Ann. Henri Poincaré 1 (2000) 543 – 567

1424-0637/00/030543-25 $ 1.50+0.20/0

**Annales Henri Poincaré**

# Maximum Principles for Null Hypersurfaces and Null Splitting Theorems

Gregory J. Galloway

## I Introduction

The geometric maximum principle for smooth (spacelike) hypersurfaces, which is a consequence of Alexandrov's [1] strong maximum for second order quasi-linear elliptic operators, is a basic tool in Riemannian and Lorentzian geometry. In [2], extending earlier work of Eschenburg [7], a version of the geometric maximum principle in the Lorentzian setting was obtained for rough ($C^0$) spacelike hypersurfaces which obey mean curvature inequalities in the sense of support hypersurfaces. In the present paper we establish an analogous result for *null* hypersurfaces (Theorem III.4) and consider some applications. For the applications, it is important to have a version of the maximum principle for null hypersurfaces which does not require smoothness. The reason for this, which is described in more detail in Section 3, is that the null hypersurfaces which arise most naturally in spacetime geometry and general relativity, such as black hole event horizons, are in general $C^0$ but not $C^1$. To establish our basic approach, we first prove a maximum principle for smooth null hypersurfaces (Theorem II.1), and then proceed to the $C^0$ case. The general $C^0$ version is then applied to study some rigidity properties of spacetimes which contain null lines (inextendible globally maximal null geodesics). The standard Lorentzian splitting theorem, which is the Lorentzian analogue of the Cheeger-Gromoll splitting theorem of Riemannian geometry, describes the rigidity of spacetimes which contain timelike lines (inextendible globally maximal timelike geodesics); see [3, Chapter 14] for a nice exposition. Here we show how the maximum principle for rough null hypersurfaces can be used to obtain a general "splitting theorem" for spacetimes with null lines (Theorem IV.1). We then consider an application of this null splitting theorem to asymptotically flat spacetimes. We prove that a nonsingular asymptotically flat (in the sense of Penrose [17]) vacuum (i.e., Ricci flat) spacetime which contains a null line is isometric to Minkowski space (Theorem IV.3).

In Section 2 we review the relevant aspects of the geometry of null hypersurfaces and present the maximum principle for smooth null hypersurfaces. In Section 3 we present the maximum principle for $C^0$ null hypersurfaces. In Section 4 we consider the aforementioned applications. For basic notions used below from Lorentzian geometry and causal theory, we refer the reader to the excellent references, [3], [13], [16] and [18].

# II The maximum principle for smooth null hypersurfaces

## II.1 The geometry of null hypersurfaces

Here we review some aspects of the geometry of null hypersurfaces. For further details, see e.g. [14] which is written from a similar point of view.

Let $M$ be a spacetime, i.e., a smooth time-oriented Lorentzian manifold, of dimension $n \geq 3$. We denote the Lorentzian metric on $M$ by $g$ or $\langle\,,\rangle$. A *(smooth) null hypersurface* in $M$ is a smooth co-dimension one embedded submanifold $S$ of $M$ such that the pullback of the metric $g$ to $S$ is degenerate. Because of the Lorentz signature of $g$, the null space of the pullback is one dimensional at each point of $S$. Hence, every null hypersurface $S$ admits a smooth nonvanishing future directed null vector field $K \in \Gamma TS$ such that the normal space of $K$ at $p \in S$ coincides with the tangent space of $S$ at $p$, i.e., $K_p^{\perp} = T_pS$ for all $p \in S$. It follows, in particular, that tangent vectors to $S$ not parallel to $K$ are spacelike. It is well-known that the integral curves of $K$, when suitably parameterized, are null geodesics. These integral curves are called the *null geodesic generators* of $S$. We note that the vector field $K$ is unique up to a positive (pointwise) scale factor.

Since $K$ is orthogonal to $S$ we can introduce the null Weingarten map and null second fundamental form of $S$ with respect $K$ in a manner roughly analogous to what is done for spacelike hypersurfaces or hypersurfaces in a Riemannian manifold.

We introduce the following equivalence relation on tangent vectors : For $X, X' \in T_pS$, $X' = X \bmod K$ if and only if $X' - X = \lambda K$ for some $\lambda \in \mathbb{R}$. Let $\overline{X}$ denote the equivalence class of $X$. Simple computations show that if $X' = X \bmod K$ and $Y' = Y \bmod K$ then $\langle X', Y'\rangle = \langle X, Y\rangle$ and $\langle \nabla_{X'}K, Y'\rangle = \langle \nabla_X K, Y\rangle$, where $\nabla$ is the Levi-Civita connection of $M$. Hence, for various quantities of interest, components along $K$ are not of interest. For this reason one works with the tangent space of $S$ modded out by $K$, i.e., let $T_pS/K = \{\overline{X} : X \in T_pS\}$ and $TS/K = \cup_{p\in S} T_pS/K$. $TS/K$ is a rank $n-2$ vector bundle over $S$. This vector bundle does not depend on the particular choice of null vector field $K$. There is a natural positive definite metric $h$ in $TS/K$ induced from $\langle\,,\rangle$: For each $p \in S$, define $h : T_pS/K \times T_pS/K \to \mathbb{R}$ by $h(\overline{X}, \overline{Y}) = \langle X, Y\rangle$. From remarks above, $h$ is well-defined.

The *null Weingarten map* $b = b_K$ of $S$ with respect to $K$ is, for each point $p \in S$, a linear map $b : T_pS/K \to T_pS/K$ defined by $b(\overline{X}) = \overline{\nabla_X K}$. It is easily verified that $b$ is well-defined. Note if $\widetilde{K} = fK$, $f \in C^\infty(S)$, is any other future directed null vector field tangent to $S$, then $\nabla_X \widetilde{K} = f\nabla_X K \bmod K$. It follows that the Weingarten map $b$ of $S$ is unique up to positive scale factor and that $b$ at a given point $p \in S$ depends only on the value of $K$ at $p$.

A standard computation shows, $h(b(\overline{X}), \overline{Y}) = \langle \nabla_X K, Y\rangle = \langle X, \nabla_Y K\rangle = h(\overline{X}, b(\overline{Y}))$. Hence $b$ is self-adjoint with respect to $h$. The *null second fundamental form* $B = B_K$ of $S$ with respect to $K$ is the bilinear form associated to $b$ via $h$: For each $p \in S$, $B : T_pS/K \times T_pS/K \to \mathbb{R}$ is defined by $B(\overline{X}, \overline{Y}) = h(b(\overline{X}), \overline{Y}) =$

$\langle \nabla_X K, Y\rangle$. Since $b$ is self-adjoint, $B$ is symmetric. We say that $S$ is *totally geodesic* iff $B \equiv 0$. This has the usual geometric consequence: A geodesic in $M$ starting tangent to a totally geodesic null hypersurface $S$ remains in $S$. Null hyperplanes in Minkowski space are totally geodesic, as is the event horizon in Shwarzschild spacetime.

The *null mean curvature* of $S$ with respect to $K$ is the smooth scalar field $\theta \in C^\infty(S)$ defined by $\theta = \mathrm{tr}\, b$. Let $e_1, e_2, ...e_{n-2}$ be $n-2$ orthonormal spacelike vectors (with respect to $\langle,\rangle$) tangent to $S$ at $p$. Then $\{\overline{e}_1, \overline{e}_2, ...\overline{e}_{n-2}\}$ is an orthonormal basis (with respect to $h$) of $T_pS/K$. Hence at $p$,

$$\begin{aligned}\theta &= \mathrm{tr}\, b = \sum_{i=1}^{n-2} h(b(\overline{e}_i), \overline{e}_i) \\ &= \sum_{i=1}^{n-2} \langle \nabla_{e_i} K, e_i\rangle. \end{aligned} \tag{II.1}$$

Let $\Sigma$ be the *properly transverse* intersection of a hypersurface $P$ in $M$ with $S$. By *properly transverse* we mean that $K$ is not tangent to $P$ at any point of $\Sigma$. Then $\Sigma$ is a smooth $(n-2)$-dimensional spacelike submanifold of $M$ contained in $S$ which meets $K$ orthogonally. From Equation II.1, $\theta|_\Sigma = \mathrm{div}_\Sigma K$, and hence the null mean curvature gives a measure of the divergence of the null generators of $S$. Note that if $\widetilde{K} = fK$ then $\widetilde{\theta} = f\theta$. Thus the null mean curvature inequalities $\theta \geq 0$, $\theta \leq 0$, are invariant under positive rescaling of $K$. In Minkowski space, a future null cone $S = \partial I^+(p) \setminus \{p\}$ (resp., past null cone $S = \partial I^-(p) \setminus \{p\}$) has positive null mean curvature, $\theta > 0$ (resp., negative null mean curvature, $\theta < 0$).

The null second fundamental form of a null hypersurface obeys a well-defined comparison theory roughly similar to the comparison theory satisfied by the second fundamental forms of a family of parallel spacelike hypersurfaces (cf., Eschenburg [6], which we follow in spirit).

Let $\eta : (a, b) \to M$, $s \to \eta(s)$, be a future directed affinely parameterized null geodesic generator of $S$. For each $s \in (a, b)$, let

$$b(s) = b_{\eta'(s)} : T_{\eta(s)}S/\eta'(s) \to T_{\eta(s)}S/\eta'(s)$$

be the Weingarten map based at $\eta(s)$ with respect to the null vector $K = \eta'(s)$. The one parameter family of Weingarten maps $s \to b(s)$, obeys the following Ricatti equation,

$$b' + b^2 + R = 0. \tag{II.2}$$

Here $'$ denotes covariant differentiation in the direction $\eta'(s)$; if $X = X(s)$ is a vector field along $\eta$ tangent to $S$, we define,

$$b'(\overline{X}) = b(\overline{X})' - b(\overline{X'}). \tag{II.3}$$

$R : T_{\eta(s)}S/\eta'(s) \to T_{\eta(s)}S/\eta'(s)$ is the curvature endomorphism defined by $R(\overline{X}) = \overline{R(X, \eta'(s))\eta'(s)}$, where $(X, Y, Z) \to R(X, Y)Z$ is the Riemann curvature tensor of $M$, $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$.

We indicate the proof of Equation II.2. Fix a point $p = \eta(s_0)$, $s_0 \in (a, b)$, on $\eta$. On a neighborhood $U$ of $p$ in $S$ we can scale the null vector field $K$ so that $K$ is a geodesic vector field, $\nabla_K K = 0$, and so that $K$, restricted to $\eta$, is the velocity vector field to $\eta$, i.e., for each $s$ near $s_0$, $K_{\eta(s)} = \eta'(s)$. Let $X \in T_pM$. Shrinking $U$ if necessary, we can extend $X$ to a smooth vector field on $U$ so that $[X, K] = \nabla_X K - \nabla_K X = 0$. Then, $R(X, K)K = \nabla_X \nabla_K K - \nabla_K \nabla_X K - \nabla_{[X,K]} K = -\nabla_K \nabla_K X$. Hence along $\eta$ we have, $X'' = -R(X, \eta')\eta'$ (which implies that $X$, restricted to $\eta$, is a Jacobi field along $\eta$). Thus, from Equation II.3, at the point $p$ we have,

$$\begin{aligned} b'(\overline{X}) &= \overline{\nabla_X K}\,' - b(\overline{\nabla_K X}) = \overline{\nabla_K X}\,' - b(\overline{\nabla_X K}) \\ &= \overline{X''} - b(b(\overline{X})) = -\overline{R(X, \eta')\eta'} - b^2(\overline{X}) \\ &= -R(\overline{X}) - b^2(\overline{X}), \end{aligned}$$

which establishes Equation II.2.

