# On singularity theorems and curvature growth

L. B. Szabados
*Central Research Institute for Physics, P.O. Box 49, H-1525 Budapest 114, Hungary*

(Received 29 April 1986; accepted for publication 27 August 1986)

It is shown that the proofs of a series of classical singularity theorems of general relativity can be modified such that these theorems also state the maximality of the incomplete nonspacelike geodesics. Since along maximal incomplete nonspacelike geodesics with affine parameter $u$ certain parts of the tidal curvature cannot blow up faster than $(\bar{u}-u)^{-2}$, where $\bar{u}$ is the parameter value until which the geodesics cannot be extended, the classical singularity theorems do restrict the behavior of the curvature.

## I. INTRODUCTION

In the general theory of relativity the singularity theorems of Hawking, Penrose, and others state the existence of incomplete nonspacelike geodesics under very general conditions; i.e., space-time is singular.[1,2] However, these theorems do not show which geodesics are incomplete and in many cases they do not tell us explicitly if these geodesics are future or past incomplete.

In other parts of classical physics the notion of singularities is connected with the misbehavior of a characteristic quantity; e.g., field strength. Thus, based on this intuitive notion of singularities, certain misbehavior of the curvature tensor is expected along the incomplete nonspacelike geodesics. Unfortunately, such a misbehavior does not follow from incompleteness, furthermore the singularity theorems say nothing about the curvature's behavior.[3]

Surprisingly, while we are expecting a lower bound to the rate of growth of some parts of the curvature along incomplete nonspacelike geodesics approaching a "true, physical" singularity, in certain cases the existence of an upper bound can be proved, which is due to the causal structure.[4–7] The basic idea[1,2] is the fact that a nonspacelike curve is maximal iff it is geodesic without any pair of conjugate points. But the occurrence of conjugate points is a very general phenomenon: if a nonspacelike geodesic gathers enough curvature in some sense, then conjugate points must occur. Null geodesics are maximal if they lie in an achronal set, furthermore there exists a maximal nonspacelike geodesic between causally related points in a globally hyperbolic set. Thus maximality, in certain situations, follows from causality. Consequently, if $\gamma$ is a maximal nonspacelike geodesic, then $\gamma$ cannot gather arbitrarily large curvature; and in particular if $\gamma$ is a maximal incomplete nonspacelike geodesic (and so it hits a singularity), we obtain an upper bound to the rate of the blowing up of certain parts of the curvature. This argument was used to obtain an upper bound to the rate of the divergence of the Ricci and Weyl part of the tidal force along incomplete null geodesics lying in an achronal set by Tipler[5] and the author,[6] respectively: they cannot blow up faster than $(\bar{u}-u)^{-2}$, where $u$ is the affine parameter and $\bar{u}$ is the parameter value until which the geodesic cannot be extended.

Unfortunately, incomplete nonspacelike geodesics are not necessarily maximal, so we do not have any upper bound in general. For example, if $\gamma$ is a future incomplete null geodesic [so the TIP $P := I^-\gamma$ represents a singular point of the causal boundary $\partial_c$ (see Ref. 3)], then $\gamma$ is not necessarily in an achronal set, therefore $\gamma$ is not necessarily maximal. On the other hand, though the boundary $\partial P$ is achronal and is generated by future endless null geodesics, these geodesics are not necessarily incomplete, even if $\gamma$ is. Therefore one may ask the question: Is there any "physically realistic" situation in which the incomplete nonspacelike geodesics are maximal; i.e., in which the existence of an upper bound can be proved?

In this paper we show that the proofs of a number of classical singularity theorems can be modified such that they state not the existence of incomplete nonspacelike geodesics only, but that these geodesics must be maximal, therefore these theorems do restrict the curvature's behavior. We conclude that in "physically realistic" situations, defined by the conditions of the classical singularity theorems, certain parts of the curvature cannot blow up faster than $(\bar{u}-u)^{-2}$ along the geodesics approaching the singularity. Furthermore, in many cases they show where the incomplete geodesics lie.

