PHYSICAL REVIEW D VOLUME 28, NUMBER 6 15 SEPTEMBER 1983

Stellar collapse without singularities?

Thomas A. Roman
Department of Physics, Syracuse University, Syracuse, New York 13210
and Department of Mathematics and Physics, University of Hartford,
*West Hartford, Connecticut 06117**

Peter G. Bergmann
Department of Physics, Syracuse University, Syracuse, New York 13210
and Department of Physics, New York University, New York, New York 10003
(Received 29 November 1982)

For the singularity theorems of Hawking and Penrose to hold, the stress-energy tensor of matter must satisfy certain restrictions. A model is developed representing the interior of a collapsing, spherically symmetric cloud of matter, which is everywhere free of singularities, due to a relaxation of the so-called "weak energy condition." The regions of the model in which the condition must be violated, and the properties of matter and energy which result, are determined. The indications are that, at least in spherically symmetric cases involving very large masses, the energy condition must be violated in a region where the density is no larger than normal matter densities, which seems physically implausible. Hence suspending the energy conditions may not be a reasonable approach to the avoidance of singularities.

I. INTRODUCTION

This paper is concerned with one aspect of the singularity theorems due to R. Penrose, S. Hawking, and R. Geroch, which assert that, under fairly general conditions, singularities of spacetime cannot be avoided. These conditions include some causality requirements, and also certain restrictions on the stress-energy tensor of matter. We shall examine the possible circumvention of the impact of these theorems by a relaxation of what are known as "the energy conditions."

The "weak energy condition" with which we shall be concerned requires that $T_{\mu\nu}W^{\mu}W^{\nu}\geq 0$ for every timelike vector W^{μ}. By continuity, this inequality also holds when W^{μ} is a null vector. This assumption is equivalent to saying that the energy density as measured by any observer is non-negative. The "strong energy condition" requires that $(T_{\mu\nu}-\frac{1}{2}g_{\mu\nu}T)W^{\mu}W^{\nu}\geq 0$, with W^{μ} again any timelike vector; this inequality implies that gravity is always an attractive force. These conditions ensure that matter always has a converging effect on null and timelike geodesics, respectively. (For details, see Hawking and Ellis,[1] which will be referred to as HE.) Formally, the convergence of a congruence of timelike or null geodesics, c, is defined as $c=-W^{\mu}{}_{;\mu}$, with W^{μ} representing the field of tangent vectors, taken to be unit vectors when timelike, and indicating the increase in the affine parameter when null.

One can initiate convergence in bundles of geodesics by various methods. For example, a theorem involving the gravitational collapse of a star, such as Penrose's theorem (Penrose,[2] HE, p. 263), assumes the existence of a *closed trapped surface,* which is defined to be a closed spacelike two-surface on which both the ingoing and outgoing null geodesics orthogonal to the surface are converging. This concept is intended to make rigorous the intuitive notion of light being "dragged back" by a massive body. (The prototype closed trapped surfaces are the surfaces $t=$const, $r=$const for $r<2M$ in the Schwarzschild solution.) The energy conditions guarantee that convergence, once initiated, will continue until conjugate points form in the geodesic congruence, a conjugate point being a point where infinitesimally neighboring geodesics intersect. This result can then be shown to contradict the assumed global causal properties of the spacetime, indicating that the spacetime is geodesically incomplete.

To evaluate the implications of the singularity theorems we may ask the following: Can the assumptions involved be relaxed sufficiently so that singularities are avoided without producing even more objectionable physical situations? Attempts to circumvent the predicted singularities have usually consisted of modifying either the causality requirements or the energy conditions. Work by Tipler[3] seems to indicate that relaxation of the causality requirements (by allowing, for example, the existence of closed timelike curves) is more likely to facilitate the formation of singularities than to prevent them. Bekenstein[4] has created a closed homogeneous and isotropic model universe which "bounces," under certain circumstances, after reaching a critical density upon recollapse. The "bounce" results from a violation of the strong energy condition by a classical conformal massless scalar field which constitutes part of the total stress-energy tensor of the model. A paper of Tipler's[5] suggests that, in fact, a more realistic version of Bekenstein's model (i.e., one that allows the development of inhomgeneities) would evolve singularities. Tipler has shown that singularities predicted by the Hawking-Penrose theorem will still occur if the strong energy condition is replaced by the weak energy condition and the requirement that the strong energy condition hold on the average. (For a review of this and other results concerning singularities and horizons, see Tipler, Clarke, and Ellis.[6]) In considering whether the weak energy condition is violated, Tipler shows that if $T_{\mu\nu}W^{\mu}W^{\nu}$ is bounded from below for all unit timelike vectors W^{μ}, and if $T_{\mu\nu}$ is type I (see HE, p. 89), then $T_{\mu\nu}K^{\mu}K^{\nu}\geq 0$ for all null vectors K^{μ}.

In our investigation, we work exclusively with *Penrose's theorem,* which states the following.

Spacetime (M,g) cannot be null geodesically complete if

(1) $R_{\mu\nu}K^{\mu}K^{\nu}\geq 0$ for all null vectors K^{μ} [(1) is implied by the weak energy condition],

(2) there is a noncompact Cauchy surface H in M,

(3) there is a closed trapped surface T in M.

We ask how badly must one violate the weak energy condition (while maintaining the other assumptions of Penrose's theorem) in order to obtain a spacetime which is singularity free. Our motivation for considering violations of the weak energy condition (which is certainly satisfied by all known forms of classical matter) is the possibility that under the extreme conditions of high density and large spacetime curvature, as would exist in the late stages of gravitational collapse, one's notion of what constitutes "reasonable" behavior of matter may have to be modified.

In Sec. II, we formulate a general line element for a singularity-free, spherically symmetric spacetime, containing trapped surfaces, that might represent a collapsing "star." Differentiating that line element leads to an expression for $G_{\mu\nu}$, the Einstein tensor, and hence to $T_{\mu\nu}$, the stress-energy tensor. If the spacetime is to be singularity-free, in the presence of a trapped surface, the weak energy condition must be violated somewhere, since we assume that the other condition of the theorem, i.e., the noncompact-Cauchy-surface requirement, is satisfied. The goal of Sec. III is to discover where in spacetime the weak energy condition must be violated. The resulting properties of the stress-energy tensor are examined in Sec. IV. Difficulties with our model do arise, in that it predicts that for large masses the weak energy condition must be violated even in regions of low density, a result that appears physically implausible. These difficulties are discussed in Sec. V, and very briefly, the question of whether the energy condition violation required by the model can be provided by particle production processes, due to the nonstationary character of the metric, is considered. Our general conclusion is that, at least in the spherically symmetric case, violation of the weak energy condition does not seem to be a reasonable way of avoiding singularities.

Units and signs. We work in units of $G=c=1$ and our metric signature is chosen to be $(-,+,+,+)$.

II. THE METRIC AND THE GEOMETRIC PICTURE

Our model is to have the following properties. (a) Absence of singularities, (b) spherical symmetry, (c) presence of trapped surfaces, and (d) the existence of a noncompact Cauchy surface. Conditions (c) and (d) are required for Penrose's theorem to hold. We will assume that the spacetime is asymptotically simple and empty (HE, pp. 222 and 223). Such a spacetime will be asymptotically flat and will admit a Cauchy surface; thus condition (d) is satisfied. As we require freedom from singularities, the weak energy condition must be violated somewhere in the model. We have restricted ourselves to a spherically symmetric model in order to simplify the situation, both calculationally and intuitively. We will consider these four physical properties in turn.