By taking the trace of II.2 we obtain the following formula for the derivative of the null mean curvature $\theta = \theta(s)$ along $\eta$,

$$\theta' = -\mathrm{Ric}(\eta', \eta') - \sigma^2 - \frac{1}{n-2}\theta^2, \tag{II.4}$$

where $\sigma$, the shear scalar, is the trace of the square of the trace free part of $b$. Equation II.4 is the well-known Raychaudhuri equation (for an irrotational null geodesic congruence) of relativity. This equation shows how the Ricci curvature of spacetime influences the null mean curvature of a null hypersurface.

## II.2 The maximum principle for smooth null hypersurfaces

The aim here is to prove the geometric maximum principle for smooth null hypersurfaces. Because of its natural invariance, we restrict attention to the zero mean curvature case. In the statement we make use of the following intuitive terminology. Let $S_1$ and $S_2$ be null hypersurfaces that meet at a point $p$. We say that $S_2$ *lies to the future (resp., past) side of* $S_1$ *near* $p$ provided for some neighborhood $U$ of $p$ in $M$ in which $S_1$ is closed and achronal, $S_2 \cap U \subset J^+(S_1 \cap U, U)$ (resp., $S_2 \cap U \subset J^-(S_1 \cap U, U)$).

**Theorem II.1** *Let $S_1$ and $S_2$ be smooth null hypersurfaces in a spacetime $M$. Suppose,*

*(1) $S_1$ and $S_2$ meet at $p \in M$ and $S_2$ lies to the future side of $S_1$ near $p$, and*

*(2) the null mean curvature scalars $\theta_1$ of $S_1$, and $\theta_2$ of $S_2$, satisfy, $\theta_1 \le 0 \le \theta_2$.*

*Then $S_1$ and $S_2$ coincide near $p$ and this common null hypersurface has null mean curvature $\theta = 0$.*

The proof is an application of Alexandrov's [1] strong maximum principle for second order quasi-linear elliptic operators. It will be convenient to state the precise form of this result needed here.

For connected open sets $\Omega \subset \mathbb{R}^n$ and $U \subset \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, we say $u \in C^2(\Omega)$ is *$U$-admissible* provided $(x, u(x), \partial u(x)) \in U$ for all $x = (x^1, x^2, ..., x^n) \in \Omega$, where $\partial u = (\partial_1 u, \partial_2 u, ..., \partial_n u)$, and $\partial_i u = \frac{\partial u}{\partial x^i}$.

Let $Q = Q(u)$ be a second order quasi-linear operator, i.e., for $U$-admissible $u \in C^2(\Omega)$, consider

$$Q(u) = \sum_{i,j=1}^{n} a^{ij}(x, u, \partial u)\partial_{ij} u + b(x, u, \partial u), \tag{II.5}$$

where $a^{ij}, b \in C^1(U)$, $a^{ij} = a^{ji}$, and $\partial_{ij} = \frac{\partial^2}{\partial u^j \partial u^i}$. The operator $Q$ is *elliptic* provided for each $(x, r, p) \in U$, and for all $\xi = (\xi^1, ...\xi^n) \in \mathbb{R}^n$, $\xi \neq 0$,

$$\sum_{i,j=1}^{n} a^{ij}(x, r, p)\xi^i \xi^j > 0.$$

We now state the form of the strong maximum principle for second order quasi-linear elliptic operators most suitable for our purposes.

**Theorem II.2** *Let $Q = Q(u)$ be a second order quasi-linear elliptic operator as described above. Suppose the $U$-admissible functions $u, v \in C^2(\Omega)$ satisfy,*

*(1) $u \leq v$ on $\Omega$ and $u(x_0) = v(x_0)$ for some $x_0 \in \Omega$, and*

*(2) $Q(v) \leq Q(u)$ on $\Omega$.*

*Then $u = v$ on $\Omega$.*

The idea of the proof of Theorem II.1 is to intersect, in a properly transverse manner, the null hypersurfaces $S_1$ and $S_2$ with a *timelike* (i.e., Lorentzian in the induced metric) hypersurface through the point $p$, and to show that the spacelike intersections agree. Analytically, intersecting the null hypersurfaces in this manner reduces the problem to a nondegenerate elliptic one. In order to apply Theorem II.2 we need a suitable analytic expression for the null mean curvature, which we now derive.

Let $p$ be a point in a spacetime $M$, and let $P$ be a timelike hypersurface passing through $p$. Let $V$ be a connected spacelike hypersurface in $P$ (and hence a co-dimension two spacelike submanifold of $M$) passing through $p$. Via the normal exponential map along $V$ in $P$, we can assume, by shrinking $P$ if necessary, that $P$ can be expressed as,

$$P = (-a, a) \times V, \tag{II.6}$$

and that the induce metric on $P$ takes the form,

$$ds^2 = -dt^2 + \sum_{i=1}^{n-2} g_{ij}(t,x)dx^i dx^j, \tag{II.7}$$

where $x = (x_1, ..., x_{n-2})$ are coordinates in $V$ centered on $p$.

Let $S$ be a null hypersurface which meets $P$ properly transversely in a spacelike hypersurface $\Sigma$ in $P$. By adjusting the size of $P$ and $S$ if necessary, we may assume that $\Sigma$ can be expressed as a graph over $V$. Thus, there exists $u \in C^\infty(V)$ such that $\Sigma = \operatorname{graph} u = \{(u(x), x) \in P : x \in V\}$. Let $H(u)$ denote the mean curvature of $\Sigma = \operatorname{graph} u$. (By our sign conventions the mean curvature of $\Sigma$ is + the divergence of the future pointing normal along $\Sigma$.) To describe $H(u)$ we introduce the following notation. Let $h$ be the Riemannian metric on $V$ whose components are given by $h_{ij}(x) = g_{ij}(u(x), x)$, and let $h^{ij}$ be the $i,j$th entry of the inverse matrix $[h_{ij}]^{-1}$. Let $\nabla u$ denote the gradient of $u$. In terms of coordinates, $\nabla u = \sum_j u^j \partial_j$, where $u^j = \sum_i h^{ij}\partial_i u$. Finally, introduce the quantity,

$$\nu = \frac{1}{\sqrt{1 - |\nabla u|^2}}.$$

The positivity of $\nu$ is equivalent to $\Sigma = \operatorname{graph} u$ being spacelike. With respect to these quantities, we have (cf., [2]),

$$H(u) = \sum_{i,j=1}^{n-2} a^{ij}(x,u,\partial u)\partial_{ij}u + b(x,u,\partial u), \tag{II.8}$$

where $a^{ij} = \nu h^{ij} + \nu^3 u^i u^j$ and $b$ is a polynomial expression in $\partial_i u$, $h_{ij}$, $h^{ij}$, $\partial_t g_{ij}(u(x),x)$ and $\nu$. From the form of $a^{ij}$, it is clear that $H = H(u)$ is a second order quasi-linear elliptic operator.

Let $K$ be a future directed null vector field on $S$. Since $K$ is orthogonal to $\Sigma$, by rescaling we may assume that along $\Sigma$ , $K = Z + N$, where $Z$ is the future directed unit normal vector field to $\Sigma$ in $P$ and $N$ is one of the two unit spacelike normal vector fields to $P$ in $M$. Let $\theta$ be the null mean curvature of $S$ with respect to $K$. We obtain an expression for $\theta$ along $\Sigma$. Let $B_P$ denote the second fundamental form of $P$ with respect to $N$, and let $B_\Sigma$ denote the second fundamental form of $\Sigma$ in $P$ with respect to $Z$. Then for $q \in \Sigma$ and vectors $X, Y \in T_q\Sigma$, $B_P(X,Y) = \langle \nabla_X N, Y\rangle$, and $B_\Sigma(X,Y) = \langle \overline{\nabla}_X Z, Y\rangle = \langle \nabla_X Z, Y\rangle$, where $\overline{\nabla}$ is the induced Levi-Civita connection on $P$.

Now let $\{e_1, e_2, ..., e_{n-2}\}$ be an orthonormal basis for $T_q\Sigma$. Then the value of $\theta$ at $q$ is given by,

$$\begin{aligned}\theta &= \sum_{i=1}^{n-2}\langle\nabla_{e_i}K, e_i\rangle = \sum_{i=1}^{n-2}\langle\nabla_{e_i}Z, e_i\rangle + \sum_{i=1}^{n-2}\langle\nabla_{e_i}N, e_i\rangle \\ &= \sum_{i=1}^{n-2}B_\Sigma(e_i, e_i) + \sum_{i=1}^{n-2}B_P(e_i, e_i) \\ &= H_\Sigma + B_P(Z, Z) + H_P, \end{aligned} \tag{II.9}$$

where $H_\Sigma$ is the mean curvature of $\Sigma$ and $H_P$ is the mean curvature of $P$. In the notation introduced above,

$$\begin{aligned} Z &= \nu(\partial_0 + \nabla u) \\ &= \sum_{i=0}^{n-2}\nu u^i\partial_i, \end{aligned} \tag{II.10}$$

where $\partial_0 = \frac{\partial}{\partial t}$, $u^0 = 1$, and as above, $u^i = \sum_{j=1}^{n-2} h^{ij}\partial_j u$, $i = 1, ..., n-2$. Hence,

$$\begin{aligned} B_P(Z, Z) &= B_P(\sum_{i=0}^{n-2}\nu u^i\partial_i, \sum_{i=0}^{n-2}\nu u^i\partial_i) \\ &= \sum_{i,j=1}^{n-2}\nu^2\beta_{ij}(u)u^iu^j, \end{aligned} \tag{II.11}$$

where for $x \in V$, $\beta_{ij}(u)(x) = B_P(\partial_i, \partial_j)|_{(u(x),x)}$.

Now let $\theta(u)$ denote the null mean curvature of $S$ along $\Sigma = \operatorname{graph} u$. Equations II.9 and II.11 give,

$$\theta(u) = H(u) + \sum_{i,j=1}^{n-2}\nu^2\beta_{ij}(u)u^iu^j + \alpha(u),$$

where $\alpha(u)$ is the function on $V$ defined by $\alpha(u)(x) = H_P(u(x), x)$. Thus, by II.8, we finally arrive at,

$$\theta(u) = \sum_{i,j=1}^{n-2} a^{ij}(x, u, \partial u)\partial_{ij}u + b_1(x, u, \partial u) \tag{II.12}$$

where,

$$b_1(x, u, \partial u) = b(x, u, \partial u) + \sum_{i,j=1}^{n-2}\nu^2\beta_{ij}(u)u^iu^j + \alpha(u), \tag{II.13}$$

and where $a^{ij}$ and $b$ are as in Equation II.8. In particular, $\theta = \theta(u)$ is a second order quasi-linear elliptic operator with the same symbol as the mean curvature operator for spacelike hypersurfaces in $P$.