This paper consists of four parts. In the first one we review what restrictions on the growth of the curvature can be obtained from the maximality of incomplete nonspacelike geodesics. The second part contains the auxiliary statements we need for the new proofs. Four of the most important classical singularity theorems with modified proofs are contained in the third part. Finally, in the fourth part some remarks are given.

This paper is based on the matter given in Refs. 1 and 2, so the continuous references to well-known statements are omitted. Our conventions and notations are the same as those of the book of Hawking and Ellis[1] except that the chronological, causal, etc. future of the set $K$ is denoted by $I^+K$, $J^+K$, etc., respectively.

## II. MAXIMAL NONSPACELIKE CURVES

A nonspacelike curve $\gamma$ from $p$ to $q$ is said to be maximal[1] if $p$ and $q$ cannot be joined by any nonspacelike curve, obtained from $\gamma$ by small deformation, longer than $\gamma$. Since, by definition, incomplete nonspacelike curves are inextendable, we are interested in nonspacelike curves without end points. A future inextendable nonspacelike curve $\gamma$ starting at $p$ is said to be maximal if for every $q \in \gamma$, the segment of $\gamma$ between $p$ and $q$ is maximal. Based on the statements of Chap. 4 of Ref. 1, one can state that a future inextendable nonspacelike curve $\gamma$ from $p$ is maximal iff it is geodesic and

 0022-2488/87/010142-04$02.50  

contains no point conjugate to $p$ along $\gamma$.

Ruling out the possibility of the occurrence of conjugate points along future directed incomplete nonspacelike geodesics, Tipler has a restriction on the growth of the $R_{ab}K^aK^b$ part of the curvature, where $K^a$ is the tangent to the geodesics.[4–7]

*Proposition 2.1*: Let $\gamma$: $[0,\bar{u})\to M$ be a maximal incomplete nonspacelike geodesic with affine parameter $u$ and tangent $K^a$. If the convergence condition $R_{ab}K^aK^b \geqslant 0$ holds, then, for $\forall\alpha>0$, the inequality

$$\liminf_{\Delta u\to 0}\left(-\Delta u\int_{\bar{u}-\Delta u}^{\bar{u}-(1+\alpha)\Delta u} R_{ab}K^aK^b\,du\right)\leqslant 2n$$

must be satisfied, where $n=2$ for null and $n=3$ for timelike geodesics.

This proposition implies that the energy-density-like expression $R_{ab}K^aK^b$ cannot blow up faster than $(\bar{u}-u)^{-2}$ as we approach the singularity at $\bar{u}$, provided the metric is $C^2$. For timelike geodesics, the equations describing the Jacobi fields (and so the conjugate points) are very complicated, therefore obtaining restrictions on further parts of the curvature seems to be almost hopeless.

Along maximal incomplete null geodesics, however, a restriction on the eigenvalue $\epsilon$ of $C_{manb}K^aK^b$ can be obtained in certain situations. Before considering the statement giving us this restriction, we have to examine the possible behaviors of $C_{manb}K^aK^b$ as $u\to\bar{u}$. Since $C_{manb}K^aK^b$ is a symmetric and traceless $2\times 2$ matrix, it has the form

$$\epsilon\begin{pmatrix}\cos 2\chi & \sin 2\chi\\ \sin 2\chi & -\cos 2\chi\end{pmatrix}.$$

Thus it is completely characterized by the functions $\epsilon(u)$ and $\chi(u)$. Consequently, the behavior of $C_{manb}K^aK^b$ is determined by those of $\epsilon$ and $\chi$ (see Ref. 7). From physical points of view the most important case is that in which the components of $C_{manb}K^aK^b$ diverge; i.e., there is a definite limiting eigenframe of $C_{manb}K^aK^b$ [$\lim_{u\to\bar{u}}\chi(u)$ exists] and $\epsilon$ tends to infinity. Introducing the notation

$$\begin{aligned}W_{u_1,u-u_1} &:= \frac{1}{2}\int_{u_1}^{u}\left(\int_{u_1}^{u'} C_{manb}K^aK^b\,du''\right)\\ &\quad\times\left(\int_{u_1}^{u'} C_{menf}K^eK^f\,du''\right)du'\\ &=\int_{u_1}^{u}\left[\left(\int_{u_1}^{u'}\epsilon\cos 2\chi\,du''\right)^2\right.\\ &\quad\left.+\left(\int_{u_1}^{u'}\epsilon\sin 2\chi\,du''\right)^2\right]du',\end{aligned}$$

the next proposition gives us a restriction on the diverging $C_{manb}K^aK^b$.[6]

