

# Singularities in Globally Hyperbolic Space-Time

C. J. S. Clarke

Department of Applied Mathematics and Theoretical Physics, Cambridge, U.K.

Received April 7, 1974

**Abstract.** A singularity reached on a timelike curve in a globally hyperbolic space-time must be a point at which the Riemann tensor becomes infinite (as a curvature or intermediate singularity) or is of type  $D$  and electrovac.

## 1. Introduction

It is known ([1] § 8.2) that, in certain physically realistic situations, general relativity predicts either the occurrence of a singularity in spacetime or the violation of some sort of causality condition. Hitherto little has been known as to the nature of the singularities which might arise in this context: in particular, it has not been known whether or not the Riemann tensor must become “infinite”. The aim of the present paper is to show that, in the situations envisaged in the singularity theorems, the Riemann tensor cannot be well-behaved.

By ‘singularity’ I mean a point  $p$  on the  $b$ -boundary  $\dot{M}$  of a space-time  $M$  [2]. Such a point is the end-point of a curve in  $M$  which has finite length according to a generalised affine parameter defined by a parallelly propagated (p.p.) tetrad (see § 2). In particular, the singularities predicted by the singularity theorems are the end-points of incomplete timelike geodesics. The precise construction of the  $b$ -boundary defines when two finite-length curves have the same end point, which enables one to divide singularities into three classes, as follows ([4], with slight changes).

(i) Curvature singularity: there is a curve running to the singularity on which  $R_{\alpha\beta\gamma\delta}$  does not tend to a limit, in whatever tetrad it is evaluated.

(ii) Intermediate singularity: not (i); but there is a curve on which  $R_{\alpha\beta\gamma\delta}$  does not tend to a limit in a p.p. tetrad.

(iii) Locally extensible: not (i) or (ii).

The justification for basing the classification on p.p. tetrads, and the reason for the name in (iii), lie in the result [5] that a curve on which  $R_{\alpha\beta\gamma\delta}$  *does* tend to limit in a p.p. tetrad has a neighbourhood isometric to a neighbourhood in a singularity-free space time. If  $p$  is a singularity of this third type, then there are two sub-cases.

(a)  $p$  is inessential: there is an isometry  $\psi : M \rightarrow M'$  into a larger ( $C^2$ , Hausdorff) space-time which carries  $p$  into an interior point, i.e.  $\bar{\psi}p \in M'$  (where  $\bar{\psi}$  is  $\psi$  extended to the  $b$ -boundary, [3]).

(b) Otherwise  $p$  is essential.

An inessential singularity is one created by “cutting out” a portion from a larger space-time  $M'$ . There would seem to be no conceivable physical mechanism

for such surgery, and so this class must be excluded: in a physical model, if the space-time can continue past  $p$ , then it will.

In classes (i) and (ii) (with the exception of the behaviour discussed in example III below) it seems likely that the non-convergence of the Riemann tensor in a p.p. frame would cause any test system to be subjected to indefinitely large tidal forces as it approached the singularity, and that quantum effects would at some point intervene. But in class (iii)  $b$ , the essential locally extensible singularity, one has the anomalous situation that a true singularity is apparently encountered without any associated infinite tidal forces. This is the class which, in the context of the singularity theorems, is effectively ruled out by the theorem of the present paper.

Three examples of class (iii)  $b$  singularities serve to define the subsidiary conditions that must be imposed.

*Example I.* The cone. Take the cylindrical polar Minkowski metric

$$ds^2 = -dt^2 + dz^2 + dr^2 + r^2 d\phi^2$$

but identify  $\phi = 0$  with  $\phi = \alpha \mp 2\pi$ . The 2-surfaces  $t = \text{const.}$ ,  $z = \text{const.}$  are ordinary cones, having a singularity at the point. (Alternatively, one may take  $-\infty < \phi < +\infty$ ).

The singularity at  $r=0$  in this space-time exists for all  $t$ : in cosmological terms, it is a primordial singularity, explicitly present in the initial data. The singularity theorems are not concerned with these, but with singularities that arise from regular data through the collapse of matter. To express this dependence on the data we demand that  $M$  be globally hyperbolic. In applications this does not restrict the entire space-time: we need only a part of the space-time to be globally hyperbolic and contain the singularity.

This restriction is still not quite enough, since a "conical" singularity could still be present on an initial surface  $t=0$  but at no other time: for example, take the flat metric given above, but restrict  $r$  to  $r > |t|$ . These singularities are ruled out if we consider only those which are accessible along timelike curves (as are the singularities predicted by the theorems).

*Example II.* "2-dimensional Taub-NUT" [7] (see Fig. 2). Write the past null cone of Minkowski space in hyperbolic coordinates as

$$ds^2 = -dq^2 + q^2 d\phi^2 + dy^2 + dz^2$$

( $-\infty < \phi < +\infty$ ,  $-\infty < q < 0$ ) and then identify  $\phi = 0$  with  $\phi = k$ ; the surfaces  $y = \text{const.}$ ,  $z = \text{const.}$  are then "cones" with timelike generators. The singularity is at  $q = 0$ . This space is globally hyperbolic ( $q = \text{const.}$  is a Cauchy surface). But it has nonetheless come from a larger space-time which has been identified under an isometry ( $\phi \rightarrow \phi + k$ ) with a fixed point: this is indicated by the Riemann tensor's admitting a boost-like isotropy at the singularity. This is only possible with "reasonable" matter if it is of type  $D$  and vacuum there (Definition 2, below, and scholium). We must thus rule this possibility out by a special condition (#3).

*Example III.* The metric

$$ds^2 = -dt^2 + dx^2 + dy^2 + \left(1 + \frac{1}{2} \tanh(r^2 t^4)\right) \sin(1/t) dz^2$$

$$(t < 0, r^2 = x^2 + y^2 + z^2).$$

This has a locally extendible singularity at the coordinate origin which is made essential by a pathological  $\sin(1/t)$  behaviour of the Riemann tensor nearby. Such behaviour should not be physically significant, since a system with a finite response time will, for small enough  $t$ , rest unaffected by the oscillations in  $R$ . In the case of this example, such a system would be progressively less affected by the tidal force as  $t \rightarrow 0$ , suggesting that it should be possible to define a "continuation" of the metric as Minkowsky space in  $t > 0$ . Hence we must consider space-times which are not  $C^2$ . Clearly this condition cannot be suspended entirely, since a Riemann tensor must be definable in some sense. We shall therefore work in a weaker differentiability class, requiring the metric to be  $C^{1-}$  and the Riemann tensor locally bounded (Definition 1 below).