*Proof of Theorem 2.1:* Let $P$ be a timelike hypersurface passing through $p$, as described by equations II.6 and II.7. $S_1$ and $S_2$ have a common null tangent at $p$. Choose $P$ so that it is transverse to this tangent. Then, by choosing $P$ small enough the intersections $\Sigma_1 = S_1 \cap P$ and $\Sigma_2 = S_2 \cap P$ will be properly transverse, and hence $\Sigma_1$ and $\Sigma_2$ will be spacelike hypersurfaces in $P$. Let $K_i$, $i = 1, 2$, be the null vector field on $S_i$ with respect to which the null mean curvature function $\theta_i$ is defined. Let $N$ be the unit normal to $P$ pointing to the same side of $P$ as $K_1|_{\Sigma_1}$ and $K_2|_{\Sigma_2}$. By rescaling we can assume $K_i|_{\Sigma_i} = Z_i + N|_{\Sigma_i}$, where $Z_i$ is the future directed unit normal to $\Sigma_i$ in $P$.

Let $u_i = u_i(x)$, $i = 1, 2$, be the smooth function on $V$ such that $\Sigma_i = \operatorname{graph} u_i$. From the hypotheses of Theorem II.1 we know,

(1) $u_1 \le u_2$ on $V$ and $u_1(p) = u_2(p)$, and

(2) $\theta(u_2) \le \theta(u_1)$ on $V$.

By Theorem II.2, we conclude that $u_1 = u_2$ on $V$, i.e. $\Sigma_1 = \Sigma_2$. Now, $S_i$, $i = 1, 2$, is obtained locally by exponentiating out from $\Sigma_i$ in the orthogonal direction $K_i|_{\Sigma_i} = Z_i + N|_{\Sigma_i}$. It follows that $S_1$ and $S_2$ agree near $p$, i.e., there is a spacetime neighborhood $\mathcal{O}$ of $p$ such that $S_1 \cap \mathcal{O} = S_2 \cap \mathcal{O} = S$, and $S$ has null mean curvature $\theta = 0$. □

# III The maximum principle for $C^0$ null hypersurfaces

## III.1 $C^0$ null hypersurfaces

The usefulness of the maximum principle for smooth null hypersurfaces obtained in the previous section is limited by the fact that the most interesting null hypersurfaces arising in general relativity, such as black hole event horizons and Cauchy horizons, are rough, i.e., are $C^0$ but in general not $C^1$. The aim of this section is to present a maximum principle for rough null hypersurfaces, similar in spirit to the maximum principle for rough spacelike hypersurfaces obtained in [2].

Horizons and other null hypersurfaces commonly occurring in relativity arise essentially as the null portions of *achronal boundaries* which are sets of the form $\partial I^{\pm}(A)$, $A \subset M$. Achronal boundaries are always $C^0$ hypersurfaces, but simple examples illustrate that they (and their null portions) may fail to be differentiable at certain points. Consider, for example, the set $S = \partial I^{-}(A) \setminus A$, where $A$ consists of two disjoint closed disks in the $t = 0$ slice of Minkowski 3-space. This surface, which represents the merger of two truncated cones, has a "crease", i.e., a curve of nondifferentiable points (corresponding to the intersection of the two cones) but which otherwise is a smooth null hypersurface.

An important feature of the null portion of an achronal boundary is that it is ruled by null geodesics which are either past or future inextendible within the hypersurface. This is illustrated in the example above. $S$ is ruled by null geodesics which are future inextendible in $S$, but which are in general not past inextendible in $S$. Null geodesics in $S$ that meet the crease when extended toward the past leave $S$ when extended further, and hence have past end points on $S$.

We now formulate a general definition of $C^0$ null hypersurface which captures the essential features of these examples.

A set $A \subset M$ is said to be achronal if no two points of $A$ can be joined by a timelike curve. $A \subset M$ is locally achronal if for each $p \in A$ there is a neighborhood $U$ of $p$ such that $A \cap U$ is achronal in $U$.

A $C^0$ *nontimelike hypersurface* in $M$ is a topological hypersurface $S$ in $M$ which is locally achronal. We remark that for each $p \in S$, there exist arbitrarily small neighborhoods $U$ of $p$ such that $S \cap U$ is closed and achronal in $U$, and for each $q \in U \setminus S$, either $q \in I^+(S \cap U, U)$ or $q \in I^-(S \cap U, U)$.

**Definition III.1** *A* $C^0$ future null hypersurface *in* $M$ *is a nontimelike hypersurface* $S$ *in* $M$ *such that for each* $p \in S$ *and any neighborhood* $U$ *of* $p$ *in which* $S$ *is achronal, there exists a point* $q \in S$, $q \neq p$, *such that* $q \in J^+(p, U)$.

Since $q \in J^+(p, U) \setminus I^+(p, U)$, there is a null geodesic segment $\eta$ from $p$ to $q$. The segment $\eta$ must be contained in $S$, for otherwise at some point $\eta$ would enter either $I^+(S, U)$ or $I^-(S, U)$, which would contradict the achronality of $S$ in $U$. The geodesic $\eta$ can be extended further to the future in $S$: Choose $r$ in $S \cap J^+(q, U)$, $r \neq q$. The null geodesic from $q$ to $r$ in $S$ must smoothly extend the one from $p$ to $q$, otherwise there would be an achronality violation of $S$ in $U$. Thus, for each point $p \in S$ there is a future directed null geodesic in $S$ starting at, or passing through, $p$ which is future inextendible in $S$, i.e., which does not have a future end point in $S$. These null geodesics are called the *null geodesic generators* of $S$. They may or may not have past end points in $S$. Summarizing, a $C^0$ future null hypersurface is a locally achronal topological hypersurface $S$ of $M$ which is ruled by future inextendible null geodesics. A $C^0$ past null hypersurface is defined in a time-dual manner.

Let $S$ be a $C^0$ future null hypersurface. Adopting the terminology introduced in [4] for event horizons, a *semi-tangent* of $S$ is a future directed null vector $K$ which is tangent to a null generator of $S$. We do not want to distinguish between semi-tangents based at the same point and pointing in the same null direction, so we assume the semi-tangents of $S$ have been uniformly normalized in some manner, e.g., by requiring each semi-tangent to have unit length with respect to some auxilliary Riemannian metric on $M$. Then note that the local achronality of $S$ implies that at each interior point (non-past end point) of a null generator of $S$ there is a unique semi-tangent at that point. Techniques from [4] can be adapted to prove the following.

**Lemma III.1** *Let* $S$ *be a* $C^0$ *future null hypersurface in a spacetime* $M$.

1. *If $p_n \to p$ in $S$ and $X_n \to X$ in $TM$, where, for each $n$, $X_n$ is a semi-tangent of $S$ at $p_n$ then $X$ is a semi-tangent of $S$.*

2. *Suppose $p$ is an interior point of a null generator of $S$, and let $X$ be the unique semi-tangent of $S$ at $p$. Then semi-tangents of $S$ at points near $p$ must be close to $X$, i.e., if $X_n$ is any semi-tangent of $S$ at $p_n$, and $p_n \to p$ then $X_n \to X$.*

The proof of the maximum principle for $C^0$ null hypersurfaces will proceed in a fashion similar to the smooth case. Thus we will need to consider the intersection of a $C^0$ null hypersurface $S$ with a smooth timelike hypersurface $P$.

**Lemma III.2** *Let $S$ be a $C^0$ future null hypersurface and let $p \in S$ be an interior point of a null generator $\eta$ of $S$. Let $P$ be a smooth timelike hypersurface passing through $p$ transverse to $\eta$. Then there exists a neighborhood $\mathcal{O}$ of $p$ in $P$ such that $\Sigma = S \cap P$ is a partial Cauchy surface in $\mathcal{O}$, i.e., $\Sigma$ is a closed acausal $C^0$ hypersurface in $\mathcal{O}$.*

*Proof.* The proof uses the edge concept, in particular the result that an achronal set is a closed $C^0$ hypersurface if and only if it is edgeless; see e.g., Corollary 26, p. 414 in [16]. Let $U$ be a neighborhood of $p$ in $M$ in which $S$ is achronal and edgeless. Then $V = U \cap P$ is a neighborhood of $p$ in $P$ in which $\Sigma$ is achronal and edgeless in $P$. Hence, $\Sigma$ is a closed achronal $C^0$ hypersurface in $V$, and it remains to show that it is actually acausal in a perhaps smaller neighborhood. Suppose there exists a sequence of neighborhoods $V_n \subset V$ of p, which shrink to $p$, such that $\Sigma$ is not acausal in $V_n$ for each $n$. Then, for each $n$, there exists a pair of points $p_n, q_n \in \Sigma \cap V_n$ such that $p_n \to p$ and $q_n \in J^+(p_n, V_n)$. Hence for each $n$, there exists a $P$-null geodesic $\eta_n$ from $p_n$ to $q_n$. Now, $\eta_n$ is a causal curve in $U$, and, in fact, must be a null geodesic in $U$, since otherwise we would have $q_n \in I^+(p_n, U)$, which would violate the achronality of $S$ in $U$. Hence $\eta_n \subset S$, and the initial tangent $X_n$ to $\eta_n$, when suitably normalized, is a semi-tangent of $S$ at $p_n$. By the second part of Lemma III.1, $X_n \to X$, where $X$ is tangent to $\eta$ at $p$. But $X$ is tangent to $P$, since each $X_n$ is, which contradicts the assumption that $P$ is transverse to $\eta$ at $p$.

We now extend the meaning of mean curvature inequalities to $C^0$ null hypersurfaces. The idea, motivated by previous work involving spacelike hypersurfaces ([7], [2]) is to use smooth null support hypersurfaces. *Henceforth we set the scale for all null vectors on $M$ by requiring that they have unit length with respect to a fixed Riemannian metric on $M$.*

**Definition III.2** *Let $S$ be a $C^0$ future null hypersurface in $M$. We say that $S$ has null mean curvature $\theta \geq 0$* in the sense of support hypersurfaces *provided for each $p \in S$ and for each $\epsilon > 0$ there exists a smooth (at least $C^2$) null hypersurface $S_{p,\epsilon}$ such that,* *(1) $S_{p,\epsilon}$ is a past support hypersurface for $S$ at $p$, i.e., $S_{p,\epsilon}$ passes through $p$ and lies to the past side of $S$ near $p$, and*

*(2) the null mean curvature of $S_{p,\epsilon}$ at $p$ satisfies $\theta_{p,\epsilon} \geq -\epsilon$.*

For example, if $p$ is a point in Minkowski space, the future null cone $\partial I^+(p)$ has null mean curvature $\theta \geq 0$ in the sense of support hypersurfaces. One can use null hyperplanes, even at the vertex, as support hypersurfaces. Another, less trivial example is that of a black hole event horizon $H = \partial I^-(\mathcal{I}^+)$ in an asymptotically flat black hole spacetime $M$. Here $\mathcal{I}^+$ refers to future null infinity; see Section 4. Assuming the generators of $H$ are future complete and $M$ obeys the null energy condition, $\mathrm{Ric}(X,X) \geq 0$, for all null vectors $X$, it follows from Lemma IV.2 in Section 4 that $H$ has null mean curvature $\theta \geq 0$ in the sense of support hypersurfaces. This observation and other consequences of the existence of smooth null support hypersurfaces provided the initial impetus for the development of a proof of the black hole area theorem under natural regularity conditions, i.e. the regularity implicit in the fact that $H$ is an achronal boundary, cf. [5].

