GEOMETRY IN POTSDAM

Christian Bär

# Lorentzian Geometry

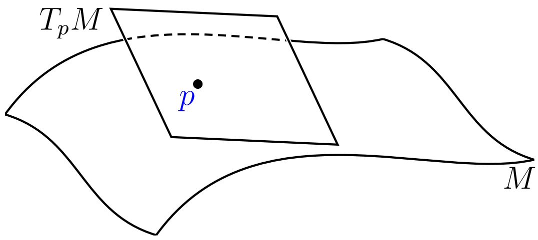
## Lecture Notes, Summer Term 2004

Version of January 7, 2020

# Contents

# Preface

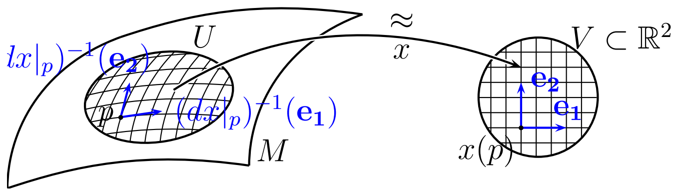
A Lorentzian manifold is a pair $(M^{n+1}, g)$, where $M^{n+1}$ is a $(n+1)$-dimensional differentiable manifold and $g$ a Lorentz metric such that $g$ assigns to each point $p \in M$ a non-degenerate symmetric bilinear form $g_p$ of index $1$ on $T_pM$.




Here "index $1$" means that there is a basis $e_0, \ldots, e_n$ of $T_pM$ such that

$$g_p(e_i, e_j) := \begin{cases} -1, & i = j = 0, \\ 1, & \text{falls } i = j = 1, \ldots, n, \\ 0, & \text{otherwise.} \end{cases}$$

The metric $g$ depends smoothly ($C^\infty$) on $p$, i.e. for local coordinates $x^0, \ldots, x^n$ on $M$, the functions $g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) =: g_{ij}$ are smooth.




One writes

$$g = \sum_{i,j=0}^{n} g_{ij} dx^i \otimes dx^j.$$

In general relativity, the spacetime (i.e. the set of all events) is modelled by a four-dimensional Lorentzian manifold. The Einstein equations on such spacetimes are of the form:

$$\text{curvature expression} = \text{energy-momentum-tensor.}$$

This equation provides the following correspondence:

$$\text{world lines of massless particles (e.g. photons)} = \text{lightlike geodesics}$$

$$\text{world lines of particles with mass} = \text{timelike geodesics}$$

The aim of the lecture is to investigate the global properties of Lorentz manifolds.

# 1 Important Examples

## 1.1 The Minkowski space

**Notation 1.1.** On $\mathbb{R}^n$, we define the **Euclidean scalar product** via

$$\langle x, y\rangle := \sum_{i=1}^{n} x^i y^i.$$

On $\mathbb{R}^{n+1}$, we define the **Minkowski scalar product** via

$$\langle\!\langle x, y\rangle\!\rangle := -x^0 y^0 + \sum_{i=1}^{n} x^i y^i.$$

We then denote the "Euclidean part" by $\hat{x} := (x^1, \dots, x^n)$.

**Definition 1.2.** An $(n+1)$-dimensional Lorentzian manifold is called **Minkowski space** if it is isometric to $\left(\mathbb{R}^{n+1}, \langle\!\langle \cdot, \cdot \rangle\!\rangle\right)$.

**Remark 1.3.** The Minkowski space is the spacetime of special relativity. Here, all $g_{ij}$ are constant, that is all Christoffel symbols and thus the curvature tensor $R$ vanish. The geodesic equation reads $\ddot{c}^i = 0$, where $c^i := x^i \circ c$, so the geodesics are the affin-linearly parametrized straight lines.

**Definition 1.4.** The **light cone** is defined as the set

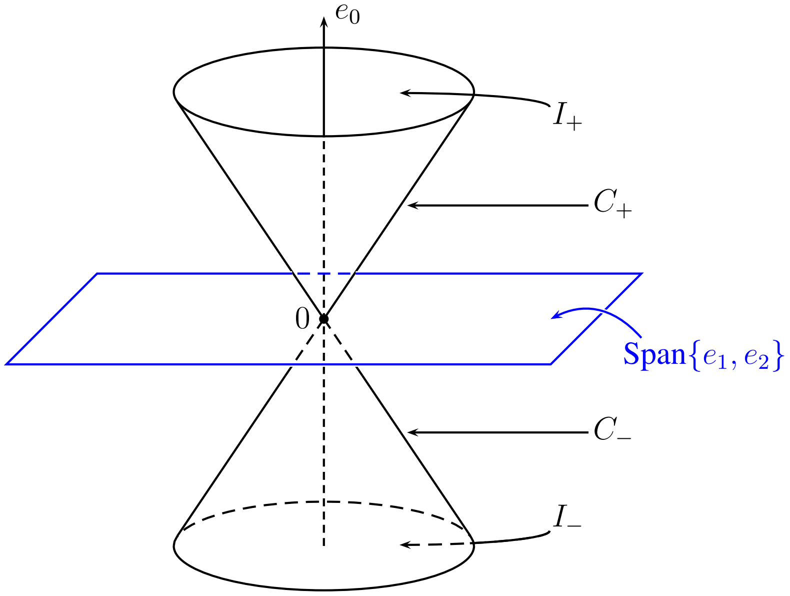
$$C := \left\{x \in \mathbb{R}^{n+1} \mid \langle\!\langle x, x\rangle\!\rangle = 0\right\}.$$

**Definition 1.5.** A vector $x \in \mathbb{R}^{n+1} \backslash \{0\}$ is called **timelike** or **lightlike** or **spacelike** if $\langle\!\langle x, x\rangle\!\rangle < 0$ or $\langle\!\langle x, x\rangle\!\rangle = 0$ or $\langle\!\langle x, x\rangle\!\rangle > 0$. We call $x$ **causal** if it is timelike or lightlike. The zero vector is considered as spacelike.

**Definition 1.6.** The set of timelike vector $I$ consists of two connected components and we choose a **time-orientation** by picking one component $I_+$ and call its members **future directed**. The members of the other component $I_-$ are called **past directed**. Correspondingly, we define

$$C_\pm := \partial I_\pm, \qquad J_\pm := \overline{I_\pm}.$$

The sets of future/past directed lightlike and causal vectors are given by $C_\pm \backslash \{0\}$ and $J_\pm \backslash \{0\}$, respectively. Moreover, the set of spacelike vectors equals $\mathbb{R}^{n+1} \backslash J \cup \{0\}$.




**Lemma 1.7.** *Let $x, y \in \mathbb{R}^{n+1}$ and $t > 0$.*

*(i) For $A \in \{I_\pm, C_\pm, J_\pm, \mathbb{R}^{n+1} \backslash J_\pm\}$, we have:* $\quad x \in A \implies tx \in A.$

*(ii) For $A \in \{I_\pm, J_\pm\}$, we have:* $\quad x, y \in A \implies x + y \in A.$

*Proof.* (i) Since $\langle\!\langle tx, tx \rangle\!\rangle = t^2 \langle\!\langle x, x \rangle\!\rangle$, the sign of $\langle\!\langle x, x \rangle\!\rangle$ and hence the causal type of $x$ is preserved. Moreover, due to $t > 0$, the orientation of causal vectors does not change.

