PHYSICAL REVIEW D VOLUME 22, NUMBER 12 15 DECEMBER 1980

Singularities from colliding plane gravitational waves

Frank J. Tipler
Center for Theoretical Physics, University of Texas, Austin, Texas 78712
(Received 23 June 1980)

A simple geometrical argument is given which shows that a collision between two plane gravitational waves must result in singularities. The argument suggests that these singularities are a peculiar feature of plane waves, because singularities are also a consequence of a collision between self-gravitating plane waves of other fields with arbitrarily small energy density.

I. INTRODUCTION

Gravitational-wave astronomy is a burgeoning field of research. However, because of the complexity of the Einstein field equations, the theory behind the experiments is generally based on a linear approximation, and such an aaproximation may not always give reliable results.[1,2] It is therefore important to study exact solutions corresponding to gravitational waves in order to understand their actual behavior. But the analysis of exact solutions is also fraught with perils. Again, because of the complexity of field equations, exact solutions necessarily have a very high degree of symmetry which would not be present in the real world, and the properties of the exact solutions might be artifacts of the high symmetry and have no analog in realistic solutions.[1,2]

In the case of gravitational waves, the exact solutions most often studied are plane waves. In linear field theories such as electromagnetism, plane waves are known to illustrate all the features of more realistic waves, and the plane symmetry does not give rise to misleading ideas about the structure of actual waves. The hope that a similar situation would persist in a nonlinear theory, such as general relativity in spite of the lack of a principle of superposition has been the strongest motivation in the study of plane gravitational waves. I shall argue, however, in the concluding section of this paper that the global properties of plane gravitational waves are quite different from the properties one would expect to see in realistic gravitational waves; in particular, colliding plane gravitational waves must necessarily develop singularities which are a consequence of the plane symmetry acting through a nonlinear field. A more realistic gravitational-wave collision should not, in general, give rise to singularities.

The question of whether a collision between plane gravitational waves would necessarily give rise to singularities has itself been disputed. Penrose and Khan[3] and, independently, Szekeres[4] found a solution interpreted as a collision between collinearly polarized impulsive gravitational waves, and Szekeres[5] later obtained the general solution corresponding to a collision between collinearly polarized gravitational waves. Nutku and Halil[6] then considered a collision between impulsive linearly polarized plane gravitational waves with arbitrary relative polarization, and Sbytov[7] derived the solution for colliding plane waves with arbitrary polarization. All of these authors concluded that a collision between plane gravitational waves would necessarily result in singularities. However, in a recent paper Stoyanov argued that the singularities which occurred in the solutions given by the above authors are consequences not of the planar nature of the wave collision, but rather of nonsmooth wave fronts. He justified this contention by giving a "solution" to the empty-space field equations with different junction conditions, a "solution" which had no singularity and which is claimed to be smoother than the above-mentioned solutions. (The word "solution" is in quotes because Nutku[8] has questioned whether the Stoyanov metric satisfies the Einstein equations at the junction.)

I shall prove in this paper that in fact any solution to the field equations with metric differentiability at least C^2 and corresponding to a collision between plane gravitational waves will necessarily have a singularity somewhere. Furthermore, a similar result holds for a collision between self-gravitating plane waves of *every* physically realistic field. In Sec. II this result will be stated and proved in the form of a singularity theorem. The proof will be a simple geometrical argument based on the proof of Penrose's singularity theorem.[9,10] A discussion of the physical significance of the assumptions in the theorem will be found in Sec. III, together with an analysis of the Stoyanov result. It will be emphasized in Sec. III that the singularity result depends crucially on the assumption of plane symmetry: There is absolutely no reason to believe, as suggested by Penrose and Khan[3], that a collision between weak nonplanar gravitational waves would result in singularities.

In this paper it will be assumed that colliding plane waves correspond to a spacetime in which there exists everywhere a pair of commuting spacelike Killing vectors, s_1^a and s_2^a. It should be emphasized that this is an assumption; if such Killing vectors are assumed to exist on an initial partial Cauchy surface, then it can be shown[11] that such Killing vectors will also exist in the domain of dependence of this partial Cauchy surface, but this will not establish global existence, since spacetimes with plane waves will not, in general, possess Cauchy surfaces.[12,13] At each point of such a spacetime there will exist two null directions orthogonal to the two spacelike Killing vectors. We let l^a and n^a be null vectors in these two directions, with $l^a n_a = 1$, and m^a the complex vector which spans the two-surface generated by the Killing vector fields s_1^a and s_2^a. The vectors l^a and n^a will be tangent to null geodesics since s_1^a and s_2^a are Killing vectors; u and v will be affine parameters along the geodesics defined by l^a and n^a, respectively. The tetrad (l^a, n^a, m^a, $\bar{m}^a$) defines the usual Newman-Penrose scalars, for which I will use the notation of Hawking,[14] and of Hawking and Ellis.[10] We will say that a necessary condition for a plane-wave collision to have occurred is that at least *one* of the quantities $\psi_0 \equiv C_{abcd} l^a m^b l^c m^d$, $\psi_4 \equiv C_{abcd} n^a \bar{m}^b n^c \bar{m}^d$, $\phi_{00} \equiv \frac{1}{2} R_{ab} l^a l^b$, $\phi_{44} \equiv \frac{1}{2} R_{ab} n^a n^b$, $\sigma \equiv l_{a;b} m^a m^b$, $\lambda \equiv -n_{a;b} \bar{m}^a \bar{m}^b$ is nonzero somewhere in the spacetime.