If we thus allow discontinuity in the Riemann tensor, we cannot expect it to have a good limiting behaviour at the singularity. We must therefore consider a wider class than that of the locally extendible singularities, namely the class of those where the Riemann tensor is merely *bounded*.

I can now state my primary definitions and the theorem, whose full proof will follow in subsequent sections. It will be convenient to formulate these in terms of  $L(M)$ , the bundle of all pseudo-orthonormal frames on  $M$  ([1] p. 52). Thus a point  $u \in L(M)$  is a tetrad  $(u_0, u_1, u_2, u_3)$  of (contravariant) vectors at a point  $x \equiv \pi(u) \in M$  with  $g(u_\alpha, u_\beta) = \eta_{\alpha\beta}$ ,  $\eta = \text{diag}(-1, +1, +1, +1)$ . Tetrad indices (Greek) are raised and lowered by  $\eta$ . The connection in  $L(M)$  defines *horizontal curves*; these are simply curves formed by parallelly propagating a frame along a curve of  $M$ . (Here, and throughout, all curves are assumed  $C^0$ , piecewise  $C^1$ ). If  $t$  is a totally covariant tensor I shall denote its components in a frame  $u$  by  $t_{\alpha\beta\dots\kappa}(u) \equiv t(u_\alpha, u_\beta, \dots, u_\kappa)$ , also writing  $c_{\alpha\beta\dots\kappa;\lambda}(u) = (\nabla t)_{\alpha\beta\dots\kappa\lambda}(u)$ .

In what follows " $C^{r-}$ " means that the  $(r-1)$ st. derivatives satisfy Lipschitz conditions, while " $C^{0-}$ " means "locally bounded and locally integrable".

*Definition 1.* A space-time  $(M, g)$  is of *strong curvature differentiability class*  $C^{2-}$  ( $sc - C^{2-}$ ) if  $M$  has a  $C^\infty$  atlas with respect to which the components of the metric are  $C^{1-}$  and those of the Riemann tensor are  $C^{0-}$ .

*Scholium.* The qualification "strong" here is to distinguish this definition from a similar one in [8]. It does not appear to be known whether the two definitions are really distinct, or whether  $C^{2-}$  and  $sc - C^{2-}$  are distinct. The important point about the definition used here is that it implies the existence and uniqueness of geodesics. Indeed, an  $sc - C^{2-}$  space-time can be regarded as the limit of a sequence of  $C^2$  space-times. For, let  $\{g_n\}$  be a sequence of smoothed metrics tending to  $g$ , defined by

$$g_{n\mu\nu}(X) = \int S(n(x^\rho - y^\rho)) n^4 g_{\mu\nu}(y^\rho) d^4 y$$

where  $S: \mathbb{R}^4 \rightarrow \mathbb{R}^+$  is a suitably normalised smoothing kernel. Then, almost everywhere,  $g_{n\alpha\beta,\lambda} \rightarrow g_{\alpha\beta,\lambda}$ , a locally bounded generalised derivative of  $g_{\alpha\beta}$ ; moreover  $R_{\alpha\beta\gamma\delta}(g_n) \rightarrow R_{\alpha\beta\gamma\delta}(g)$ . The boundedness of  $R_{\alpha\beta\gamma\delta}$  implies that round each point  $x$  there is a neighbourhood  $U_x$  in which geodesics are unique for all the  $g_n$ . In  $U_x$  we can *define* the geodesic with initial tangent vector  $\xi^\mu \partial/\partial x^\mu|_x$  as the limit of the geodesics with this tangent vector in the  $g_n$ , the limit existing

as a  $C^{2-}$  curve and being unique, both by virtue of the limiting behaviour of  $g_{n\alpha\beta,\lambda}$  and  $R_{\alpha\beta\gamma\delta}(g_n)$ . Any result on geodesics which can be formulated in terms of the Riemann tensor alone will then hold on an  $sc - C^{2-}$  space-time as the limiting case of such a sequence of  $g$ 's. Thus in the following sections I shall prove lemmas only for the  $C^2$  case, the extensions to  $sc - C^{2-}$  space-times being understood.

Now denote the spinor components of the Weyl tensor [9] by  $\Psi_{ABCD}(u)$  and those of the Ricci tensor by  $\Phi_{ABX'Y'}(u)$ , where  $u$  is a (spin-entangled) frame.

*Definition 2.* A space-time  $(M, g)$  is *D-specialised* at  $p$ , for  $p \in \bar{M}$ , if every curve  $\kappa : [0, 1) \rightarrow M$  terminating at  $p$  has a horizontal lift  $\mu$  in  $L(M)$  such that the components  $\Psi_{ABCD}(\mu(s))$  and  $\Phi_{ABX'Y'}(\mu(s))$  all tend to zero except for  $(ABCD) = (0011)$  and  $(ABX'Y') = (010'1')$  or  $(001'1')$  or  $(110'0')$  and the components related to them by symmetry.

*Scholium.* If  $p \in M$ , “*D-specialised*” has the following meaning: there is a tetrad  $u \in L_p(M)$  such that  $R_{\alpha\beta\gamma\delta}(L(v)u)$  is independent of  $v$ ,  $L(v)$  being a boost along the  $z$ -axis. Since the spinor equivalent of  $L(\tanh \theta)$  is the matrix  $\text{diag}(e^{\theta/2}, e^{-\theta/2}) \in SL(2, C)$  this clearly happens if, and only if, all spinor components vanish in the frame  $u$  except for those referred to in the definition. This in turn means that the Weyl tensor is of type *D* while the Ricci tensor, in an appropriately rotated tetrad, takes the form  $\text{diag}(\varrho, p, q, -\varrho)$ . For this to happen we must have either negative pressure ( $\varrho > 0$ ), negative density ( $\varrho < 0$ ), energy-less matter ( $\varrho = 0, p^2 + q^2 \neq 0$ ) or vacuum ( $\varrho = p = q = 0$ ) – this last being the only physically realistic possibility. Thus “*D-specialised*” is physically equivalent to “type *D* and vacuum”.

**Theorem.** Let  $(M, g)$  be a globally hyperbolic  $sc - C^{2-}$  spacetime and  $p \in \dot{M}$  a singularity for which

- (1) there is a future-directed timelike curve  $\kappa : [0, 1) \rightarrow M$  terminating at  $p$ ,
- (2) on any horizontal curve  $\mu : [0, 1) \rightarrow L(M)$  for which  $\pi \circ \mu$  terminates at  $p$  the numbers  $R_{\alpha\beta\gamma\delta}(\mu(s))$  are bounded,
- (3)  $(M, g)$  is not *D-specialised* at  $p$ . Then there is an  $sc - C^{2-}$  space-time  $M'$  and an isometry  $\phi : M \rightarrow M'$  for which  $\bar{\phi}(p) \in M'$ .