(ii) Clearly, if $x, y$ are both future/past directed, then so is $x + y$, so we just check the causal type. For $x, y$ timelike and equally oriented, the Cauchy-Schwarz-inequality provides

$$\begin{aligned}\left(x^0 + y^0\right)^2 &= \left(x^0\right)^2 + 2x^0y^0 + \left(y^0\right)^2 > \|\hat{x}\|^2 + 2 \cdot \|\hat{x}\| \cdot \|\hat{y}\| + \|\hat{y}\|^2 \\ &\geq \|\hat{x}\|^2 + 2\langle \hat{x}, \hat{y} \rangle + \|\hat{y}\|^2 = \|\hat{x} + \hat{y}\|^2 = \|\widehat{x+y}\|^2.\end{aligned}$$

The proof for causal vectors is obtained by replacing ">" by "$\geq$". $\square$

**Corollary 1.8.** *The subsets $I_\pm, J_\pm$ of $\mathbb{R}^{n+1}$ are convex.*

*Proof.* For $A \in \{I_\pm, J_\pm\}$, let $x, y \in A$ and $t \in (0,1)$, i.e. $t, (1-t) > 0$. Then $tx, (1-t)y \in A$ by (i), and $tx + (1-t)y \in A$ by (ii) of Lemma 1.7. □

**Definition 1.9.** A map $\phi\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is called **Lorentz transformation** if

$$\langle\!\langle \phi(x), \phi(y) \rangle\!\rangle = \langle\!\langle x, y \rangle\!\rangle\,, \qquad x, y \in \mathbb{R}^{n+1}.$$

**Definition 1.10.** A basis $b_0, \dots, b_n$ of $\mathbb{R}^{n+1}$ is called **Lorentz-orthonormal**, if

$$\langle\!\langle b_0, b_0 \rangle\!\rangle = -1, \qquad \langle\!\langle b_0, b_i \rangle\!\rangle = 0, \qquad \langle\!\langle b_i, b_j \rangle\!\rangle = \delta_{ij}, \qquad i, j = 1, \dots, n.$$

**Example 1.11.** The standard basis $e_0 = (1, 0, \dots, 0)^t, \dots, e_n = (0, \dots, 0, 1)^t$ is a Lorentz orthonormal basis of $\mathbb{R}^{n+1}$.

**Proposition 1.12.** *A map $\phi\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is a Lorentz transformation if and only if it is linear and $\phi(e_0), \dots, \phi(e_n)$ is a Lorentz orthonormal basis.*

*Proof.* "$\Longrightarrow$": Let $\phi$ be a Lorentz transformation, then by definition, we have that $\phi(e_0), \dots, \phi(e_n)$ is a Lorentz orthonormal basis, so we just have to show linearity. Let $v = \sum_{i=0}^n v^i e_i \in \mathbb{R}^{n+1}$ and $\phi(v) = \sum_{i=0}^n w^i \phi(e_i)$, so that

$$\langle\!\langle v, e_j \rangle\!\rangle = \sigma_j v^j, \qquad \langle\!\langle \phi(v), \phi(e_j) \rangle\!\rangle = \sigma_j w^j,$$

where $\sigma_0 = -1$ and $\sigma_j = 1$ if $j = 1, \dots, n$. Therefore,

$$\phi(v) = \sum_{i=0}^n \sigma_i \langle\!\langle \phi(v), \phi(e_i) \rangle\!\rangle\, \phi(e_i) = \sum_{i=0}^n v^i \phi(e_i).$$

"$\Longrightarrow$": For $\phi$ linear and $\phi(e_0), \dots, \phi(e_n)$ a Lorentz orthonormal basis directly follows

$$\langle\!\langle \phi(x), \phi(y) \rangle\!\rangle = \sum_{i,j=0}^n x^i y^i \langle\!\langle \phi(e_i), \phi(e_j) \rangle\!\rangle = -x^0 y^0 + \sum_{i=1}^n x^i y^i = \langle\!\langle x, y \rangle\!\rangle\,.$$ □

Let $J_n := \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1}_n \end{pmatrix}$, and hence $\langle\langle x, y\rangle\rangle = \langle x, J_n y\rangle$ for all $x, y \in \mathbb{R}^{n+1}$. It follows that a matrix $\Lambda$ represents a Lorentz transformation if

$$\langle x, J_n y\rangle = \langle\langle x, y\rangle\rangle = \langle\langle \Lambda x, \Lambda y\rangle\rangle = \langle \Lambda x, J_n \Lambda y\rangle = \left\langle x, \Lambda^t J_n \Lambda y\right\rangle,$$

that is, if and only if $J_n = \Lambda^t J_n \Lambda$. Thus, we just showed

**Proposition 1.13.** *Let $\Lambda \in \mathrm{Mat}(n+1, \mathbb{R})$. Then the following statements are equivalent:*

- *$\Lambda$ represents a Lorentz transformation.*
- *The columns of $\Lambda$ yield a Lorentz orthonormal basis of $\mathbb{R}^{n+1}$.*
- *$J_n = \Lambda^t J_n \Lambda$.*

**Proposition 1.14.**

*(i) Let $\mathscr{L}(n+1)$ denote the set of all Lorentz transformations on $\mathbb{R}^{n+1}$. Then $(\mathscr{L}, \circ)$ is a group.*

*(ii) The matrix $\Lambda$ of a Lorentz transformation satisfies $\det \Lambda = \pm 1$.*

*Proof.* (i): By Proposition 1.12, every Lorentz transformation is an isomorphism and simple calculations show that $\mathrm{id}, \phi^{-1}, \phi \circ \psi \in \mathscr{L}$ for all $\phi, \psi \in \mathscr{L}$. (ii): Proposition 1.13 (iii) implies

$$1 = -\det J_n = -\det\left(\Lambda^t J_n \Lambda\right) = \left(\det \Lambda\right)^2 \qquad \Longrightarrow \qquad \det \Lambda = \pm 1. \qquad \square$$

**Example 1.15.**

1. Let $A$ be an orthogonal $n \times n$-matrix, i.e. $A^t A = \mathbb{1}_n$. Then $\Lambda := \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$ represents a Lorentz transformation, since

$$\Lambda^t J_n \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & A^t \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1}_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & A^t A \end{pmatrix} = J_n.$$

2. Let $\Lambda := \begin{pmatrix} \cosh\alpha & \sinh\alpha & 0 \\ \sinh\alpha & \cosh\alpha & 0 \\ 0 & 0 & \mathbb{1}_{n-1} \end{pmatrix}$, that is the matrix of a **Lorentz boost**. Due to $\cosh^2\alpha - \sinh^2\alpha = 1$, this represents a Lorentz transformation.

3. The matrices $J_n$ and $-\mathbb{1}_{n+1}$ represent Lorentz transformations.

**Theorem 1.16.** *For all $x, y \in \mathbb{R}^{n+1}$, the following statements are equivalent:*

*(i)* $\langle\langle x, x\rangle\rangle = \langle\langle y, y\rangle\rangle$

*(ii) There is some $\phi \in \mathscr{L}(n+1)$ such that $\phi(x) = y$.*

*Proof.* "$\Leftarrow$": $\langle\langle y, y\rangle\rangle = \langle\langle \phi(x), \phi(x)\rangle\rangle = \langle\langle x, x\rangle\rangle$.
"$\Rightarrow$": Let $x$ and thus $y$ be timelike, and set $-c^2 := \langle\langle x, x\rangle\rangle = \langle\langle y, y\rangle\rangle$, where $c > 0$. It suffices to consider $y = ce_0$. Without loss of generality, let $x$ be future directed (otherwise replace it by $J_n x$), i.e. $x = \left(x^0, \hat{x}\right)^t$ with $x^0 > 0$. Since there is some $A \in O(n)$ such that $A\hat{x} = \lambda e_1$, we can assume that $x = \left(x^0, x^1, 0, \dots, 0\right)^t$, so we have

$$-(x^0)^2 + (x^1)^2 = \langle\langle x, x\rangle\rangle = \langle\langle y, y\rangle\rangle = -c^2 \qquad \Longrightarrow \qquad -\left(\frac{x^0}{c}\right)^2 + \left(\frac{x^1}{c}\right)^2 = -1.$$

The solutions of this equation yield a hyperbola, for which we have the parametrization $\alpha \mapsto \left(\sinh\alpha, \cosh\alpha\right)$. Hence, there is some $\alpha \in \mathbb{R}$ such that $\left(\frac{x^0}{c}, \frac{x^1}{c}\right) = \left(\sinh\alpha, \cosh\alpha\right)$, i.e.

$$\Lambda x = \begin{pmatrix} \cosh\alpha & \sinh\alpha & 0 \\ \sinh\alpha & \cosh\alpha & 0 \\ 0 & 0 & \mathbb{1}_{n-1} \end{pmatrix} \begin{pmatrix} c\cosh\alpha \\ c\sinh\alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix} = y.$$