II. THE SINGULARITY THEOREM AND ITS PROOF

Theorem

Let (M,g) be a spacetime with g at least C^2, and suppose (M,g) has two globally defined commuting spacelike Killing vector fields s_1^a and s_2^a, which together generate plane symmetry. If (1) the null convergence condition holds; (2) at least one of the six Newman-Penrose quantities ψ_0, ψ_4, ϕ_{00}, ϕ_{44}, σ, λ is nonzero at some point p in (M,g); and (3) through the point p there is a spacelike partial Cauchy surface S, which is everywhere tangent to s_1 and s_2, and S is noncompact in the spacelike direction normal to s_1 and s_2; then (M,g) is null incomplete.

Proof of the theorem

The proof is a straightforward generalization of Penrose's theorem (Ref. 15; Ref. 10, p. 263). Suppose, on the contrary, that (M,g) is null complete, and suppose it is one of $(\psi_0, \phi_{00}, \sigma)$ which is nonzero at p. Then along the null geodesic tangent to l^a we have

$$\frac{d\rho}{dv} = \rho^2 + \sigma\bar{\sigma} + \phi_{00}\,, \qquad (1)$$

$$\frac{d\sigma}{dv} = 2\rho\sigma + \psi_0\,. \qquad (2)$$

These equations and the null convergence condition (i.e., that $\phi_{00} \geq 0$) imply that there is a conjugate point to p an affine distance v_0 to the future or the past of p. We may suppose that it is to the future of p. Let Q be the two-surface generated by sweeping p under the action of s_1^a and s_2^a, and consider $\dot{J}^+(Q)$. This achronal boundary will be generated by the null geodesics tangent to l^a or n^a, and having end points at Q. [This structure of $\dot{J}^+(Q)$ is guaranteed by the plane symmetry and the fact that S is a partial Cauchy surface through Q.] Using the plane symmetry and Proposition 4.5.14 of Ref. 10, we see that each null geodesic generator of $\dot{J}^+(Q)$ tangent to l^a leaves $\dot{J}^+(Q)$ at the same affine parameter distance v_1 from Q, and $v_1 \leq v_0$. Thus $\dot{J}^+(Q)$ must be compact in the l^a direction. Let $\dot{J}_l^+(Q)$ be that portion of $\dot{J}^+(Q)$ which is generated by null geodesics tangent to l^a. Since $\dot{J}_l^+(Q) \subset \dot{J}^+(S)$ by construction, and since the plane symmetry is global, we can define a past-directed C^1 timelike vector field which is normal to s_1^a and s_2^a, and which intersects $\dot{J}_l^+(Q)$ and S. Since $\dot{J}_l^+(Q)$ and S are achronal, each integral curve of this field will intersect $\dot{J}^+(Q)$ and S exactly once. Thus they will define a continuous one-to-one map α: $\dot{J}_l^+(Q) \to S$. If $\dot{J}_l^+(Q)$ were compact in the l^a direction, its image $\alpha(\dot{J}_l^+(Q))$ would also be compact in the direction normal to s_1^a and s_2^a. Since by assumption S is noncompact in this direction, $\alpha(\dot{J}_l^+(Q))$ would have to have a boundary in $S \cap [J^-(\dot{J}_l^+(Q)) - Q]$. But this is impossible since by the plane symmetry and Proposition 6.3.1 of Ref. 10, $\dot{J}_l^+(Q)$ and hence $\alpha(\dot{J}_l^+(Q))$ has no boundary. Thus the assumption of null completeness [which was made to prove $\dot{J}_l^+(Q)$ compact in the l^a direction] is incorrect. If the conjugate point to p were to the past of p rather than to the future as we assumed, we obtain a contradiction by repeating the above argument with past null cones rather than future ones. If it were one of $(\psi_4, \phi_{44}, \lambda)$ which were nonzero at p, we can use equations analogous to (1) and (2) to obtain conjugate points to the future or the past of p along the null geodesics tangent to n^a, and then repeat the above argument. In any case, the spacetime (M,g) must be null incomplete. Q.E.D.