The line element representing the interior of a collapsing spherically symmetric cloud of matter is taken to be

$$ds^2=-2F(u,v)du\,dv+r^2(u,v)d\Omega^2\ , \tag{2.1}$$

where $d\Omega^2=d\theta^2+\sin^2\theta\,d\phi^2$, $r=r(u,v)$ is the luminosity radius, and u and v are null coordinates. We work in null coordinates to avoid the types of coordinate problems that one encounters in the Schwarzschild metric at $r=2M$, and the interpretational ambiguities that arise for the region $r<2M$. F is an undetermined function of u,v; the only *a priori* requirements are that $F>0$ and finite everywhere, and that F is C^2. [In order to satisfy (a), F must be nonzero everywhere, hence it has the same sign throughout; it is positive if both u and v are chosen to be future-pointing.] The luminosity radius $r(u,v)$ is a nonvanishing positive C^2 function of u and v (except at the center where $r=0$).

Because the functions F and r, and their derivatives, will appear later in the stress-energy tensor, we shall examine their transformation properties. As relative signs of terms in $T_{\mu\nu}$ will be important in consideration of the energy conditions, let us list the quantities that have covariant signs. The only coordinate transformations preserving the character of our null coordinates (i.e., future-pointing, linearly independent) must be of the form

$$\bar{u}=\bar{u}(u)\ ,\quad \frac{du}{d\bar{u}}>0\ ,$$

and (2.2)

$$\bar{v}=\bar{v}(v)\ ,\quad \frac{dv}{d\bar{v}}>0\ .$$

The only quantities whose signs are invariant under the transformations (2.2) are r, its first partial derivatives $r_{,u}$ and $r_{,v}$, its second mixed derivative $r_{,uv}$, and F. The derivatives $r_{,uu}$ and $r_{,vv}$ may change signs under (2.2), except when the first derivatives of r are zero. Although the sign of F is invariant, those of its first and higher derivatives are not. These fairly obvious invariance properties will be used later in this section and in Sec. III.

Let us now focus our attention on property (c), the presence of a trapped surface. The usual diagrams illustrating the gravitational collapse of a star that one finds in the literature (e.g., Hawking[7]) show a trapped surface being formed when the surface of the star passes through its Schwarzschild radius at $r=2M$. The first trapped surface is depicted as appearing in the vacuum outside the collapsing matter. Although such pictures are useful in emphasizing the main features of gravitational collapse, they do not illustrate the general case. Clearly there is no need to synchronize the formation of the first trapped surface with the passage of the star's boundary through $r=2M$. A real star has no such well-defined "surface" in any case, since its mass distribution tapers off gradually. Consider a massive spherically symmetric collapsing cloud of matter with mass M. A trapped surface will still form if some portion of the cloud with mass ΔM, passes through *its* Schwarzschild radius, $r=2\Delta M$. This portion will be surrounded by the rest of the matter distribution; hence, in general, the first trapped surface forms inside the body of the collapsing object. As the collapse continues, a region of trapped surfaces will develop and grow both outward from the center and inward toward the center. When the entire mass M has fallen inside $r=2M$, the region of trapped surfaces extends into the vacuum surrounding the star, and a black hole of mass M is formed.

Although trapping *need not* begin at the surface of the collapsing object, it *cannot* begin at $r=0$. In the spherically symmetric case, trapping at $r=0$ is coincident with the presence of a singularity. If $r=0$ is spacelike and therefore trapped, timelike and null geodesics cannot pass through the origin and remain causal. Such geodesics will have end points on the curve $r=0$, and thus the spacetime will be geodesically incomplete. (This is also true in the limiting case when the curve $r=0$ has an inflection point, where its tangent is null. A timelike geodesic encountering this point could not pass through the origin without becoming null or spacelike.)

An instructive example illustrating the relationship between trapped surfaces and singularities is the case of a collapsing thin spherical shell. Outside the shell, the metric is that of Kruskal. The metric inside the shell is Minkowskian, since the spacetime is flat. When the shell passes through its Schwarzschild radius, trapped surfaces will form in the vacuum outside the shell. The vacuum spacetime interior to the shell, however, is still Minkowskian, and will remain so until the imploding shell reaches $r=0$. Thus, in this case, the trapped region grows inward as the shell collapses, eventually intersecting $r=0$, with the formation of a singularity.

As a prelude to the discussion of the formation of trapped surfaces in our model, we now consider some features of Kruskal's extension of the Schwarzschild line element (Kruskal,[8] HE, p. 155). The Kruskal metric may be cast in the form (2.1)

$$ds^2=-32\frac{M^3e^{-r/2M}}{r}du\,dv+r^2d\Omega^2\ . \tag{2.3}$$

The function F of Eq. (2.1) is $F=(16M^3/r)e^{-r/2M}$ and r, in turn is given implicitly by

$$\left[\frac{r}{2M}-1\right]e^{r/2M}=-uv\ .$$

Conversely, here $u=-e^{-(t-r^*)/4M}$ and $v=e^{(t+r^*)/4M}$ with

$$r^*=r+2M\ln\left|\frac{r}{2M}-1\right|\ ,$$

and t being the Schwarzschild time. (The relationship between u, v, r, and t given above holds for the region $r>2M$, $-\infty<t<+\infty$. For further details, see Misner, Thorne, and Wheeler.[9])

On a Kruskal diagram we can examine the $r=\text{const}$ curves for the three regions $r<2M$, $r=2M$, and $r>2M$ (see Fig. 1). In the region $r>2M$, the $r=\text{const}$ curves are timelike hyperbolas with null asymptotes $r=2M$, $t=+\infty$ and $r=2M$, $t=-\infty$. In the region $r<2M$, the $r=\text{const}$ curves are spacelike hyperbolas with the same asymptotes. The line $r=2M$, $t=+\infty$ is the future event horizon, i.e., it is the boundary of the region from which it is possible to escape to infinity along a future-directed causal curve. The transition of the $r=\text{const}$ curves from timelike to spacelike is characteristic of the presence of trapped surfaces. Figure 1 is the diagram representing an extension of the Schwarzschild line element which is everywhere matter-free. Note that in this diagram, each $r=\text{const}$ curve is timelike, null, or spacelike in its entirety. The situation is somewhat different in the case of a spherically symmetric cloud of pressureless dust collapsing to form a black hole.

One can represent the interior metric of the dust cloud by a Friedmann geometry which is joined to a Kruskal vacuum exterior. A Kruskal-type diagram illustrates the situation in Fig. 2. (taken, with minor modifications, from Rees, Ruffini, and Wheeler[10]). The curve AF represents the world line of the surface of the cloud, which collapses from radius $3M$ to $r=0$. As before, the line $r=2M$, $t=+\infty$, is the event horizon. All points to the left of AF are inside the dust cloud, all points to the right of AF are in the exterior vacuum region. Note that Fig. 2 differs from Fig. 1 in that each of the $r=\text{const}$ curves for $r<2M$ makes a transition from timelike-to-null-to-spacelike; this transition occurs inside the body of the cloud. Each point where an $r=\text{const}$ curve is spacelike, in both Figs. 1 and 2, represents a trapped two-sphere.

There is an important distinction between a trapped surface and event horizon. The event hor-

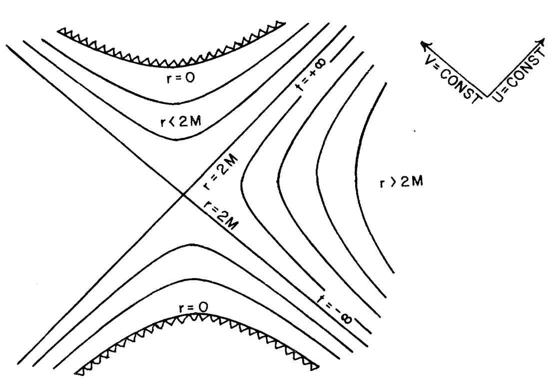

FIG. 1. Kruskal's vacuum extension of the Schwarzschild metric, in null coordinates. Each $r=\text{const}$ curve is timelike, null, or spacelike, in its entirety.