$\square$

**Lemma 1.17 (Cauchy-Schwarz inequality).** *Let $z$ be timelike and $x, y \in \mathbb{R}^{n+1}$ such that $\langle\langle x, z\rangle\rangle = 0 = \langle\langle y, z\rangle\rangle$. Then the Cauchy Schwarz inequality holds:*

$$|\,\langle\langle x, y\rangle\rangle\,| \leq \sqrt{|\,\langle\langle x, x\rangle\rangle\,|}\sqrt{|\,\langle\langle y, y\rangle\rangle\,|}.$$

*Furthermore, equality holds if and only if $x, y$ are linearly dependent.*

*Proof.* Without loss of generality, we can assume that $z = ce_0$ for some $c \neq 0$. Then by assumption, we have $x^0 = 0 = y^0$, and hence $\langle\langle x, y\rangle\rangle = \langle\hat{x}, \hat{y}\rangle$, so the statement follows from the familiar Cauchy Schwarz inequality for $\langle\cdot,\cdot\rangle$. $\square$

**Lemma 1.18 (inverse Cauchy-Schwarz inequality).** *For $x, y \in I_+$, we have*

$$|\,\langle\langle x, y\rangle\rangle\,| \geq \sqrt{|\,\langle\langle x, x\rangle\rangle\,|}\sqrt{|\,\langle\langle y, y\rangle\rangle\,|}.$$

*Furthermore, equality holds if and only if $x, y$ are linearly dependent.*

*Proof.* By application of some Lorentz transformation, without loss of generality, we can assume that $y = ce_0$ for some $c > 0$. Then we have

$$\langle\!\langle x,y\rangle\!\rangle^2 - |\,\langle\!\langle x,x\rangle\!\rangle\,\langle\!\langle y,y\rangle\!\rangle\,| = c^2\big((x^0)^2 - |\,\langle\!\langle x,x\rangle\!\rangle\,|\big) = c^2\|\hat{x}\|^2 \geq 0.$$

Moreover, linear dependence is equivalent to $\hat{x} = 0$, for which equality holds. □

Note that if $\langle\!\langle x,y\rangle\!\rangle = 0$ for $x$ timelike, the inverse Cauchy Schwarz inequality particularly implies that $y$ is spacelike. On the other hand for $n \geq 3$, there are spacelike $x, y$ such that $\langle\!\langle x,y\rangle\!\rangle = 0$.

**Proposition 1.19.** *Let $\Lambda = (\lambda_{ij})_{i,j=0,\dots,n}$ represent a Lorentz transformation. The the following statements are equivalent:*

*(i)* $\lambda_{00} > 0$

*(ii)* $\Lambda e_0 \in I_+$

*(iii)* $\Lambda(I_+) \subset I_+$

*Proof.* By definition, Lorentz transformations do not change the causal type of vectors since $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ is preserved, and hence $\Lambda e_0 \in I$. On the other hand, $\big(\Lambda e_0\big)^0 = a_{00}$, that is, $(i)$ and $(ii)$ are equivalent. Furthermore, the implication $(iii) \Rightarrow (ii)$ is trivial.
$(ii) \Rightarrow (iii)$: Let $x \in I_+$ and hence $c(t) := tx + (1-t)e_0 \in I_+$ for all $t \in [0,1]$ due to convexity of $I_+$ (Corollary 1.8). Then from $(ii)$ follows $\Lambda c(t) \in I_+$ and thus $\Lambda x \in I_+$ again by convexity. □

In the following, we will not distinguish between Lorentz transformations and the matrices representing them.

**Definition 1.20.** The elements of the subset $\mathscr{L}^{\uparrow}(n+1) := \{\lambda_{00} > 0\} \subset \mathscr{L}(n+1)$ are called time-orientation preserving Lorentz transformations.

**Corollary 1.21.** *$\mathscr{L}^{\uparrow}(n+1)$ yields a subgroup of $\mathscr{L}(n+1)$, and for all $\Lambda \in \mathscr{L}^{\uparrow}(n+1)$, we have $\Lambda(I_\pm) = I_\pm$.*

*Proof.* Let $\Lambda_1, \Lambda_2 \in \mathscr{L}^{\uparrow}(n+1)$, then $\Lambda_1\big(\Lambda_2(I_+)\big) \subset \Lambda_1(I_+) \subset I_+$ by Proposition 1.19, and hence $\Lambda_1\Lambda_2 \in \mathscr{L}^{\uparrow}(n+1)$. Furthermore, if $\Lambda^{-1}e_0 \in I_-$ for $\Lambda \in \mathscr{L}^{\uparrow}(n+1)$, we had $\Lambda^{-1}(-e_0) = -\Lambda^{-1}(e_0) \in I_+$, that is, $-e_0 = \Lambda\big(-\Lambda^{-1}e_0\big) \in I_+$, which is a contradiction due

to $I_+ \cap I_- = \emptyset$. It follows that $\Lambda^{-1}e_0 \in I_+$, i.e. $\Lambda^{-1} \in \mathscr{L}^\uparrow(n+1)$, and hence $\mathscr{L}^\uparrow(n+1)$ is a subgroup of $\mathscr{L}(n+1)$ since clearly $\mathbb{1}_{n+1} \in \mathscr{L}^\uparrow(n+1)$.
Moreover, the equality can be directly seen via

$$I_+ = \Lambda^{-1}\Lambda(I_+) \subset \Lambda^{-1}(I_+) \subset I_+,$$

and similarly $\Lambda(I_-) = I_-$ by $I_- = -I_+$ and linearity of $\Lambda$. □

**Example 1.22.** 1. Let $A \in O(n)$, then $\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \in \mathscr{L}^\uparrow(n+1)$.

2. Let $A \in O(n)$, then $\begin{pmatrix} -1 & 0 \\ 0 & A \end{pmatrix} \in \mathscr{L}^\downarrow(n+1) := \mathscr{L}(n+1)\backslash\mathscr{L}^\uparrow(n+1)$.

3. All Lorentz-Boosts (Example 1.15, 2) are in $\mathscr{L}^\uparrow(n+1)$.

We define the following subsets of $\mathscr{L}(n+1)$:

$$\mathscr{L}_\pm(n+1) := \left\{\Lambda \in \mathscr{L}(n+1) \;\middle|\; \det\Lambda = \pm 1\right\},$$

$$\mathscr{L}_\pm^\uparrow := \mathscr{L}^\uparrow(n+1) \cap \mathscr{L}_\pm(n+2), \qquad \mathscr{L}_\pm^\downarrow(n+1) := \mathscr{L}^\downarrow(n+1) \cap \mathscr{L}_\pm(n+1).$$

**Definition 1.23.** A map $\psi\colon\ \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is called **Poincaré transformation**, if it is of the form

$$\psi(x) = \Lambda x + b$$

for some $\Lambda \in \mathscr{L}(n+1)$ and $b \in \mathbb{R}^{n+1}$.
We denote the set of all Poincaré transformations of $\mathbb{R}^{n+1}$ by $\mathscr{P}(n+1)$.

