane-wave spacetimes (Sec. III B), it is too naive even to identify the set Q of rational numbers as a nongeneric subset within the real line R. Similarly, it fails to describe the nongenericity of the subset $\bigcup_{\delta>0} H_\delta$ of *all* horizon-producing initial data within the Banach space of all initial data for colliding plane waves (Sec. III B). Clearly, a more sophisticated generalization of the above notion of nongenericity is needed to avoid these drawbacks; in this appendix we will describe such a generalization. Just as the above notion of genericity applies not only to a Banach space but more generally to arbitrary topological spaces, so also here we will formulate our generalization for a broad class of topological spaces called *Baire spaces* (see the definitions below). Any complete metric space (hence in particular any Banach space) is a Baire space; thus our notions would be applicable to most function spaces that arise naturally in mathematical physics. In the following, we will omit the full proofs of many of the standard results that we use; more detailed discussions on these results can be found in any textbook on general topology, e.g., in Ref. 45.

We first review some of the basic definitions: A topological space X is called "of the first category" if X is the union of *countably many* closed subsets with empty interiors; otherwise, X is called of the second category. These definitions apply to a subset $S \subset X$ by regarding S as a topological space under the topology induced from X. [Thus: $Q \subset R$ is of the first category; { irrational numbers } $\subset R$ is of the second category.] The space X is said to be a *Baire space*[45] if every nonempty open subset of X is of the second category. It is not very difficult to prove[45] that X is a Baire space if and only if for every countable collection of nonempty closed subsets $\{A_n \subset X\}$ with empty interiors, $\bigcup_{n=1}^{\infty} A_n \subset X$ is a subset with empty interior. (Thus: Q is *not* a Baire space; R is a Baire space.) A fundamental result[45] is that *every complete metric space* *is a Baire space*.

Our definition of "thin" subsets: Let B be a Baire space (or more specifically a complete metric space). A subset $S \subset B$ is called *thin* if and only if there exists a family of subsets $\{H_\delta \subset B\}$ with the following properties (here $\delta > 0$ ranges over all positive real numbers):

(i) For each $\delta > 0$, H_δ is a closed subset with empty interior in B.

(ii) If $\delta_2 < \delta_1$, then $H_{\delta_2} \supset H_{\delta_1}$.

(iii) $\bigcup_{\delta > 0} H_\delta = S$.

In particular, if $S \subset B$ is a closed subset with empty interior then it is thin: just take $H_\delta \equiv S$ for all $\delta > 0$. Hence the notion of "thin" subsets generalizes the naive notion of nongenericity that we introduced in Ref. 6. In fact, this is an intuitively plausible generalization: It follows from the properties (i)–(iii) that the thin subset S is essentially the "limit" as $\delta \to 0$ of the "nongeneric" subsets H_δ; therefore, intuitively a thin subset is just the "limit" of a continuous family of subsets which are all nongeneric in the sense of Ref. 6. Some of the other properties that thin subsets have according to the above definition are described in the following paragraph.

The first important property is the following alternative characterization: A subset $S \subset B$ in a Baire space is thin if and only if there exists a countable family $\{A_n \subset B\}$ of closed subsets of B, each with empty interior, such that $S = \bigcup_{n=1}^{\infty} A_n$. [To prove the if part, given the countable family $\{A_n\}$ of closed subsets with empty interiors satisfying $\bigcup_{n=1}^{\infty} A_n = S$, put $H_\delta \equiv \bigcup_{n=1}^{[1/\delta]} A_n$, where $[1/\delta]$ denotes the smallest integer $\geq 1/\delta$. The family $\{H_\delta\}$ satisfies property (i) since B is a Baire space; the other properties (ii) and (iii) are satisfied by construction. To prove the only if part, given the family $\{H_\delta\}$ satisfying properties (i)–(iii), put $A_n \equiv H_{1/n}$.] As a consequence, the subset Q of rationals is thin in R, whereas the subset of irrational numbers is *not* thin. Also, if $S \subset B$ is thin and $P \subset S$ is closed *in* S, then P is a thin subset in B. Notice that our notion of a thin subset is essentially different from the notion of a subset of the first category: A thin subset is not necessarily of the first category (any closed subset with empty interior in a complete metric space is thin but not of the first category), and conversely a subset of the first category is not necessarily thin (the subset $S \subset R^2$ given by

$$S \equiv \{\, (x,y) \in R^2 \mid 0 < x < 1 \,, x \text{ is irrational} \,,\ 0 \leq y \leq 1 \,, y \text{ is rational} \,\}$$

is of the first category but not thin in R^2). Nevertheless, it follows from the above alternative characterization of thin subsets that just as the subsets of the first category of a Baire space have empty interiors, so also its thin subsets have empty interiors; in other words the complement of any thin subset is dense in B.

Although it presents a more general alternative to our older, more naive concept of a nongeneric subset, the notion of a thin subset is nevertheless inappropriate as a concept of nongenericity. The reason is that subsets of a thin set are not necessarily thin unless they are closed (see above), whereas intuitively one expects that any subset of a nongeneric set should itself be nongeneric. To satisfy this requirement and at the same time to preserve the remaining plausible characteristics of "thinness," we therefore adopt the following most straightforward derivative of the notion of a thin subset as our generalized concept of nongenericity:

Our notion of nongeneric subsets: A subset $P \subset B$ of a Baire space B is called *nongeneric* if and only if P is contained in a thin subset of B.

It becomes obvious that any subset of a nongeneric subset is itself nongeneric. It also follows that although a nongeneric subset is not necessarily of the first category, any subset of the first category in a Baire space is nongeneric.

APPENDIX E: PROOF OF THE LEMMA THAT FUTURE NULL CONES IN A NONDEGENERATE KASNER SPACETIME START TO RECONVERGE

In this appendix we will prove the following result which is used in the proof of Theorem 2 in Sec. IV B:

Lemma 3: In a nondegenerate Kasner spacetime [Eq. (3.18)], the future null cone $\dot{J}^+(q)$ of any point q starts to reconverge near the singularity $\{t=0\}$, i.e., on each future-directed null geodesic from q the convergence $\hat{\theta}$ (Sec. 4.2 of Ref. 14) of the null generators of $\dot{J}^+(q)$ becomes negative near $t=0$.

Consider a general nondegenerate Kasner spacetime with the metric (3.18):

$$g = -a\,dt^2 + b\,t^{2p_3} d\beta^2 + c\,t^{2p_1} dx^2 + d\,t^{2p_2} dy^2 , \tag{E1}$$

where a, b are positive constants having the dimensions of (length)^2, c, d are dimensionless positive constants, t, β are dimensionless coordinates, and the exponents p_k, $k=1,2,3$ satisfy the Kasner relations

$$p_1 + p_2 + p_3 = p_1^2 + p_2^2 + p_3^2 = 1 . \tag{E2}$$

It follows from Eqs. (E2) that if the metric (E1) is nondegenerate [i.e., if all exponents p_k are different from 0 (or equivalently if all exponents are different from 1)], then precisely two exponents are strictly positive and precisely one is strictly negative. Thus we will assume, without loss of generality, that

$$1 > p_1 \geq p_2 > 0 \,, \qquad -1 < p_3 < 0 \,. \tag{E3}$$

[In fact Eqs. (E2) imply that $p_3 > -(\sqrt{5}-1)/2$ in this case, but we will not use this sharper inequality below.] Consider now an arbitrary point q in the Kasner spacetime (E1) with coordinates $t_0, x_0, y_0, \beta_0, t_0 > 0$. We will explore the behavior of the future null cone $\dot{J}^+(q)$ of this point q; in particular, we are interested in evaluating the asymptotic behavior (as $t \to 0$) of the convergence $\hat{\theta}$ for the null geodesics which generate $\dot{J}^+(q)$. Let the integrals of motion $g(\gamma_*, \partial/\partial x)$, $g(\gamma_*, \partial/\partial y)$, and $g(\gamma_*, \partial/\partial\beta)$ (associated with the Killing vector fields $\partial/\partial x$, $\partial/\partial y$, and $\partial/\partial\beta$) along a future-directed null geodesic γ from q be denoted by

$$g(\gamma_*, \partial/\partial x) \equiv C_x \,, \qquad g(\gamma_*, \partial/\partial y) \equiv C_y \,, \qquad g(\gamma_*, \partial/\partial\beta) \equiv C_\beta \,. \tag{E4}$$