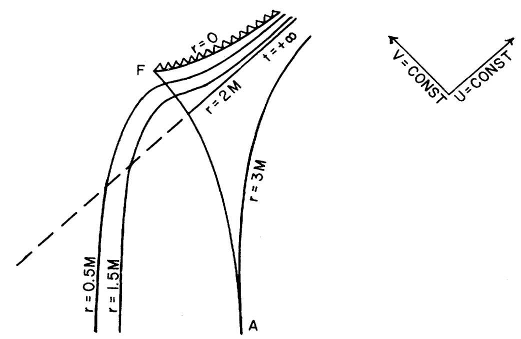

FIG. 2. Collapse of a dust cloud to a black hole, in null coordinates. The $r=\text{const}$ curves for $r<2M$ undergo a transition from timelike to null to spacelike (see Ref. 10).

izon, defined to be the boundary of the past of $\mathscr{I}^+$, is a global concept in the sense that in order to locate the event horizon one must have access to the entire evolution of the spacetime. By contrast, a trapped surface is a local concept, which implies the existence of an event horizon only if the weak energy condition is satisfied. Penrose's theorem applies to a spacetime that contains at least one trapped surface. It states that once a trapped surface has formed in spacetime, and provided the weak energy condition holds, then the degree of convergence increases, until eventually a singularity forms.

Our model is shown in Fig. 3. The crisscrossed lines at 45° are the null lines of constant u and v. The heavy lines represent the congruence of $r=\text{const}$ curves; these are timelike outside the boundary δ, null upon intersecting δ, and spacelike inside the boundary δ. The region interior to δ will be hereafter referred to as the domain of trapped surfaces $\mathscr{D}$. The intrinsic points A,B,C,D are the points where the tangent to δ becomes null. The interior metric for the spherically symmetric collapsing cloud of matter, illustrated by Fig. 3 is described by Eq. (2.1); we require that the spacetime is locally Minkowskian at the origin, which precludes intersection of $r=0$ by $\mathscr{D}$, and asymptotically flat (i.e., "scri" is null).

This model represents a situation in which a region of trapped surfaces develops, exists for some finite time, and then disappears. Kodama[11,12] has recently claimed that the occurrence and disappearance of a trapped region between two partial Cauchy surfaces, satisfying certain conditions, is necessarily accompanied by the formation of a naked singularity. Our model does not exhibit this property, as, by construction it is everywhere singularity-free. Kodama's use of what he refers to as a "locally trapped region" suggests that in his terminology a trapped surface is invisible from $\mathscr{I}^+$, which is true only if the weak energy condition holds (HE propo-

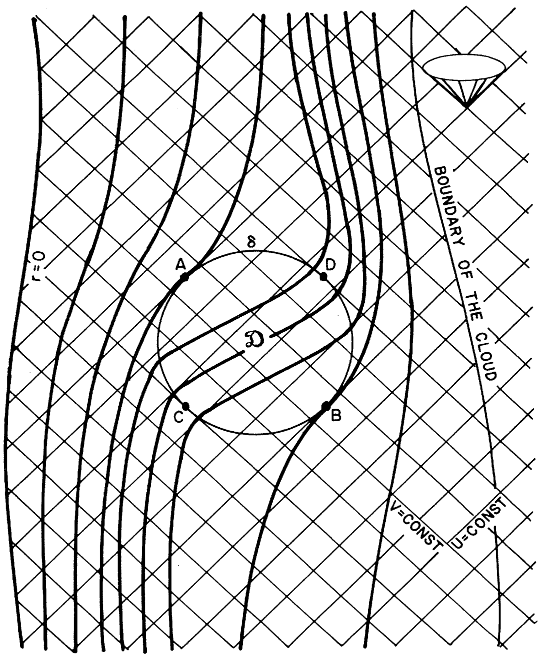

FIG. 3. Our geometric model, showing the domain of trapped surfaces, $\mathscr{D}$. The heavy black lines are $r=\text{const}$ curves. A given $r=\text{const}$ curve is timelike outside $\mathscr{D}$, null upon intersecting $\mathscr{D}$, and spacelike inside $\mathscr{D}$.

sition 9.2.1). Notice that there is *no* event horizon in this model—all world points inside $\mathscr{D}$ are connected with $\mathscr{I}^+$ by null geodesics. Outgoing null rays starting at the center, $r=0$, cross the $r=\text{const}$ curves in an increasing direction until they reach δ; once inside the domain of the trapped surfaces, $\mathscr{D}$, they cross the $r=\text{const}$ lines in a decreasing direction; on leaving $\mathscr{D}$ they again cross the $r=\text{const}$ curves in an increasing direction.

In the u,v plane, we wish to require that $\mathscr{D}$ be connected and simply connected, which means that δ is S^1 (topologically equivalent to a circle); the purpose of these assumptions to keep the model as simple as possible. For example, allowing the $r=\text{const}$ curves to oscillate in character between timelike and spacelike inside $\mathscr{D}$ (i.e., $\mathscr{D}$ nonsimply connected) would needlessly complicate the physical situation.

To view the evolution of the trapped surface region, we may look at a series of spacelike slices through $\mathscr{D}$ in Figs. 4(a) and 4(b). A single "marginally trapped" sphere is present on S_1 (a marginally trapped surface is a closed spacelike two-surface on which the outgoing null geodesics orthogonal to the surface have zero convergence); on S_2 the trapped surfaces occupy a region between two concentric spheres. This region continues to grow inward and outward until S_3; by S_4, it has begun to shrink, disappearing altogether after S_5. Figure 4(b) illus-

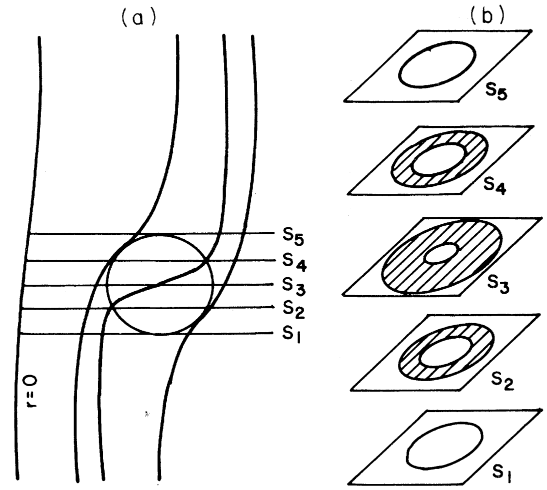

FIG. 4. Time evolution on a series of two-dimensional spacelike slices.

trates this evolution in two dimensions.

As we have previously listed those derivatives of r whose signs are invariant, we will determine the signs of these derivatives in the different regions of Fig. 3 for use in the next section.

In our notation, a $u=\text{const}$ line is parametrized by v; i.e., $\partial/\partial v$ is a directional derivative along a $u=\text{const}$ line. An ingoing null ray ($v=\text{const}$ line) always crosses the $r=\text{const}$ lines in a decreasing order, both inside and outside $\mathscr{D}$. Therefore, $\partial r/\partial u<0$ everywhere in Fig. 3. An outgoing null ray ($u=\text{const}$ line) crosses the $r=\text{const}$ lines in an increasing order outside $\mathscr{D}$, but in a decreasing order inside $\mathscr{D}$. Therefore, $\partial r/\partial v>0$ outside $\mathscr{D}$, $\partial r/\partial v<0$ inside $\mathscr{D}$, and $\partial r/\partial v=0$ on δ, the boundary of $\mathscr{D}$. If we examine $r_{,v}\equiv\partial r/\partial v$ along an ingoing null ray ($v=\text{const}$ line), we see that $r_{,v}$ is positive outside $\mathscr{D}$, and negative inside $\mathscr{D}$, then $r_{,vu}<0$ below a certain line, $r_{,vu}>0$ above that line, and $r_{,vu}=0$ on that line. This line must connect points C and D. (One sees this by choosing the ingoing null ray to be the one passing through C or D. Since $r_{,v}>0$ outside $\mathscr{D}$ and $r_{,v}<0$ inside $\mathscr{D}$, then $r_{,vu}=0$ at points C and D.)