**Proposition 1.24.** *The isometry group of Minkowski space is given by* $\mathscr{P}(n+1)$.

*Proof.* Let $\eta := -\,\mathrm{d}x^0 \otimes \mathrm{d}x^0 + \sum_{i=1}^n \mathrm{d}x^i \otimes \mathrm{d}x^i$ the standard Minkowski metric on $\mathbb{R}^{n+1}$. For all $\psi \in \mathscr{P}(n+1)$, we have $\mathrm{d}\psi = \Lambda$ and each Lorentz transformation is an isometry of the Minkowski space by definition. Therefore, $\mathscr{P}(n+1) \in \mathrm{Isom}\big(\mathbb{R}^{n+1}_{\mathrm{Mink}}\big)$.
Let $\psi \in \mathrm{Isom}\big(\mathbb{R}^{n+1}_{\mathrm{Mink}}\big)$ and $c(t) = \exp_x(tX)$, i.e. the unique geodesic with $c(0) = p$ and $\dot{c}(0) = X$. Then $\psi\circ c$ is a geodesic as well with $(\psi\circ c)(0) = \psi(x)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}(\psi\circ c) = \mathrm{d}\psi_x(X)$, that is

$$\psi\big(\exp_x(tX)\big) = \exp_{\psi(x)}\big(t\cdot \mathrm{d}\psi_x(X)\big). \tag{1.1}$$

Recall that $X \mapsto x + X$ represents the canonical isomorphism $T_x\mathbb{R}^{n+1} \cong \mathbb{R}^{n+1}$, so (1.1) provides for $x = 0$:

$$\psi(X) = \psi\big(\exp_0(X)\big) = \exp_{\psi(0)}\big(\mathrm{d}\psi_0(X)\big) = \mathrm{d}\psi_0(X) + \psi(0) =: \Lambda X + b.$$

Note that $\mathrm{d}\psi_0 \in \mathscr{L}(n+1)$ since $\langle\!\langle\, \mathrm{d}\psi_0(X),\, \mathrm{d}\psi_0(X)\rangle\!\rangle = \langle\!\langle X, X\rangle\!\rangle$ as an isometry. □

We define the analogous subsets $\mathscr{P}_{\pm}^{\uparrow,\downarrow}(n+1)$ of $\mathscr{P}(n+1)$ by correspondingly restricting $\mathrm{d}\psi$ to $\mathscr{L}_{\pm}^{\uparrow,\downarrow}$.

**Example 1.25.** Apart from Minkowski space, there are several other flat Lorentzian manifolds:

- Open subsets of Minkowski space
- The quotients

$$T^{n+1} := \mathbb{R}^{n+1}/\mathbb{Z}^{n+1}, \quad S^1 \times \mathbb{R}^n = \mathbb{R}^{n+1}/\mathbb{Z}e_0, \quad \mathbb{R} \times T^n = \mathbb{R}^{n+1}/\mathbb{Z}e_1 \oplus \ldots \oplus \mathbb{Z}e_n.$$

  Here the quotient $\mathbb{R}^{n+1}/\sim$ is equipped with the unique Lorentz metric, with respect to which $\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}/\sim$ is a local isometry.

## 1.2 The de Sitter space

Consider the smooth function $x \mapsto \gamma(x) := \langle\langle x, x \rangle\rangle$ with $\mathrm{d}\gamma_x = -2x^0\,\mathrm{d}x^0 + 2\sum_{i=1}^n x^i\,\mathrm{d}x^i$. In particular, all $x \neq 0$ are regular points for $\gamma$ and all $c \in \mathbb{R}\backslash\{0\}$ regular values.

**Definition 1.26.** For fixed $r > 0$ the hypersurface

$$S_1^n(r) := \gamma^{-1}(r^2) \subset \mathbb{R}^{n+1}$$

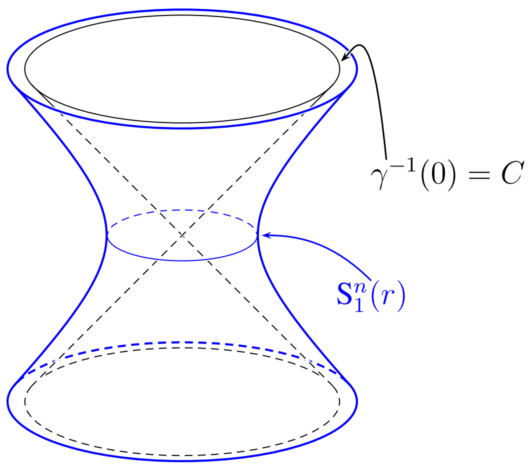
is called $n$**-dimensional de Sitter space**.

As a differentiable manifold, $S_1^n(r)$ is diffeomorphic to $\mathbb{R} \times S^{n-1}$ via

$$S_1^n(r) \longrightarrow \mathbb{R} \times S^{n-1},$$

$$(x^0, \hat{x}) \longmapsto \left( x^0, \frac{\hat{x}}{\sqrt{r^2 + (x^0)^2}} \right),$$

where $S^{n-1}$ denotes the $(n-1)$-dimensional standard sphere. The inverse map is then given by $(y^0, \hat{y}) \mapsto \left(y^0, \sqrt{(y^0)^2 + r^2} \cdot \hat{y}\right)$.




As a hypersurface, $S_1^n(r)$ possesses a trivial normal bundle spanned by

$$\operatorname{grad}_x \gamma = -2x^0 \left( -\frac{\partial}{\partial x^0} \right) + 2\sum_{i=1}^n x^i \frac{\partial}{\partial x^i} = 2\sum_{i=0}^n x^i \frac{\partial}{\partial x^i}. \tag{1.2}$$

The normal bundle is spacelike since

$$\eta\big(\mathrm{grad}_x\gamma, \mathrm{grad}_x\gamma\big) = 4\gamma(x) = 4r^2 > 0,$$

so the Minkowski metric $\eta$ induces a Lorentz metric $g$ on $S_1^n(r)$. Furthermore, a unit normal field is given by $\nu := \frac{1}{2r}\mathrm{grad}\gamma$, which by (1.2) induces a map

$$S_1^n(r) \longrightarrow S_1^n(1), \qquad \nu(x) = \frac{x}{r}. \tag{1.3}$$

By identifying $T_{\nu(x)}S_1^n(1) = \nu(x)^\perp = T_xS_1^n(r)$, we consider $\mathrm{d}\nu_x$ as the endomorphism $\frac{1}{r}\mathrm{id}\colon T_xS_1^n(r) \to T_xS_1^n(r)$, which is the **shape operator of** $T_xS_1^n(r)$ **at** $x$ given by $W(X) = \frac{X}{r}$. Here and in the following $X, Y, Z$ denote vector fields on $S_1^n(r)$. For the curvature, Gauß' formula provides

$$\begin{aligned} R^{S_1^n(r)}(X,Y)Z &= \overbrace{R^{\mathbb{R}^{n+1}_{\mathrm{Mink}}}(X,Y)Z}^{=0} + g(W(X),Z)W(X) - g(W(X),Z)W(Y) \\ &= \frac{1}{r^2}\big(g(Y,Z)X - g(X,Z)Y\big), \end{aligned}$$

and for $E$ a non-degenerate plane spanned by the vectors $X, Y$, we obtain that the sectional curvature of $S_1^n(r)$ is constant:

$$K(E) = \frac{g\big(R(X,Y)Y,X\big)}{g(X,X)g(Y,Y) - g(X,Y)^2} = \frac{\frac{1}{r^2}g\big(g(Y,Y)X - g(X,Y)Y, X\big)}{g(X,X)g(Y,Y) - g(X,Y)^2} = \frac{1}{r^2}.$$

Let $(e_0, \dots, e_{n-1})$ be a Lorentz orthonormal basis of $T_xS_1^n(r)$ and set $\varepsilon_i := g(e_i, e_i) \in \{-1, 1\}$. It follows for the Ricci curvature

$$\begin{aligned} \mathrm{ric}(X,Y) &= \sum_{i=0}^{n-1} \varepsilon_i g(R(X,e_i)e_i, Y) = \frac{1}{r^2}\sum_{i=0}^{n-1}\varepsilon_i g\big(g(e_i,e_i)X - g(X,e_i)e_i, Y\big) \\ &= \frac{1}{r^2}\big(ng(X,Y) - g(X,Y)\big) = \frac{n-1}{r^2}g(X,Y) \end{aligned}$$

and therefore

$$\mathrm{ric} = \frac{n-1}{r^2}g, \qquad \mathrm{scal} = \frac{n(n-1)}{r^2}.$$

**Remark 1.27.** For $n = 4$, we obtain the Einstein tensor $G = \mathrm{ric} - \frac{1}{2}\mathrm{scal}\cdot g = -\frac{3}{r}g$, so the Lorentzian manifold $S_1^4(r)$ is a vacuum solution of Einstein's field equations with cosmological constant $\frac{3}{r^2}$.