III. DISCUSSION

One of the most interesting things about the theorem from the point of view of singularity studies is that it indicates the Cauchy surface condition of Penrose's theorem is not a necessary condition in certain situations of high symmetry. In Penrose's theorem the Cauchy surface condition was used for two purposes: first, to ensure that the generators of $\dot{J}^+(Q)$ have end points at Q, and second, to ensure that the map α was continuous and carried every point of $\dot{J}^+(Q)$ into the Cauchy surface S. (In Penrose's theorem, Q would be a trapped surface.) In cases like plane-wave geometries, in which the essential features of the causal structure have been reduced to two dimensions – one space and one time – the above-mentioned conditions are ensured by the existence of just a *partial* Cauchy surface through Q. Similarly, null incompleteness would be guaranteed in a spherically symmetric spacetime with the global structure of the Reissner-Nordström solution, i.e., with the global topology $S^2 \times R^2$. This is not to say that one cannot construct a nonsingular Reissner-Nordström-type spacetime with matter which obeys the null convergence condition. Indeed Bardeen (Ref. 10, p. 265) has shown that such a matter field can be used to smooth out the $r=0$ Reissner-Nordström singularities into origins of polar coordinates. However, such a smoothing operation changes the topology; in the corresponding Penrose diagram the $r=0$ "points" no longer represent two-spheres but just points—coordinates built on the spherical symmetry are degenerate there and thus the C^1 map α fails to exist. Such a situation cannot arise with the plane symmetry assumed in the theorem of this paper. (See Ref. 16 for other examples of null incompleteness in spacetime with spherical symmetry but not necessarily Cauchy surfaces.)

It is fortunate that a Cauchy surface condition is not required to prove that singularities result from plane-wave collisions, for in general plane-wave spacetimes do not possess Cauchy surfaces.[12,13] This makes it impossible to obtain the global structure of the spacetime by solving evolution equations, unless some tacit global continuity assumptions are made. (It will be shown below that this is the origin of the dispute mentioned in Sec. I about the inevitability of singularities.) The partial Cauchy surface existence assumption is rather innocuous, serving mainly to eliminate identifications which would make the generators of $\dot{J}^+(Q)$ never intersect Q. The requirement that the partial Cauchy surface be noncompact in directions normal to the plane-symmetry Killing vectors ensures that the plane waves can be regarded as coming "from infinity," but it is probably not a necessary requirement since the Penrose theorem holds for any nonspherical topology Cauchy surface,[17] and spherical topology is inconsistent with the planar symmetry.

In view of the singularity theorem of Sec. II, what are we to make of Stoyanov's claim that singularities are *not* inevitable? Stoyanov's "solution" definitely satisfies conditions 2 and 3 of the theorem (both ψ_0 and ψ_4 are nonzero in the interaction region), and Stoyanov "claims" it satisfies (1) also. I say "claims" because Nutku[8] has asserted that Stoyanov's "solution" has a Ricci tensor with an infinite discontinuity across the hypersurface $u=0$, and such a discontinuity may violate null convergence. However this may be, there is no question but that Stoyanov's "solution" has discontinuities in the first derivatives of the metric across this hypersurface, and thus the metric is not C^2 everywhere, as required in the singularity theorem of Sec. II. It is the breakdown of this assumption which is the basic reason why Stoyanov's "solution" has no singularities. It is also why I have placed the word "solution" in quotes. I tend to regard a true solution of the Einstein equations as one which is at least C^2 so that differential geometry techniques can be used to study it, though there are others, notably A.H. Taub,[18] who feel such a differentiability requirement too restrictive. In the Stoyanov "solution" one has the following curvature invariant in the $u>0$, $v>0$ region:

$$R_{abcd}R^{abcd}=\frac{1}{16}(a^2-1)^2(4a^2+7)(1+u+v)^{-(a^2+3)} . \quad (3)$$

If the constant a were $\neq 1$, and the metric analytically continued into the $u<0$, $v<0$ region, it is clear that one would obtain a singularity; the affine parameters u and v could not be continued to minus infinity. The singularity theorem of Sec. II asserts that in the plane-wave collision a singularity must occur somewhere either to the future or the past of the collision if the metric is C^2. What Stoyanov has done is to impose conditions so as to obtain a nonsingular future, and then relax the differentiability requirements sufficiently to allow him to replace the singular past which would otherwise result with flat space. Such a procedure could be used to remove not only the singularities of plane-wave collisions but also the singularities of cosmology[19] and gravitational collapse.[20] Thus rather than being removed by increased smoothness across the wave fronts as Stoyanov claims, the singularity is actually being removed by *decreased* smoothness across certain null hypersurfaces. Singularity-free C^2 plane gravitational waves do exist,[21] but they violate condition (2) of the theorem, and hence do not correspond to a plane-wave collision.