Then, a short computation shows that as functions of the time coordinate t, $t \leq t_0$, (i) the coordinates $x(t)$, $y(t)$, and $\beta(t)$ of any point $q(t)$ along the null geodesic γ are given by

$$x(t) = x_0 + \int_t^{t_0} \frac{C_x}{c\, s^{2p_1}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, s^{2p_3}} + \dfrac{C_x^2}{c\, s^{2p_1}} + \dfrac{C_y^2}{d\, s^{2p_2}} \right]^{1/2}} ds \,, \tag{E5a}$$

$$y(t) = y_0 + \int_t^{t_0} \frac{C_y}{d\, s^{2p_2}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, s^{2p_3}} + \dfrac{C_x^2}{c\, s^{2p_1}} + \dfrac{C_y^2}{d\, s^{2p_2}} \right]^{1/2}} ds \,, \tag{E5b}$$

$$\beta(t) = \beta_0 + \int_t^{t_0} \frac{C_\beta}{b\, s^{2p_3}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, s^{2p_3}} + \dfrac{C_x^2}{c\, s^{2p_1}} + \dfrac{C_y^2}{d\, s^{2p_2}} \right]^{1/2}} ds \,, \tag{E5c}$$

and (ii) the tangent vector $\gamma_*(t)$ to the null geodesic γ is given by

$$\gamma_*(t) = \left[\frac{C_x}{c\, t^{2p_1}} \right] \frac{\partial}{\partial x} + \left[\frac{C_y}{d\, t^{2p_2}} \right] \frac{\partial}{\partial y} + \left[\frac{C_\beta}{b\, t^{2p_3}} \right] \frac{\partial}{\partial \beta}$$

$$- \frac{1}{\sqrt{a}} \left[\frac{C_\beta^2}{b\, t^{2p_3}} + \frac{C_x^2}{c\, t^{2p_1}} + \frac{C_y^2}{d\, t^{2p_2}} \right]^{1/2} \frac{\partial}{\partial t} . \tag{E6}$$

[Note that $\partial/\partial t$ is a past-directed timelike vector.]

In the following, we will assume for simplicity that $p_1 > p_2$ [cf. Eqs. (E3)]. After trivial modifications, all arguments below are also valid for the case $p_1 = p_2$.

Consider a null geodesic generator of $\dot{J}^+(q)$ along which $C_x \neq 0$. It follows from Eq. (E6) that asymptotically as $t \to 0$

$$\gamma_*(t) \sim \frac{C_x}{c\, t^{2p_1}} \frac{\partial}{\partial x} - \frac{|C_x|}{\sqrt{ac}} \frac{1}{t^{p_1}} \frac{\partial}{\partial t} \tag{E7}$$

along such a generator. Now recall that given any null hypersurface $\mathcal{S}$ like $\dot{J}^+(q)$, the tangent vectors $\gamma_*(t)$ of the null geodesic generators of $\mathcal{S}$ define a null, geodesic vector field *on* $\mathcal{S}$. If this vector field on $\mathcal{S}$ is extended to *any* vector field $\vec{V}$ which is null (but not necessarily geodesic outside $\mathcal{S}$) on a neighborhood of $\mathcal{S}$, then the divergence of $\vec{V}$ restricted to $\mathcal{S}$, $(\nabla \cdot \vec{V})|_{\mathcal{S}}$, is equal to the convergence $\hat{\theta}$ of $\mathcal{S}$'s null generators; i.e., the quantity $(\nabla \cdot \vec{V})|_{\mathcal{S}}$ is independent of the null extension $\vec{V}$ and equals $\hat{\theta}$. [For a proof of this well-known fact see Sec. 4.2 of Ref. 14.] Thus, consider the null, geodesic vector field γ_* on $\dot{J}^+(q)$ defined by those generators of $\dot{J}^+(q)$ which lie in the vicinity of our generator with $C_x \neq 0$; all these neighboring generators similarly have $C_x \neq 0$. Applying the general formula (valid in a coordinate basis)

$$\nabla \cdot \vec{V} = \frac{1}{\sqrt{-g}} (\sqrt{-g}\, V^\mu)_{,\mu} \tag{E8}$$

to *any* null extension $\vec{V}$ of this field γ_*, we find that Eq. (E7) implies, asymptotically as $t \to 0$

$$\hat{\theta}(t) = \nabla \cdot \vec{V}(t) \sim \frac{C_{x,x}}{c} \frac{1}{t^{2p_1}} - \frac{|C_x|}{\sqrt{ac}} (1-p_1) \frac{1}{t^{1+p_1}} \tag{E9}$$

along our generator, provided $C_{x,x}$ is finite as $t \to 0$. On the other hand, it is obvious that in the vicinity of any generator with $C_x \neq 0$ we can find a null extension $\vec{V}$ of γ_* which satisfies $C_{x,x} \equiv 0$. [To see this, observe that the vector field $\partial / \partial x$ intersects $\dot{J}^+(q)$ transversally in the vicinity of such a generator. Also, although one might worry about the terms of the form $C_{y,y} / t^{2p_2}$ and $C_{\beta,\beta} / t^{2p_3}$ in $\nabla \cdot \vec{V}(t)$ which are not included in Eq. (E9), it similarly follows that whenever $C_y \neq 0$ and $C_\beta \neq 0$ one can find an extension with $C_{y,y} \equiv C_{\beta,\beta} \equiv 0$ and thus make these terms identically zero. On the other hand, a straightforward application of the arguments we present below shows that along the generators on which $C_y = 0$ or $C_\beta = 0$ the quantities $C_{y,y}(t)$ and $C_{\beta,\beta}(t)$ remain finite as $t \to 0$.] Therefore, it follows from Eq. (E9) that:

Along any generator of $\dot{J}^+(q)$ with $C_x \neq 0$ the convergence $\hat{\theta}$ diverges to $-\infty$ as $t \to 0$.

Now consider a generator of $\dot{J}^+(q)$ along which $C_x = 0$ but $C_y \neq 0$. It is easy to see that on such a generator we have, instead of Eq. (E9),

$$\hat{\theta}(t) = \nabla \cdot \vec{V}(t) \sim \frac{C_{x,x}}{c} \frac{1}{t^{2p_1}} - \frac{|C_y|}{\sqrt{ad}} (1-p_2) \frac{1}{t^{1+p_2}} . \tag{E10}$$

Now, by using Eq. (E5a), we can actually compute the asymptotic behavior of the quantity $C_{x,x}(t)$ *along* this generator on which $C_x = 0$ and $C_y \neq 0$. Differentiating both sides of Eq. (E5a) with respect to x and putting $C_x = 0$, we obtain

$$1 = \int_{t}^{t_0} \frac{C_{x,x}(t)}{c\, s^{2p_1}} \frac{\sqrt{a}}{\left[\dfrac{C_\beta^2}{b\, s^{2p_3}} + \dfrac{C_y^2}{d\, s^{2p_2}} \right]^{1/2}} ds \; . \tag{E11}$$

The asymptotic ($t \rightarrow 0$) limit of Eq. (E11) is easily computed; it gives

$$1 \sim \int_{t}^{t_0} C_{x,x}(t) \frac{\sqrt{ad}}{c\, |C_y|} s^{p_2 - 2p_1} ds \; . \tag{E12}$$