We proceed with the geodesics of $S_1^n(r)$. For given $x \in S_1^n(r)$ and $X \in T_xS_1^n(r)\backslash\{0\}$, i.e. $\langle\langle x, X\rangle\rangle = 0$ due to (1.3), let $E \subset \mathbb{R}^{n+1}$ be the plane spanned by $x$ and $X$. Let $X$ be non-lightlike, that is, $E$ is non-degenerate and hence $\mathbb{R}^{n+1} = E \oplus E^\perp$, and $A$ the reflection about $E$, i.e. $A\big|_E = \mathrm{id}_E$ and $A\big|_{E^\perp} = -\mathrm{id}_{E^\perp}$. This yields a Lorentz transformation since

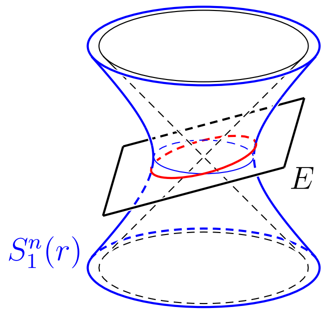
$$\begin{aligned}\langle\!\langle Ax, Ay\rangle\!\rangle &= \langle\!\langle A(x_E + x_{E^\perp}), A(y_E + y_{E^\perp})\rangle\!\rangle = \langle\!\langle x_E - x_{E^\perp}, y_E - y_{E^\perp}\rangle\!\rangle\\ &= \langle\!\langle x_E, y_E\rangle\!\rangle + \langle\!\langle x_{E^\perp}, y_{E^\perp}\rangle\!\rangle = \langle\!\langle x, y\rangle\!\rangle\,.\end{aligned}$$

It follows that $\gamma(Ax) = \gamma(x)$ and hence $A\big(S_1^n(r)\big) \subset S_1^n(r)$, so $A\big|_{S_1^n(r)}$ is an isometry on $S_1^n(r)$ with fixed point set $S_1^n(r) \cap E$. Therefore, the components of $S_1^n(r) \cap E$, considered as point sets, are geodesics of $S_1^n(r)$.
For $X \in T_x S_1^n(r)$ lightlike, we choose a sequence $(X_j)_{j\in\mathbb{N}} \subset T_x S_1^n(r)$ of non-lightlike vectors such that $X_j \to X$. Then $E_j := \operatorname{span}(x, X_j)$ converges to the plane $E$ spanned by $x$ and $E$, and we have $\exp_x(tX_j) \to \exp_x(tX)$ for all $t$. Hence, the connected components of $S_1^n(r) \cap E$, again considered as point sets, are lightlike geodesics of $S_1^n(r)$. Let us interpret this geometrically:

Let $X$ be spacelike. Since $x$ is spacelike as well, $\gamma\big|_E$ is positive definite, so the intersection

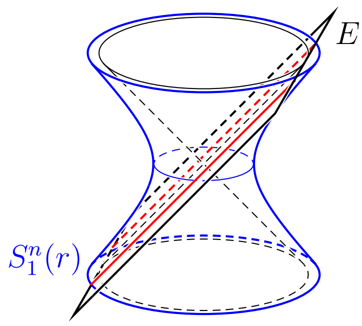
$$S_1^n(r) \cap E = \big\{y \in E \;\big|\; \gamma(y) = r^2\big\}$$

is a closed spacelike geodesic.




Let $X$ be lightlike. Then $\gamma\big|_E$ is positive semidefinite and degenerate, and $S_1^n(r) \cap E$ consists of two parallel straight lines:

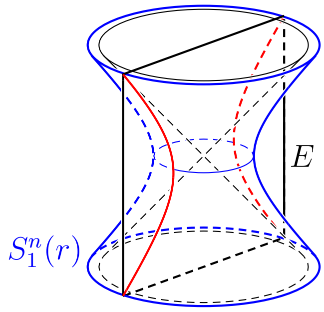
$$\begin{aligned}S_1^n(r) \cap E &= \big\{\alpha x + \beta X \;\big|\; \gamma(\alpha x + \beta X) = \alpha^2\gamma(x) \overset{!}{=} r^2\big\}\\ &= \big\{\alpha x + \beta X \;\big|\; \alpha = \pm 1, \beta \in \mathbb{R}\big\}\\ &= \big\{\pm x + \beta X \;\big|\; \beta \in \mathbb{R}\big\}.\end{aligned}$$




Let $X$ be timelike, so $\gamma\big|_E$ is indefinite and non-degenerate. Then $S_1^n(r) \cap E$ is given by two hyperbola components:

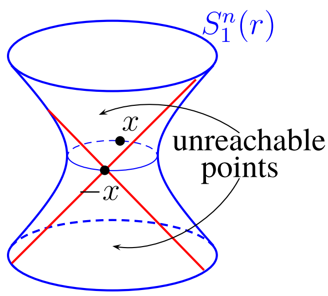
$$\begin{aligned}S_1^n(r) \cap E &= \big\{\alpha x + \beta X \;\big|\; \alpha^2\gamma(x) + \beta^2\gamma(X) \overset{!}{=} r^2\big\}\\ &= \big\{\alpha x + \beta X \;\big|\; \alpha^2 + \beta^2 \underbrace{\frac{\gamma(X)}{r^2}}_{<0} = 1\big\}.\end{aligned}$$

Hence, the two hyperbolas correspond to $\alpha > 1, \alpha < -1$.



Note that the second case ($X$ lightlike) is contained in the third one ($X$ timelike) since $\gamma(X) = 0$ implies $\alpha = \pm 1$ and arbitrary choice of $\beta \in \mathbb{R}$.

The determination of the geodesics particularly implies geodesical completeness of $S_1^n(r)$, i.e. $\exp_x$ is defined on all of $T_x S_1^n(r)$ for all $x \in S_1^n(r)$. Now for some fixed $x \in S_1^n(r)$, we investigate, which $y \in S_1^n(r)$ can be reached from $x$ via some geodesic.
The cases $y = \pm x$ are trivial since for any spacelike $X$, we have $\pm x \in S_1^n(r) \cap \operatorname{span}\{x, X\}$. In all other cases, $x, y$ are linearly independent, and hence they uniquely determine some plane $E := \operatorname{span}\{x, y\}$. The geodesic connecting $x$ and $y$ therefore has to run in $S_1^n(r) \cap E$, which yields two connected components if the starting vector of the geodesic is causal. It follows that the points $y$, which can not be reached via some geodesic from $x$, are exactly those points, which can be reached from $-x$ via some causal geodesic, i.e. lie in the other connected component. For $y = \alpha x + \beta X$ with causal $X$, that is $|\alpha| \geq 1$, we have $\langle\langle x, x + y\rangle\rangle = (1 + \alpha)r^2$, and thus the set points on $S_1^n(r)$, which can not be reached from $x$ via some geodesic, is given by




$$\left\{y \in S_1^n(r) \;\middle|\; \langle\langle x, x + y\rangle\rangle \leq 0,\; y \neq \pm x\right\}.$$

**Remark 1.28.** This property yields an important distinction from the Riemannian world. On a connected and geodesically complete Riemannian manifold, every pair of points can be joined by some geodesic according to the Hopf-Rinow-Theorem.