With the notation as in Definition III.2, let $B_{p,\epsilon}$ denote the null second fundamental form of $S_{p,\epsilon}$ at $p$. We say that the collection of null second fundamental forms $\{B_{p,\epsilon} : p \in S, \epsilon > 0\}$ is *locally bounded from below* provided that for all $p \in S$ there is a neighborhood $\mathcal{W}$ of $p$ in $S$ and a constant $k > 0$ such that

$$B_{q,\epsilon} \geq -kh_{q,\epsilon} \quad \text{for all } q \in \mathcal{W} \text{ and } \epsilon > 0, \tag{III.14}$$

where $h_{q,\epsilon}$ is the Riemannian metric on $T_qS_{q,\epsilon}/K_{q,\epsilon}$, as defined in Section 2.1. This technical condition arises in the statement of the maximum principle for $C^0$ null hypersurfaces, and is satisfied in many natural geometric situations for essentially a priori reasons.

**Lemma III.3** *Let $S$ be a $C^0$ future null hypersurface and let $W$ be a smooth null hypersurface which is a past support hypersurface for $S$ at $p$. If $K \in T_pW$ is the future directed (normalized) null tangent of $W$ at $p$ then $K$ is a semi-tangent of $S$ at $p$.*

*Proof.* Let $U$ be a neighborhood of $p$ such that $S \cap U$ is closed and achronal in $U$ and $W \cap U \subset J^-(S \cap U, U)$. For simplicity, we may assume that $S \subset U$ and $W \subset J^-(S,U)$. Let $\eta \subset U$ be an initial segment in $U$ of the future directed null generator of $W$ starting at $p$ with initial tangent $K$. Then $\eta \subset J^-(S,U) \cap J^+(S,U)$. By the remark before Definition III.1, if $\eta$ leaves $S$ at some point, it will enter either $I^-(S,U)$ or $I^+(S,U)$. Either case leads to an achronality violation. Hence, $\eta$ must be a null generator of $S$, which implies that $K$ is a semi-tangent of $S$.

If $S$ is a $C^0$ past null hypersurface, one defines in a time-dual fashion what it means for $S$ to have null mean curvature $\theta \leq 0$ in the sense of support hypersurfaces. In this case one uses smooth null hypersurfaces which lie locally to the *future* of $S$. Although, in principle, one can also consider future null hypersurfaces with nonpositive null mean curvature, this appears to be a less useful notion, as future support hypersurfaces cannot typically be constructed at past end points of generators.

### III.2 The maximum principle for $C^0$ null hypersurfaces

The aim now is to present a proof of the geometric maximum principle stated below. Unless otherwise stated, we continue to assume that all null vectors are normalized to unit length with respect to a fixed background Riemannian metric.

**Theorem III.4** *Let $S_1$ be a $C^0$ future null hypersurface and let $S_2$ be a $C^0$ past null hypersurface in a spacetime $M$. Suppose,*

*(1) $S_1$ and $S_2$ meet at $p \in M$ and $S_2$ lies to the future side of $S_1$ near $p$,*

*(2) $S_1$ has null mean curvature $\theta_1 \geq 0$ in the sense of support hypersurfaces, with null second fundamental forms $\{B_{p,\epsilon} : p \in S_1, \epsilon > 0\}$ locally bounded from below, and*

*(3) $S_2$ has null mean curvature $\theta_2 \leq 0$ in the sense of support hypersurfaces.*

*Then $S_1$ and $S_2$ coincide near $p$, i.e., there is a neighborhood $\mathcal{O}$ of $p$ such that $S_1 \cap \mathcal{O} = S_2 \cap \mathcal{O}$. Moreover, $S_1 \cap \mathcal{O} = S_2 \cap \mathcal{O}$ is a smooth null hypersurface with null mean curvature $\theta = 0$.*

The proof proceeds in a similar fashion to the proof of Theorem II.1. Instead of Theorem II.2, we use the strong maximum principle for weak (in the sense of support functions) sub and super solutions of second order quasi-linear elliptic equations obtained in [2]. We will state here a restricted form of this result, adapted to our purposes.

Let $\Omega$ be a domain in $\mathbb{R}^n$ and let $U$ be an open set in $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$ of the form $U = \Omega \times I \times B$, where $I$ is an open interval and $B$ is an open ball in $\mathbb{R}^n$. Consider the second order quasi-linear elliptic operator $Q = Q(u)$ as defined in Equation II.5 for $U$-admissible functions $u \in C^2(\Omega)$, where now we assume $a^{ij}, b \in C^\infty(U)$. We now also assume $Q = Q(u)$ is uniformly elliptic, by which we mean (1) the quantity $\sum_{i,j=1}^n a^{ij}(x,r,p)\xi^i\xi^j$ is uniformly positive and bounded away from infinity for all $(x,r,p) \in U$ and all unit vectors $\xi = (\xi_1, ..., \xi_n)$, and (2) $a^{ij}, b$ and their first order partial derivatives are bounded on $U$.

We need the notion of a support function. Given $u \in C^0(\Omega)$ and $x_0 \in \Omega$, $\phi$ is an upper (resp., lower) support function for $u$ at $x_0$ provided $\phi(x_0) = u(x_0)$ and $\phi \geq u$ (resp., $\phi \leq u$) on some neighborhood of $x_0$. We say that a function $u \in C^0(\Omega)$ satisfies $Q(u) \geq 0$ *in the sense of support functions* iff for all $\epsilon > 0$ and all $x \in \Omega$ there is a $U$-admissible lower support function $\phi_{x,\epsilon}$, which is $C^2$ in a neighborhood of $x$, such that $Q(\phi_{x,\epsilon})(x) \geq -\epsilon$. We say that $u$ satisfies $Q(u) \geq 0$ *in the sense of support functions with Hessians locally bounded from below* iff, in addition, there is a constant $k > 0$, independent of $\epsilon$ and $x$, such that $\mathrm{Hess}(\phi_{x,\epsilon})(x) \geq -kI$, where

$I$ is the identity matrix. For $u \in C^0(\Omega)$, we define $Q(u) \le 0$ *in the sense of support functions* in an analogous fashion.

The following theorem is a special case of Theorem 2.4 in [2]

**Theorem III.5** *Let $Q = Q(u)$ be a second order quasi-linear uniformly elliptic operator as described above. Suppose $u, v \in C^0(\Omega)$ satisfy,*

*(1) $u \le v$ on $\Omega$ and $u(x_0) = v(x_0)$ for some $x_0 \in \Omega$,*

*(2) $Q(u) \ge 0$ in the sense of support functions with Hessians locally bounded from below, and*

*(3) $Q(v) \le 0$ in the sense of support functions.*

*Then $u = v$ on $\Omega$ and $u = v \in C^\infty(\Omega)$.*

We now proceed to the proof of the maximum principle for $C^0$ null hypersurfaces.

*Proof of Theorem III.4:* We first observe that $p$ is an interior point of null generators for both $S_1$ and $S_2$, and that these two null generators agree near $p$. To show this, let $\eta_i$ be a null generator of $S_i$ starting at $p$, $i = 1, 2$; $\eta_1$ is future directed and $\eta_2$ is past directed. Let $U$ be a neighborhood of $p$ in which $S_1$ is closed achronal , such that $S_2 \cap U \subset J^+(S_1 \cap U, U)$. We may assume $\eta_1$ and $\eta_2$ are contained in $U$. Since $\eta_2$ is past pointing, $\eta_2 \subset J^-(S_1 \cap U, U) \cap J^+(S_1 \cap U, U)$. Then, as in Lemma III.3, the achronality of $S_1$ in $U$ forces $\eta_2 \subset S_1$. To avoid an achronality violation, $\eta = -\eta_2 \cup \eta_1$ must be an unbroken null geodesic, and hence a null generator of $S_1$ passing through $p$. Similarly, $-\eta = -\eta_1 \cup \eta_2$ is a null generator of $S_2$ passing through $p$.

Let $P$ be a timelike hypersurface passing through $p$ transverse to $\eta$. Let $K_p$ be the normalized semi-tangent of $S_1$ at $p$; $K_p$ is tangent to $\eta$. Let $N$ be the spacelike unit normal vector field to $P$ that points to the same side of $P$ as $K_p$. By making a homothetic change in the background Riemannian metric we may assume $\langle K_p, N_p \rangle = 1$. Hence, $K_p$ is of the form, $K_p = Z_p + N_p$, where $Z_p \in T_pP$ is a future directed unit timelike vector.

As in the proof of Theorem II.1, $P$ in the induced metric can be expressed as in Equations II.6 and II.7. Moreover, $V$ can be constructed so that $Z_p$ is perpendicular to $V$. Then $K_p = (\partial_0 + N)_p$. By Lemma III.2, provided $P$ is taken small enough, $\Sigma_1 = S_1 \cap P$, and $\Sigma_2 = S_2 \cap P$ will be partial Cauchy surfaces in $P$ passing through $p$, with $\Sigma_2$ to the future of $\Sigma_1$. Thus, shrinking $P$ further if necessary, there exist functions $u_i \in C^0(V)$, $i = 1, 2$, such that $\Sigma_i = \mathrm{graph}(u_i)$ and

(1) $u_1 \le u_2$ on $V$ and $u_1(p) = u_2(p) = 0$.

Let $\{S_{q,\epsilon}\}$ be the family of smooth null lower support hypersurfaces for $S_1$. Restrict attention to those members of the family for which $q \in \Sigma_1$. Let $B_{q,\epsilon}$ be the null second fundametal form of $S_{q,\epsilon}$ at $q$ with respect to the null vector

$K_{q,\epsilon}$. By Lemmas III.1 and III.3, the collection of null vectors $\{K_{q,\epsilon}\}$ can be made arbitrarily close to $K_p$ by taking $P$ sufficiently small. This has several implications. It implies, in particular, for $P$ sufficiently small, that $K_{q,\epsilon}$ is transverse to $P$. Hence, by shrinking $S_{q,\epsilon}$ about $q$, if necessary, $S_{q,\epsilon}$ meets $P$ in a properly transverse manner, and thus $\Sigma_{q,\epsilon} = S_{q,\epsilon} \cap P$ is a smooth spacelike hypersurface in $P$. Thus for each $\epsilon > 0$ and $q \in \Sigma_1$, there exists $\phi_{q,\epsilon} \in C^2(W_{q,\epsilon})$, $W_{q,\epsilon} \subset V$, such that $\Sigma_{q,\epsilon} = \mathrm{graph}(\phi_{q,\epsilon})$.