*Outline of Proof.* (i) The boundedness of the Riemann tensor in (2) is used (§ 2) to derive inequalities on the lengths of timelike curves, characterising the extent to which the geometry departs from the Minkowskian. These are then used (§ 3) to construct a neighbourhood  $N$  of the form  $I^-(|\kappa|) \cap I^+(O)$ , (where  $O \in |\kappa| \equiv \kappa([0, 1))$ ) in which no geodesic attains sufficient length to develop conjugate points, and where uniform bounds can be set on all relevant quantities.

(ii) Within  $N$  a geodesic is constructed (§ 4) running to  $p$ . This is then “thickened” to a bundle of approximately parallel geodesics which terminate at points which cover a neighbourhood of  $p$  in  $\bar{M}$ . The generality condition (3) is used to show that these geodesics are non-intersecting, and cover a cone-shaped region  $H$  in which good coordinates can be defined. It is extended to a larger space  $K$ , and finally  $K$  is joined back to  $M$  to define the full extension.

## 2. Results on Families of Geodesics

First note that a frame  $u \in L(M)$  can be identified with a map  $u : \mathbb{R}^4 \rightarrow T_{\pi(u)}(M)$ ; i.e.  $u(\xi)$  means  $\xi^\alpha u_\alpha$ . In  $\mathbb{R}^4$  we use the Euclidean norm  $\| \cdot \|$ . Then the *generalised*

affine parameter length ([1] p. 259) (or simply, the length)  $l(\lambda, u; s)$  of a curve  $\lambda: [0, \tau_\lambda] \rightarrow M$  (defined to be the  $b$ -metric [2] length, up to  $s$ , of that horizontal lift  $\mu$  of the curve which starts at  $u \in L_{\lambda(0)}(M)$ ) is given by

$$l(\lambda, u; s) = \int_0^s \|(\mu(s'))^{-1}(\dot{\lambda}(s'))\| ds'.$$

The argument  $u$  will be omitted if no confusion thereby arises, and  $l(\lambda; \tau_\lambda)$ , where defined, will be denoted by  $l(\lambda)$ .

Let  $\lambda_t(s) \equiv \lambda(s, t)$  ( $0 \leq t \leq t_0, 0 \leq s \leq \tau_\lambda^t$ ) be a one-parameter family of geodesics such that the curves  $\lambda(s, \cdot)$ , for each  $s$ , are  $C^0$ , piecewise  $C^1$ , and choose  $\hat{u} \in L_{\lambda(0,0)}(M)$ . This frame can be propagated either first along  $s=0$  and then along  $t = \text{const.}$ , yielding a lift  $\mu$  of  $\lambda$ ; or first along  $t=0$  and then  $s = \text{const.}$ , yielding  $\mu'$ . Thus:  $\mu(0, 0) = \mu'(0, 0) = \hat{u}$ ,  $\pi \circ \mu = \pi \circ \mu' = \lambda$ ,  $\omega(\mu_*(\partial/\partial s)) = \omega(\mu'_*(\partial/\partial t)) = \mathbf{0}$ , where  $\omega$  is the connection form in  $L(M)$ .

Define Lorentz transformations  $L(s, t): \mathbb{R}^4 \rightarrow \mathbb{R}^4$  by  $\mu' = \mu \circ L$  and set  $V = \lambda'_*(\partial/\partial s)$ ,  $W = \lambda'_*(\partial/\partial t)$ ,  $V = \mu^{-1}(V)$ ,  $W = \mu^{-1}(W)$ ,  $W' = \mu'^{-1}(W)$ . From  $[\partial/\partial s, \partial/\partial t] = 0$  we obtain  $[V, W] = 0$  and hence

$$\nabla_V \nabla_W P = \nabla_W \nabla_V P + R(V, W)P$$

(for standard notations here, see [1]). Applying this to  $W$  gives Jacobi's equation

$$\partial^2 W_\delta^0 / \partial s^2 = R_{\alpha\beta\gamma\delta}(\mu(s, t)) V^\alpha V^\gamma W^\beta; \quad (1)$$

while application to  $\mu'_\alpha$  gives

$$\frac{\partial^2 L_{\alpha}{}^\beta}{\partial s \partial t} - \frac{\partial L_{\alpha}{}^e}{\partial t} \frac{\partial L_{\sigma}{}^\beta}{\partial s} L^\sigma_e + L^{\delta\beta} R_{e\sigma\alpha\delta}(\mu(s, t)) W^e V^\sigma = 0. \quad (2)$$

Our proofs will be based on the fundamental idea that, if  $R_{\alpha\beta\gamma\delta}(\mu)$  is limited by some bound, then (1) and (2) give rise to corresponding bounds on  $W$  and  $L$ . We bound  $R$  by setting, for  $v \in L(M)$

$$\bar{r}(v) = \sup \{ \|R_{\alpha\beta\gamma\delta}(v) A^\alpha B^\gamma C^\delta\| : A, B, C \in \mathbb{R}^4, \|A\| = \|B\| = \|C\| = 1 \}$$

and  $r(\lambda; u) = \sup \{ \bar{r}(\mu(s)) : 0 \leq s < \tau_\lambda \}$ , where  $u, \mu$  and  $\tau_\lambda$  are as in the definition of  $l$ . Finally, if  $\lambda_t$  and  $\mu_t \equiv \mu(\cdot, t)$  are as just described, with  $\tau_{\lambda'} \equiv 1$ , define

$$\hat{r}(t) = \sup \{ r(\lambda_{t'}; \mu(0, t')) : 0 \leq t' \leq t \}$$

and

$$\hat{l}(t) = l(\lambda_t, \mu(0, t); 1)$$

$$d(t) = l(\lambda(1, \cdot), \mu'(1, 0); t).$$

**Proposition 1.** *Let  $\lambda: [0, 1] \times [0, 1] \rightarrow M$  define a 1-parameter family of geodesics and let  $\hat{u} \in L_{\lambda(0,0)}(M)$  be chosen so as to define  $\mu$  and  $\mu'$ . Suppose that  $W(0, t) = 0$ , that  $\|W'(1, t)\| = \text{const.}$  ( $= d(1)$ ) and that the functions  $\hat{r}$ ,  $\hat{l}$  and  $d$  satisfy*

$$\hat{l}(t) \leq 1, \quad (3a)$$

$$d(t) \leq 3/10, \quad (3b)$$

$$\hat{r}(t) \hat{l}(t) \leq 1/2 \quad (3c)$$

for all  $t \in [0, 1]$ . Then

$$\|W(s, t) - s(\partial W/\partial s)(0, t)\| \leq 1/2 \hat{l}^2 \hat{r} s^3 \|(\partial W/\partial s)(0, t)\| \quad (4a)$$

$$\|W(s, t) - sW(1, t)\| \leq 2\hat{l}^2 \hat{r} s \|W(1, t)\| \quad (4b)$$

$$\|W(1, t)\| \leq 1/3 t \quad (4c)$$

$$\|L_\alpha^\beta - \delta_\alpha^\beta\| \leq 2st\hat{r}\hat{l} \|W(1, t)\| \quad (4d)$$