After evaluating the integral in Eq. (E12) and combining the result with Eq. (E10), we reach the following final conclusions: (i) When $p_2 - 2p_1 + 1 < 0$,

$$C_{x,x}(t) \sim (2p_1 - p_2 - 1) \frac{c\, |C_y|}{\sqrt{ad}} \frac{1}{t^{1+p_2-2p_1}} \; , \tag{E13a}$$

and

$$\hat{\theta}(t) \sim \frac{2(p_1 - 1)}{\sqrt{ad}} |C_y| \frac{1}{t^{1+p_2}} \rightarrow -\infty \; . \tag{E13b}$$

(ii) When $p_2 - 2p_1 + 1 > 0$,

$$C_{x,x}(t) \sim (p_2 - 2p_1 + 1) \frac{c\, |C_y|}{\sqrt{ad}} \frac{1}{t_0^{p_2-2p_1+1}} \; , \tag{E14a}$$

and

$$\hat{\theta}(t) \sim \frac{|C_y|}{\sqrt{ad}} \left[\frac{(p_2 - 2p_1 + 1)}{t_0^{p_2-2p_1+1}\, t^{2p_1}} - \frac{(1 - p_2)}{t^{1+p_2}} \right] \rightarrow -\infty \; . \tag{E14b}$$

(iii) When $p_2 - 2p_1 + 1 = 0$,

$$C_{x,x}(t) \sim \frac{c\,|C_y|}{\sqrt{ad}}\,\frac{1}{\ln(t_0/t)}\,, \tag{E15a}$$

and

$$\hat{\theta}(t) \sim \frac{|C_y|}{\sqrt{ad}}\left[\frac{1}{t^{2p_1}\ln(t_0/t)} - \frac{(1-p_2)}{t^{2p_1}}\right] \to -\infty\,. \tag{E15b}$$

Consequently, our overall conclusion is:

Along any generator of $\dot{J}^+(q)$ *with* $C_x = 0$ *and* $C_y \neq 0$ *the convergence* $\hat{\theta}$ *diverges to* $-\infty$ *as* $t \to 0$.

For a generator of $\dot{J}^+(q)$ along which $C_x = C_y = 0$ but $C_\beta \neq 0$, we have

$$\hat{\theta}(t) = \nabla \cdot \vec{V}(t) \sim \frac{C_{x,x}}{c}\,\frac{1}{t^{2p_1}} + \frac{C_{y,y}}{d}\,\frac{1}{t^{2p_2}} - \frac{|C_\beta|}{\sqrt{ab}}\,(1-p_3)\,\frac{1}{t^{1+p_3}}\,. \tag{E16}$$

The quantities $C_{x,x}$ and $C_{y,y}$ of Eq. (E16) can be computed along this generator in exactly the same way as before, i.e., by differentiating Eqs. (E5a) and (E5b) with respect to x and y, respectively, and then putting $C_x = C_y = 0$. Evaluating the asymptotic forms of the resulting integrals and proceeding in precisely the same manner as we did in Eqs. (E11)–(E15), we obtain the conclusion that:

Along any generator of $\dot{J}^+(q)$ *with* $C_x = C_y = 0$ *and* $C_\beta \neq 0$ *the convergence* $\hat{\theta}$ *diverges to* $-\infty$ *as* $t \to 0$.

Combined with the two previous conclusions, this last result completes the proof of Lemma 3.

REFERENCES FOR CHAPTER 7

[1] K. Khan and R. Penrose, Nature (London) **229**, 185 (1971).

[2] R. A. Matzner and F. J. Tipler, Phys. Rev. D **29**, 1575 (1984).

[3] R. Penrose, Rev. Mod. Phys. **37**, 215 (1965).

[4] P. Szekeres, J. Math. Phys. **13**, 286 (1972).

[5] U. Yurtsever, Phys. Rev. D **37**, 2790 (1988) (Chapter 4).

[6] U. Yurtsever, Phys. Rev. D **38**, 1706 (1988) (Chapter 5).

[7] S. Chandrasekhar and B. Xanthopoulos, Proc. R. Soc. Lond. **A410**, 311 (1987).

[8] U. Yurtsever, Phys. Rev. D **36**, 1662 (1987) (Chapter 2).

[9] U. Yurtsever, Phys. Rev. D **38**, 1731 (1988) (Chapter 6).

[10] E. Kasner, Am. J. Math. **43**, 217 (1921); E. M. Lifshitz and I. M. Khalatnikov, Usp. Fiz. Nauk **80**, 391 (1964) [Sov. Phys. Uspekhi **6**, 495 (1963)]; C. W. Misner, Phys. Rev. Lett. **22**, 1071 (1969); V. A. Belinsky and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **57**, 2163 (1969) [Sov. Phys. JETP 30, 1174 (1970)]; V. A. Belinsky, I. M. Khalatnikov, and E. M. Lifshitz, Adv. Phys. **19**, 525 (1970).

[11] F. J. Tipler, Phys. Rev. D **22**, 2929 (1980).

[12] U. Yurtsever, Phys. Rev. D. **37**, 2803 (1988) (Chapter 3).

[13] S. W. Hawking and R. Penrose, Proc. R. Soc. Lond. **A314**, 529 (1970).

[14] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973).

[15] A. Fischer and J. E. Marsden in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England 1979); see also Sec. 7.6 of Ref. 14.

[16] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).

[17] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups* (Springer, New York, 1983).

[18] F. John, *Partial Differential Equations* (Springer, New York, 1982).

[19] P. Garabedian, *Partial Differential Equations* (Chelsea, New York, 1986).

[20] J. Leray, *Hyperbolic Differential Equations* (The Institute of Advanced Study, Princeton, 1953).

[21] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam, 1982).

[22] F. John, Comm. Pure Appl. Math. **29**, 649 (1976).

[23] S. Klainerman, Comm. Pure Appl. Math. **33**, 43 (1980).

[24] T. Kato, Comm. Pure Appl. Math. **33**, 501 (1980).

[25] J. Shatah, J. Diff. Eqs. **46**, 409 (1982).

[26] F. John, Comm. Pure Appl. Math. **36**, 1 (1983).

[27] S. Klainerman and G. Ponce, Comm. Pure Appl. Math. **36**, 133 (1983).

[28] S. Klainerman, Comm. Pure Appl. Math. **36**, 325 (1983).

[29] T. C. Sideris, J.Diff. Eqs. **52**, 378 (1984).

[30] W. Strauss, An. Acad. Brasil Cienc. **42**, 645 (1970).

[31] S. Klainerman, Comm. Pure Appl. Math. **38**, 321 (1985).

[32] D. Christodoulou, Comm. Pure Appl. Math. **39**, 267 (1986).

[33] The eigenvalues of the matrix (in x, y coordinates) of the $x-y$ part of (2.31) are $\alpha e^{\pm \hat{V}}$, where $\hat{V}$ is defined in Eq. (3.13e).

[34] B. Xanthopoulos, private communication.

[35] R. Penrose in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).

[36] F. J. Tipler, C. J. S. Clarke, and G. F. R. Ellis, in *Einstein Centenary Volume*, edited by A. Held and P. Bergmann (Plenum, NewYork, 1979).

[37] S. Chandrasekhar and B. Xanthopoulos, Proc. R. Soc. Lond. **A414**, 1 (1987).

[38] V. Moncrief and J. Isenberg, Commun. Math. Phys. **89**, 387 (1983); J. Math. Phys. **26**, 1024 (1985).

[39] K. Yosida, *Functional Analysis* (Springer, Berlin, 1985).

[40] E. Phillips, *An Introduction to Analysis and Integration Theory* (Dover, New York, 1984).

[41] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, New York, 1983).

[42] R. Geroch, private communication.

[43] Y.Choquet-Bruhat and J. York in *Einstein Centenary Volume*, edited by A. Held and P. Bergmann (Plenum, New York, 1979).

[44] V. I. Arnold, *Geometric Methods in the Theory of Ordinary Differential Equations* (Springer, New York, 1983).

[45] S. Willard, *General Topology* (Addison-Wesley, Massachusetts, 1970).