We now consider the null mean curvature operator $\theta = \theta(u)$, as described in equations II.12 and II.13, on the set $U = V \times (-a, a) \times B$, where $B$ is an open ball in $\mathbb{R}^{n-2}$ centered at the origin. By choosing $V$, $a$ and $B$ sufficiently small, $\theta = \theta(u)$ will be uniformly elliptic on $U$, in the sense described above. Since the vectors $\{K_{q,\epsilon}\}$ can be made uniformly close to $K_p$, the inner products $\langle K_{q,\epsilon}, N_q\rangle$ can be made uniformly close to the value one. Hence, we can rescale the vectors $K_{q,\epsilon}$ so that $\langle K_{q,\epsilon}, N_q\rangle = 1$ without altering the validity of the assumed null mean curvature inequality $\theta_1 \geq 0$ in the sense of support hypersurfaces, at points of $S_1$ in $P$. Then $K_{q,\epsilon}$ can be expressed as, $K_{q,\epsilon} = Z_{q,\epsilon} + N_q$, where $Z_{q,\epsilon} \in T_qP$ is a future directed unit timelike vector. Moreover, the vectors $Z_{q,\epsilon}$ can be made uniformly close to $Z_p = \partial_0|p$ by taking $P$ small enough. Equation II.10 then implies that the Euclidean vectors $\partial\phi_{q,\epsilon}(q) = (\partial_1\phi_{q,\epsilon}(q), ..., \partial_{n-2}\phi_{q,\epsilon}(q))$ can be made to lie in the ball $B$.

We conclude that by taking $P$ sufficiently small, each function $\phi_{q,\epsilon}$ is $U$-admissible. Now, $\phi_{q,\epsilon}$ is a $C^2$ lower support function for $\Sigma_1$ at $q$. By assumption, the null mean curvature of $S_{q,\epsilon}$ at $q$ satisfies, $\theta_{q,\epsilon}(q) \geq -\epsilon$, which, in the analytic setting, translates into, $\theta(\phi_{q,\epsilon})(q) \geq -\epsilon$. Thus, $u_1$ satisfies, $\theta(u_1) \geq 0$ in the sense of support functions.

For each $q \in \Sigma_1$, $Z_{q,\epsilon}$ is the future directed timelike unit normal to $\Sigma_{q,\epsilon}$. Let $\beta_{q,\epsilon}$ be the second fundamental form of $\Sigma_{q,\epsilon} \subset P$ at $q$ with respect to the normal $Z_{q,\epsilon}$. Let $B_{P,q}$ be the second fundamental form of P at $q$ with respect to $N$. The second fundamental forms $B_{q,\epsilon}$, $\beta_{q,\epsilon}$, and $B_{P,q}$ are related by

$$B_{q,\epsilon}(\overline{X}, \overline{X}) = \beta_{q,\epsilon}(X, X) + B_{P,q}(X, X), \tag{III.15}$$

for all unit vectors $X \in T_q\Sigma_{q,\epsilon} = [Z_{q,\epsilon}]^\perp \subset T_qP$. In a sufficiently small relatively compact neighborhood $P_0$ of $p$ in $P$, the collection of vectors $\{Z_{q,\epsilon} : q \in \Sigma_1 \cap P_0\}$ has compact closure in $TP$. It follows that the collection of vectors $\mathcal{X} = \{X_q \in T_q\Sigma_{q,\epsilon} : q \in \Sigma_1 \cap P_0, |X_q| = 1\}$ has compact closure in $TP$, as well. Hence the set of numbers $\{B_{P,q}(X_q, X_q) : X_q \in \mathcal{X}\}$ is bounded. Coupled with the assumption that the second fundamenal forms $\{B_{q,\epsilon}\}$ are locally bounded from below, we conclude, by shrinking $P$ further if necessary, that the second fundamental forms $\{\beta_{q,\epsilon} : q \in \Sigma_1\}$ are locally bounded from below, i.e., for each $q_0 \in \Sigma_1$ there is a neighborhood $\mathcal{W}$ of $q_0$ in $\Sigma_1$ and a constant $k$ such that

$$\beta_{q,\epsilon} \geq -kg_{q,\epsilon} \quad \text{for all } q \in \mathcal{W} \text{ and } \epsilon > 0, \tag{III.16}$$

where $g_{q,\epsilon}$ is the induced metric on $\Sigma_{q,\epsilon}$ at $q$. For $P$ sufficiently small, the induced metrics $g_{q,\epsilon}$ will be close to the metric of $V$ at $p$. Using the relationship between

$\beta_{q,\epsilon}$ and Hess $\phi_{q,\epsilon}$, worked out, for example, in Section 3.1 in [2], inequality III.16 translates into the analytic statement that for each $q_0 \in \Sigma_1$ there is a neighborhood $\mathcal{W}$ of $q_0$ in $\Sigma_1$ and a constant $k_1$ such that Hess $\phi_{q,\epsilon}(q) \geq -k_1 I$ for all $q \in \mathcal{W}$ and $\epsilon > 0$. Thus, we finally conclude that,

(2) $u_1$ satisfies $\theta(u_1) \geq 0$ in the sense of support functions with Hessians locally bounded from below.

By similar reasoning, adjusting the size of $P$ as necessary, we have that

(3) $u_2$ satisfies $\theta(u_2) \leq 0$ in the sense of support functions.

In view of (1), (2), and (3), Theorem III.5 applied to the operator $\theta = \theta(u)$ implies that $u_1 = u_2$ on $V$ and $u_1 = u_2$ is $C^\infty$. Hence, $\Sigma_1$ and $\Sigma_2$ are smooth spacelike hypersurfaces in $P$ which coincide near $p$. Then near $p$, $S_1$ and $S_2$ are obtained by exponentiating normally out along a common smooth null orthogonal vector field along $\Sigma_1 = \Sigma_2$. The conclusion to Theorem III.4 now follows. □

For simplicity we have restricted attention to the null mean curvature inequalities $\theta_1 \leq 0 \leq \theta_2$. However, Theorem III.4, with an obvious modification of Definition III.2, remains valid under the null mean curvature inequalities $\theta_1 \leq a \leq \theta_2$, for any $a \in \mathbb{R}$.

## IV The null splitting theorem

We now consider some consequences of Theorem III.4. The proof of many global results in general relativity, such as the classical Hawking-Penrose singularity theorems and more recent results such as those concerning topological censorship (see e.g., [11, 12]) involve the construction of a *timelike line* or a *null line* in spacetime. A timelike geodesic segment is *maximal* if it is longest among all causal curves joining its end points, or equivalently, if it realizes the Lorentzian distance between its end points. A timelike line is an inextendible timelike geodesic each segment of which is maximal. Similarly, a null line is an inextendible null geodesic each segment of which is maximal. But since each segment of a null geodesic has zero length, it follows that an inextendible null geodesic is a null line iff it is achronal, i.e., iff no two points of it can be joined by a timelike curve. In particular, null lines, like timelike lines, must be free of conjugate points. The standard Lorentzian splitting theorem [3, Chapter 14], considers the rigidity of spacetimes which contain timelike lines. Recall, it asserts that a timelike geodesically complete spacetime $(M, g)$ which obeys the strong energy condition, $\mathrm{Ric}(X, X) \geq 0$ for all timelike vectors $X$, and which contains a timelike line splits along the line, i.e., is isometric to $(\mathbb{R} \times V, -dt^2 \oplus h)$, where $(V, h)$ is a complete Riemannian manifold. We now consider the analogous problem for spacetimes with null lines. The theorem stated below establishes the rigidity of spacetime in this null case. Unless otherwise stated, we continue to assume that all null vectors are normalized to unit length with respect to a fixed background Riemannian metric.

**Theorem IV.1** *Let $M$ be a null geodesically complete spacetime which obeys the null energy condition,* $\mathrm{Ric}(X,X) \geq 0$ *for all null vectors $X$ and contains a null line $\eta$. Then $\eta$ is contained in a smooth closed achronal totally geodesic ($B \equiv 0$) null hypersurface $S$.*

The simplest illustration of Theorem IV.1 is Minkowski space: Each null geodesic in Minkowski space is contained in a null hyperplane. De Sitter space, anti-de Sitter, and gravitational plane wave solutions furnish other illustrations.

**Remark IV.1** Theorem IV.1 may be viewed as a "splitting" theorem of sorts, where the splitting takes place in $S$. Let $K$ be the unique (up to scale) future pointing null vector field of $S$. The vanishing of the null second fundamental form $B$ of $S$ means that the standard kinematical quantities associated with $K$, i.e., the expansion $\theta$, and shear $\sigma$ (as well as vorticity) of $K$ vanish. In a similar vein, the vanishing of $B$ implies that the metric $h$ on $TS/K$ defined in Section 2.1 is invariant under the flow generated by $K$; see [14] for a precise statement and proof. To take this a step further, suppose, in the setting of Theorem IV.1, that $M$ is also globally hyperbolic. Let $V$ be a smooth Cauchy surface for $M$. Since $V$ is Cauchy, and the generators of $S$ are inextendible in $S$, each null generator of $S$ meets $V$ exactly once, and this intersection is properly transverse. Hence, $\Sigma = S \cap V$ is a smooth codimension two spacelike submanifold of $M$, and the map $\Phi : \mathbb{R} \times \Sigma \to S$, defined by, $\Phi(s,x) = \exp_x sK(x)$, is a diffeomorphism. The invariance of the metric $h$ with respect to the flow generated by $K$, or equivalently, by $\Phi_*(\frac{\partial}{\partial s})$, implies that $(i \circ \Phi)^* g = \pi^* g_0$, where $i : S \hookrightarrow M$ is inclusion, $\pi : \mathbb{R} \times \Sigma \to \Sigma$ is projection, and $g_0$ is the induced metric on $\Sigma$. In more heuristic terms, $S \approx \mathbb{R} \times \Sigma$, and $i^* g = 0dt^2 + g_0$. It is in this sense that one may view Theorem IV.1 as a splitting theorem.

**Remark IV.2** The proof of Theorem IV.1 shows that $S = \widehat{\partial I^+(\eta)} = \widehat{\partial I^-(\eta)}$, where $\widehat{\partial I^\pm(\eta)}$ is the component of $\partial I^\pm(\eta)$ containing $\eta$. In more heuristic terms $S$ is obtained as a limit of future null cones $\partial I^+(\eta(t))$ (resp., past null cones $\partial I^-(\eta(t))$) as $t \to -\infty$ (resp., $t \to \infty$). The proof also shows that the assumption of null completeness can be weakened. It is sufficient to require that the generators of $\partial I^-(\eta)$ be future geodesically complete and the generators of $\partial I^+(\eta)$ be past geodesically complete. (As discussed in the proof, the generators of $\partial I^-(\eta)$ (resp., $\partial I^+(\eta)$) are in general future (resp., past) inextendible in $M$.)

The proof of Theorem IV.1 is an application of Theorem III.4, and relies on the following basic lemma.