*Proof.* (4a) follows directly from (3a), (3c) and (1). This then implies (4b). Then (4b) and (2) can be used to show that so long as  $t$  is so small that (4c) holds, then (4d) will hold. But  $\|W(1, t)\| \leq d(1) \|L(1, t)\|$ , whence it follows from (3b) and (4d) that equality cannot hold in (4c) if  $t \in [0, 1]$ . Hence (4c) and (4d) are always valid.  $\square$

**Proposition 2.** *Under the same conditions, suppose that the curve  $\kappa: t \mapsto \eta(1, t)$  and the geodesics  $\lambda_t$  are all timelike. Define  $\tau_1(t) = |V^0(1, t)|$  and*

$$\tau_2(t) = \int_0^t |W'^0(1, t')| dt'.$$

Then if  $(i, j) = (1, 2)$  or  $(2, 1)$  we have

$$\left| \frac{d\tau_i}{d\tau_j} - 1 \right| < \frac{11}{2} \hat{r}(1) \hat{l}(1) \quad (5)$$

*Proof.* If we note that

$$\frac{d\tau_2}{dt} = W_2(1, t) L^2_0, \quad \frac{d\tau_1}{dt} = \frac{\partial W_0}{\partial s} \Big|_{0,t}$$

then (5) follows immediately from the definitions and Eqs. (3) and (4).  $\square$

A number of other minor results will be used, which follow straightforwardly from the inequalities established above.

### 3. Definition of N

Let  $Q, O \in |\kappa|$  with  $Q \in I^+(O)$ . Define two sets of curves:  $A_O$ , the set all future-directed causal curves  $\lambda: [0, \tau_\lambda] \rightarrow M$  through  $O$  that are homotopic to  $\kappa$  through causal curves; and  $A_{OQ}$ , the subset of  $A_O$  consisting of curves homotopic to the part of  $\kappa$  between  $O$  and  $Q$  through causal curves terminating at  $Q$ .

Choose a horizontal lift  $v$  of  $\kappa$ , and for  $\lambda \in A_{OQ}$  a lift  $\mu$  coinciding with  $v$  at  $O$ . Then at  $Q$  there is a Lorentz matrix  $L(\lambda)$  such that  $v_Q = \mu_Q \circ L$ , (writing for  $v(\kappa^{-1}(Q))$ ,  $v_Q$ , and so on).

**Lemma 1.** *There are positive constants  $\varepsilon, N_1, N_2, N_3, N_4$  such that for any  $t_2$  having  $d = l(\kappa, v_O; t_2) < \varepsilon$  we have either*

A) *For all  $\lambda \in A_{OQ}$ , where  $Q = \kappa(t_2)$ ,*

$$l(\lambda) < N_1 d < 1, \quad (6a)$$

$$r(\lambda) l(\lambda) < 1 \quad (6b)$$

or

B) There is a curve  $\lambda_0 \in \Lambda_{OQ}$  with

$$l(\lambda_0) < N_4 d, \quad (7a)$$

$$r(\lambda_0) > N_2 d^{-1/2}, \quad (7b)$$

$$\|L(\lambda_0) - I\| < N_3 d^{1/2}. \quad (7c)$$

*Proof.* Suppose that  $\kappa$  be parametrised so that  $\kappa(0) = O$ . Set  $\hat{u} = v(0)$ . Every point to the future of  $O$  on  $|\kappa|$  can be joined to  $O$  by a geodesic, since  $M$  is globally hyperbolic ([1] p. 213). Hence  $\exp_0$  covers  $|\kappa|$ . If  $U \subset T_O(M)$  is the domain of regularity of  $\exp_0$  then there is a curve  $X : [0, t_1] \rightarrow U$  such that  $X(0) = \mathbf{0}$ ,  $\exp X(t) = \kappa(t)$ . Choose  $t_1$  maximally: either  $t_1 = 1$  or  $|\overline{X}| \cap \dot{U} \neq \emptyset$  ( $|X| = X([0, t_1])$ ).

Define  $\lambda_t(s) = \exp_s X(t) = \lambda(s, t)$ . Let  $t_0$  be the largest number such that  $t_0 \leq t_1$  and (3) holds for  $t \leq t_0$ . Then  $t_0 > 0$ . We can now reparameterise  $\kappa$  so that all the conditions of Proposition 1 hold,  $t_0$  becoming 1 under reparameterisation. In  $[0, 1)$ ,  $X(t)$  will remain causal since

$$\frac{d}{dt} [-g(X(t), X(t))]^{\frac{1}{2}} = -\frac{(V_\alpha W^\alpha)|_{s=1}}{[-g(X(t), X(t))]^{\frac{1}{2}}} > 0 \quad (8)$$

(cf. [1] p. 107). Thus Proposition 2 holds as well.

Note now that either  $t_0$  is the parameter of the endpoint of  $\kappa$ , or else equality must hold in one of Eq. (3). For, if  $t_0$  were not the endpoint and if (3) held with strict inequalities up to  $t_0$ , then  $t_0$  would equal  $t_1$ . Thus a vector  $X(t_0)$  would be definable, in the closure of  $|X|$ , on which there were conjugate points. But from (4a) this would require  $\frac{1}{2} \hat{l}^2 \hat{r} \geq 1$ , so that (3c) would be violated.

Set  $\hat{l}(t_0) = l_1$ ,  $d(t_0) = d_1$ ,  $\hat{r}(t_0) = r_1$ , and return to the original parameterisation of  $\kappa$ , noting that (5) does not depend on the parameterisation used.