FIGURE CAPTIONS FOR CHAPTER 7

FIG. 1. The two-dimensional geometry of the characteristic initial-value problem for colliding plane waves. The null surfaces $\{u=0\}$ and $\{v=0\}$ are the past wave fronts of the incoming plane waves 1 and 2. Initial data corresponding to waves 1 and 2 are posed, respectively, on the upper portions of the surfaces $\{v=0\}$ and $\{u=0\}$ that are adjacent to the interaction region I. The geometry in region IV is flat, and the geometry in regions II and III is given by the metric describing the incoming waves 1 and 2, respectively. The geometry of the interaction region I is uniquely determined by the solution of the above initial-value problem. The directions in which the various lines of constant coordinates u, v, α, β, r, and s run are also indicated, along with the descriptions of the initial null surfaces in these different coordinate systems.

FIG. 2. The geometry of the initial-value problem described by Eqs. (2.32), (2.37), and (A1). The problem is posed in the ordinary Euclidean space R^2 determined by the coordinates (α,β). The characteristic initial surface $\mathcal{N}$ is given by $\mathcal{N} \equiv \{r=1, -1<s\leq 1\} \cup \{s=1, -1<r\leq 1\}$, where $r\equiv\alpha-\beta$ and $s\equiv\alpha+\beta$. The domain of dependence $D^+(\mathcal{N})$ is given by $D^+(\mathcal{N})=\{\alpha-\beta\leq 1, \alpha+\beta\leq 1\}\cap\{\alpha>0\}$.

FIG. 3. The two-dimensional geometry of the Minkowski-space initial-value problem (A9)–(A11) with the Z and Y directions suppressed. The characteristic initial surface $\mathcal{C}$ consists of the two null hypersurfaces $\{ (T^2-X^2)^{1/2}-Z=1, T<0,\ 0<(T^2-X^2)\leq 1 \}$ and $\{ (T^2-X^2)^{1/2}+Z=1, T<0,\ 0<(T^2-X^2)\leq 1 \}$ which intersect at the spacelike two-surface $\mathcal{Z}$; in fact $\mathcal{C}$ is generated by null geodesics that are orthogonal to this spacelike two-surface $\mathcal{Z}\equiv\{\alpha=(T^2-X^2)^{1/2}=1, \beta=Z=0\}$ inside the Minkowski wedge, i.e.,

by those null generators of $\dot{J}^+(\mathcal{Z})$ that have their past endpoints on $\mathcal{Z}$. The domain of dependence of the initial surface $\mathcal{C}$ is $D^+(\mathcal{C}) = \{\, |T| > |X|, T < 0 \,\} \cap J^+(\mathcal{C})$, and the horizon $\{\, |T| = |X|, T \leq 0\} \equiv \{\alpha = 0\}$ of the Minkowski wedge is the future Cauchy horizon $H^+(\mathcal{C})$ of $\mathcal{C}$. The region $J^-(q) \cap \mathcal{C}$ becomes unboundedly large in the ξ direction as any arbitrary point q of the wedge approaches the horizon. As a result, when the initial data posed on $\mathcal{C}$ have a plane-symmetric [(ξ,η)–independent] structure, the data "seen" by any field point q become infinitely extended in the ξ direction as q approaches the horizon $\{\alpha=0\}$. This effect in the formalism (A9)–(A11) is the geometric counterpart of the focusing effect caused by the singular $1/\alpha$ terms in Eqs. (2.32).

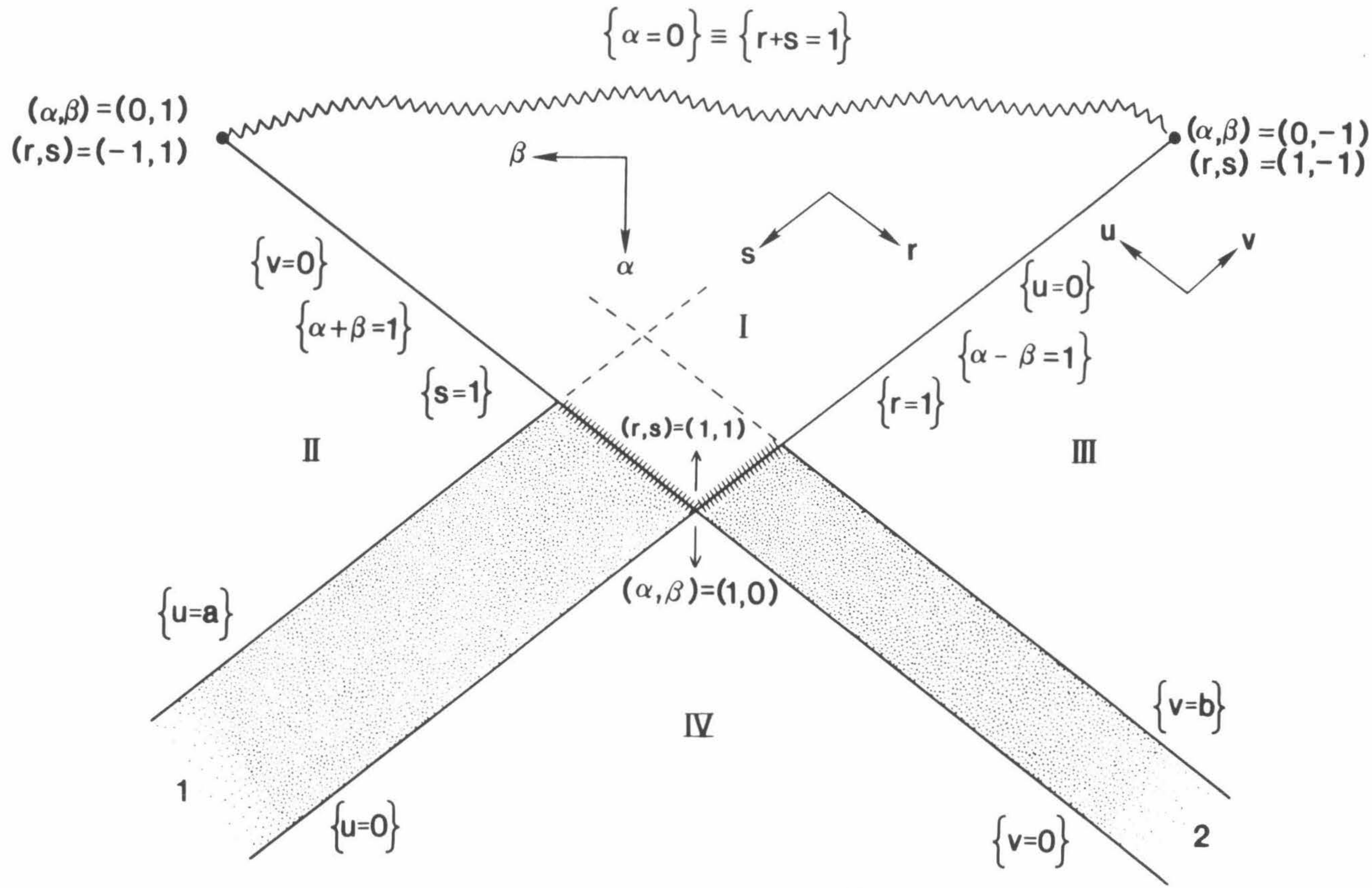
FIG. 4. (a): If reduced global existence (RGE) holds for Eqs. (A9), then this suffices to prove general global existence (GE) (see the precise formulations given in the text): Given an arbitrary partial Cauchy surface Σ and arbitrary data d on it, for any point $q \in D^+(\Sigma)$ we can introduce a $\{T=0\}$ surface [with some suitable Lorentz coordinates (T, X, Y, Z)] in the vicinity of the compact region $J^-(q) \cap \Sigma$. This reduces the global existence problem for Σ to the problem of RGE, provided the data on Σ can be transferred onto $\{T=0\}$ by means of local existence (LE). (b): If this fails, then we iteratively apply the construction described in (a) to the points of $\{T=0\}$, and we continue this iteration until the new smaller $\{T=0\}$ surfaces fall into that small neighborhood of Σ on which local existence is guaranteed by LE. Tracing our steps backwards by means of RGE after this last step is achieved, we see that the data on Σ can indeed be transferred to the first $\{T=0\}$ surface depicted in (a).