**Proposition 1.29.** *Let $M$ be some connected semi-Riemannian manifold, $p \in M$ and $\psi_1, \psi_2$ isometries of $M$ with $\psi_1(p) = \psi_2(p)$ and $\mathrm{d}\psi_1|_p = \mathrm{d}\psi_2|_p$. Then we have $\psi_1 = \psi_2$.*

*Proof.* Let $\psi := \psi_2^{-1} \circ \psi_1 \in \operatorname{Isom}(M)$, so that $\psi(p) = p$ and $\mathrm{d}\psi|_p = \mathrm{id}_{T_pM}$. We show that then $\psi = \mathrm{id}_M$. The subset

$$U := \left\{q \in M \;\middle|\; \psi(q) = q,\; \mathrm{d}\psi|_q = \mathrm{id}_{T_qM}\right\} \subset M$$

is non-empty since $p \in U$, and is closed due to continuity of $\psi$ and $\mathrm{d}\psi$. For $q \in U$, $\exp_q$ is a local diffeomorphism, i.e. $\exp_q \colon \Omega' \to \Omega$ is a diffeomorphism for some neighborhoods $\Omega \subset M$, $\Omega' \in T_qM$ of $q$ and $0$. Recalling (1.1) provides

$$\psi\big(\exp_p(tX)\big) = \exp_{\psi(p)}\left(t \cdot \mathrm{d}\psi\big|_p(X)\right) = \exp_p(tX),$$

that is $V \subset U$. Since $M$ is connected, we obtain $U = M$. $\square$

**Proposition 1.30.** *The map*

$$\mathscr{L}(n+1) \longrightarrow \operatorname{Isom}\big(S_1^n(r)\big), \qquad \Lambda \longmapsto \Lambda\big|_{S_1^n(r)} \tag{1.4}$$

*is a group isomorphism. Moreover,* $\operatorname{Isom}\big(S_1^n(r)\big)$ *acts transitively on* $S_1^n(r)$*, that is, for all* $x, y \in S_1^n(r)$*, there is some* $\psi \in \operatorname{Isom}\big(S_1^n(r)\big)$ *such that* $\psi(x) = y$*.*
*Furthermore, for all* $X, Y \in TS_1^n(r)$ *with* $\langle\!\langle X, X\rangle\!\rangle = \langle\!\langle Y, Y\rangle\!\rangle$*, there is some* $\psi \in \operatorname{Isom}\big(S_1^n(r)\big)$ *such that* $\mathrm{d}\psi(X) = Y$*.*

*Proof.* Clearly, we have $\Lambda\big|_{S_1^n(r)} \in \operatorname{Isom}\big(S_1^n(r)\big)$ and injectivity of the map (1.4). For all $x, y \in S_1^n(r)$, Theorem 1.16 implies existence of some $\Lambda \in \mathscr{L}(n+1)$ such that $\phi(x) = y$. Hence, without loss of generality, we can assume that $X, Y \in T_pS_1^n(r)$ (otherwise replace $Y$ by $\Lambda^{-1}Y$). Let $A \in \mathscr{L}(n)$ such that $AX = Y$ and define

$$B := \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \in \mathscr{L}(n+1).$$

Due to $T_xS_1^n(r) = x^\perp$, the map $\psi := B\big|_{S_1^n(r)}$ satisfies $\psi(x) = x$ and $\mathrm{d}\psi(X) = BX = Y$, which proves the second claim.
For the first property, it remains to show surjectivity of (1.4). Let $\psi \in \operatorname{Isom}\big(S_1^n(r)\big)$, fix some $x \in S_1^n(r)$ and set $y := \psi(x)$. For $\Lambda \in \mathscr{L}(n+1)$ satisfying $\Lambda x = y$, $x$ is a fixed point of $\psi_1 := \Lambda^{-1} \circ \psi \in \operatorname{Isom}\big(S_1^n(r)\big)$, so we obtain a linear isometry $\mathrm{d}\psi_1 \colon T_xS_1^n(r) \to T_xS_1^n(r)$. Choose $B \in \mathscr{L}(n+1)$ such that $Bx = x$ and $\mathrm{d}B\big|_x = \mathrm{d}\psi_1\big|_x$, then $\psi_2 := B^{-1}\Lambda^{-1} \circ \psi \in \operatorname{Isom}\big(S_1^n(r)\big)$ fulfills $\psi_2(x) = x$ and $\mathrm{d}\psi_2\big|_x = \mathrm{id}_{T_xS_1^n(r)}$. Now from Proposition 1.29 follows $\psi_2 = \mathrm{id}_{S_1^n(r)}$, that is $\psi = \Lambda B\big|_{S_1^n(r)}$. $\square$

**Remark 1.31.** Proposition 1.30 states that $S_1^n(r)$ is a homogeneous (first property) and isotropic (second property) space.

## 1.3 The anti-de Sitter space

We equip $\mathbb{R}^{n+1}$ with the semi-Riemannian metric $g := -\sum_{i=0}^{1} \mathrm{d}x^i \otimes \mathrm{d}x^i + \sum_{i=2}^{n} \mathrm{d}x^i \otimes \mathrm{d}x^i$ of index 2 and consider the smooth function $x \mapsto g(x, x) =: \widetilde{\gamma}(x)$.

**Definition 1.32.** For fixed $r > 0$, the $n$-dimensional anti-de Sitter space is defined by

$$H_1^n(r) := \widetilde{\gamma}^{-1}(-r^2).$$

Just like $S^n_1(r)$, $H^n_1(r)$ is a smooth hypersurface of $\mathbb{R}^{n+1}$, which is diffeomorphic to $R \times S^{n-1}$ via

$$H^n_1(r) \longrightarrow \mathbb{R} \times S^{n-1}, \qquad (x^0, x^1, \tilde{x}) \longmapsto \left( \frac{x^0}{\sqrt{\|\tilde{x}\|^2 + r^2}}, \frac{x^1}{\sqrt{\|\tilde{x}\|^2 + r^2}}, \tilde{x} \right).$$

The inverse map is then given by $y \longmapsto \left(\sqrt{\|\tilde{y}\|^2 + r^2} \cdot y^0, \sqrt{\|\tilde{y}\|^2 + r^2} \cdot y^1, \tilde{y}\right)$. The normal bundle of $H^n_1(r)$ is similarly generated by

$$\mathrm{grad}_x \widetilde{\gamma} = 2 \sum_{i=0}^{n} x^i \frac{\partial}{\partial x^i},$$

which, considered as an element of the Minkowski space, is equal to $2x$. It follows

$$\eta\big(\mathrm{grad}_x \widetilde{\gamma}, \mathrm{grad}_x \widetilde{\gamma}\big) = 4g(x, x) = 4\widetilde{\gamma}(x) = -4r^2 < 0,$$

so unlike $S^n_1(r)$, the normal bundle of $H^n_1(r)$ is timelike. Hence, the induced metric $g$ on $H^n_1(r)$ has index $1$ and therefore is a Lorentz metric. Furthermore, the curvature quantities of $H^n_1(r)$ coincides with those of $S^n_1(r)$ up to a sign:

$$R(X, Y)Z = -\frac{1}{r^2}\big(g(Y, Z)X - g(X, Z)Y\big),$$

$$K = -\frac{1}{r^2}, \qquad \mathrm{ric} = -\frac{n-1}{r^2} g, \qquad \mathrm{scal} = -\frac{n(n-1)}{r^2},$$

for vector fields $X, Y, Z$ on $H^n_1(r)$. The universal cover $\widetilde{H}^n_1(r)$ of $H^n_1(r)$, which is diffeomorphic to $\mathbb{R}^n$, is also called anti-de Sitter space.

**Remark 1.33.** The manifolds $H^4_1(r)$, $\widetilde{H}^4_1(r)$ are vacuum solutions of Einstein's field equations with cosmological constant $-\frac{3}{r^2}$.

Similarly to $S^n_1(r)$, geodesics are obtained as point sets $H^n_1(r) \cap E$, where $E \subset \mathbb{R}^{n+1}$ some two-dimensional subspace. Also in anti-de Sitter space, there are pairs of points, which can not be joined by some geodesic. Finally, the isometry group

$$\mathrm{Isom}\big(H^n_1(r)\big) \cong O(2, n-1)$$

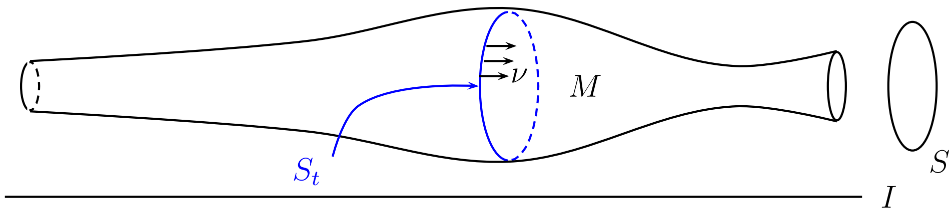
acts transitively on $H^n_1(r)$ and isometries of $H^n_1(r)$ can be lifted to isometries of $\widetilde{H}^n_1(r)$.