III. THE STRESS-ENERGY TENSOR AND THE WEAK ENERGY CONDITION

Appendix I contains the components of the stress-energy tensor, which were obtained by deriving $G_{\mu\nu}=R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R$ from the line element (2.1), and setting this expression equal to $8\pi T_{\mu\nu}$. We wish to examine the weak energy condition (referred to as WEC) for various timelike and null vectors. The weak energy condition (HE, p. 89) states that

$$T_{\mu\nu}W^{\mu}W^{\nu}\geq 0 \text{ for every timelike vector } W\ . \tag{3.1}$$

By continuity, this inequality also holds when W is a null vector.

In coordinates (u,v,θ,ϕ), the requirement for an arbitrary radial vector $W^{\mu}=(c,K,0,0)$ to be timelike is that c and K have the same sign. Hence for timelike radial vectors the WEC turns into

$$T_{\mu\nu}W^{\mu}W^{\nu}=c^2T_{uu}+2cKT_{uv}+K^2T_{vv}\geq 0\ , \tag{3.2}$$

which in turn is equivalent to a set of inequalities involving the stress-energy tensor alone.

Assume first that $T_{uu}\neq 0$, $T_{vv}\neq 0$. Define a function $f(x)\equiv T_{uu}+T_{uv}+x^2T_{vv}$, where $x\equiv K/c$ with range $0\leq x\leq\infty$. For the WEC to be satisfied, $f(x)$ must be non-negative throughout this range. Looking both at the extremum of $f(x)$, if any, and at the boundaries of the range of x, we obtain the three inequalities

$$T_{uv}\geq-(T_{uu}T_{vv})^{1/2}\ , \tag{3.3}$$

$$T_{vv}\geq 0\ , \tag{3.4}$$

$$T_{uu}\geq 0\ . \tag{3.5}$$

These inequalities can be shown to include the limiting cases in which T_{uu} or T_{vv} vanish, or both. Equation (3.4) is equivalent to the statement of the WEC along an arbitrary null vector in the u direction $(c=0)$, while (3.5) is the same expression along a null vector in the v direction $(K=0)$.

Analogous results can be obtained for nonradial timelike/null vectors. For such timelike vectors with u,v,θ components, a necessary (but not sufficient) condition for the WEC to hold is that

$$T_{\theta\theta}\geq-\frac{r^2}{F}[T_{uv}+(T_{uu}T_{vv})^{1/2}]\ . \tag{3.6}$$

Equation (3.6) is a sufficient condition for null vectors, and therefore only a necessary condition for timelike vectors, since the WEC is violated for some timelike vectors if it is violated for null vectors. (This is simply the contrapositive of the earlier statement that if the WEC is satisfied for timelike vectors, then it is satisfied for null vectors.) An expression similar to (3.6) may be obtained for the timelike/null vectors with u,v,ϕ components.

Using the relevant expressions from Appendix I, (3.4) and (3.5) may be rewritten as

$$\frac{-r_{,vv}}{r}+\frac{r_{,v}F_{,v}}{rF}\geq 0\ , \tag{3.7}$$

$$\frac{-r_{,uu}}{r}+\frac{r_{,u}F_{,u}}{rF}\geq 0\ , \tag{3.8}$$

respectively. Equations (3.3) and (3.6) may be recast in a similar fashion, but because of their length, we do not include them explicitly. (Note that the apparent singularities that occur in certain terms, at $r=0$, are merely coordinate singularities, due to our choice of spherical coordinates. Recall that our metric is intrinsically singularity-free, by construction.)

As our next step we shall examine the inequalities stemming from (3.3) to (3.6) in different regions of Fig. 3, so as to ascertain where the WEC violation occurs. Our approach is to assume initially that the energy condition is satisfied everywhere and then attempt to arrive at a contradiction. The discussion of such equations as (3.7) and (3.8) is hampered by the occurrence of derivatives of r and F whose signs are not covariant. However, for Eq. (3.7) *only* these derivatives drop out on the boundary δ, so this is where we shall focus our attention. On the boundary, Eq. (3.7) reduces to

$$-r_{,vv}\geq 0\ , \tag{3.9}$$

an invariant statement as $r_{,v}=0$ on δ.

Consider (3.9) on different portions of δ. At A, (3.9) is trivially satisfied; since A is an inflection point of an $r=$const curve, $r_{,vv}=0$. The same is true at B. Along the open interval ABC, $r_{,vv}<0$; therefore (3.9) is satisfied; along open interval ADB, $r_{,vv}>0$; therefore (3.9) cannot be satisfied. This violation of the WEC along ADB is *independent of F and its derivatives;* there is no way to adjust F for its derivatives in order to preserve the energy condition in this region—its violation is unavoidable because of the geometry of the model. The only way to remove it would be to allow the $r=$const curves emerging from $\mathscr{D}$ to be timelike past-pointing. (Deforming the shape of $\mathscr{D}$ will not remove this feature: it is a characteristic of the transition of the $r=$const curves from spacelike to timelike, i.e., of trapped surfaces becoming "untrapped.") Placing the curve $r=0$ to the right of $\mathscr{D}$, instead of to the left, in Fig. 3, is not a viable alternative either; in that case an outgoing future-pointing null ray from the origin always crosses the $r=$const curves in an increasing order; therefore, there are no (future) trapped surfaces.

A little more information as to the extent of the region of energy condition violation can be obtained by examining the Raychaudhuri equation (HE, Secs. 4.1 and 4.2) along a radially outgoing null geodesic that passes through the interior of $\mathscr{D}$. That equation is

$$\frac{dc}{d\lambda}=R_{\mu\nu}U^{\mu}U^{\nu}-2\omega^2+2\sigma^2+\frac{c^2}{2}\ , \tag{3.10}$$

where c is the convergence of infinitesimally neighboring null geodesics, ω is the vorticity, σ is the shear (see HE, Secs. 4.1 and 4.2 for details; in our notation $c=-\theta$, where θ is the expansion used in HE), and λ is the affine parameter along the geodesic. Because of the spherical symmetry of our model, the shear and vorticity vanish for radial geodesics and the Raychaudhuri equation reduces to

$$\frac{dc}{d\lambda}=R_{\mu\nu}U^{\mu}U^{\nu}+\frac{c^2}{2}\ , \tag{3.11}$$

or, using the field equations and setting $U=(0,\alpha,0,0)$, to

$$\frac{dc}{d\lambda}=8\pi\alpha^2T_{vv}+\frac{c^2}{2} \tag{3.12}$$

along radially outgoing null geodesics. In fact, the Raychaudhuri equation (3.12) is simply the field equation (see Appendix II):

$$8\pi T_{vv}=-\frac{2r_{,vv}}{r}+\frac{2r_{,v}F_{,v}}{rF}\ . \tag{3.13}$$

Outside $\mathscr{D}$, the convergence is negative. On the interval ACB, the convergence along U^{μ} is zero, as it is on interval ADB. In the interior of $\mathscr{D}$, the convergence is positive, since each point in this region represents a trapped surface. Therefore, the convergence must have at least one maximum somewhere inside $\mathscr{D}$. Consider the simplest case of a single maximum. Let λ_1 be the point where the null geodesic intersects ACB (i.e., enters $\mathscr{D}$), λ_3 the point where it intersects ADB (i.e., exits $\mathscr{D}$), and λ_2 the point in the interior of $\mathscr{D}$ where the convergence reaches its maximum.