*Proposition 2.2*: Let $\gamma$: $[0,\bar{u})\to M$ be an incomplete null geodesic lying in an achronal set, $u$ be its affine parameter, and $K^a$ be its tangent. If the convergence condition $R_{ab}K^aK^b\geqslant 0$ holds, $\lim_{u\to\bar{u}}\chi(u)$ exists and $\exists\theta>0$ such that $\epsilon(u)$ does not change sign on $(\bar{u}-\theta,\bar{u})$, then for $\forall\alpha>0$, the inequality

$$\liminf_{\Delta u\to 0}\left(-\Delta u W_{(\bar{u}-\Delta u)-\Delta u,-\alpha\Delta u}\right)\leqslant(2+\alpha)(1+\alpha)^3$$

must be satisfied.

This statement implies that $\epsilon$ cannot blow up faster than $(\bar{u}-u)^{-2}$ along maximal incomplete null geodesics, provided the metric is $C^2$ and $\lim\chi(u)$ exists. Otherwise, assuming for example that $\epsilon(u)=b(\bar{u}-u)^{-2-\nu}$ for some $b\neq 0$ and $\nu>0$, we would obtain that

$$-\Delta u W_{(\bar{u}-2\Delta u),-\alpha\Delta u}=b^2G(\alpha,\nu)(\Delta u)^{-2},$$

$\forall\alpha>0$, where $G(\alpha,\nu)$ is a nowhere zero expression of $\alpha$ and $\nu$; which would tend to infinity as $\Delta u\to 0$.

## III. AUXILIARY STATEMENTS

In this section we state four lemmas that will be used in the modification of the proofs of the classical singularity theorems.

*Lemma 3.1*: If (1) the null convergence condition holds on $M$, (2) the null generic condition holds on $M$ (i.e., on each inextendable null geodesic with tangent $K^a$ there is a point where $K_{[a}R_{b]ef[c}K_{d]}K^eK^f\neq 0$), and (3) the chronology condition holds on $M$, then the strong causality condition holds on $M$ or there is a point $p$ where the strong causality condition is violated and there is an incomplete null geodesic through $p$ lying in $E^-\{p\}\cup E^+\{p\}$.

This lemma is a modified form of Proposition 6.4.6 of Ref. 1 and its proof is almost the same.

The following lemmas state compactness of certain subsets of space-time, provided the maximal nonspacelike geodesics generating them, in some sense, leave these sets. The proofs are based on the standard matter given in Chap. 6 of Ref. 1. The first one has been published yet,[8] so its proof will be omitted here.

*Lemma 3.2*: Let $K$ be a nonempty set. If each future directed null geodesic $\gamma$ generating $E^+K-K$ leaves $E^+K$ in the future direction (i.e., each $\gamma$ has a point $r$ such that the points of $\gamma$ following $r$ do not belong to $E^+K$ and the points of $\gamma$ preceding $r$ belong to $E^+K$), then $E^+K$ is compact.

*Lemma 3.3*: If there is no past directed past endless nonspacelike geodesic $\gamma$ from $p$ such that its segment $\gamma-\{p\}$ is maximal, or each such a nonspacelike geodesic leaves $D^-E^-\{p\}$ in the past direction, then $\overline{D^-E^-\{p\}}$ is compact.

*Sketch of proof*: Using the technique developed in Refs. 1 and 2, one can show that for each point $q$ of $\overline{D^-E^-\{p\}}$ there exists a past directed nonspacelike geodesic $\gamma$ from $p$ through $q$, such that the segment $(q,p)$ of $\gamma$ is maximal.