Fix  $Q = \kappa(t_2)$  to the future of  $O$  for which  $l(\kappa; t_2) \leq 1/5$  so that (3b) is satisfied for  $t \leq t_2$ . Then for  $t \leq \min(t_0, t_2)$  the inequalities

$$l/\sqrt{2} \leq \tau_1 \leq l, \quad d/\sqrt{2} \leq \tau_2 \leq d \quad (9)$$

with (5) imply that  $l_1 \leq 15d_1/2\sqrt{2} < 1$ , giving strict inequality in (3a). Hence either  $t_0 \geq t_2$  or there is equality in (3c):

$$r_1 l_1 = 1/2.$$

Next, suppose we have a homotopy  $\kappa^h, h \in [0, 1]$  between  $\kappa^0$  and  $\kappa^1$  through timelike curves keeping  $O$  and  $Q$  fixed. So long as  $\exp$  remains diffeomorphic this lifts to a homotopy  $X^h$  of  $X$ , and so to a homotopy of  $\lambda$ .

We now repeat the preceding argument to show that  $t_0^h$ , the value of  $t_0$  on  $\kappa^h$ , either (α) remains throughout at  $t_2$ , or else (β) there is a maximal value  $h'$  at which  $\kappa^{h'}(t_0^h) = Q$  and  $r_1^{h'} l_1^{h'} = 1/2$ .

In case (α) the final geodesic  $\lambda_{t_2}^h$ , having no conjugate points, is independent of  $h$ . Comparing it with both  $\kappa^0$  and  $\kappa^1$  via (5) and (9) we find

$$d_1^i < 16\sqrt{2} d_1^i. \quad (10)$$

If this holds throughout  $A_{OQ}$  then case  $A$  of the lemma is proved: (10) implies (6a) when  $N_1 = 16\sqrt{2}$  and  $\varepsilon < 1/16\sqrt{2}$ , while (6b) follows from (4d), (10), (3a) and (3c) provided  $\varepsilon < 1/50$ , say.

On the other hand, suppose there is a  $\lambda$  for which case  $(\beta)$  holds. Then (10) holds for all  $h$  up to  $h'$  and we can find a  $t'$  at which

$$\hat{r}^{h'}(t') \hat{l}^{h'}(t') = \frac{1}{2}(d_1^{h'})^{\frac{1}{2}} < 5(d_1^1)^{\frac{1}{2}}$$

by (10). Then, from (5) and (9) again,

$$\hat{r}^{h'}(t') > 1/(60(d_1^1)^{\frac{1}{2}})$$

$$\|L' - I\| < 5/3(d_1^1)^{\frac{1}{2}}$$

and the curve formed by  $\lambda_t^{h'}$  followed by  $\kappa^{h'}|[t', t_2]$  has length

$$D < 167 d_1^1$$

Case B of the lemma then follows on taking  $\varepsilon = 1/50$ ,  $N_2 = 1/60$ ,  $N_3 = 5/3$ ,  $N_4 = 167$ .  $\square$

**Lemma 2.** *Under the assumptions of the theorem, we can find a point  $O$  on  $|\kappa|$  such that alternative (A) holds in Lemma 1 for every  $Q$  on  $|\kappa|$  to the future of  $O$ .*

*Proof.* Suppose not; i.e. that for every  $O$  there is a  $Q$  such that (B) holds. Then we can set up sequences  $O_1, O_2, \dots$  and  $Q_1, Q_2, \dots$ , with  $O_i = \kappa(a_i)$ ,  $Q_i = \kappa(b_i)$  having causal curves  $\lambda_i$  from  $O_i$  to  $Q_i$  as in (B), and such that  $a_i < b_i < a_{i+1}$  ( $i = 1, 2, \dots$ ) and  $l(\kappa_{(i)}) < \min(\varepsilon, 1/(N_3 2^{i+1})^2)$ , where  $\kappa_{(i)}(t) \equiv \kappa(t + a_i)$ . Let  $\tilde{\lambda}$  be the curve formed from the curve-segments  $\lambda_1, \kappa|[b_1, a_2]$ ,  $\lambda_2, \dots, \lambda_i, \kappa|[b_i, a_{i+1}]$ , ... and let  $\tilde{\mu}$  be its horizontal lift through  $v(a_1)$ . This defines at each  $s \in [b_i, a_{i+1}]$  a Lorentz matrix  $\tilde{L}(s)$  such that  $\tilde{\mu}(s) = L(s)v(s)$ . If  $L_{(i)}$  is the value of the  $L(\lambda_0)$  in Lemma 1 associated with  $\lambda_i$ , then  $\tilde{L}(b_i) = \prod_{k=1}^i L_{(k)}$ . Consequently  $\|\tilde{L}(b_i)\| < 2$  by (7c). The components of the Riemann tensor at  $R \in |\lambda_i|$  measured by  $\tilde{\mu}$  are obtained from those measured in  $\mu_i$ , the lift of  $\lambda_i$  through  $v(a_i)$ , by an application of  $\tilde{L}(b_{i-1})$ . This Lorentz matrix is bounded, while the  $\mu_i$ -components, by (7b), are unbounded. Thus the  $\tilde{\mu}$ -components are unbounded. Moreover  $\tilde{\lambda}$  has finite length and runs to  $p$ , again using the boundedness of  $\tilde{L}(b_i)$  and (7a). But this contradicts (2).  $\square$

We now choose  $O$  as in this lemma, and define  $N$  to be  $I^- (|\kappa|) \cap I^+ (O)$ .

#### 4. Construction of $K$

The situation in this section is depicted in Fig. 1.

We first show that there is a geodesic running to  $p$ .

Having fixed  $O$ , we can suppose that  $O = \kappa(0)$  and that  $\kappa$  has a generalised affine parameterisation. In Lemma 1 we constructed a map  $X : [0, b] \rightarrow T_0^+(M)$  (the set of future causal vectors in  $M$ ), where  $\exp_0 \circ X = \kappa|[0, b]$  for some  $b \in (0, 1)$ ,



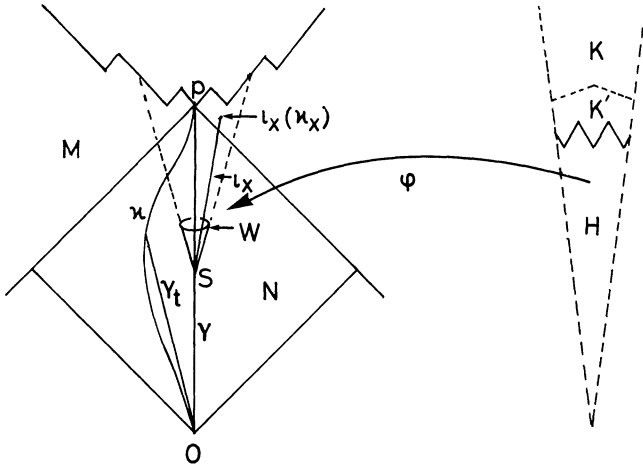


Fig. 1. The construction of  $K$ .  $\gamma$  is the limit of geodesic  $\gamma_t$  as its endpoint runs up  $\kappa$

and our choice of 0 now ensures that this map extends to  $[0, 1)$ . Let  $\gamma_t(s) = \exp_s X(t)$ . Then the next lemma will construct the required geodesic as the limit of the  $\gamma_t$ .

