PHYSICAL REVIEW D VOLUME 15, NUMBER 4 15 FEBRUARY 1977

On the nature of singularities in general relativity*†

Frank J. Tipler
Department of Mathematics, University of California, Berkeley, California 94720
(Received 18 August 1976)

It is shown that there is an upper bound to the rate of growth of the Ricci curvature near a singularity.

Singularities in the form of incomplete causal geodesics appear to be a necessary property of any spacetime which could be a realistic model of the universe; the standard singularity theorems imply such a conclusion provided we assume causality is not violated, and I have shown elsewhere[1,2] that causality violation could not in general prevent the singularities. In fact, causality violation arising from regular initial data would in general *create* singularities. Once the inevitability of singularities is proven, we should study the nature of these predicted singularities: What is the structure of a singularity in the generic case? There are several approaches one could take to answer this question. First, one could analyze the singularities in exact (and approximate) solutions of greater and greater generality. This method of attack is very popular; examples can be found in the work of Ryan and Shepley,[3] Gowdy,[4] King,[5] Ellis and King,[6] Khalatnikov and Lifshitz,[7] Collins and Hawking,[8] Eardley, Liang, and Sachs.[9] Second, one could use global techniques to see if the causal and differential structure of spacetime imposed limits on the nature of the singularity. In my opinion, this second approach, though not used as extensively as the first, is superior to the first because of the generality of the conclusions. With the first method, one is never certain that the singularity structure derived in a particular model has any features in common with the singularity structure of a generic spacetime, since the singularity structure in the model could be due almost entirely to the simplifying assumptions necessary to construct the model. Thus, no confidence can be placed in the conclusions of the first method until they are confirmed by the second. As an example of the generality possible when the second approach is used, Clarke has been able to show[10] that, in a generic spacetime, at least one component of the Riemann tensor must diverge in a frame which is parallel-propagated along any incomplete timelike curve contained in $D(S)$, where S is a partial Cauchy surface. This is an important result as there are models in which this does not occur.[11,12]

I shall follow Clarke and use the second approach; in this paper I shall derive some very general restrictions on the rate of growth of $R_{ab}K^aK^b$ along certain classes of incomplete causal geodesics. (K^a is the tangent vector to the geodesic.) My notations and conventions will be the same as those of Hawking and Ellis,[12] who are hereafter referred to as HE.

The theorems of this paper will be based on the properties of conjugate points along a causal geodesic. Recall that a point p along a causal geodesic $\gamma(t)$ is said to be conjugate to a point q along $\gamma(t)$ if there is a Jacobi field along $\gamma(t)$, not identically zero, which vanishes at q and p. It is well known (HE, pp. 96–102) that such a Jacobi field will exist if and only if there is a function y, defined by the equations

$$\theta = \frac{1}{y}\frac{dy}{dt}, \tag{1}$$

$$\frac{d\theta}{dt} = -R_{ab}K^aK^b - 2\sigma^2 - \frac{1}{n}\theta^2, \tag{2}$$

such that $y \not\equiv 0$, but $y=0$ at q and p. [t is an affine parameter along $\gamma(t)$, $\sigma^2 \geq 0$, and n equals 2 or 3 depending on whether $\gamma(t)$ is a null or timelike geodesic, respectively.] If we define a new function x by $x^n = y$, then Eq. (2) becomes

$$\frac{d^2x}{dt^2} + F(t)x = 0, \tag{3}$$

where $F(t) \equiv (1/n)(R_{ab}K^aK^b + 2\sigma^2)$, and q and p will be conjugate if and only if $x=0$ at q and p. Thus the presence (or absence) of a pair of conjugate points in a given affine-parameter interval is equivalent to the presence (or absence) of a pair of zeros of the function x in the same interval.

The latter is a well-known problem in the theory of ordinary differential equations, and the theorems which I quote without proof can be found in the literature.[13]

Sturm comparison theorem. Consider the two equations

$$u''(t) + F(t)u(t) = 0, \tag{4}$$

$$v''(t) + G(t)v(t) = 0, \tag{5}$$

where $F(t)$ and $G(t)$ are non-negative, continuous, and $G(t) \geq F(t)$ in the closed interval $[a, b]$ and $G(t) > F(t)$ for at least one point in $[a, b]$. Suppose that (4) has a solution $u(t)$ having two consecutive zeros at a and b. Then *all* solutions $v(t)$ of (5) have a zero in the open interval (a, b).