Let $\{q_n\}$ be an infinite sequence of points of $\overline{D^-E^-\{p\}}$. Because of Lemma 3.2, $E^-\{p\}$ is compact. Thus without loss of generality one may assume that $p$ is not limit point of $\{q_n\}$ and no point of this sequence belongs to $E^-\{p\}$. Let $\gamma_n$ be the past directed nonspacelike geodesic from $p$ through $q_n$ such that its segment $(q_n,p)$ is maximal. $\{\gamma_n\}$ has a limit curve $\gamma$ from $p$ which is geodesic. Furthermore, the segment of $\gamma$ consisting of those points that are limit points of the maximal segments $(q_n,p)$ of the $\gamma_n$'s is maximal. Consequently, there is a point $q\in\gamma-\{p\}$ such that this maximal segment is $[q,p)$ or $(q,p)$. Of course, $q\in\overline{D^-E^-\{p\}}$, and $q$ is a limit point of $\{q_n\}$, i.e., $\overline{D^-E^-\{p\}}$ is compact.

In a similar way one can prove our fourth lemma.

*Lemma 3.4*: Let $S$ be a compact $C^2$ partial Cauchy surface. If there is no future inextendable maximal timelike geodesic orthogonal to $S$, or each such a timelike geodesic leaves $D^+S$ in the future direction, then $\overline{D^+S}$ is compact.

## IV. CLASSICAL SINGULARITY THEOREMS MODIFIED

In this section the four classical singularity theorems contained in Ref. 1 will be reexamined and it will be clear that the original conditions of these theorems guarantee the maximality of incomplete nonspacelike geodesics.

The first one is Penrose's theorem[9] and, since its modified proof has been published elsewhere,[8] its new proof, which is based on Lemma 3.2, will not be repeated here.

**Theorem 4.1**: If (1) $R_{ab}K^aK^b \geqslant 0$ for every null vector $K^a$; (2) $(M,g)$ admits a noncompact Cauchy surface; and (3) there exists at least one of the following: (a) a closed trapped surface $T$, (b) a point $t$ such that along each future directed null geodesic from $t$ the expansion $\hat{\theta}$ becomes negative ($t$ may be called a future trapped point); then there exists a future incomplete null geodesic lying in $\partial J^+T$ or in $\partial J^+\{t\}$, respectively.

Possibility (b) in condition (3) is due to Tipler[10] and we note that this concept of trapped point differs from the trapped point of Królak.[11,12]

The second theorem is that of Hawking and Penrose.[13]

**Theorem 4.2**: If (1) $R_{ab}K^aK^b \geqslant 0$ for every nonspacelike vector $K^a$; (2) the chronology condition holds on $M$; (3) the generic condition holds on $M$; and (4) there exists at least one of the following: (a) a compact achronal set $S$ without edge, (b) a closed trapped surface $T$, (c) a future trapped point $t$; then at least one of the following statements holds: ($\alpha$) there exists a compact set $C \neq 0$ and an incomplete null geodesic lying in $E^+C \cup E^-C$, and/or ($\beta$) there exists an open globally hyperbolic set $D$ and an incomplete maximal nonspacelike geodesic in $D$.

*Proof:* Let $K$ be $S$ or $T$ or $\{t\}$ in case (a), or (b), or (c), respectively. In cases (b) and (c), $E^+K - K$ is generated by future directed null geodesics with past end points on $K$. If each null geodesic generator of $E^+K - K$ leaves $E^+K$, then $E^+K$ is compact (Lemma 3.2). Thus, if $E^+K$ is not compact then there must be a null geodesic generator $\gamma$ which does not leave $E^+K$ in the future direction. However, $\gamma$ cannot be future complete, as otherwise a point conjugate to $K$ would occur on $\gamma$, thus statement ($\alpha$) holds with $C = T$ or $\{t\}$. Since $S$ is achronal without edge, one has $E^{\pm}S = S$ and so, because of the compactness of $S$, $E^+S$ is compact. Hence one may assume that $E^+K$ is compact (i.e., $K$ is a future trapped set).