At the point λ_1, $c=0$ and $dc/d\lambda>0$, as the null geodesic is passing from a region of negative convergence to a region of positive convergence. Hence, $T_{vv}>0$ at λ_1. This argument holds at all points on the open interval ADB. At λ_2, $dc/d\lambda=0$ and $\alpha^2T_{vv}=-c^2/2$, implying $T_{vv}<0$ at λ_2. At λ_3, $c=0$ and $dc/d\lambda<0$; therefore $T_{vv}<0$ everywhere on open interval ADB in accordance with our earlier result. As T_{vv} is positive at λ_1 and negative at λ_2, T_{vv} must have a zero between λ_1 and λ_2. It follows directly from the field equation (3.13) that $T_{vv}=0$ at the points A and B, since $r_{,v}=0$ at all points of δ, the boundary of $\mathscr{D}$, and $r_{,vv}=0$ at A and B, as both are inflection points for the $r=$const curves that pass through them.

If we connect the points along each null geodesic where T_{vv} changes sign, and connect the points along each null geodesic where $dc/d\lambda$ changes sign, we obtain two "spines" connecting points A and B, as illustrated in Fig. 5. The solid dark line connects the turning points of T_{vv}, and the dash-dotted line connects the turning points of $dc/d\lambda$. We again note that along each null geodesic (with the exception of those passing through points A and B), T_{vv} changes sign before (i.e., at a smaller value of the affine parameter) the convergence has reached its maximum value.

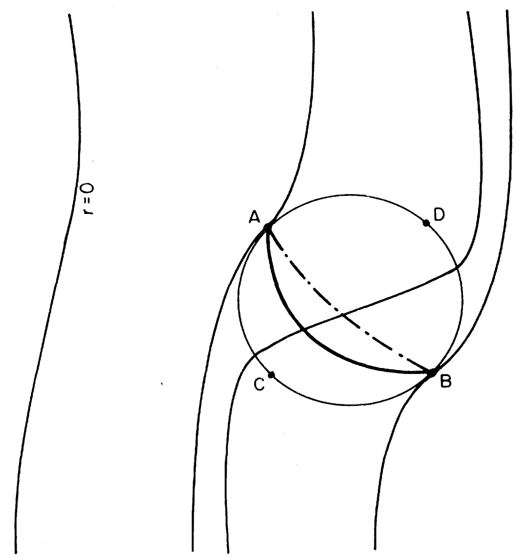

FIG. 5. Zero lines of T_{vv} and $dc/d\lambda$. The solid dark line connects the turning points of T_{vv}, and the dash-dotted lines connects the turning points of $dc/d\lambda$.

Thus the region of unavoidable WEC violation extends beyond a narrow neighborhood of the boundary ADB, and penetrates deep into the interior of $\mathscr{D}$ as well. (We have shown that the WEC is violated for null vectors in these regions; therefore, by continuity, it is also violated for some timelike vectors.) Violations may occur in other regions of Fig. 3, however such violations do not seem to be forced upon us, in contrast to the cases just discussed.

IV. CONSEQUENCES OF THE VIOLATION OF THE WEAK ENERGY CONDITION

In order to examine the physical consequences of the WEC violation, we will express the components of $T_{\mu\nu}$ in a local Lorentz frame. Defining new timelike and spacelike coordinates, t and x, by

$$t=\frac{1}{\sqrt{2}}(v+u)=x^0\ ,\quad x=\frac{1}{\sqrt{2}}(v-u)=x^1\ , \tag{4.1}$$

we transform $T_{\mu\nu}$ to obtain

$$\begin{aligned} T_{00}&=\tfrac{1}{2}(T_{uu}+T_{vv}+2T_{uv})\ ,\\ T_{11}&=\tfrac{1}{2}(T_{uu}+T_{vv}-2T_{uv})\ ,\\ T_{01}&=T_{10}=\tfrac{1}{2}(T_{vv}-T_{uu})\ . \end{aligned} \tag{4.2}$$

A straightforward calculation shows that, in a local Lorentz frame, an arbitrary two-dimensional stress-energy tensor can be diagonalized by a Lorentz transformation (i.e., the T_{10} component vanishes in the new frame of reference) only if the following inequality holds:

$$4T_{10}{}^2 \leq (T_{00}+T_{11})^2 \ . \tag{4.3}$$

Equation (4.3) is a necessary (though not sufficient) condition for diagonalization. [This is the condition that the usual special-relativistic velocity parameter β, of the transformation, be real. Since β must also be less than 1, (4.3) is not sufficient.] Substituting expressions (4.2) in (4.3) we arrive at the condition

$$T_{uu}T_{vv} \geq 0 \ , \tag{4.4}$$

which is equivalent to (4.3). $T_{uu}T_{vv}>0$ is satisfied only if T_{uu} and T_{vv} have the same sign.

We have seen in Sec. III that there is no reasonable way to avoid a violation of the weak energy condition in the form of $T_{vv}<0$ in a neighborhood of δ, the boundary of $\mathscr{D}$, and in a portion of the interior of $\mathscr{D}$. However, we are not forced to conclude that $T_{uu}<0$ in these regions. The only reason for requiring $T_{uu}<0$ would be to avert the focusing of ingoing null rays. Since ingoing null rays converge even in Minkowski spacetime, there does not seem to be sufficient justification for demanding that $T_{uu}<0$ (Tipler[13]). We will subsequently assume that $T_{uu}\geq 0$, for although it is possible that $T_{uu}<0$, we are not driven to such a conclusion, in contrast to the case of $T_{vv}<0$.

Therefore, in the regions of Fig. 3 where $T_{uu}>0$ and $T_{vv}<0$, the inequality (4.4) cannot be satisfied. There exist no local observers who see zero energy flux.

What are the conditions for $T'_{00}<0$ in some frame of reference, where T'_{00} is the Lorentz-transformed energy density? $T'_{00}<0$ implies that

$$T'_{00}=\gamma^2 T_{00}-2\beta\gamma^2 T_{10}+\beta^2\gamma^2 T_{11}<0 \ , \tag{4.5}$$

where β is the usual special-relativistic velocity parameter, and $\gamma\equiv(1-\beta^2)^{-1/2}$. Assume $T_{00}>0$, $T_{11}>0$. Define $f(\beta)\equiv T_{00}-2\beta T_{10}+\beta^2 T_{11}$, with $-1\leq\beta\leq 1$. Since γ^2 is positive, $f(\beta)$ has the same sign as T'_{00}. If $f(\beta)<0$ at either end of the range, then there will be frames of reference (with $|\beta|$ close to 1) where $T'_{00}<0$. At the ends of the range,

$$\begin{aligned} &T_{00}+T_{11}<2T_{01} \ , \\ &T_{00}+T_{11}<-2T_{01} \end{aligned} \tag{4.6}$$

are the conditions for $f(\beta)<0$. Equation (4.6) can be rewritten as

$$|T_{01}| > \tfrac{1}{2}(T_{00}+T_{11}) \ , \tag{4.7}$$

or

$$4T_{01}{}^2 > (T_{00}+T_{11})^2 \ .$$

Equation (4.7) is the condition that $f(\beta)<0$, on at least one end of the range. Comparing with (4.3), we see that (4.7) is equivalent to the *non*-diagonalizability requirement.

Does $f(\beta)$ have a minimum? Setting $f'(\beta)=0$, we find that

$$\beta_{\min}=\frac{T_{01}}{T_{11}} \ . \tag{4.8}$$

Since $|\beta|$ must be <1,

$$|T_{01}| < T_{11} \ . \tag{4.9}$$

For $f(\beta)<0$ at the minimum, the condition

$$T_{00}T_{11} < T_{01}{}^2 \tag{4.10}$$

must be satisfied along with (4.9). Equation (4.10) can be rewritten as

$$4T_{01}{}^2 > (T_{00}+T_{11})^2-(T_{00}-T_{11})^2 \ . \tag{4.11}$$

Comparing (4.7) with (4.11), we see that a sufficient condition for (4.11) to hold is that (4.7) is satisfied.

Thus in regions of our model where $T_{vv}<0$ and $T_{uu}>0$, $T_{\mu\nu}$ is nondiagonalizable in a local Lorentz frame; this implies that there will also exist local Lorentz observers for whom the energy density is negative.