Lemma 1. A sufficient condition that *every* solution to (3) have at least n zeros in (a, b) is that

$$\inf_{a<t<b} F(t) > \left(\frac{n\pi}{b-a}\right)^2 . \tag{6}$$

Definition. Equation (3) will be called *oscillatory* on (a, b) if every solution ($\neq 0$) has infinitely many zeros on (a, b). When every solution ($\neq 0$) has at most a finite number of zeros on (a, b), (3) is said to be *nonoscillatory* on (a, b).

Proposition 1. In Eq. (3) let $F(t)$ be continuous and positive in the interval (a, b). If either

$$\lim_{t\to b} \inf[(t-b)^2 F(t)] > \tfrac{1}{4}$$

or (7)

$$\lim_{t\to a} \inf[(a-t)^2 F(t)] > \tfrac{1}{4} ,$$

then (3) is oscillatory on (a, b). (Needless to say, both a and b are finite.)

Proof. Consider the equation

$$u'' + \left[\frac{m}{(b-t)^2}\right] u = 0 , \tag{8}$$

where m is a constant. If $\lim_{t\to b} \inf[(t-b)^2 F(t)] > m$, then there exists a number $c \in (a, b)$ for which $[(t-b)^2 F(t)] > m$ for all $t \in (c, b)$; i.e., $F(t) > m/(t-b)^2$ for all $t \in (c, b)$. By the Sturm comparison theorem, between the zeros of any solution to (8), there is a zero in *every* solution to (3) in the interval (c, b). We now show that all solutions to (8) have infinitely many zeros in the interval (c, b) if $m > \frac{1}{4}$. Consider the substitution

$$u = (b-t)^{1/2} y, \quad b - t = e^s$$

so that

$$s = \ln(b-t), \quad u = e^{s/2} y .$$

This gives

$$u'' = e^{-3s/2}\left(\frac{d^2y}{ds^2} - \tfrac{1}{4}y\right) .$$

Thus Eq. (8) becomes

$$u'' + \left[\frac{m}{(b-t)^2}\right] u = e^{-3s/2}\left(\frac{d^2y}{ds^2} - \tfrac{1}{4}y\right) + m e^{-2s}(e^{s/2}y)$$
$$= e^{-3s/2}\left[\frac{d^2y}{ds^2} + (m - \tfrac{1}{4})y\right] = 0$$

or

$$\frac{d^2y}{ds^2} + (m - \tfrac{1}{4})y = 0 .$$

Now since $u = (b-t)^{1/2} y$, it is clear that in the interval (c, b) u will have a zero if and only if y does, and in this interval s varies from $s = \ln(b-c)$ = finite number to $s = -\infty$. If $m > \frac{1}{4}$, y will have infinitely many zeros in this interval, and hence so will u. The same procedure can be used with the second inequality of (7). Therefore, if either of the inequalities of (7) hold, all solutions of (3) have infinitely many zeros on (a, b); (3) is oscillatory on that interval.

Using the above results from the theory of ordinary differential equations and several propositions found in HE concerning relationships between the causal structure and causal geodesics, we can prove the following:

Theorem 1. Let S be a partial Cauchy surface. To each point $p \in \text{int}D^+(S)$ there is a timelike geodesic $\lambda(t)$ from S such that *no* affine-parameter interval (t_1, t_2) of $\lambda(t)$ satisfies

$$\inf_{t_1<t<t_2} R_{ab}K^aK^b > \frac{12\pi^2}{(t_2-t_1)^2} , \tag{9}$$

where K^a is the tangent vector to $\lambda(t)$.

Proof. Suppose that every timelike geodesic from S to p satisfies (9). Then we have

$$\inf_{t_1<t<t_2} F(t) \equiv \inf_{t_1<t<t_2} \tfrac{1}{3}(R_{ab}K^aK^b + 2\sigma^2)$$
$$\geq \tfrac{1}{3}\left[\frac{12\pi^2}{(t_2-t_1)^2}\right] = \frac{(2\pi)^2}{(t_2-t_1)^2}$$

for some interval (t_1, t_2) along every timelike geodesic from S to p. By lemma 1, every timelike geodesic from S to p would have a pair of conjugate points between S and p. However, by the corollary on p. 217 of HE, this is impossible.

Theorem 2. Let S be a partial Cauchy surface, and let $\lambda(t)$ be a timelike geodesic from S to a point p on the b-boundary of $\tilde{D}^+(S)$ such that $\lambda(t)$ is at each point a curve of maximal length from S. Then, along $\lambda(t)$, $R_{ab}K^aK^b$ satisfies

$$\lim_{t\to t_1} \inf(t-t_1)^2 R_{ab}K^aK^b \leq \tfrac{3}{4} , \tag{10}$$

where t_1 is the limit of the affine parameter along $\lambda(t)$ as $\lambda \to p$, and K^a is the unit tangent vector to λ. (We assume that $R_{ab}K^aK^b \geq 0$ for all causal vectors K^a.)