Since $K$ is a future trapped set in a strongly causal space-time, there exists a future inextendable timelike curve $\mu$ in $\operatorname{int} D^+E^+K$. The set $F := E^+K \cap \overline{J^-\mu}$ is compact and achronal. Furthermore, $E^-F = F \cup G$, where $G$ is a connected subset of $\partial J^-\mu$. Thus through each point of $G$ there is a future inextendable null geodesic. Here $E^-F$ may be compact or noncompact. If $E^-F$ is not compact, then by Lemma 3.2 there is a null geodesic generator $\gamma$ of $E^-F - F$ that does not leave $E^-F$ in the past direction; i.e., $\gamma$ is an inextendable null geodesic in $\partial J^-\mu$ through $p := \gamma \cap F$. But $\gamma$ must be incomplete, as otherwise conditions (1) and (2) would imply the existence of a pair of conjugate points along $\gamma$. Thus with $C = \{p\}$, statement ($\alpha$) holds, therefore one can assume that $E^-F$ is compact.

$F$ is a past trapped set in a strongly causal space-time, thus there exists a past inextendable timelike curve $\lambda$ in $\operatorname{int} D^-E^-F$. From this point the proof is the same one given in Ref. 1: one can show that there is a point $q \in E^+K$ and a maximal inextendable nonspacelike geodesic $\gamma$ through $q$ in $D := \operatorname{int} DE^-F$. This geodesic must be incomplete, because if it were complete then a pair of conjugate points would occur, which would contradict its maximality.

The original version of the following theorems were published by Hawking.[14,1]

**Theorem 4.3**: If (1) $R_{ab}K^aK^b \geqslant 0$ for every nonspacelike vector $K^a$; (2) the strong causality condition holds on $M$; and (3) there is a point $p$, a past directed unit timelike vector $W$ at $p$ and a positive number $b$ such that the expansion $\theta = V^a{}_{;a}$ of the past directed timelike geodesics from $p$ with unit tangent $V^a$ becomes less than $-3|W^aV_a|b^{-1}$ within parameter distance $\Delta t = |W^aV_a|^{-1}b$; then at least one of the following statements holds: ($\alpha$) there is a past directed incomplete null geodesic from $p$ in $E^-\{p\}$, and/or ($\beta$) there is a past directed maximal incomplete nonspacelike geodesic from $p$ in $D^-E^-\{p\}$.

*Sketch of proof:* If $E^-\{p\}$ is not compact, then by Lemma 3.2 there is a null geodesic from $p$ in $E^-\{p\}$. If $E^-\{p\}$ is compact, then $\overline{D^-E^-\{p\}}$ must be noncompact, as otherwise there would be a past imprisioned timelike curve in $\overline{D^-E^-\{p\}}$. Thus, because of Lemma 3.3, there is a past endless maximal nonspacelike geodesic in $D^-E^-\{p\}$. However, these curves must be incomplete, according to conditions (1) and (2).

**Theorem 4.4**: If (1) $R_{ab}K^aK^b \geqslant 0$ for every timelike vector $K^a$, (2) there is a compact $C^2$ partial Cauchy surface, (3) the unit normals to $S$ are everywhere converging then there exists a maximal future incomplete timelike geodesic orthogonal to $S$ in $D^+S$.

Based on Lemma 3.4, a similar argument can be used to prove this statement too.

## V. DISCUSSION AND FINAL REMARKS

Since we summarized the results in the Abstract and the Introduction, we do not repeat them, but we have some final remarks.

The second condition of Theorem 4 of Ref. 1 is weaker than that of Theorem 4.4: for the proof of the existence of incomplete timelike geodesic, only a compact spacelike hypersurface without any edge is needed. Of course, in Hawking's covering space $M_H$, the maximality of the incomplete timelike geodesic $\gamma$ orthogonal to an $S$-homeomorphic preimage $S_H$ of $S$ (Ref. 15) can be proved. Thus, it would be interesting to see whether or not the geodesic might lose its maximality under the projection $\pi: M_H \to M$. If not then, of course, the second condition of Theorem 4.4 can be weakened to that of Theorem 4 of Ref. 1 (see *Note added in proof*).