FIG. 5. If global existence for Eqs. (A9) is proved for spacelike initial surfaces, then it also holds when Σ is a characteristic initial surface consisting of two null hypersurfaces that intersect transversally: Given a characteristic surface Σ and data d posed on it, there is a neighborhood (dotted region) of Σ where local existence is guaranteed (by LE; see text). We can find a spacelike initial surface Σ' that lies entirely in this neighborhood, and thereby transfer the data d posed on Σ onto new data d' posed on Σ'. If global existence and uniqueness hold for Σ' and d', then they also hold for Σ and d.

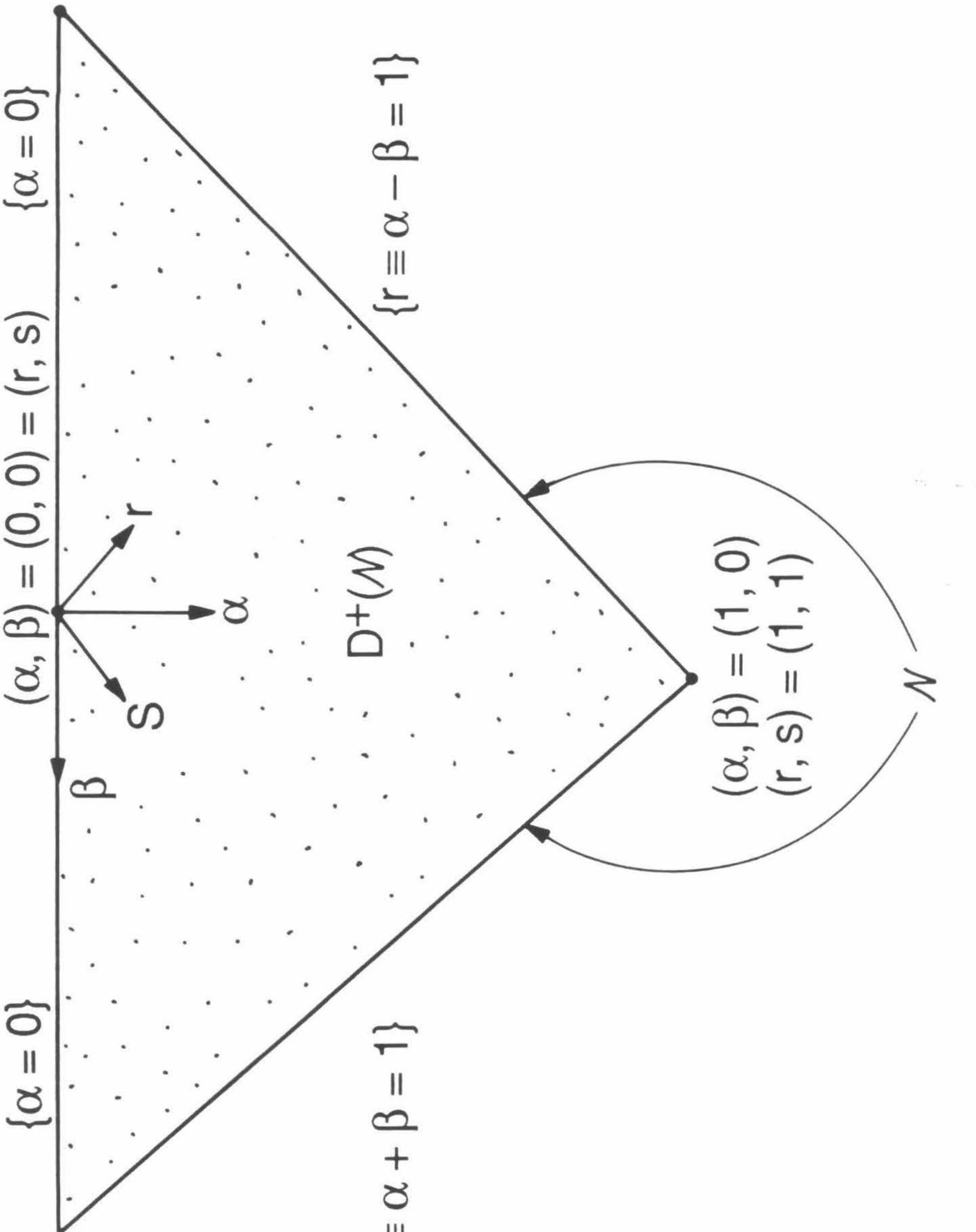
FIG. 6. The geometry of the energy inequality (A21). Initial data d are posed on $\{T=0\}$ and are compact supported in the open ball S_0. The domain of dependence $D^+(S_0)$ of S_0 is the interior of the null cone $H^+(\overline{S_0})$, and S_τ denotes the compact set $\{T=\tau\}\cap\overline{D^+(S_0)}$.

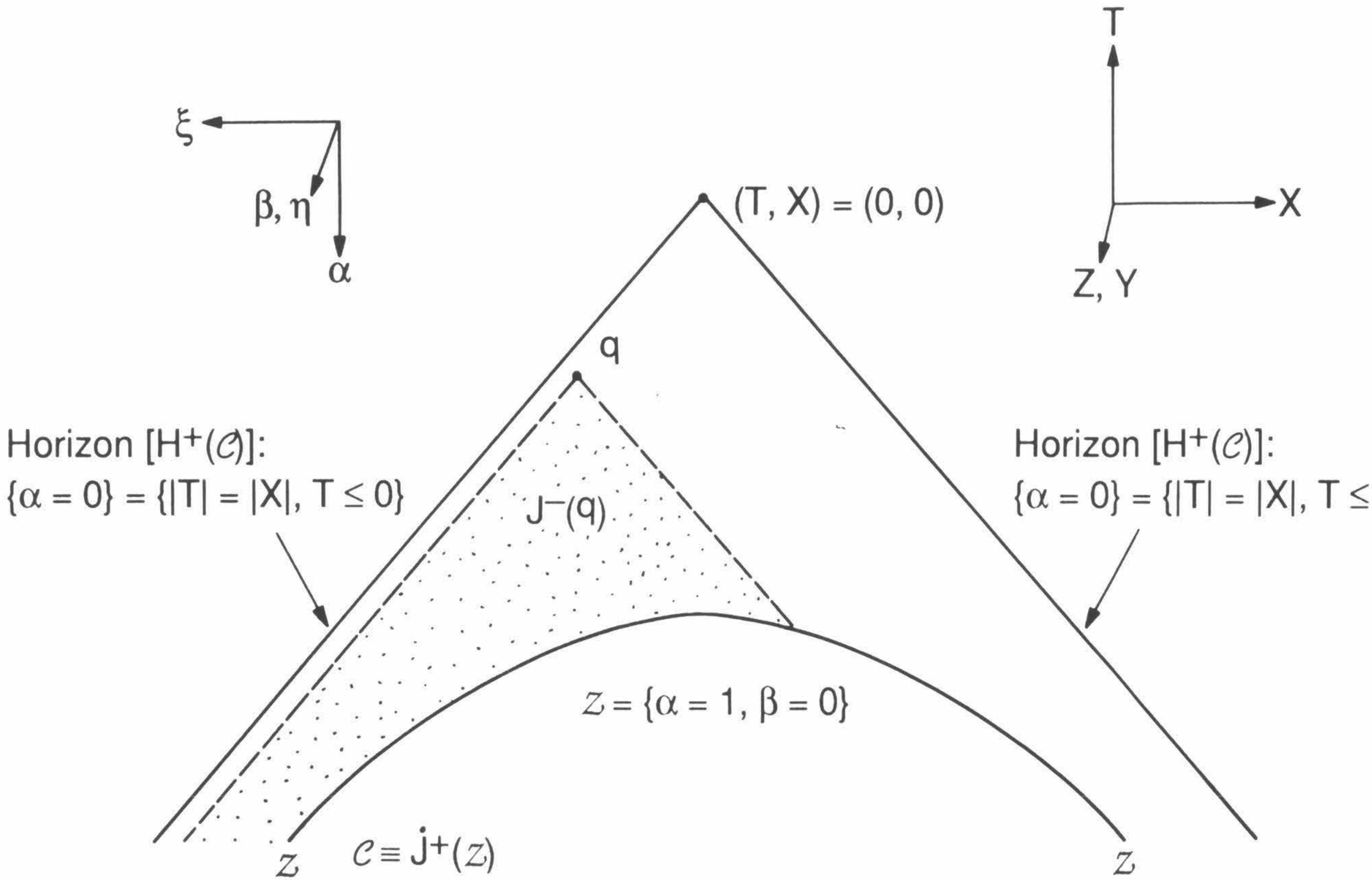
FIG. 7. Geometry of the initial-value problem for Eqs. (A9) where the initial data given by Eqs. (B8) are posed on the characteristic surface $\mathcal{C}$ (see Fig. 3). The initial data $\{V_0{}^{(L)}, W_0{}^{(L)}\}$ [Eqs. (B8)] are obtained by smoothly cutting off the plane-symmetric initial data $\{V_0{}^{(\infty)}, W_0{}^{(\infty)}\}$ at a distance $2L$ in the ξ and η directions. To prove that the solution $(V^{(L)}, W^{(L)})$ that evolves from these data is smooth across the horizon $H^+(\mathcal{C})$, a new initial surface Π is constructed by (i) choosing an $R>0$ with $R>2L$, (ii) adjoining a smooth spacelike hypersurface Σ to the initial surface $\mathcal{C}$ at the cylindrical cross-section $\mathcal{C}\cap\{\xi^2+\eta^2=R^2\}$ *through* $\mathcal{C}$, and finally (iii) discarding the portion of $\mathcal{C}$ that remains in the past of Σ. On the new initial surface Π, new initial data d for (V,W) are posed by leaving the data as they are on $\mathcal{C}$ [i.e. $\{d \text{ on } \mathcal{C}\}\equiv\{V_0{}^{(L)}, W_0{}^{(L)}\}$] and by putting $d\equiv 0$ on Σ. Throughout the subset of the Minkowski wedge that corresponds to the dotted region, the global solution of the