The fact that $T_{\mu\nu}$ is nondiagonalizable does not necessarily imply noncausal propagation of information. As an example, we consider a (very naive) two-dimensional model in Minkowski spacetime. Take a system whose total stress-energy tensor consists of the stress-energy of two oppositely directed streams of noninteracting particles, one whose particles have positive rest mass, the other consisting of particles with negative rest mass. The particles move with equal but opposite velocities in some Lorentz frame of reference, and in that frame let $\rho_-=-\rho_+$ (i.e., equal but opposite mass densities). The stress-energy tensor for each stream has the form

$$T_{\mu\nu}=\pm\rho_0\xi_\mu\xi_\nu \ , \ (\text{“cold dust”}) \text{ with } \begin{cases}\xi_0=-\gamma \ , \\ \xi_1=\pm\beta\gamma \text{ with } 0<\beta<1 \text{ and } \rho_0>0 \ ,\end{cases}$$

where ρ_0 is the absolute value of the mass density in the rest frame of the particles and β is the magnitude of the velocity of the stream, in units of c, in the observer's Lorentz frame. The total stress-energy tensor for the entire system has the form

$$T_{\mu\nu}=\begin{bmatrix} 0 & -2\rho_0\beta\gamma^2 \\ -2\rho_0\beta\gamma^2 & 0 \end{bmatrix} . \tag{4.12}$$

Applying condition (4.3) to (4.12), we see that (4.3) is violated, i.e., (4.12) cannot be diagonalized by a Lorentz transformation. (In the Lorentz frame we started with, because of the opposite signs of the mass densities, $T_{10}\neq 0$.) However, nondiagonalizability of $T_{\mu\nu}$ does not, in itself, imply a violation of the WEC (e.g., the stress-energy tensor for electromagnetic radiation, all of which is moving in one direction, is nondiagonalizable, but obviously does not violate the WEC).

For a timelike vector $W^{\mu}=(1,0)$, the WEC applied to the stress-energy tensor of (4.12) yields

$$T_{\mu\nu}W^{\mu}W^{\nu}=0 .$$

Similarly for an ingoing null vector $V^{\mu}=(1,-1)$,

$$T_{\mu\nu}V^{\mu}V^{\nu}=4\rho_0\beta\gamma^2>0 .$$

However, for an outgoing null vector $U^{\mu}=(1,1)$ we obtain

$$T_{\mu\nu}U^{\mu}U^{\nu}=-4\rho_0\beta\gamma^2<0 .$$

Hence, the WEC is violated. Interestingly enough, in our model also the WEC is violated along outgoing null rays.

For our model, T^{tx} represents the local energy flux, in terms of local timelike and spacelike coordinates t and x. If $T^{tx}>0$ or equivalently $T_{tx}<0$ (since $T^{tx}=-T_{tx}$), then the flux is outwards (i.e., positive energy is transported away from the center, or negative energy is transported toward the center of the matter cloud). If $T^{tx}<0$ or equivalently $T_{tx}>0$, then the flux is inwards (i.e., positive energy is transported toward the center, or negative energy is transported away from the center of the cloud).

Recalling that $T_{tx}=\frac{1}{2}(T_{vv}-T_{uu})=-T^{tx}$ (note that in null coordinates $T^{uu}=T_{vv}$), we see that if $T_{vv}<0$ and $T_{uu}>0$, then $T^{tx}>0$. Therefore, one consequence of the WEC violation can be interpreted as the existence of an outgoing energy flux (or equivalently, negative-energy transport toward the center), which cannot be reduced to zero in any local Lorentz frame.

The stress-energy tensor (4.12) and that of our model are both examples of a Type-IV stress-energy tensor (HE, p. 90), i.e., $T_{\mu\nu}$ has no timelike or null eigenvector.

V. PHYSICAL PLAUSIBILITY OF THE MODEL

Originally it was planned to develop this scenario. A spherically symmetric cloud of matter is undergoing gravitational collapse. The collapse proceeds until trapped surfaces begin to form inside the body of the cloud. Meanwhile, the density of matter in the core is steadily increasing, to become ultimately many orders of magnitude larger than any densities that have been observed in the laboratory. We postulate that under these "exotic" conditions the weak energy condition may be violated sufficiently that no singularity forms.

We have shown that the WEC is violated at least in a neighborhood of the interval ADB, on the boundary of $\mathscr{D}$, including a region interior to $\mathscr{D}$. Our model will have these same physical properties even for the case of very large masses. That this is so can be seen by considering a "scaling" argument.

In a quasi-Cartesian coordinate system (i.e., one in which the coordinates all have the same dimension, that of length) the Christoffel symbols have the dimension of $(\text{length})^{-1}$. The Einstein tensor, $G_{\mu\nu}$, being formed from derivatives of the Christoffel symbols, has the dimension of $(\text{length})^{-2}$. If one "scales up" the coordinates (i.e., let $x^{\sigma}\to\alpha x^{\sigma}$, $\alpha>0$), then the Christoffel symbols are scaled by a factor of α^{-1}, and the components of $T_{\mu\nu}$ are scaled by a factor of α^{-2}, while the total mass increases by a factor of α. Therefore, if one forms a new solution of the Einstein equations by scaling up the original solution, then the new solution will have the same physical properties (e.g., trapped surfaces) as the original solution, since the new right-hand side of the field equations only differs from the original right-hand side by a factor of α^{-2}.

The $r=$const curve that passes through A, $r_{\min}$, is the smallest value of r intercepted by $\mathscr{D}$; the $r=$const curve that passes through B, $r_{\max}$, is the largest value of r intercepted by $\mathscr{D}$. B represents the largest value of r at which a trapped surface begins to form. For a large enough mass, $r_{\max}$ can be quite large, and the density in this neighborhood may be quite small. If the mass undergoing collapse is of the order of the mass of a galaxy, for example, the density of matter in the region where a trapped surface first forms may be less than the density of air. An observer falling freely through this region would notice no gross deviation from flat spacetime; he would experience no excessive tidal forces.

Therefore, in the case of large masses, a neighborhood of B could be a region of normal matter density; whereas a neighborhood of A, being closer to the core, could be a region of high density. The results of our model imply that the weak energy condition must be violated even in regions that may be of very low density. We saw earlier that one consequence of this violation was that any local inertial observer in this region must see an energy flux. How does the matter of the cloud "know" to behave in a bizarre fashion, since this region would not otherwise appear to exhibit any exotic properties to a freely falling observer? Such behavior would be more acceptable if it occurred in regions of high density because matter there would experience enormous curvatures and tidal forces.

One might consider the possibility that the WEC violation observed in the low-density region is a result of a violation which first began in the high-density region and propagated outward. We can ask: Can $\mathscr{D}$ be stretched or distorted in such a way that all observers will see the WEC violation begin first in the high-density (i.e., close to the core) regime? This idea can be made more precise. Recall from Fig. 4(b) that if S is a spacelike slice through $\mathscr{D}$, then $S \cap \mathscr{D}$ has the form of an annulus when restricted to two dimensions. Referring to Fig. 6, the trapped region on S is confined to the annulus, and the outer boundary of this region, which is defined to be the apparent horizon (HE, p. 320) on S, is indicated by the heavy circle. An observer having S as one of his "$t=$const slices" would see the violation begin initially in the low-density region, i.e., in a neighborhood of the apparent horizon.

Suppose now that $\mathscr{D}$ is distorted, as in Fig. 7. An observer with S' as one of his $t=$const slices would see the violation begin initially in the high-density region, i.e., in a neighborhood of the dashed circle marking the inner boundary of the annulus. In both Fig. 4(a) and Fig. 7, A and B are defined as the two inflection points of $r_{\min}$ and $r_{\max}$, respectively.