**Lemma IV.2** *Let $M$ be a spacetime which satisfies the null energy condition. Suppose $S$ is an achronal $C^0$ future null hypersurface whose null generators are future geodesically complete. Then $S$ has null mean curvature $\theta \geq 0$ in the sense of support hypersurfaces, with null second fundamental forms locally bounded from below.*

*Proof.* The basic idea for constructing past support hypersurfaces for $S$ is to consider the "past null cones" of points on generators of $S$ formed by past null geodesics emanating from these points. For the purpose of establishing certain properties about these null hypersurfaces, it is useful to relate them to achronal boundaries of the form $\partial J^-(q)$, $q \in S$, which are defined purely in terms of the causal structure of $M$. For this reason we assume initially that $M$ is globally hyperbolic. At the end we will indicate how to remove this assumption.

For each $p \in S$ and normalized semi-tangent $K$ at $p$, let $\eta : [0, \infty) \to M$, $\eta(s) = \exp_p sK$, be the affinely parameterized null geodesic generator of $S$ starting at $p$ with initial tangent $K$. Since $S$ is achronal each such generator is a null *ray*, i.e., a maximal null half-geodesic. Since $\eta$ is maximal, no point on $\eta|_{[0,r)}$ is conjugate to $\eta(r)$. For each $r > 0$, consider the achronal boundary $\partial J^-(\eta(r))$, which is a $C^0$ hypersurface containing $\eta|_{[0,r]}$. By standard properties of the null cut locus (see especially, Theorems 9.15 and 9.35 in [3], which assume global hyperbolicity) there is a neighborhood $U$ of $\eta|_{[0,r)}$ such that $S_{p,K,r} = \partial J^-(\eta(r)) \cap U$ is a smooth null hypersurface diffeomorphic under the exponential map based at $\eta(r)$ to a neighborhood of the line segment $\{-\tau\eta'(r) : 0 < \tau \leq 1\}$ in the past null cone $\Lambda^-_{\eta(r)} \subset T_{\eta(r)}M$.

From the achronality of $S$ we observe, $\partial J^-(\eta(r)) \cap I^+(S) = \emptyset$. This implies that $S_{p,K,r}$ is a past support hypersurface for $S$ at $p$. Let $\theta_{p,K,r}$ denote the null mean curvature of $S_{p,K,r}$ at $p$ with respect to $K$. We use Equation II.4 to obtain the lower bound

$$\theta_{p,K,r} \geq -\frac{n-2}{r}. \tag{IV.17}$$

The argument is standard. In the notation of Section 2.1, let $\theta(s)$, $s \in [0, r)$, be the null mean curvature of $S_{p,K,r}$ at $\eta(s)$ with respect to $\eta'(s)$. Equation II.4 and the energy condition imply,

$$\frac{d\theta}{ds} \leq -\frac{1}{n-2}\,\theta^2. \tag{IV.18}$$

Without loss of generality we may assume $\theta(0) = \theta_{p,K,r} < 0$. Then $\theta = \theta(s)$ is strictly negative on $[0, r)$, and we can devide IV.18 by $\theta^2$ to obtain,

$$\frac{d}{ds}\,\theta^{-1} \geq \frac{1}{n-2}. \tag{IV.19}$$

Integrating IV.19 from 0 to $r-\delta$, and letting $\delta \to 0$ we obtain the lower bound IV.17.

Thus, since $r$ can be taken arbitrarily large, we have shown, in the globally hyperbolic case, that $S$ has null mean curvature $\theta \geq 0$ in the sense of support hypersurfaces with respect to the collection $\{S_{p,K,r}\}$ of smooth null hypersurfaces.

By Lemma III.3, $K$ is tangent to $S_{p,K,r}$ at $p$. Let $B_{p,K,r}$ denote the null second fundamental form of $\{S_{p,K,r}\}$ at $p$ with respect to $K$. We now argue that the collection of null second fundamental forms $\{B_{p,K,r} : r \geq r_0\}$, for some $r_0 > 0$, is locally bounded from below. Recall, "locally", means "locally in the point $p$"; the lower bound must hold for all $r > r_0$, cf., inequality III.14. This lower bound follows from a continuity argument and an elementary monotonicity result, as we now discuss.

Fix $p \in S$. Let $U$ be a convex normal neighborhood of $p$. Thus, for each $q \in U$, $U$ is the diffeomorphic image under the exponential map based at $q$ of a neighborhood of the origin in $T_qM$. Provided $U$ is small enough, we have that for each $q \in U$, $\partial J^-(q) \cap U = \partial(J^-(q) \cap U) = \exp_q(\Lambda_q^- \cap \exp_q^{-1}(U))$, where $\Lambda_q^-$ is the past null cone in $T_qM$. In particular, for each $q \in U$, $W(q) = \partial J^-(q) \cap U \setminus \{q\}$ is a smooth null hypersurface in $U$ such that if $q_n \to q$ in $U$, $W(q_n)$ converges smoothly to $W(q)$.

There exists a neighborhood $V$ of $p$, $V \subset U$, and $r_0 > 0$ such that for each $q \in V$, and each normalized null vector $K \in T_qM$, the null geodesic segment $s \to \eta(s) = \exp_q sK$, $s \in [0, r_0]$ is contained in $U$. Let $B(q, K)$ be the null second fundamental form of $W(\exp_q r_0K)$ at $q$ with respect to $K$. Since, as $(q_n, K_n) \to (q, K)$, $W(\exp_{q_n} r_0K_n)$ converges smoothly to $W(\exp_q r_0K)$, we have that $B(q_n, K_n) \to B(q, K)$ smoothly. Returning to the original family of support hypersurfaces, $\{S_{p,K,r}\}$, with associated family of null second fundamental forms $\{B_{p,K,r}\}$, observe that when $q \in S \cap V$, $S_{q,K,r_0}$ agrees with $W(\exp_q r_0K)$ near $q$. Hence, if $(q_n, K_n) \to (q, K)$ in $S \cap V$, $B_{q_n,K_n,r_0} \to B_{q,K,r_0}$ smoothly. It follows that the collection of null second fundamental forms $\{B_{p,K,r_0}\}$ is locally bounded from below.

Consider as in the beginning, for $p \in S$ and $K$ a normalized semi-tangent at $p$, the null geodesic generator, $\eta : [0, \infty) \to M$, $\eta(s) = \exp_p sK$, $s \ge 0$. For $0 < r < t < \infty$, $J^-(\eta(r)) \subset J^-(\eta(t))$. Then, since $\partial J^-(\eta(t))$ is achronal, $\partial J^-(\eta(r))$ cannot enter $I^+(\partial J^-(\eta(t))$. It follows that for $r < t$, $S_{p,K,r}$ lies to the past side of $S_{p,K,t}$ near $p$. By an elementary comparison of null second fundamental forms at $p$ we obtain the monotonicity property,

$$B_{p,K,t} \ge B_{p,K,r} \quad \text{for all } 0 < r < t < \infty. \tag{IV.20}$$

This monotonicity now implies that the entire family of null second fundamental forms $\{B_{p,K,r} : r \ge r_0\}$ is locally bounded from below.

This concludes the proof of Lemma IV.2 under the assumption that $M$ is globally hyperbolic. We now describe how to handle the general case. In general, $M$ may have bad causal properties. In particular the generators of $S$ could be closed. Nevertheless, the past support hypersurfaces are formed in the same manner, as the "past null cones" of points on generators of $S$ formed by past null geodesics emanating from these points. But as an intermediary step, to take advantage of the arguments in the globally hyperbolic case, we pull back each generator to a spacetime having good causal properties.

Again, consider, for $p \in S$ and $K$ a normalized semi-tangent at $p$, the null geodesic generator, $\eta : [0, \infty) \to M$, $\eta(s) = \exp_p sK$, $s > 0$. Restrict attention to the finite segment $\eta|_{[0,r]}$. Roughly speaking, we introduce Fermi type coordinates near $\eta|_{[0,r]}$. Let $\{e_1, e_2, ...e_{n-1}\}$ be an orthonormal frame of spacelike vectors in $T_{\eta(0)}M$. Parallel translate these vectors along $\eta$ to obtain spacelike orthonormal vector fields $e_i = e_i(s)$, $1 \le i \le n-2$ along $\eta|_{[0,r]}$. Consider the map $\Phi : \bar{M} \subset \mathbb{R}^n \to M$ defined by $\Phi(s, x^1, x^2, ..., x^{n-1}) = \exp_{\eta(s)}(\sum_{i=1}^{n-1} x^i e_i)$. Here $\bar{M}$ is an open set containing the line segment $\{(s, 0, 0, ..., 0) : 0 \le s \le r\}$. By the inverse function theorem we can choose $\bar{M}$ so that $\Phi$ is a local diffeomorphism. Equip $\bar{M}$ with the pullback metric $\bar{g} = \Phi^* g$, where $g$ is the Lorentzian metric on $M$, thereby making $\bar{M}$ Lorentzian and $\Phi$ a local isometry. Let $t \in C^\infty(\bar{M})$ be the 0th coordinate function, $t(s, x^1, x^2, ..., x^{n-1}) = s$. Since the slices $t = s$ are spacelike, $\nabla t$ is timelike, and hence $t$ is a time function on $\bar{M}$. Thus, $\bar{M}$ is a strongly causal spacetime.

The curve $\bar{\eta} : [0, \bar{r}+\delta] \to \bar{M}$, $\bar{\eta}(s) = (s, 0, 0, ..., 0)$, defined for $\delta > 0$ sufficiently small, is a maximal null geodesic in $\bar{M}$. Let $\mathcal{K}$ be a compact neighborhood of $\bar{\eta}$ in $\bar{M}$. Then by Corollary 7.7 in [18], any two causally related points in $\mathcal{K}$ can be joined by a longest causal curve $\gamma$ in $\mathcal{K}$, and each segment of $\gamma$ contained in the interior of $\mathcal{K}$ is a maximal causal geodesic. This property is sufficient to push through, with only minor modifications, all relevant results concerning the null cut locus of $\bar{\eta}(0)$ on $\bar{\eta}$. In view of the above discussion, there is a neighborhood $\bar{U}$ of $\bar{\eta}|_{[0,r)}$ such that $\bar{S}_{p,K,r} = \partial J^-(\eta(r)) \cap \bar{U}$ is a smooth null hypersurface diffeomorphic under the exponential map based at $\bar{\eta}(r)$ to a neighborhood of the line segment $\{-\tau\bar{\eta}'(r) : 0 < \tau \le 1\}$ in the past null cone $\Lambda^-_{\bar{\eta}(r)} \subset T_{\bar{\eta}(r)}\bar{M}$. Let $\bar{V} \subset \bar{U}$ be a neighborhood of $\bar{\eta}(0)$ on which $\Phi$ is an isometry onto its image, and let $S_{p,K,r} = \Phi(\bar{S}_{p,K,r} \cap \bar{V})$. Then $\{S_{p,K,r}\}$ is the desired collection of past support hypersurfaces for $S$, having all the requisite properties. In particular, the mean curvature inequality IV.17 and the monotonicity property IV.20 hold for the family $\{\bar{S}_{p,K,r}\}$, and hence the family $\{S_{p,K,r}\}$, by just the same arguments as in the globally hyperbolic case.