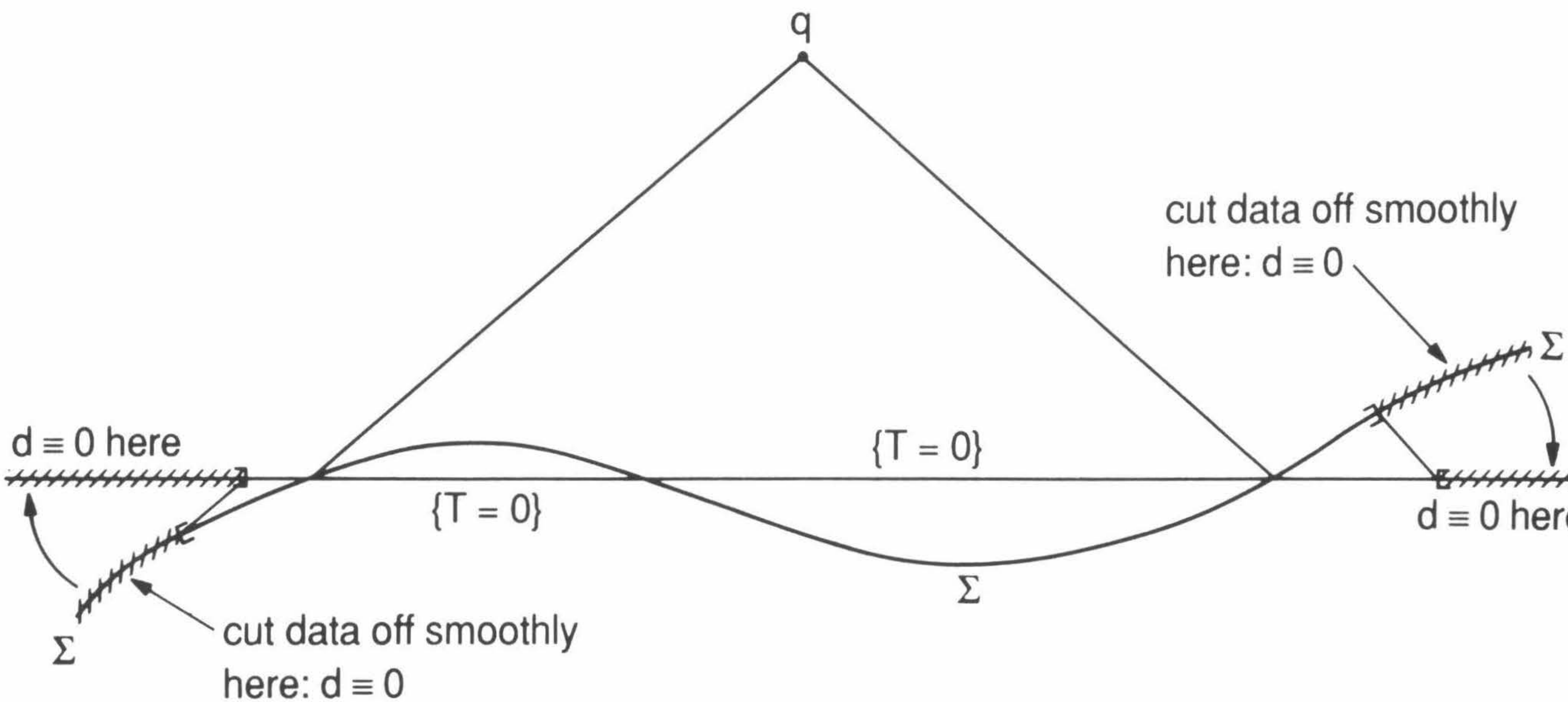
initial-value problem (Π, d) for Eqs. (A9) [or for Eqs. (B4)] is precisely equal to the solution $(V^{(L)}, W^{(L)})$. Moreover, the domain of dependence of Π includes the horizon $H^+(\mathcal{C})$. Since by Appendix A the solution of the initial-value problem (Π, d) exists smoothly throughout $D^+(\Pi)$, we conclude that the solution $(V^{(L)}, W^{(L)})$ of the problem $(\mathcal{C}, d^{(L)})$ is also smooth at and across the horizon $H^+(\mathcal{C})$.