**Lemma 3.** *A geodesic  $\gamma$  is defined by  $\gamma(s) = \lim_{t \rightarrow 1} \gamma_t(s)$ , for all  $s \in (0, 1)$ ;  $\gamma$  is timelike.*

*Proof.* Let  $\lambda$  be as in the proof of Lemma 1. The same arguments as in Proposition 1 show that  $X(t)$  is bounded and uniformly continuous in  $t$ . Application of (8) then shows that  $X_1 = \lim_{t \rightarrow 1} X(t)$  is timelike. It thus suffices to show that  $\lim_{t \rightarrow 1} \lambda(s, t)$  exists for  $s \in (0, 1)$ .

Suppose that it failed to exist for  $s = q$ . Choose  $p, q < p < 1$ . Because  $X_1$  is timelike and  $W$  is bounded we see that for  $t_2$  close enough to 1 the curve  $\eta_{t'}$ , below, is timelike for all  $t' \in (t_2, 1)$ :

$$\eta_{t'}(s) = \begin{cases} \lambda(s, t_2(s - q)/(p - q) + t'(p - s)/(p - q)) & (q \leq s \leq p) \\ \lambda(s, t_2(q - s)/q + st'/q) & (0 \leq s \leq q) \end{cases}$$

These, for varying  $t'$ , constitute a family of timelike curves from  $\lambda(p, t_2)$  to  $\lambda(0, t_2)$  having no limit curve, contradicting the global hyperbolicity of  $M$ .  $\square$

Suppose in what follows that  $v$  be chosen so that

$$\overset{0}{u}^{-1}(X_1) = (l_0, 0, 0, 0).$$

and, for  $p \in |\gamma|$ , define  $v_p$  by parallel propagation of  $\overset{0}{u}$  up  $\gamma$ . Choose a point  $S$  on  $|\gamma|$ , a distance  $\tau$  from  $p$  ( $S = \exp(1 - \tau/l_0) X_1$ ).

Consider the field of geodesics  $l_x : [0, k_x) \ni s \mapsto \exp_s x \in M$  where  $x \in W' = \{x \in T_S^+(M) | v_S^{-1}(x)_0 = 1\}$  and where  $k_x$  is a function to be chosen in Lemma 4 below. Define  $\mu(s', x) \in L_{l_x(s')}(M)$  by parallelly propagating  $v_S$  up  $l_x$ . Let  $x_0 = v_S((1, 0, 0, 0))$ .

**Lemma 4.** For small enough  $\tau$  there is a neighbourhood  $W \subset W'$  containing  $x_0$  such that, for each  $x$  in  $W$ , the geodesic  $\iota_x$  terminates at a  $b$ -boundary point  $p^x$  at which conditions (1) and (2) of the theorem are fulfilled. Moreover

$$R_{\alpha\beta\gamma\delta}(\mu(s, x)) \quad \text{is bounded near} \quad (\tau, x_0).$$

*Proof.* Consider the one-parameter family  $\iota_{\xi(t)}$  defined by a curve  $\xi : [0, 1] \rightarrow W'$  with  $\xi(0) = x_0$ . The arguments of § 3 can be used with  $\lambda(s, t) = \iota_{\xi(t)}(s)$ , restricting the domain of  $(s, t)$  by a choice of  $k_x$  so that (3) is satisfied. The  $\mu$  of § 3 is then  $(s, t) \mapsto \mu(s, \xi(t))$ , a function which I will still call  $\mu$  where no confusion can arise.

Now  $\limsup k_{\xi(t)} = \tau$  ( $t \rightarrow 0$ ), by an argument similar to that in Lemma 3: if  $y = \lambda(\tau + \varepsilon, t)$  existed for  $t$  sufficiently close to 0, then one could construct a family of timelike geodesics from  $y$  to  $\iota_{x_0}(s)$ , for various  $s$ , such that the family of curves formed by these geodesics and a part of  $\iota_{x_0}$  would have no limiting curve, contradicting global hyperbolicity.

Also  $\liminf k_{\xi(t)} = \tau$ , since otherwise  $\iota_{\xi(t)}$  would terminate in  $N$  for some  $t$  near enough to 0. It is easily seen that such a geodesic could be prolonged in  $A_{OQ}$  for some  $Q$ , contradicting the construction of  $N$ . Hence  $k_{\xi(t)} \rightarrow \tau$ .

Moreover, for small enough  $t$  and  $\tau$ ,  $k_{\xi(t)}$  must delimit the actual endpoint of  $\iota_{\xi(t)}$ . For suppose it were otherwise: then for each  $\tau$  there would be a sequence  $\lambda(k_{\xi(t_i)}, t_i)$  at which equality obtained in (3c), but up to which propositions (1) and (2) would be valid. Define curves  $\varrho_i$  by

$$\varrho_i(s) = \begin{cases} \mu(k_{\xi(t_i)}(1 - 2s) + 2ss_i, t_i) & (0 \leq s \leq \frac{1}{2}) \\ \mu(s_i, (2 - 2s)t_i) & (\frac{1}{2} \leq s \leq 1) \end{cases}$$

where the  $s_i$  are a sequence tending to zero more slowly than  $|\tau - k_{\xi(t_i)}|$  so that  $\mu(s_i, (2 - 2s)t_i)$  is defined. From Proposition 1, the  $b$ -metric lengths of these curves tend to zero. Now  $R_{\alpha\beta\gamma\delta}(\mu(k_{\xi(t_i)}, t_i)) \geq \frac{1}{4}k_{\xi(t_i)}$ , for some  $(\alpha\beta\gamma\delta)$ , if there is equality in (3c). So, if this holds for arbitrarily small  $\tau$ ,  $R_{\alpha\beta\gamma\delta}$  is not bounded at the endpoint  $w$  of  $v$  in  $\overline{L(M)}$ , contradicting (2ii).