*Proof of Theorem IV.1:* Since $\eta$ is achronal, it follows that $\eta \subset \partial I^-(\eta)$, and hence $\partial I^-(\eta) \ne \emptyset$. Then, by standard properties of achronal boundaries, $\partial I^-(\eta)$ is a closed achronal $C^0$ hypersurface in $M$. We claim that $\partial I^-(\eta)$ is a $C^0$ future null hypersurface. By standard results on achronal boundaries, e.g., [18, Theorem 3.20], $\partial I^-(\eta) \setminus \bar{\eta}$ (where $\bar{\eta}$ is the closure of $\eta$ as a subset of $M$) is ruled by future inextendible null geodesics. However, since we do not assume $M$ is strongly causal, it is possible, in the worst case scenario, that $\bar{\eta} = \partial I^-(\eta)$, in which case [18, Theorem 3.20] gives no information. To show that $\partial I^-(\eta)$ is ruled by future inextendible null geodesics we apply instead the more general [18, Lemma 3.19]. Let $N \subset U$ be a convex normal neighborhood of $p$. $N$ as a spacetime in its own right is strongly causal. Let $K$ be a compact neighborhood of $p$ contained in $N$. Then for each $t \in \mathbb{R}$, $\eta|_{[t,\infty)}$ cannot remain in $K$ if it ever meets it. Thus there exists a sequence of $p_i = \eta(t_i)$, $t_i \nearrow \infty$, $p_i \notin K$. It follows that for each $x \in K \cap I^-(\eta)$, there exists a future directed timelike curve from $x$ to a point on $\eta$ not in $K$. We may then apply [18, Lemma 3.19] to conclude that $p$ is the past end point of a future directed null geodesic segment contained in $\partial I^-(\eta)$. Hence, according to Definition III.1, $\partial I^-(\eta)$ is a $C^0$ future null hypersurface. Moreover, because $\partial I^-(\eta)$ is closed, the null generators of $\partial I^-(\eta)$ are future inextendible in $M$, and hence future complete.

Let $S_-$ be the component of $\partial I^-(\eta)$ containing $\eta$. From the above, $S_-$ is an achronal $C^0$ future null hypersurface whose null generators are future geodesically complete. Thus, by Lemma IV.2, $S_-$ has null mean curvature $\theta_- \geq 0$ in the sense of support hypersurfaces, with null second fundamental forms locally bounded from below. Let $S_+$ be the component of $\partial I^+(\eta)$ containing $\eta$. Arguing in a time-dual fashion, $S_+$ is an achronal $C^0$ past null hypersurface whose null generators are past geodesically complete. By the time-dual of Lemma IV.2, $S_+$ has null mean curvature $\theta_+ \leq 0$ in the sense of support hypersurfaces. Moreover at any point $p$ on $\eta$, $S_+$ lies to the future side of $S_-$ near $p$. Theorem III.4 then implies that $S_- = S_+ = S$ is a smooth null hypersurface containing $\eta$ with vanishing null mean curvature, $\theta \equiv 0$. Equation II.4 and the null energy condition then imply that the shear scalar vanishes along each generator, which in turn implies that the null second fundamental form of $S$ vanishes. □

We conclude the paper with an application of Theorem IV.1. The application we consider is concerned with asymptotically simple (e.g., asymptotically flat and nonsingular) spacetimes as defined by Penrose [17] in terms of the notion of conformal infinity.

Consider a 4-dimensional connected chronological spacetime $M$ with metric $g$ which can be conformally included into a spacetime-with-boundary $M'$ with metric $g'$ such that $M$ is the interior of $M'$, $M = M' \setminus \partial M'$. Regarding the conformal factor, it is assumed that there exists a smooth function $\Omega$ on $M'$ which satisfies,

(i) $\Omega > 0$ and $g' = \Omega^2 g$ on $M$, and

(ii) $\Omega = 0$ and $d\Omega \neq 0$ along $\partial M'$.

The boundary $\mathcal{I} := \partial M'$ is assumed to consist of two components, $\mathcal{I}^+$ and $\mathcal{I}^-$, future and past null infinity, respectively, which are smooth null hypersurfaces. $\mathcal{I}^+$ (respectively, $\mathcal{I}^-$) consists of points of $\mathcal{I}$ which are future (resp., past) endpoints of causal curves in $M$. A spacetime $M$ satisfying the above is said to be *asymptotically flat at null infinity.* If, in addition, $M$ satisfies,

(iii) Every inextendible null geodesic in $M$ has a past end point on $\mathcal{I}^-$ and a future end point on $\mathcal{I}^+$

then $M$ is said to *asymptotically simple* with null conformal boundary. Condition (iii) is imposed to ensure that $\mathcal{I}$ includes all of the null infinity of $M$. It also implies that $M$ is free of singularities which would prevent some null geodesics from reaching infinity.

The notion of asymptotic simplicity was introduced by Penrose in order to facilitate the study of the asymptotic behavior of isolated gravitating systems. When restricting to vacuum (i.e., Ricci flat) spacetimes, asymptotic simplicity provides an elegant and rigorous setting for the study of gravitational radiation far from the radiating source. See the papers [9, 10] of Friedrich for discussion of the problem of global existence of asymptotically simple vacuum spacetimes. Here we prove the following rigidity result.

**Theorem IV.3** *Suppose $M$ is an asymptotically simple vacuum spacetime which contains a null line. Then $M$ is isometric to Minkowski space.*

*Proof.* The first and main step of the proof is to show that $M$ is flat (i.e., has vanishing Riemann curvature). Then certain global arguments will show that $M$ is isometric to Minkowski space.

For technical reasons, it is useful to extend $M'$ slightly beyond its boundary. In fact, by smoothly attaching a collar to $\mathcal{I}^+$ and to $\mathcal{I}^-$, we can extend $M'$ to a spacetime $P$ without boundary such that $M'$ is a retract of $P$ and both $\mathcal{I}^+$ and $\mathcal{I}^-$ separate $P$. It follows that $\mathcal{I}^+$ and $\mathcal{I}^-$ are globally achronal null hypersurfaces in $P$.

Let $M^- = M \cup \mathcal{I}^-$. A straight forward limit curve argument, using the asymptotic simplicity of $M'$, shows that $M^-$ is causally simple. This means that the sets of the form $J^\pm(x, M^-)$, $x \in M^-$, are closed subsets of $M^-$. (The limit curve lemma, in the form of Lemma 14.2 in [3], for example, is valid in $P$.)

Let $\eta$ be a null line in $M$ which has past end point $p \in \mathcal{I}^-$ and future end point $q \in \mathcal{I}^+$. Consider the "future null cone" at $p$, $N_p := \partial I^+(p, M^-)$. From the causal simplicity of $M^-$ it follows that $\partial I^+(p, M^-) = J^+(p, M^-) \setminus I^+(p, M^-)$. Hence each point in $N_p$ can be joined to $p$ by a null geodesic segment in $M^-$. In particular $N_p$ is connected. ¿From the simple equality $I^+(p, M^-) = I^+(\eta, M^-)$, it follows that,

$$N_p = \partial I^+(\eta, M^-) = \partial I^+(\eta, M) \cup \gamma_p = S \cup \gamma_p \,,$$

where $S = \partial I^+(\eta, M)$ and $\gamma_p$ is the future directed null generator of $\mathcal{I}^-$ starting at $p$.

Since asymptotically simple spacetimes are null geodesically complete, Theorem IV.1 implies that $S = \partial I^+(\eta, M)$ is a smooth null hypersurface in $M$ which is totally geodesic with respect to $g$. Since it is smooth and closed the generators of $S$ never cross and never leave $S$ in $M^-$ to the future. Moreover, since $N_p = \partial I^+(p, M^-)$ is achronal, there are no conjugate points to $p$ along the generators of $N_p$. It follows that $N_p \setminus \{p\}$ is the diffeomorphic image under the exponential map $\exp_p : T_pP \to P$ of $(\Lambda_p^+ \setminus \{0\}) \cap \exp_p^{-1}(M^-)$, where $\Lambda_p^+$ is the future null cone in $T_pP$. We are now fully justified in referring to $N_p$ as the future null cone in $M^-$ at $p$.

Given a smooth null hypersurface, with smooth future pointing null normal vector field $K$, the shear tensor $\sigma_{ab}$ is the trace free part of the null second fundamental form $B_{ab}$, $\sigma_{ab} = B_{ab} - \frac{\theta}{2} h_{ab}$. Since $S$ is totally geodesic in the physical metric $g$ and the shear scalar $\sigma = (\sigma_{ab}\sigma^{ab})^{\frac{1}{2}}$ is a conformal invariant, it follows by continuity that the shear tensor of $N_p \setminus \{p\}$ (with respect to an appropriately chosen $g'$- null normal $K'$) vanishes in the metric $g'$, $\sigma'_{ab} = 0$. The trace free part of equation II.2 then implies that the components $C'_{a0b0}$ (with respect to an appropriately chosen pseudo-orthonormal frame in which $e_0 = K'$, cf., Section 4.2 in [13]) of the conformal tensor of $g'$ vanish on $N_p \setminus \{p\}$. An argument of Friedrich in [8], which makes use of the Bianchi identities and, in the present case, the vacuum field equations expressed in terms of his regular conformal field equations, then shows that the so-called rescaled conformal tensor, and hence, the conformal tensor of the physical metric $g$ must vanish on $D^+(N_p, M^-) \cap M$. Since $M$ is Ricci flat, we conclude that $M$ is flat on $D^+(N_p, M^-) \cap M$. In a precisely time-dual fashion $M$ is flat on $D^-(N_q, M^+) \cap M$, where $N_q$ is the past null cone of $M^+ = M \cup \mathcal{I}^+$ at $q$.

To conclude that $M$ is everywhere flat we show that $M \subset D^+(N_p, M^-) \cup D^-(N_q, M^+)$. Consider the set $V = I^+(S, M) \cup S \cup I^-(S, M)$. It is clear from the fact that $S$ is an achronal boundary that $V$ is open in $M$. As $S = \partial I^+(\eta, M)$, it follows that $I^+(S, M) \subset J^+(p, M^-)$, and time-dually, that $I^-(S, M) \subset J^-(q, M^+)$. Using the fact that $J^+(p, M^-)$ and $J^-(q, M^+)$ are closed subsets of $M^-$ and $M^+$, respectively, it follows that $V$ is closed in $M$. Hence, $M = I^+(S, M) \cup S \cup I^-(S, M)$. We show that each term in this union is a subset of $D^+(N_p, M^-) \cup D^-(N_q, M^+)$. Trivially, $S \subset N_p \subset D^+(N_p, M^-)$. Consider $I^+(S, M) \subset J^+(p, P) \cap M = J^+(N_p, P) \cap M$. We claim that $J^+(N_p, P) \cap M \subset D^+(N_p, P) \cap M$. If not, then $H^+(N_p, P) \cap M \neq \emptyset$. Choose a point $x \in H^+(N_p, P) \cap M$, and let $\nu$ be a null generator of $H^+(N_p, P)$ with future end point $x$. Since $N_p$ is edgeless, $\nu$ remains in $H^+(N_p, P)$ as it is extended into the past. By asymptotic simplicity, $\nu$ must meet $\mathcal{I}^-$. In fact, since no portion of $\nu$ can coincide with a generator of $\mathcal{I}^-$, $\nu$ must meet $\mathcal{I}^-$ transversely and then enter $I^-(\mathcal{I}^-, P)$. But this means that $\nu$ has left $J^+(N_p, P)$, which is a contradiction. Since $D^+(N_p, P) \cap M = D^+(N_p, M^-) \cap M$, we have shown that $I^+(S, M) \subset D^+(N_p, M^-)$. By the time-dual argument, $I^-(S, M) \subset D^-(N_q, M^+)$.