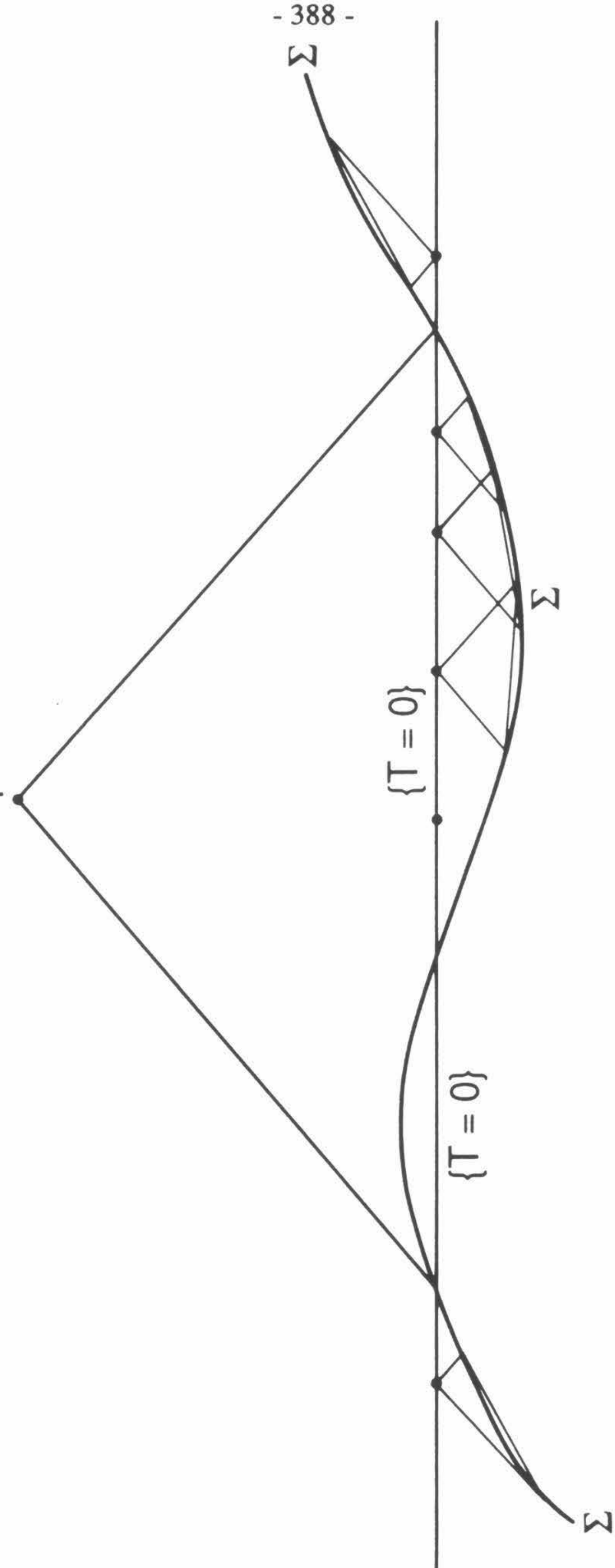


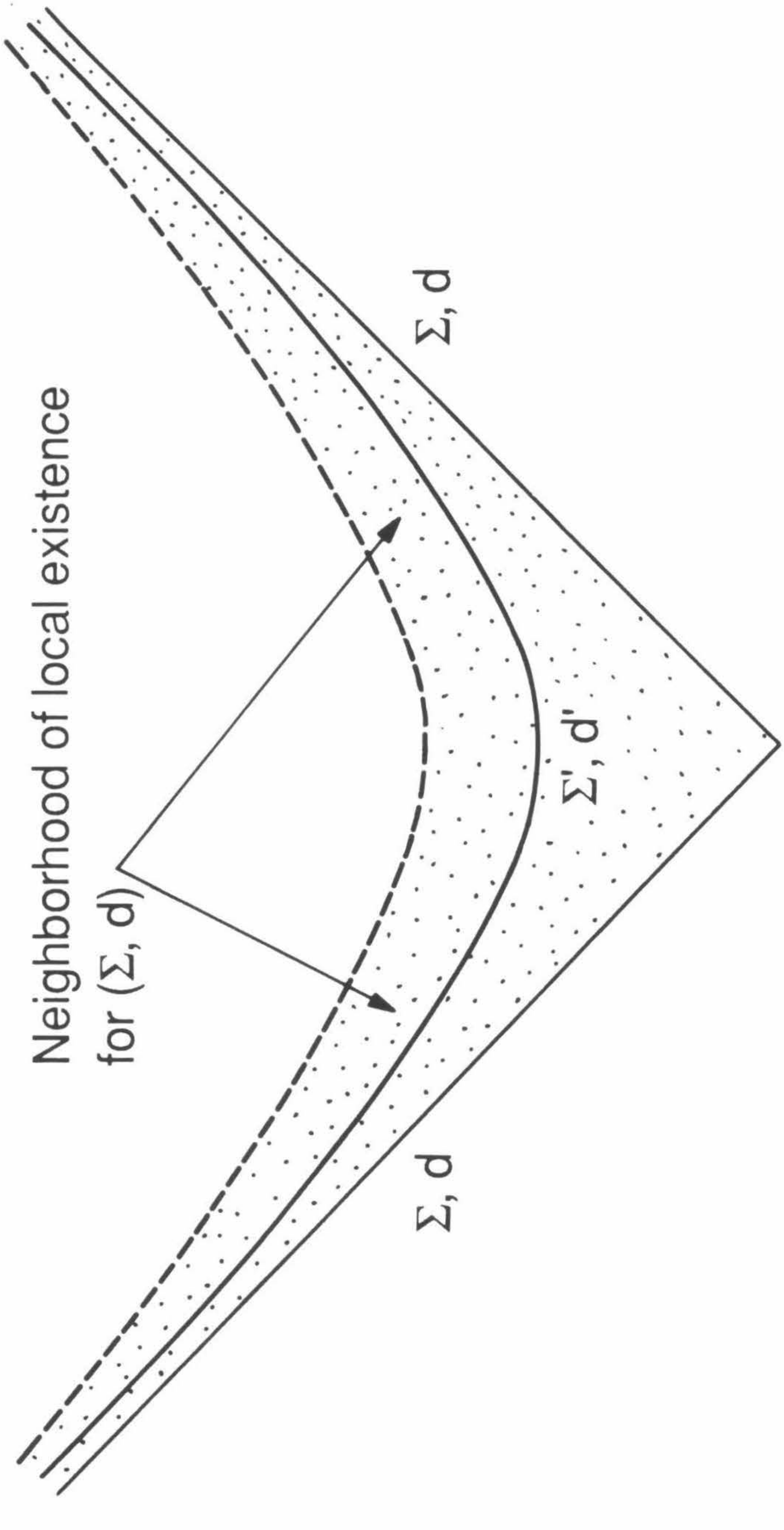
$\{\alpha = 0\} \equiv \{r+s=1\}$
$(\alpha,\beta) = (0,1)$
$(r,s) = (-1,1)$
$(\alpha,\beta) = (0,-1)$
$(r,s) = (1,-1)$
β
α
s
r
u
v
$\{v=0\}$
$\{\alpha+\beta=1\}$
$\{s=1\}$
I
$\{u=0\}$
$\{\alpha-\beta=1\}$
$\{r=1\}$
$(r,s)=(1,1)$
II
III
$(\alpha,\beta)=(1,0)$
$\{u=a\}$
IV
$\{v=b\}$
1
$\{u=0\}$
$\{v=0\}$
2