To consider now the case of a collision between two self-gravitating plane waves of nongravitational fields, we first note that all physically observed matter tensors are classified by Hawking and Ellis (Ref. 10, p. 89) at each point either as type I—the case in which the matter tensor has a timelike eigenvector t^a and hence may be expressed in diagonal form with respect to an orthonormal basis — or type II, which represents radiation travelling in a single null direction. Clearly at the point of collision of two waves the tensor will not be of type II. Adopting the orthonormal basis (s_1^a, s_2^a, r^a, t^a), we have $\phi_{00}=\phi_{44}=4\pi(\mu+p_3)$ at the point of collision. For the null convergence condition to hold we must have $\mu \geq 0$ and $\mu + p_3 \geq 0$. Since for physically reasonable matter fields (in particular, for those which satisfy the dominant energy condition) we must have $\mu > 0$ at points where $T_{ab} \neq 0$, and since fields with $\mu = -p_3$ would be unphysical at least at the point of collision, it follows that at the point of collision where it is assumed that $T_{ab} \neq 0$, we must have either ϕ_{00} or ϕ_{44} strictly positive. Thus by the theorem of Sec. II, such a plane-wave collision must result in singularities.

The fact that self-gravitating colliding plane waves of any physically realistic field must give rise to singularities strongly suggests that the singularities are a consequence of the plane symmetry and not a consequence of the nonlinear nature of the gravitational field. For instance, a collision between an expanding core of spherically symmetric star with a collapsing envelope of the star would be an example of a collision of spherically symmetric self-gravitating matter waves, and an equation of state of the matter can be found which would allow a bounce (if the relative wave velocities were sufficiently small) and thus no singularities would appear. Yet the matter tensor would still satisfy the conditions placed on the matter tensor in the plane-wave singularity theorem. (At least this would be possible if the star were sufficiently small. As the mass of the star increases, the radius it must have to avoid collapse to a singularity increases also, until there is so much mass that no reasonable equation of state can prevent collapse to a black hole.) In addition, the remarks at the beginning of this section and a close study of the proof of the singularity theorem make it clear how essential the plane symmetry assumption is in proving the inevitable existence of singularities. The Bardeen example suggests that even a slight deviation from plane symmetry would make it impossible to prove singularities inevitable in a wave collision. But there is no doubt that singularities are inevitable when the colliding waves are plane, and when the evolution preserves the plane symmetry.

ACKNOWLEDGMENTS

I should like to thank Professor J.A. Wheeler, Professor W. Unruh, and Professor R. Matzner for helpful discussions. This work was supported by the National Science Foundation under Grant No. PHY 78-26592.

---

[1] A. Fischer, J. E. Marsden, and V. Moncrief, in *Essays In General Relativity*, edited by F. J. Tipler (Academic, New York, 1980), p. 79.
[2] J. D. Barrow and F. J. Tipler, Phys. Rep. 56C, 371 (1979).
[3] K. Khan and R. Penrose, Nature 229, 185 (1971).
[4] P. Szekeres, Nature 228, 1183 (1970).
[5] P. Szekeres, J. Math. Phys. 13, 286 (1972).
[6] Y. Nutku and M. Halil, Phys. Rev. Lett. 39, 1379 (1977).
[7] Y. G. Sbytov, Zh. Eksp. Teor. Fiz. 71, 2001 (1976) [Sov. Phys.—JETP 44, 1051 (1977)].
[8] B. J. Stoyanov, Phys. Rev. D 20, 2469 (1979).
[9] R. Penrose, Phys. Rev. Lett. 14, 57 (1965).
[10] S. W. Hawking and G. R. F. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, 1973).
[11] V. Moncrief, J. Math. Phys. 16, 493 (1975).
[12] R. Penrose, Rev. Mod. Phys. 37, 215 (1965).
[13] R. Penrose, in *Battelle Rencontres*, edited by C. M. DeWitt and B. S. DeWitt (Gordon and Breach, New York, 1973), p. 1.
[15] R. Penrose, Phys. Rev. Lett. 14, 57 (1965).
[16] F. J. Tipler, Phys. Rev. D 16, 3359 (1977).
[17] F. J. Tipler, Phys. Rev. D 17, 2521 (1978).
[18] A. H. Taub, Gen. Relativ. Gravit. 10, 1009 (1979).
[19] R. Tabensky and A. H. Taub, Commun. Math. Phys. 29, 61 (1973).
[20] P. E. Parker, Ph.D. thesis, Oregon State University, 1977.
[21] J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962), p. 96.