## 1.4 Robertson-Walker spacetimes

**Definition 1.34.** An $(n+1)$-dimensional Lorentzian manifold $(M,g)$ is called **Robertson-Walker spacetime** if it is of the form

$$M = I \times S, \qquad g = -\,\mathrm{d}t \otimes \mathrm{d}t + f(t)^2 g_S,$$

where $I \subset \mathbb{R}$ is an open interval, $(S, g_S)$ a complete and connected Riemannian manifold with constant sectional curvature $K^S \equiv \kappa$, and $f\colon I \to \mathbb{R}$ a smooth and positive function.




Examples for $S$ are the model spaces, i.e. the standard sphere $S^n$ ($\kappa = 1$), Euclidean space $\mathbb{R}^n$ ($\kappa = 0$) and the hyperbolic space $H^n$ ($\kappa = -1$), as well as quotients of these like real projective space $\mathbb{RP}^n = S^n/\mathbb{Z}_2$, lense spaces or the torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$. Unlike for $\kappa < 0$, there is a complete classification of complete and connected Riemannian manifold with constant and non-negative sectional curvature, for $\kappa < 0$, there is a huge amount of such manifolds and no classification available.

**Example 1.35.**

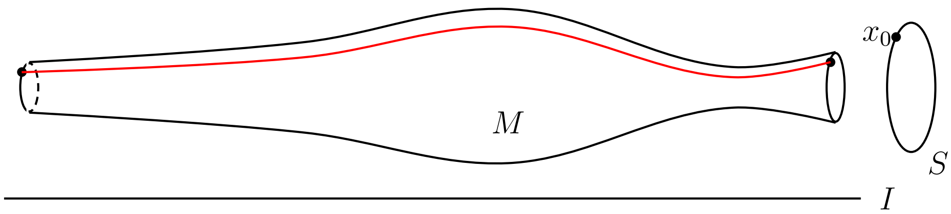
1. For $S = \mathbb{R}^n,\ I = \mathbb{R},\ f \equiv 1$, we obtain the Minkowski space $M = \mathbb{R}^{n+1}$.
2. For $S = S^n,\ I = \mathbb{R},\ f \equiv 1$, we obtain Einstein's static universe.

Let $S_t := \{t\} \times S \subset M$ and $\nu := \frac{\partial}{\partial t}$. We determine the shape operator $W$ of $S_t$ with respect to $\nu$. Let $X, Y, Z$ be vector fields along $S_t$ such that, without loss of generality, the pairwise Lie brackets of $X, Y, \nu$ vanish. Then Koszul's formula leads to

$$\begin{aligned} g\big(W(X), Y\big) = g\big(\nabla^M_X \nu, Y\big) &= \frac{1}{2}\big(\partial_X \underbrace{g(\nu, Y)}_{=0} + \partial_\nu g(X,Y) - \partial_Y \underbrace{g(X,\nu)}_{=0}\big) \\ &= \frac{1}{2}\frac{\partial}{\partial t}\big(f^2 \cdot g_S(X,Y)\big) = f' f \cdot g_S(X,Y) = \frac{f'}{f} \cdot g(X,Y), \end{aligned}$$

that is $W = \frac{f'}{f} \cdot \mathrm{id}$.

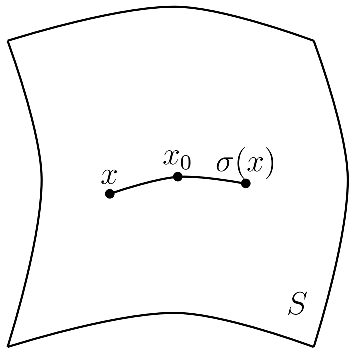
We proceed by investigating the geodesics of Robertson-Walker spacetimes and start with the special case $c(t) := (t, x_0)$ for some fixed $x_0 \in S$.




Let $\sigma$ denote the geodesic reflection in $S$ about $x_0$, which is an isometry of $S$ (at least in some neighborhood of $x_0$), so the map

$$\varphi\colon \quad M \longrightarrow M, \qquad (t,x) \mapsto \big(t, \sigma(x)\big)$$

is an isometry of $M$. Hence, the fixed point set $\mathbb{R} \times \{x_0\}$ of $\varphi$ is a geodesic as a point set. On the other hand, $c$ is the parametrization by proper time of this set and hence a geodesic.




Now let us consider a general geodesic $c$ of $(M, g)$. We write $c(s) := \big(t(s), \gamma(s)\big)$ with $t, \gamma$ mapping into $I$ and $S$, respectively. It follows $c'(s) = t'(s)\nu + \gamma'(s)$ and thus

$$\begin{aligned}
\nabla^M_{c'} c' &= \nabla^M_{c'}(t'\nu + \gamma') = t''\nu + t'\nabla^M_{t'\nu+\gamma'}\nu + \nabla^M_{t'\nu+\gamma'}\gamma' \\
&= t''\nu + (t')^2 \underbrace{\nabla^M_\nu \nu}_{=0} + \underbrace{t'\nabla^M_{\gamma'}\nu + t'\nabla^M_\nu \gamma'}_{=2t'\nabla^M_{\gamma'}\nu} + \nabla^M_{\gamma'}\gamma' \\
&= t''\nu + 2t'W(\gamma') + \nabla^S_{\gamma'}\gamma' + \nabla^S_{\gamma'}\gamma' + \underbrace{g\big(\nabla^M_{\gamma'}\gamma', \nu\big)}_{=g\left(W(\gamma'),\gamma'\right)}\nu
\end{aligned}$$

Here we used $[\gamma', \nu] = 0$ and that $\tilde{c}\colon\ t \mapsto \big(t, \gamma(s_0)\big)$ is a geodesic of the form we discussed before with $\dot{\tilde{c}}(t) = \nu$ and hence $\nabla^M_\nu \nu = 0$. Therefore, demanding $c$ to be geodesic leads to the following system of equations

$$0 = t'' + g\big(W(\gamma'), \gamma'\big) = t'' + f'fg_S\big(W(\gamma'), \gamma'\big) = \nabla^S_{\gamma'}\gamma' + 2t\frac{f'}{f}\gamma' = \nabla^S_{\gamma'}\gamma' + 2\frac{(f\circ t)'}{f\circ t}\gamma'.$$

Hence, we just showed

**Proposition 1.36.** *A smooth curve $c = (t, \gamma)$ in $M$ is a geodesic if and only if*

$$t'' = -f'fg_S\big(W(\gamma'), \gamma'\big), \qquad \nabla^S_{\gamma'}\gamma' = -2\frac{(f\circ t)'}{f\circ t}\gamma'.$$

*In particular, up to reparametrization, $\gamma$ is then a geodesic of $(S, g_S)$.*

*Proof.* Choose a parametrization $t$ such that $f\circ t = \text{const}$. □

**Corollary 1.37.** *For lightlike geodesics, the function $s \mapsto t'(s)\cdot f\big(t(s)\big)$ is constant.*

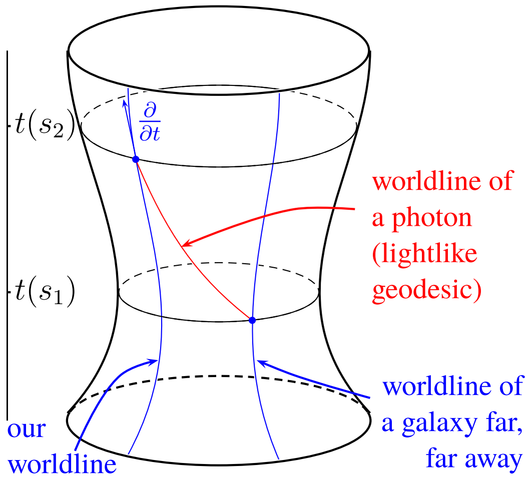
*Proof.* This is implied by the first equation in Proposition 1.36:

$$0 = f'\cdot g(c', c') = -(t')^2 f' + f'f^2 g_S(\gamma', \gamma') = -(t')^2 f' + f't'' = -(t'\cdot f)'. \qquad \square$$