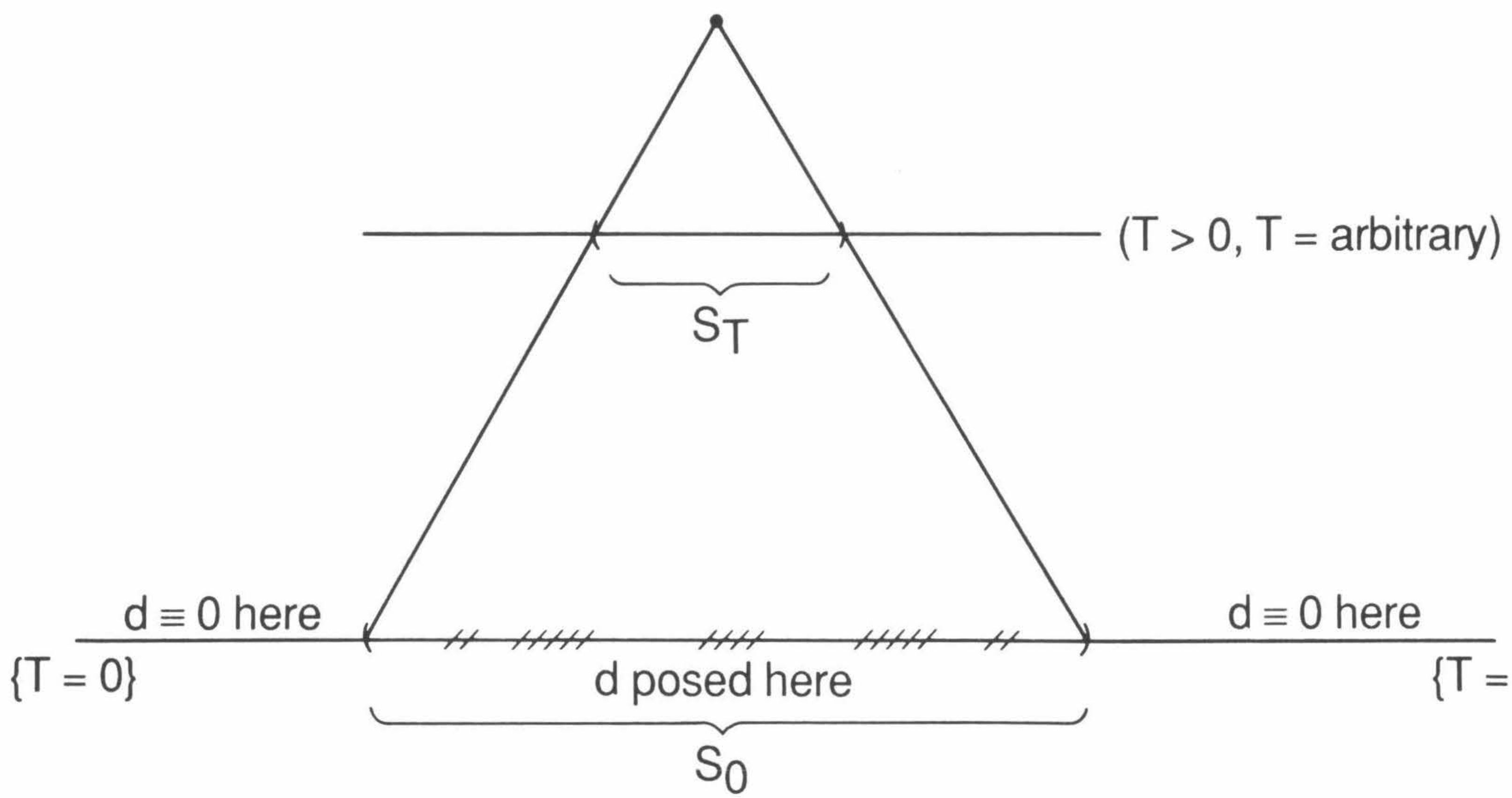

{α = 0}
(α, β) = (0, 0) = (r, s)
{α = 0}
{r ≡ α − β = 1}
{s ≡ α + β = 1}
r
α
S
β
D+(𝒩)
(α, β) = (1, 0)
(r, s) = (1, 1)
𝒩

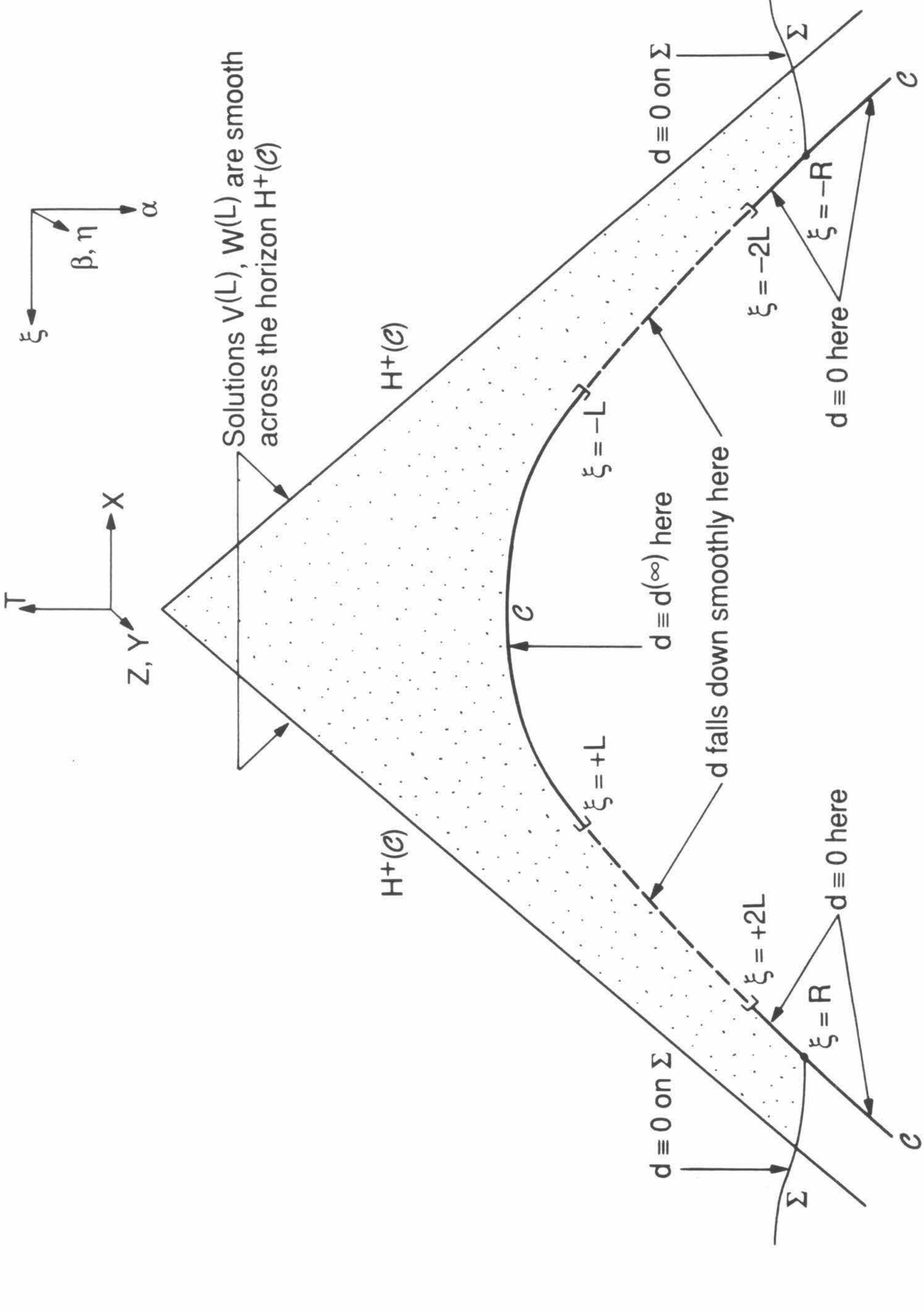

ξ
β, η
α
T
X
Z, Y
(T, X) = (0, 0)
q
J⁻(q)
Horizon [H⁺(𝒞)]:
{α = 0} = {|T| = |X|, T ≤ 0}
Horizon [H⁺(𝒞)]:
{α = 0} = {|T| = |X|, T ≤ 0}
𝒵 = {α = 1, β = 0}
𝒵
𝒞 ≡ j⁺(𝒵)
𝒵


q
cut data off smoothly
here: d ≡ 0
Σ
d ≡ 0 here
{T = 0}
{T = 0}
d ≡ 0 here
Σ
Σ
cut data off smoothly
here: d ≡ 0


q
Σ
Σ
Σ
{T = 0}
{T = 0}


Neighborhood of local existence
for (Σ, d)
Σ, d
Σ', d'
Σ, d


(T > 0, T = arbitrary)
S_T
d ≡ 0 here
{T = 0}
d posed here
d ≡ 0 here
{T = 0}
S_0


Solutions V(L), W(L) are smooth across the horizon H+(𝒞)
T
X
Z, Y
ξ
β, η
α
H+(𝒞)
H+(𝒞)
𝒞
𝒞
𝒞
Σ
Σ
d ≡ 0 on Σ
d ≡ 0 on Σ
ξ = −L
ξ = +L
ξ = −2L
ξ = +2L
ξ = −R
ξ = R
d ≡ d(∞) here
d falls down smoothly here
d ≡ 0 here
d ≡ 0 here