Proof. A timelike geodesic $\lambda(t)$ is at each point a curve of maximal length from S if, for each point $q \in \lambda(t)$, the length of λ from S to q equals $d(S, q)$ (HE, p. 288). By proposition 4.5.9 of HE, a timelike geodesic from S to any point q on the geodesic is maximal if and only if there is no point in $(S \cap \lambda, q)$ conjugate to S along λ. Thus λ can contain no conjugate points, since a pair of conjugate points would imply a point conjugate to S.

If along $\lambda(t)$ (10) were not satisfied, that is, if

$$\lim_{t\to t_1} \inf(t-t_1)^2 R_{ab}K^aK^b > \tfrac{3}{4},$$

$\lambda(t)$ would have an infinite number of conjugate points by proposition 1.

Theorem 2 assumes that there exists a geodesic normal to S to a point p on the b-boundary (HE, p. 217). There will in general not exist, for *all* points p on the b-boundary, a geodesic normal to S from S to p. For example, let S be the spacelike hypersurface $(x^1)^2+(x^2)^2+(x^3)^2-(x^4)^2=-1$, $x^4<0$, in Minkowski space. Let (M',g') be the region of Minkowski space for which S is a Cauchy surface, i.e., the region $I^-(q=(0,0,0,0))$. Then the b-boundary to (M',g') is $\dot{J}^-(q)$ (the past light cone from the origin of coordinates q) and q is the *only* point on this b-boundary for which there is a geodesic normal to S to the point, since all geodesics normal to S intersect at q (HE, p. 120). However, in many cases of interest, there will exist at least *one* maximal geodesic from S to at least *one* point on the b-boundary. For example, Hawking and Ellis have shown (HE, pp. 272, 288) that if the following conditions hold,

(a) $R_{ab}K^aK^b \geq 0$ for all causal vectors K^a,
(b) S is compact, and
(c) the unit normals to S are everywhere converging,

then there exists a timelike geodesic λ which is incomplete, which remains in $\bar{D}^+(S)$, and which is at each point a curve of maximum length from S. Along this geodesic, (10) holds.

Furthermore, in some situations there is a Gaussian normal coordinate system from S to the b-boundary. Along *each* timelike geodesic generator of this coordinate system, (10) holds. For instance, in the dust-filled Friedmann universe we have $\rho = 1/(6\pi t^2)$ near the singularity along the timelike geodesics which are normal to the surfaces of homogeneity.[14] This gives

$$\lim_{t\to 0} \inf t^2 R_{ab}K^aK^b = \lim_{t\to 0} \inf t^2 4\pi\left(\frac{1}{6\pi t^2}\right) = \tfrac{2}{3} < \tfrac{3}{4},$$

which satisfies (10).

In an orthonormal frame which is parallel-propagated along λ (with $\vec{K}=\vec{E}_4$) we have, for perfect fluids, $R_{ab}K^aK^b = 4\pi(\rho+3p)$, where ρ is the energy density and p is the pressure. Since we would expect the nongravitational forces to be repulsive (i.e., $p>0$) near the singularity (there are exceptions to this expectation-condensation phenomena, for example, but these apply only over short periods of affine-parameter time), theorems 1 and 2 physically say that the energy density cannot diverge faster than $\sim 1/t^2$ along *some* timelike geodesic $\lambda(t)$ which hits the singularity at affine-parameter time $t=0$. (This confirms a conjecture made by Misner and Taub.[11])

Theorems 1 and 2 can be used to give an upper limit to the range of validity of a perturbation calculation. For example, Zel'dovich and Starobinski[15] have calculated curvature-induced particle-production rates by considering the particle-production process as a perturbation in an empty Kasner universe. Along geodesics normal to the surfaces of homogeneity (these geodesics have no conjugate points in the empty Kasner universe), the energy density of the created particles is $\rho \sim \hbar c^{-3} t^{-4}$ near the singularity at affine-parameter value $t=0$. The principle pressures have comparable values, but neither ρ nor the other components of T_{ab} are calculated exactly; approximately (?), we have

$$R_{ab}K^aK^b \sim \hbar c^{-3} t^{-4} = (2.6\times 10^{-66}\ \text{cm}^2)t^{-4}$$

in geometrical units. It is immediately obvious from theorem 2 that the t^{-4} divergence cannot continue indefinitely, for the back reaction of the created matter on the metric would give rise to conjugate points along the geodesic congruence normal to the surfaces of homogeneity. To get an estimate of just where this occurs, we use theorem 1. Equation (9) will hold on an interval $\alpha t_1 < t < t_1$ $(\alpha<1)$ if

$$\hbar c^{-3} t_1^{-4} = 12\pi^2 t_1^{-2}(1-\alpha)^{-2}.$$

For sake of concreteness, we set $\alpha = 0.1$, obtaining a pair of conjugate points in the range $(0.1t_1, t_1)$. This gives $t_1 \sim 10^{-34}$ cm, or $\rho \sim 10^{100}$ g/cm^3. Thus the perturbation breaks down when the created-particle energy density become comparable to the Planck density (10^{94} g/cm^3). (This fact is known from other considerations; see Ref. 14, p. 804.)