We see that the question we have raised depends on the metric separation between A and B. If A and B are spacelike separated, then for some observers

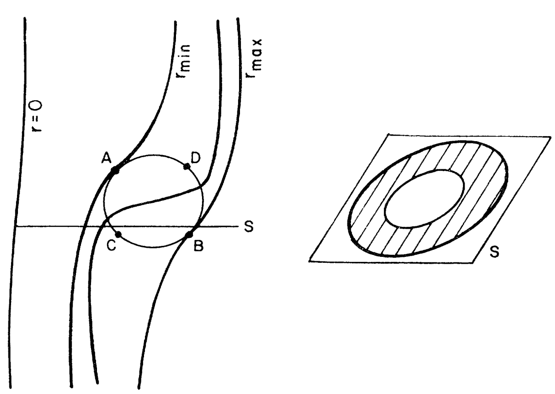

FIG. 6. Onset of WEC violation in the low-density region. The curves labeled $r_{\min}$ and $r_{\max}$ are the smallest and largest values of r intercepted by $\mathscr{D}$, respectively. S is a spacelike slice through $\mathscr{D}$.

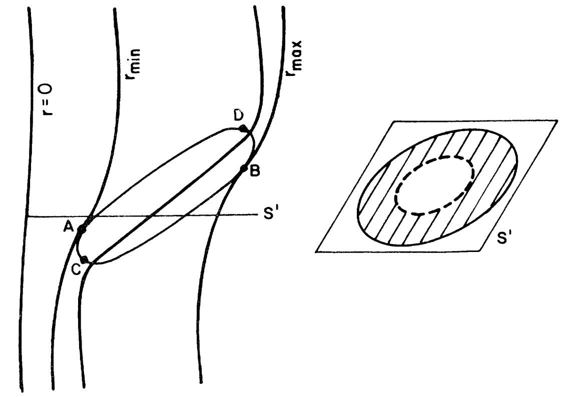

FIG. 7. Distortion of $\mathscr{D}$, with the purpose of shifting the onset of the WEC violation to the high-density region.

the WEC violation begins in the low-density region. This contingence could be avoided only if B is in the causal future of point A. If $A<B$ (see HE, Chap. 6 or Penrose[14] for notation), then all observers will see the violation occur first in the high-density (close to the core) regime. But this arrangement, $B>A$, is impossible.

If the domain $\mathscr{D}$ is compact, with smooth boundary (this is necessary in order that the derivatives of r be well defined), then in the u-v plane, the boundary of $\mathscr{D}$ must have at least two timelike segments, at least two spacelike segments, and at least four points where the tangent to the boundary becomes null. These are the points we have labeled A,B,C,D. Note, by comparing Fig. 6 and Fig. 7, that no matter how $\mathscr{D}$ is drawn, A and C each lie along a portion of a null cone such that the relationship between them is always timelike [see Fig. 8(a)]. Likewise, C and B each lie along a portion of the same future-pointing light cone, such that the relationship between them is always spacelike [Fig. 8(b)]. We notice also that $C \ll A$ (A is in the chronological future of C; i.e., C can be connected to A by a future-directed timelike curve; see HE, Chap. 6 or Penrose[14]).

Proof. Suppose that $C \ll A$ and $A<B$ (B is in the causal future of A; i.e., A can be connected to B by a

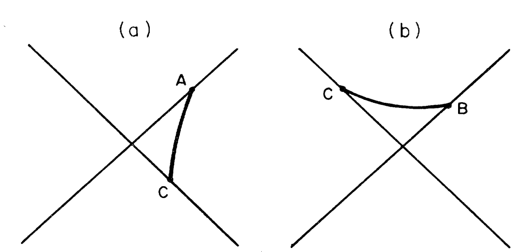

FIG. 8. Causal relations between A, B, and C. A is in the chronological future of C, while C and B are spacelike separated.

future-directed timelike or null curve). If $C \ll A < B$, then $C \ll B$ (see Geroch[15] or Geroch and Horowitz[16] for readable reviews of global techniques in general relativity), which cannot be true since we have seen previously that C and B must always be spacelike separated. Therefore, $A \not< B$, i.e., B cannot lie in the causal future of A.

By a similar argument, one can show that if $C \ll A$ and if D and A are spacelike separated, then $D \not< C$. We cannot, however, conclude from this argument that $B < A$. In fact, it is possible to have either $B < A$ or A and B spacelike separated. In conclusion, then, there will always be some observers who see the WEC violation begin in the low-density region.

One might ask whether the trapped region must be compact or whether the domain $\mathscr{D}$ could be noncompact and still not intersect the curve $r=0$. We have seen that the WEC violation arises from the transition of the $r=$const curves from spacelike to timelike. One might argue that this transition could be avoided if the *ADB* section of the boundary were moved off to infinity. This approach conflicts with the requirement that there should exist observers (stationed outside the collapsing cloud, at large values of r) who never encounter the trapped region. This requirement is implied by the supposition, adopted in this whole investigation, that the collapse of a starlike object is a local event, which does not drastically modify the structure of infinity. Accordingly, we have taken the spacetime to be asymptotically simple and empty (HE, p. 222), so that the structure of infinity is the same as in Minkowski spacetime. Constructions involving a noncompact $\mathscr{D}$ can be shown (e.g., by means of Penrose diagrams) to be incompatible with at least one of the following: (1) the assumption of asymptotic simplicity, (2) the existence of observers at large r who never encounter the trapped region, or (3) the validity of the WEC throughout the low-density region.

In the search for a model of matter compatible with the negative-energy flux needed to avoid the occurrence of a singularity following the onset of trapping, one contemplates, of course, the possibility of postulating exotic states of quantum matter, including particle creation resulting from spacetime curvature (Hawking[17] and subsequent literature). If such processes occur, our (classical, nonquantum) expressions for $T_{\mu\nu}$ might be interpreted as the expectation values of the stress-energy tensor that includes these quantum fields.

Hawking, in his paper on particle creation by black holes, states that particle production by the gravitational field arises from the uncertainty in the local energy density, for modes of the matter field whose wavelength is at least of the order of the radius of curvature of spacetime. According to Hawking, the total uncertainty in the local energy density is of order B multiplied by $(L_p{}^2 B)$, where B is the magnitude of some component of the curvature tensor and L_p denotes the Planck length. In a scaled-up version of our model, which corresponds to the case of a large mass, the local energy density of the created particles decreases by a factor of α^{-4}, where α is the scaling parameter. However, the convergence, c, decreases by a factor of α^{-1}, because c is proportional to the product of the curvature multiplied by the distance over which the curvature focuses.

As the local curvature may be small in a trapped region formed inside a collapsing body of large mass, it would seem that for a macroscopic object particle production is negligible and cannot provide the negative-energy flux required by our model. According to Hawking, the negative-energy densities involved in these processes cannot cause a breakdown of the classical singularity theorems until the radius of curvature of spacetime becomes of the order of the Planck length, $\sim 10^{-33}$ cm.

VI. CONCLUSION

We have constructed a model of a collapsing spherically symmetric cloud of matter, containing a region of trapped surfaces, $\mathscr{D}$, which forms and subsequently disappears. This disappearance of the trapped region, as well as the singularity-free character of the model, are made possible by a violation of the weak energy condition; this violation is required in a neighborhood of the outer boundary of $\mathscr{D}$ and also in a portion of the interior of $\mathscr{D}$. Any local inertial observer in these regions will see a nonzero energy flux, which can be interpreted as a transport of negative energy toward the center of the cloud. In addition, this nondiagonalizability of $T_{\mu\nu}$ by a Lorentz transformation implies that there exist local Lorentz observers for whom the local energy density is negative.

Another result of the model is that, in the case of large masses, the weak energy condition must be violated in regions of low matter density. Distortion of the trapped region $\mathscr{D}$, with the result of transferring the onset of the WEC violation to the high-density regime, was shown to be impossible.