The boundedness of the  $R_{\alpha\beta\gamma\delta}$  is clearly a local property of a neighbourhood of  $w$  and so holds for the  $p_{\xi(t)}$  (some of which may, a priori, coincide with  $p$ ).  $\square$

Choose  $W$  as in this lemma and define coordinates  $z^a$  in  $W$  by  $z^a(x) = (v_S^{-1}(x))^a$ . Set  $H_W = \{z^a(x) | x \in W\}$ , define  $\phi_W : H_W \rightarrow W$  to be the inverse of  $z$  and write

$$H = \{z \in \mathbb{R}^4 | (z^a) \in H_W \text{ \& } z^0 \in [0, k_{\phi_W(z^a)}]\}.$$

Then define the  $C^1$  map  $\phi : H \rightarrow M$  by  $\phi(s, z^a) = \iota_{\phi_W(z^a)}(s)$ .

**Lemma 5.** For  $\tau$  and  $W$  sufficiently small,  $\phi$  is 1 - 1.

*Proof.* Suppose the contrary (see Fig. 2). Then there are sequences  $\{(s_i, z_i)\}$ ,  $\{(s'_i, z'_i)\}$  in  $H$  with  $\phi(s_i, z_i) = \phi(s'_i, z'_i)$ ,  $z_i \rightarrow 0$ ,  $z'_i \rightarrow 0$ ,  $s_i \rightarrow \tau$  and  $s'_i \rightarrow \tau$ .

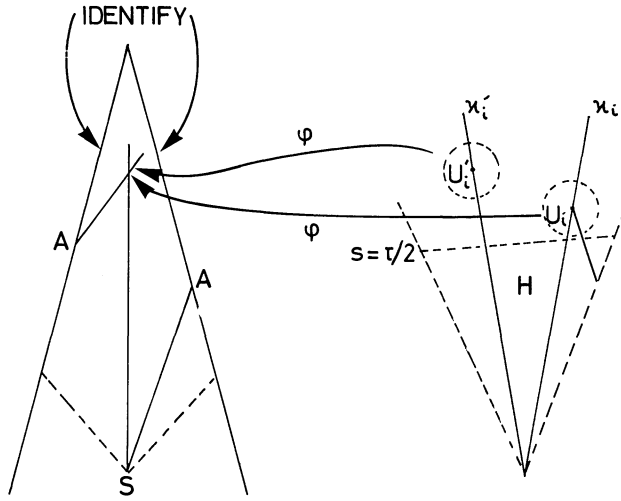


Fig. 2. The intersection of the geodesics  $\kappa_i$  in a space-time resembling Taub-NUT space. See Lemma 5

By taking a suitable subsequence we can impose in addition

$$\sum_0^\infty (\tau - s_i) < \infty \tag{11a}$$

and

$$\{\phi(s_i, z_i), \phi(s'_i, z'_i)\} \subset I^+(\phi(s_{i-1}, z_{i-1})) \cap I^+(\phi(s'_{i-1}, z'_{i-1})). \tag{11b}$$

Let  $L_i = (\mu(s'_i, z'_i))^{-1} \mu(s_i, z_i) \in \mathcal{L}_+^\dagger$  and consider the set  $A$  of accumulation points of  $(L_i)$  in the norm topology  $(A = \bigcap_{i=0}^\infty \overline{\bigcup_{k=i}^\infty L_k})$ . We shall consider two cases:

(a)  $A \subset G \simeq SO(3)$ ; (b) the rest.

*Case (a).* By hypothesis there is an  $L \in A$  and a past directed timelike eigenvector  $\xi$  such that  $L(\xi) = \xi$ . By again choosing a suitable subsequence we can suppose that  $\|L_i - L\| = \varepsilon_i \rightarrow 0$ , and so

$$\|L_i \xi - \xi\| \rightarrow 0. \tag{12}$$

Since  $\phi$  is a local homeomorphism,  $H$  can be given the structure of an  $sc - C^2$ -space-time with metric  $g^\dagger = \phi^*g$ . Let  $U_i, U'_i$  be neighbourhoods of  $(s_i, z_i), (s'_i, z'_i)$  in which  $\phi$  is a homeomorphism and set  $\phi_i = \phi|U_i, \phi'_i = \phi|U'_i$ . Then let  $\kappa_i, \kappa'_i$  be the geodesics in  $H$  defined by

$$\mu_i^{-1} \dot{\kappa}_i(0) = \mu(s_i, z_i)^{-1} \phi'_{i*} \dot{\kappa}'_i(0) = \xi$$

where  $\mu_i = \phi_{i*}^{-1} \mu(s_i, z_i)$ . Then (12) implies that  $\mu_i^{-1} \dot{\kappa}'_i(0) \rightarrow \xi$ . Thus, for any fixed  $s$ ,

$\kappa_i(s)$  and  $\kappa'_i(s)$  will tend to a common limit. But, by the definition of  $\kappa_i$  and  $\kappa'_i$ ,  $\phi\kappa_i(s) = \phi\kappa'_i(s)$ , which is then incompatible with  $\phi$  being a local homeomorphism.

Case (b). Since  $SO(3)$  is the maximal compact subgroup of  $\mathcal{L}^\dagger_\perp$ , the closure of the group generated by  $A$  is in this case not compact. Thus we can find a sequence  $\{L_j^*\}$ , where  $L_j^* = L_i$  or  $(L_i)^{-1}$  with the integers  $i_j$  tending monotonically to infinity as  $j$  goes to infinity, such that  $\left\| \prod_{k=0}^j L_k^* \right\| \rightarrow \infty$ .

Now write the pair  $((s_i, z_i), (s'_i, z'_i))$  as  $(\zeta_j, \zeta'_j)$  or  $(\zeta'_j, \zeta_j)$ , according to whether  $L_j^* = L_i$  or  $L_i^{-1}$ , respectively, and consider the union of line segments  $K = \overline{\zeta_1 \zeta'_1} \cup \overline{\zeta'_1 \zeta_2} \cup \overline{\zeta_2 \zeta'_2} \cup \overline{\zeta'_2 \zeta_3} \cup \dots$ . The  $\phi$ -image of  $K$  will be a continuous curve which, from (11), will have finite length when measured in the  $\mu$ -tetrad field, and thus will have finite generalised affine parameter length provided that  $\sum_{j=0}^\infty \left\| \prod_{k=0}^j L_k^* \right\| (\tau - s_i) < \infty$ ;

this condition also ensures that  $\phi(K)$  terminates at  $p$ . We can always choose a subsequence such that this condition is satisfied, unless  $\|L_k^*\| \rightarrow \infty$ . But that is impossible: a sequence of loops on which the Lorentz transformation became infinite could, using the product decomposition of a globally hyperbolic space, be dragged down onto such a sequence on a fixed space-like slice of  $I^-(|\kappa|)$  which, by virtue of its relation to the corresponding slice in the space analogous to  $H$  formed on all future timelike geodesics from 0, has compact closure, giving a contradiction.