Thus, $M$ is globally flat. Also, as an asymptotically simple spacetime, $M$ is null geodesically complete, simply connected and globally hyperbolic; cf., [13], [15]. We use these properties to show that $M$ is geodesically complete. It then follows from the uniqueness of simply connected space forms that $M$ is isometric to Minkowski space.

We first observe that there exists a time orientation preserving local isometry $\phi : M \to L$, where $L$ is Minkowski space. To obtain $\phi$, first construct by a standard procedure a frame $\{e_0, e_1, ..., e_{n-1}\}$ of orthonormal parallel vector fields on $M$. Then solve $dx^i = \langle e_i, \rangle$, $i = 0, ..., n-1$, for functions $x^i \in C^\infty(M)$, and set $\phi = (x^0, x^1, ..., x^{n-1})$.

From the fact that $\phi$ is a local isometry and $M$ is null geodesically complete, it follows that any null geodesic, or broken null geodesic, in $L$ can be lifted via $\phi$ to $M$. In particular it follows that $\phi$ is onto: If $\phi(M)$ is not all of $L$ then we can find a null geodesic $\bar{\eta}$ in $L$ that meets $\phi(M)$ but is not entirely contained in $\phi(M)$. Choose $p \in M$ such that $\phi(p)$ is on $\bar{\eta}$. Then the lift of $\bar{\eta}$ through $p$ is incomplete, contradicting the null geodesic completeness of $M$.

We now show that $M$ is timelike geodesically complete. If it is not, then, without loss of generality, there exists a future inextendible unit speed timelike geodesic $\gamma : [0, a) \to M$, $t \to \gamma(t)$, with $a < \infty$. Let $\bar{\gamma} = \phi \circ \gamma$; $\bar{\gamma}$ can be extended to a complete timelike geodesic in $L$ which we still refer to as $\bar{\gamma}$. Let $\bar{\eta}$ be a future directed broken null geodesic extending from $\bar{p} = \bar{\gamma}(0)$ to $\bar{q} = \bar{\gamma}(a)$. Let $\eta$ be the lift of $\bar{\eta}$ starting at $p = \gamma(0)$; $\eta$ extends to a point $q \in I^+(p)$, with $\phi(q) = \bar{q}$. Since $M$ is globally hyperbolic there exists a maximal timelike geodesic segment $\mu$ from $p$ to $q$. Then $\bar{\mu} = \phi \circ \mu$ is a timelike geodesic segment in $L$ from $\bar{p}$ to $\bar{q}$. Hence, $\bar{\mu} = \bar{\gamma}|_{[0,a]}$. It follows that $\mu$ extends $\gamma$ to $q$, which is a contradiction.

Finally, we show that $M$ is spacelike geodesically complete. If it is not, then there exists an inextendible unit speed spacelike geodesic $\gamma : [0, a) \to M$, $t \to \gamma(t)$, with $a < \infty$. Let $\bar{\gamma} = \phi \circ \gamma$; $\bar{\gamma}$ can be extended to a complete spacelike geodesic in $L$ which we still refer to as $\bar{\gamma}$. We now use the fact that timelike geodesics, and broken timelike geodesics, in $L$ can be lifted via $\phi$ to $M$. Let $\bar{\alpha} = \bar{\alpha}_1 + \bar{\alpha}_2$ be a two-segment broken timelike geodesic extending from $\bar{p} = \bar{\gamma}(0)$ to $\bar{q} = \bar{\gamma}(a)$, with $\bar{\alpha}_1$ starting at $\bar{p}$ and past pointing. Similarly, let $\bar{\beta} = \bar{\beta}_1 + \bar{\beta}_2$ be a two-segment broken timelike geodesic extending from $\bar{p}$ to $\bar{q}$, with $\bar{\beta}_1$ starting at $\bar{p}$ and future pointing. Let $\bar{x}$ be the point at the corner of $\bar{\alpha}$, and let $\bar{y}$ be the point at the corner of $\bar{\beta}$. Note that $\bar{\gamma}|_{[0,a]} \subset I^+(\bar{x}) \cap I^-(\bar{y})$. Let $\alpha$ be the lift of $\bar{\alpha}$ starting at $p = \gamma(0)$; $\alpha$ extends to a point $q$ with $\phi(q) = \bar{q}$. Similarly, let $\beta$ be the lift of $\bar{\beta}$ starting at $p$. Let $x$ be the point at the corner of $\alpha$, and let $y$ be the point at the corner of $\beta$.

Note that an initial segment of $\gamma$ is contained in $J^+(x) \cap J^-(y)$. We claim that $\gamma$ is entirely contained in $J^+(x) \cap J^-(y)$. If not then $\gamma$ either leaves $J^+(x)$ or $J^-(y)$. Suppose it leaves $J^+(x)$ at $z \in \gamma \cap \partial J^+(x)$. Since $M$ is globally hyperbolic, there exists a null geodesic segment $\eta$ from $x$ to $z$. Then $\bar{\eta} = \phi \circ \eta$ is a null geodesic in $L$ from $\bar{x}$ to $\bar{z} \in \bar{\gamma}|_{[0,a]}$. But since, by construction, $\bar{\gamma}|_{[0,a]} \subset I^+(\bar{x})$, no such null geodesic exists.

Thus, $\gamma \subset J^+(x) \cap J^-(y)$. We show that $\gamma$ extends to $q$, thereby obtaining a contradiction. Consider a sequence $\gamma(t_n)$, $t_n \to a$. Since $J^+(x) \cap J^-(y)$ is compact, there exists a subsequence $\gamma(t_{n_k})$ which converges to a point $q' \in J^+(x) \cap J^-(y)$. Since $\bar{\gamma}(t_n) \to \bar{q}$, it follows by continuity that $\phi(q') = \bar{q}$. Let $\mu$ be a causal geodesic segment from $x$ to $q'$. Then $\bar{\mu} = \phi \circ \mu$ is a causal geodesic from $\bar{x}$ to $\bar{q}$ in $L$. Thus, $\bar{\mu} = \bar{\alpha}_2$, and hence $\mu = \alpha_2$. Since $\alpha_2$ has future end point $q$, we conclude that $q' = q$. Hence, since every sequence $\gamma(t_n)$, $t_n \to a$, has a subsequence converging to $q$, it follows that $\lim_{t \to a} \gamma(t) = q$, and so $\gamma$ extends to $q$. This concludes the proof that $M$ is geodesically complete and hence, as noted above, isometric to Minkowski space.

We remark in closing that Theorem IV.3 can be generalized in various directions. For example, it is possible to formulate a version of Theorem IV.3 for asymptotically flat spacetimes which contain singularities, black holes, etc., by imposing suitable conditions on the domain of outer communications $D = I^-(\mathcal{I}^+) \cap I^+(\mathcal{I}^-)$, the conclusion then being that $D$ is flat. Also, it appears that Theorem IV.3 can be extended to vacuum spacetimes with positive cosmological constant, $\Lambda > 0$, thereby yielding a characterization of de Sitter space. Details of this will appear elsewhere.

## Acknowledgement

The author is indebted to Helmut Friedrich for many helpful discussions concerning the proof of Theorem IV.3. The author would also like to thank Lars Andersson and Piotr Chruściel for many valuable comments and suggestions. Part of the work on this paper was carried out during a visit to the Royal Institute of Technology in Stockholm, Sweden in 1999. The author wishes to thank the Institute for its hospitality and support.


## References


[1] A. D. Alexandrov, Some theorems on partial differential equations of second order, *Vestnik Leningrad. Univ. Ser. Mat. Fiz. Him.* **9** (1954), no. 8, 3–17.

[2] L. Andersson, G.J. Galloway, and R. Howard, A strong maximum principle for weak solutions of quasi-linear elliptic equations with applications to Lorentzian and Riemannian geometry, *Comm. Pure Appl. Math.* **51** (1998), 581–624.

[3] J. K. Beem, P. E. Ehrlich, and K. L.Easley, Global Lorentzian geometry, 2 ed., Pure and Applied Mathematics, vol. 202, Marcel Dekker, New York, 1996.

[4] P.T. Chruściel and G.J. Galloway, Horizons non–differentiable on dense sets, *Commun. Math. Phys.* **193** (1998), 449–470, gr-qc/9611032.

[5] P. T. Chruściel, E. Delay, G. J. Galloway, and R. Howard, The area theorem, preprint, 1999.

[6] J.-H. Eschenburg, Comparison theorems and hypersurfaces, *Manuscripta Math.* **59** (1987), no. 3, 295–323.

[7] ______, Maximum principle for hypersurfaces, *Manuscripta Math.* **64** (1989), 55–75.

[8] H. Friedrich, Existence and structure of past asymptotically simple solutions of Einstein's field eqautions with positive cosmological constant, *J. Geom. Phys.* **3** (1986) 101–117.

[9] ______, Einstein equations and conformal structure, in *The Geometric universe: Science, geometry and the work of Roger Penrose*, eds. S. A. Huggett et al., Oxford University Press, Oxford.

[10] ______, Einstein's equations and geometric asymptotics, in Gravitation and relativity: At the turn of the millennium. Proceedings of the GR-15 Conference at Pune, India, in June 1997, eds. N. Dadich and J. Narlikar, IUCAA, Pune, India.

[11] J.L. Friedman, K. Schleich, and D.M. Witt, Toplogical censorship, *Phys. Rev. Lett.* **71** (1993) 1486–1489.

[12] G.J. Galloway and E. Woolgar, The cosmic censor forbids naked topology, *Classical and Quantum Grav.* **14** (1997) L1–L7.

[13] S.W. Hawking and G.F.R. Ellis, The large scale structure of space-time, Cambridge University Press, Cambridge, 1973.

[14] D.N. Kupeli, On null submanifolds in spacetimes, *Geom. Dedicata* **23** (1987), 33–51.

[15] R.P.A.C. Newman, The global structure of simple space-times, *Commun. Math. Phys.* **123** (1989), 17–52.

[16] B. O'Neill, Semi–Riemannian geometry, Academic Press, New York, 1983.

[17] R. Penrose, Zero rest-mass fields including gravitation: asymptotic behavior, *Proc. Roy. Soc. Lond.* A **284** 159–203.

[18] ———, Techniques of differential topology in relativity, SIAM, Philadelphia, 1972, (Regional Conf. Series in Appl. Math., vol. 7).

Gregory J. Galloway
Supported in part by NSF grant # DMS-9803566.
University of Miami, Department of Mathematics and
Computer Science,
Coral Gables, FL, 33124
email :galloway@math.miami.edu

Communicated by S. Klainerman
submitted 13/09/99; accepted 25/09/99