Although we did not offer a conclusive argument, we consider it plausible that particle production processes provide an insufficiently large energy condition violation, in regions of low density, for our model.

Thus our conclusion is that, at least in spherically symmetric cases involving large masses, violation of the weak energy condition is probably not a viable method for preventing the formation of singularities.

Many workers in the field feel that the problem of singularities can be resolved (if at all) only within the framework of a satisfactory quantum theory of gravity. Alternatively, one may have to learn to live with spacetime singularities. The resolution of this problem will have to be left to the future, as a workable quantum theory of gravitation does not yet exist.

ACKNOWLEDGMENTS

One of the authors (T.A.R.) would like to thank Phillip Parker, Matthew Alexander, and Donald Salisbury, for many helpful discussions during the course of this work. P.G.B. was supported by the U. S. National Science Foundation under Grant No. PHY-8209355.

APPENDIX I.

The following expressions can be derived from the line element (2.1).

The Christoffel symbols:

$$\Gamma^v_{vv}=\frac{F_{,v}}{F}\ ,\quad \Gamma^u_{uu}=\frac{F_{,u}}{F}\ ,$$

$$\Gamma^v_{\theta\theta}=\frac{rr_{,u}}{F}\ ,\quad \Gamma^u_{\theta\theta}=\frac{rr_{,v}}{F}\ ,$$

$$\Gamma^v_{\theta\theta}=\frac{rr_{,u}}{F}\sin^2\theta\ ,\quad \Gamma^u_{\phi\phi}=\frac{rr_{,v}}{F}\sin^2\theta\ ,$$

$$\Gamma^\theta_{u\theta}=\Gamma^\theta_{\theta u}=\frac{r_{,u}}{r}\ ,\quad \Gamma^\phi_{u\phi}=\Gamma^\phi_{\phi u}=\frac{r_{,u}}{r}\ ,$$

$$\Gamma^\theta_{v\theta}=\Gamma^\theta_{\theta v}=\frac{r_{,v}}{r}\ ,\quad \Gamma^\phi_{v\phi}=\Gamma^\phi_{\phi v}=\frac{r_{,v}}{r}\ ,$$

$$\Gamma^\theta_{\phi\phi}=-\sin\theta\cos\theta\ ,\quad \Gamma^\phi_{\theta\phi}=\Gamma^\phi_{\phi\theta}=\cot\theta\ .$$

The Ricci tensor:

$$R_{uu}=-\frac{2r_{,uu}}{r}+\frac{2r_{,u}F_{,u}}{rF}\ ,\quad R_{\theta\theta}=\frac{2r_{,u}r_{,v}}{F}+\frac{2rr_{,uv}}{F}+1\ ,$$

$$R_{vv}=-\frac{2r_{,vv}}{r}+\frac{2r_{,v}F_{,v}}{rF}\ ,\quad R_{\phi\phi}=R_{\theta\theta}\sin^2\theta\ ,$$

$$R_{uv}=R_{vu}=-\frac{2r_{,uv}}{r}-\frac{F_{,uv}}{F}+\frac{F_{,u}F_{,v}}{F^2}\ ,$$

The curvature scalar:

$$R=\frac{8r_{,uv}}{rF}+\frac{2F_{,uv}}{F^2}-\frac{2F_{,u}F_{,v}}{F^3}+\frac{4r_{,u}r_{,v}}{r^2F}+\frac{2}{r^2}\ .$$

The stress-energy tensor:

$$T_{uu}=\frac{1}{8\pi}\left[-\frac{2r_{,uu}}{r}+\frac{2r_{,u}F_{,u}}{rF}\right],$$

$$T_{vv}=\frac{1}{8\pi}\left[-\frac{2r_{,vv}}{r}+\frac{2r_{,v}F_{,v}}{rF}\right],$$

$$T_{uv}=\frac{1}{8\pi}\left[\frac{2r_{,uv}}{r}+\frac{2r_{,u}r_{,v}}{r^2}+\frac{F}{r^2}\right],$$

$$T_{\theta\theta}=\frac{1}{8\pi}\left[-\frac{2rr_{,uv}}{F}-\frac{r^2F_{,uv}}{F^2}+\frac{r^2F_{,u}F_{,v}}{F^3}\right],$$

$$T_{\phi\phi}=T_{\theta\theta}\sin^2\theta\ .$$

APPENDIX II

(1) Let $U^\mu=(o,\alpha,0,0)$ be the tangent vector to an outgoing null geodesic. Then

$$U^\mu{}_{;\nu}U^\nu=0\Longrightarrow\alpha=\frac{1}{F}\ .$$

(2) The convergence c is defined to be

$$c\equiv-U^\mu{}_{;\mu}=-U^v{}_{,v}-(\Gamma^v_{vv}+\Gamma^\theta_{v\theta}+\Gamma^\phi_{v\phi})U^v$$
$$=-\frac{2r_{,v}}{rF}\ .$$

Therefore,

$$c=-\frac{2r_{,v}}{rF}\ .$$

(3) If λ is the affine parameter along the null geodesic, then

$$\frac{dc}{d\lambda}=\frac{1}{F}\frac{dc}{dv}=-\frac{2r_{,vv}}{rF^2}+\frac{2(r_{,v})^2}{r^2F^2}+\frac{2r_{,v}F_{,v}}{rF^3}$$
$$=\frac{1}{F^2}\left[-\frac{2r_{,vv}}{r}+\frac{2r_{,v}F_{,v}}{rF}\right]+\frac{2(r_{,v})^2}{r^2F^2}$$
$$=8\pi\alpha^2T_{vv}+\frac{c^2}{2}\ .$$

Therefore,

$$\frac{dc}{d\lambda}=8\pi\alpha^2T_{vv}+\frac{c^2}{2}\ ,$$

which is the Raychaudhuri equation for a radial outgoing null geodesic. Thus for the case of radial null geodesics, the Raychaudhuri equation is simply one of the field equations.

*Current address.

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

[2]R. Penrose, Phys. Rev. Lett. 14, 57 (1965).

[3]F. J. Tipler, Ann. Phys. (N.Y.) 108, 1 (1977).

[4]J. D. Bekenstein, Phys. Rev. D 11, 2072 (1975).

[5]F. J. Tipler, Phys. Rev. D 17, 2521 (1978).

[6]F. J. Tipler, C. J. S. Clarke, and G. F. R. Ellis, in *General Relativity: An Einstein Centenary Survey,* edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979), pp. 97–205.

[7]S. W. Hawking, in *Black Holes,* 1972 Les Houches Lectures, edited by C. DeWitt and B. S. DeWitt (Gordon and Breach, New York, 1973), pp. 1–56.

[8]M. D. Kruskal, Phys. Rev. 119, 1743 (1960).

[9]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), Chap. 31.

[10]M. J. Rees, R. Ruffini, and J. A. Wheeler, *Black Holes, Gravitational Waves, and Cosmology: An Introduction to Current Research* (Gordon and Breach, New York, 1974), pp. 49 and 50.

[11]H. Kodama, Prog. Theor. Phys. 62, 1434 (1979).

[12]H. Kodama, Prog. Theor. Phys. 63, 1217 (1980).

[13]F. J. Tipler, private communication.

[14]R. Penrose, *Techniques of Differential Topology in Relativity,* Society of Industrial and Applied Mathematics Regional Conference Series in Applied Mathematics (SIAM, Philadelphia, 1972), No. 7.

[15]R. Geroch, in *General Relativity and Cosmology,* proceedings of the International School of Physics, Enrico Fermi Course XLII, Varenna, 1979, edited by R. K. Sachs (Academic, New York, 1971) pp. 71–99.

[16]R. P. Geroch and G. T. Horowitz, in *General Relativity: An Einstein Centenary Survey,* Ref. 6, pp. 212–289.

[17]S. W. Hawking, Commun. Math. Phys. 43, 199 (1975).