There is another class of singularity theorems[16,17] where the maximality of the incomplete nonspacelike geodesics can be proved in certain covering spaces. However, it is not clear whether or not their maximality is preserved under the covering projection (see *Note added in proof*).

Among the oldest singularity theorems,[18,19] there is a great variety of statements that predict maximal incomplete timelike geodesics. For example, one can show easily the next theorem.

**Theorem 5.1**: If (1) $R_{ab} V^a V^b \geqslant 0$ for every timelike vector $V^a$; (2) $(M,g)$ admits a compact $C^2$ Cauchy surface $S$; and (3) the unit normals to $S$ are everywhere converging; then every future inextendible timelike curve has finite total length measured from one of its points, furthermore for every TIP $P$ there exists a maximal incomplete timelike geodesic $\gamma$ orthogonal to $S$ such that $I^- \gamma \subseteq P$.

Condition (3) can be replaced by one of the series of conditions that guarantee the expansion of the timelike geodesics becoming negative.[18,19,20] If the future causal boundary of space-time has no null part,[3,21] then Theorem 4.1 states that each point of the singular future boundary can be reached by maximal timelike geodesics.

Finally, it is worth noting that the incomplete null geodesic in Theorem 1 of Ref. 22 (which states that, roughly speaking, chronology violation creates incomplete null geodesics in an asymptotically flat space-time) is also maximal.

*Note added in proof:* Since the covering projection is a local diffeomorphism, the maximality of the inextendable nonspacelike geodesics is preserved: if there were a nontrivial Jacobi field along $\pi \circ \gamma$ with zeros then a Jacobi field describing conjugate points could be given along $\gamma$. (I am indebted to C. J. S. Clarke for suggesting this idea of the proof.)


## ACKNOWLEDGMENTS

I am indebted to Dr. Á. Sebestyén and the whole relativity group of CRIP for the continuous discussions. Special thanks to Miss M. Vida for the clarification of the properties of $G(\alpha,\nu)$.



[1]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-time* (Cambridge U. P., Cambridge, 1973).

[2]R. Penrose, *Techniques of Differential Topology in Relativity* (SIAM, Philadelphia, 1972).

[3]R. Penrose; in *Theoretical Principles in Astrophysics and Relativity*, edited by N. R. Lebovitz, W. H. Reid, and P. O. Vandervort (Chicago U. P., Chicago, IL, 1978).

[4]F. J. Tipler, C. J. S. Clarke, and G. F. R. Ellis, in *General Relativity and Gravitation*, edited by A. Held (Plenum, New York, 1980).

[5]F. J. Tipler, Phys. Rev. D **15**, 942 (1977).

[6]L. B. Szabados, Gen. Relativ. Gravit. **14**, 891 (1982).

[7]L. B. Szabados, Gen. Relativ. Gravit. **15**, 187 (1983).

[8]L. B. Szabados, in *Proceedings of the Fourth Marcel Grossmann Meeting*, edited by R. Ruffini (to be published).

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

[10]F. J. Tipler, Phys. Rev. D. **17**, 2721 (1978).

[11]A. Królak, Acta. Phys. Pol. B. **12**, 643 (1981).

[12]A. Królak, Gen. Relativ. Gravit. **14**, 793 (1982).

[13]S. W. Hawking and R. Penrose, Proc. R. Soc. London Ser. A **314**, 529 (1970).

[14]S. W. Hawking, Proc. R. Soc. London Ser. A **300**, 187 (1967).

[15]B. C. Haggman, G. W. Horndeski, and G. Mess, J. Math. Phys. **21**, 2412 (1980).

[16]D. Gannon, Gen. Relativ. Gravit. **7**, 219 (1976).

[17]D. Gannon, J. Math. Phys. **16**, 2364 (1975).

[18]S. W. Hawking, Proc. R. Soc. London Ser. A **294**, 511 (1966).

[19]S. W. Hawking, Proc. R. Soc. London Ser. A **295**, 490 (1966).

[20]F. J. Tipler, Astrophys. J. **209**, 12 (1976).

[21]R. M. Wald and P. Yip, J. Math. Phys. **22**, 2659 (1981).

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