There are theorems analogous to theorems 1 and 2 which restrict the rate of growth of $R_{ab}K^aK^b$ along the generators of $H(S)$.

Theorem 3. Let S be a partial Cauchy surface, and let $\lambda(s)$ be a null geodesic generator of $H^+(S)$. Then there is *no* affine-parameter interval (s_1, s_2) of $\lambda(s)\cap H^+(S)$ satsifying

$$\inf_{s_1<s<s_2} R_{ab}K^aK^b > \frac{8\pi^2}{(s_2-s_1)^2}, \qquad (11)$$

where K^a is the tangent vector to $\lambda(s)$.

Proof. If (11) were obeyed in some interval (s_1, s_2) along $\lambda(s)\cap H^+(S)$, then by lemma 1 $\lambda(s)$ would have a pair of conjugate points in (s_1, s_2). But by proposition 4.5.12 of HE, this would contradict the achronality of $H^+(S)$.

Theorem 4. Let S be a partial Cauchy surface, and let $\lambda(s)$ be a generator of $H^+(S)$, with $\lambda(s)$ intersecting the b-boundary in the past direction at affine-parameter value s_1. If $R_{ab}K^aK^b \geq 0$ on $\lambda(s)$, then along $\lambda(s)$ we must have

$$\lim_{s\to s_1} \inf(s_1 - s)^2 R_{ab}K^aK^b \leq \tfrac{1}{2}, \qquad (12)$$

where K^a is the tangent vector to $\lambda(s)$.

Proof. If along $\lambda(s)$ (12) were not satisfied, that is, if

$$\lim_{s\to s_1} \inf(s_1 - s)^2 R_{ab}K^aK^b > \tfrac{1}{2},$$

$\lambda(s) \cap H^+(s)$ would have an infinite number of conjugate points, by proposition 1. But this would contradict the achronality of $H^+(S)$, by proposition 4.5.12 of HE.

The most interesting thing about theorems 3 and 4 is that they restrict the rate of growth of $R_{ab}K^aK^b$ along *all* generators of $H^+(S)$.

I would like to thank Professor Dieter R. Brill for many helpful discussions and comments.

---

*Part of a Ph.D. thesis submitted to the Department of Physics and Astronomy, University of Maryland.

†Work supported in part by the National Science Foundation.

[1]F. J. Tipler, Ph.D. thesis, Department of Physics and Astronomy, University of Maryland, 1976 (unpublished).

[2]F. J. Tipler, Phys. Rev. Lett. 37, 879 (1976).

[3]M. P. Ryan and L. C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton Univ. Press, Princeton, N. J., 1975).

[4]R. H. Gowdy, Ann. Phys. (N.Y.) 83, 203 (1974).

[5]A. R. King, Commun. Math. Phys. 38, 157 (1974).

[6]G. F. R. Ellis and A. R. King, Commun. Math. Phys. 38, 119 (1974).

[7]I. M. Khalatnikov and E. M. Lifshitz, Phys. Rev. Lett. 24, 76 (1969).

[8]C. B. Collins and S. W. Hawking, Astrophys. J. 180, 317 (1973).

[9]D. Eardley, E. Liang, and R. Sachs, J. Math. Phys. 13, 99 (1972).

[10]C. J. S. Clarke, Commun. Math. Phys. 32, 205 (1973); 41, 65 (1975).

[11]C. W. Misner and A. H. Taub, Zh. Eksp. Teor. Fiz. 55, 233 (1968) [Sov. Phys.–JETP 28, 122 (1969)].

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

[13]P. Hartman, *Ordinary Differential Equations* (Wiley, New York, 1964); E. Hille, *Lectures on Ordinary Differential Equations* (Addison-Wesley, Reading, Mass., 1969); C. A. Swanson, *Comparison and Oscillation Theory of Linear Differential Equations* (Academic, New York, 1968).

[14]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), p. 735.

[15]Ya. B. Zel'dovich and A. A. Starobinski, Zh. Eksp. Teor. Fiz. 61, 2161 (1971) [Sov. Phys.—JETP 34, 1159 (1972)].