We can now see that the existence of  $\phi(K)$ , a curve to  $p$  on which the parallelly propagated tetrad undergoes an unbounded Lorentz transformation relative to  $\mu$ , requires that the space-time be  $D$ -specialised at  $p$ , in contradiction with the assumptions of the theorem. For any curve terminating at  $p$  must then have its parallelly propagated tetrad unboundedly Lorentz transformed relative either to  $\gamma$  or to  $\phi(k)$  (or both), in the sense that it can be modified by the addition of connecting curves so as to have points in common with, say,  $\gamma$  at which the p.p. tetrads on the two curves can be compared. A subsequence  $\{L_i\}$  of this sequence of Lorentz transformations will then tend to a boost which becomes infinite ( $L_i = R_i L(v_i) R'_i$ ,  $R_i$  and  $R'_i$  being rotations that converge to limits while  $v_i \rightarrow \infty$ ). From the Scholium to Definition 2 this then implies  $D$ -specialisation, giving the contradiction.  $\square$

We can now perform the extension. It will be convenient to first extend  $g^\dagger$  to an enlargement of a part of  $H$ , finally joining this extension back onto  $M$ . Set  $K = H_W \times \mathbb{R}^+$ .

**Lemma 6.** *There is a submanifold  $K'$  of  $K$  such that  $\overline{\phi^{-1}(p)} \in K'$ ,  $g^\dagger|_{K' \cap H}$  can be extended to  $K'$  and  $\overline{K'}$  (closure in  $K$ ) is a compact manifold with boundary.*

*Proof.* The extension is carried out using "radar coordinates". If  $P^\alpha (\alpha = 0, \dots, 3)$  are four fixed points and  $U^\alpha$  is a normal neighbourhood for  $P^\alpha$  we define  $w^\alpha(Q)$ , for  $Q \in \bigcap_\beta U^\beta$ , to be the (positive) geodesic distance between  $P^\alpha$  and  $Q$ . Then  $\overset{\alpha}{X}$ , the unit tangent vector field to the geodesics from  $P^\alpha$ , is given by  $g(\overset{\alpha}{X}) = dw^\alpha$

and so the metric components in these coordinates are

$$g^{\alpha\beta} = g(\overset{\alpha}{X}, \overset{\beta}{X}).$$

Since the geodesics are  $C^{2-}$  (see the scholium to Definition 1) it follows immediately that the  $g^{\alpha\beta}$  are  $C^{1-}$ . But one can show further that when the  $P^\alpha$  are chosen so that the  $w^\alpha$  form a good coordinate system in a patch extending up to  $\dot{H}$ , then the  $g^{\alpha\beta}$  are Lipschitz at points of  $\dot{H}$ . This is because the Lipschitz factor can be expressed in terms of the Riemann tensor on  $H$  which, by Lemma 4, is bounded: the proof involves extending the results of § 2 to 3-parameter families of geodesics and, being straight-forward but tedious, is omitted. This in turn means that the  $g^{\alpha\beta}$  tend to limits on  $\dot{H}$ .

Change coordinates in  $w$ -space, choosing functions  $w^{*\alpha}$ ,  $C^\infty$  with respect to the  $w^\alpha$ , such that  $\partial/\partial\omega^{*0}$  is timelike and the  $\partial/\partial\omega^{*i}$  are spacelike, in the range of  $w^\alpha$  mapping into  $H$ . Then extend the  $g^{\alpha\beta}$  as functions of the  $w^*$  by setting, for  $w^\alpha$  outside this range,  $g^{\alpha\beta}(w^{*\gamma}) = g^{\alpha\beta}(c(w^{*i}), w^{*i})$ , where  $c$  is defined so that  $(c(w^{*i}), w^{*i}) \in \dot{H}$ . The differentiability is then automatically  $sc - C^{2-}$ . The space  $K'$  is defined by mapping the domain where the  $w^\alpha$  are well-behaved into  $K$  by continuing the  $\iota$ -family of geodesics, and the conditions of the lemma can then immediately be realised.  $\square$

### Conclusion of Theorem

From Lemma 6 we derive a space-time  $K'$  and an isometry  $\phi' : H \cap K' \rightarrow M$  with the end of  $\kappa$  in the range of  $\phi'$ . This isometry (or rather, its inverse) is a local extension in the sense of [5]. But because we have constructed explicit coordinates in  $H$ , a full extension follows immediately. For the isometry  $\phi$  defines an identification between  $\dot{K}' \cap H$  and  $\widehat{\phi(K')}$ . This can be used to fix together the manifolds-with-boundaries  $K' \cup (\dot{K}' \cap H)$  and  $M \setminus \phi(K')$ , the identifications of boundaries being clearly consistent with the orientations. The identification space  $M' = (M \setminus \phi(K') \cup K') / (\phi|_{\dot{K}' \cap H})$  is then the required extension.  $\square$

*Acknowledgements.* I must record my indebtedness to Drs. Ellis and Schmidt, on whose ideas this paper is largely built, to Dr. Hawking who first brought the significance of Example I to my attention, and to Dr. King for discussions out of which arose Definition 2 and Scholium.

### References

1. Hawking, S. W., Ellis, G. F. R.: The large scale structure of space-time. Cambridge: University Press 1973
2. Schmidt, B. G.: J. General Relativity and Gravitation, **1**, 269—280 (1971)
3. Schmidt, B. G., Hajicek, P.: Commun. math. Phys. **23**, 285—295 (1971)
4. Ellis, G. F. R., King, A.: Was the big bang a whimper? Commun. math. Phys. **38**, 119—156 (1974)
5. Clarke, C. J. S.: Commun. math. Phys. **32**, 205—214 (1973)
6. Clarke, C. J. S.: Proc. Roy. Soc. Lond. A, **314**, 417—428 (1970)
7. Geroch, R.: J. Math. Phys. **9**, 450—465 (1968)

8. Clarke, C. J. S.: The Classification of singularities, to appear in *J. General Relativity and Gravitation* (1974)
9. Penrose, R.: *Ann. Phys.*, **10**, 171—201 (1960)

Communicated by J. Ehlers

C. J. S. Clarke  
University of Cambridge  
Department of Applied Mathematics and  
Theoretical Physics  
Silver Street  
Cambridge CB3 9EW, U.K.