**Remark 1.38.**
Explanation of the cosmic redshift:
The world line of a photon is a lightlike geodesic $c$. Its energy measured by an observer $\nu$ is $-g(c', \nu) = t'$. Then for the wavelength $\lambda := \frac{h}{t'}$ of the photon, we obtain

$$t'(s_1)f\big(t(s_1)\big) = t'(s_2)f\big(t(s_2)\big)$$

$$\implies \quad \frac{\lambda(s_1)}{\lambda(s_2)} = \frac{f\big(t(s_1)\big)}{f\big(t(s_2)\big)}.$$




Let us investigate the curvature of $(M, g)$. Recall that for semi-Riemannian manifolds $(S, g_S)$ with constant section curvature $\kappa$, it reduces to the simple expressions

$$R^S(X,Y)Z = \kappa\big(g_S(Y,Z)X - g_S(X,Z)Y\big),$$
$$\mathrm{ric}^S(X,Y) = \kappa(n-1)\cdot g_S(X,Y),$$
$$\mathrm{scal}^S = \kappa n(n-1).$$

From that we extract the curvature of $(M, g)$. Recall that $R^M$ is completely determined by its tangential and normal component as well as $g\big(R^M(X,\nu)\nu, Y\big)$:

- Tangential part for vector fields $X, Y, Z$ on $S_t$ via Gauß' Formula:

$$\begin{aligned}\tan\big(R^M(X,Y)Z\big) &= R^{S(t)}(X,Y)Z + g\big(W(Y),Z\big)\cdot W(X) - g\big(W(X),Z\big)\cdot W(Y)\\ &= \kappa\big(g_S(Y,Z)X - g_S(X,Z)Y\big) + \left(\frac{f'}{f}\right)^2 \cdot f^2\big(g_S(Y,Z)X - g_S(X,Z)Y\big)\\ &= \big(\kappa + (f')^2\big)\big(g_S(Y,Z)X - g_S(X,Z)Y\big)\\ &= \frac{\kappa+(f')^2}{f^2}\big(g(Y,Z)X - g(X,Z)Y\big).\end{aligned}$$

- Let $II$ denote the second fundamental form of $S_t$, which is linked to $W$ via $g\big(W(X),Y\big) = \big(II(X,Y),\nu\big)$. For vector fields $X, Y, Z$ on $S_t$ define

$$(\nabla_X II)(Y,Z) := \nabla_X II(Y,Z) - II(\nabla_X Y, Z) - II(Y, \nabla_X Z).$$

Then the Codazzi equation provide

$$\mathrm{nor}\big(R(X,Y)Z\big) = -(\nabla_X II)(Y,Z) + (\nabla_Y II)(X,Z),$$

which, due to metricity of $g$, results in

$$g\big(R(X,Y)Z,\nu\big) = g\big(\nabla_X W(Y) - W(\nabla_X Y) - \nabla_Y W(X) - W(\nabla_Y X), Z\big) = 0,$$

since $\nabla_X W(Y) - W(\nabla_X Y) = \partial_X \frac{f'}{f}\cdot Y = 0$.

- It remains the mixed term

$$\begin{aligned}R^M(X,\nu)\nu &= \nabla^M_X \underbrace{\nabla^M_\nu \nu}_{=0} - \nabla^M_\nu \underbrace{\nabla^M_X \nu}_{=W(X)} - \nabla^M_{[X,\nu]}\nu = -\nabla^M_\nu\big(W(X)\big) - \nabla^M_{W(X)-\nabla^M_\nu X}\nu\\ &= -\big(\nabla^M_\nu W\big)(X) - W\big(\nabla^M_\nu X\big) - W(W(X)) + W\big(\nabla^M_\nu X\big)\\ &= -\left(\nabla_{\frac{\partial}{\partial t}} W\right)(X) - W(W(X)) = -\frac{f''f - (f')^2}{f^2}X - \left(\frac{f'}{f}\right)^2 X = -\frac{f''}{f}X.\end{aligned}$$

We proceed with the Ricci curvature. Let $(e_1,\dots,e_n)$ be a local orthonormal frame of $TS_t$.

- For vector fields $X, Y$ on $S_t$, we obtain

$$\begin{aligned}\mathrm{ric}^M(X,Y) &= -g\big(R^M(X,\nu)\nu, Y\big) + \sum_{i=1}^n g\big(R^M(X,e_i)e_i, Y\big)\\ &= \frac{f''}{f}g(X,Y) + \frac{\kappa+(f')^2}{f^2}\underbrace{\sum_{i=1}^n \big(g(e_i,e_i)g(X,Y) - g(X,e_i)g(e_i,Y)\big)}_{=(n-1)g(X,Y)}\\ &= \left(\frac{f''}{f} + \frac{(n-1)\big(\kappa+(f')^2\big)}{f^2}\right) g(X,Y).\end{aligned}$$

- It directly follows

$$\begin{aligned}\mathrm{ric}^M(X,\nu) &= -g\big(\underbrace{R^M(X,\nu)\nu}_{=0},\nu\big) + \sum_{i=1}^n g\big(R^M(X,e_i)e_i,\nu\big)\\ &= \frac{(n-1)\big(\kappa+(f')^2\big)}{f^2}g(X,\nu) = 0.\end{aligned}$$

- For the remaining part, we obtain

$$\mathrm{ric}^M(\nu,\nu) = -g\big(\underbrace{R^M(\nu,\nu)}_{=0}\nu,\nu\big) + \sum_{i=1}^n g\big(\underbrace{R^M(e_i,\nu)\nu}_{=-\frac{f''}{f}e_i},e_i\big) = -n\frac{f''}{f}.$$

Consequently, the scalar curvature is given by

$$\mathrm{scal}^M = -\mathrm{ric}^M(\nu,\nu) + \sum_{i=1}^n \mathrm{ric}^M(e_i,e_i) = n\left(\frac{2f''}{f} + \frac{(n-1)\big(\kappa+(f')^2\big)}{f^2}\right).$$

**Remark 1.39.** From the derivation of $R^M(X,\nu)\nu$, we can somehow derive a more general principle: For any Lorentzian manifold, which is foliated by spacelike hypersurfaces with normal field $\nu$ such that $\nabla^M_\nu \nu = 0$ (we call this a **Riemannian foliation**), then the shape operator $W$ of the hypersurfaces satisfies the following Riccati equation

$$R^M(\cdot,\nu)\nu + \nabla^M_\nu W + W^2 = 0.$$

Let us now give a physical interpretation of these formulas via Einstein's field equations. Certain terms in the energy momentum tensor have a physical interpretation, which can be now expressed in geometric terms:

$$\Big(\mathrm{ric}^M - \frac{\mathrm{scal}^M}{2}g\Big)(X,Y) = \underbrace{\left\{\left(1-\frac{n}{2}\right)(n-1)\left(\frac{\kappa}{f^2}+\left(\frac{f'}{f}\right)^2\right) - (n-1)\frac{f''}{f}\right\}}_{=:p \quad \text{pressure}} g(X,Y),$$

$$\Big(\mathrm{ric}^M - \frac{\mathrm{scal}^M}{2}g\Big)(\nu,\nu) = \underbrace{\frac{n(n-1)}{2}\cdot\frac{\kappa+(f')^2}{f^2}}_{=:\varrho \quad \text{mass density}},$$

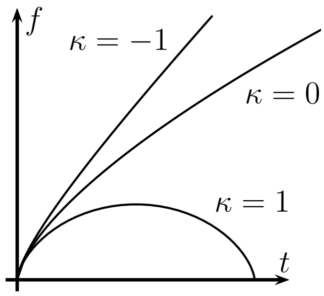
where $X, Y$ are vector field on $S_t$.

**Definition 1.40.** A Robertson-Walker spacetime is called **Friedman cosmos** if $p = 0$.

**Remark 1.41.** The condition $p = 0$ leads to a second order ODE in $f$, which can be solved. For $n = 3$, we obtain

(i) $\kappa = 0$ : Neil's parabola

$$f(t) = C(t - t_0)^{\frac{2}{3}}.$$




(ii) $\kappa = 1$ : cycloid

$$t(\theta) = C(\theta - \sin\theta), \quad f\big(t(\theta) - t_0\big) = C(1 - \cos\theta).$$

(iii) $\kappa = -1$ :

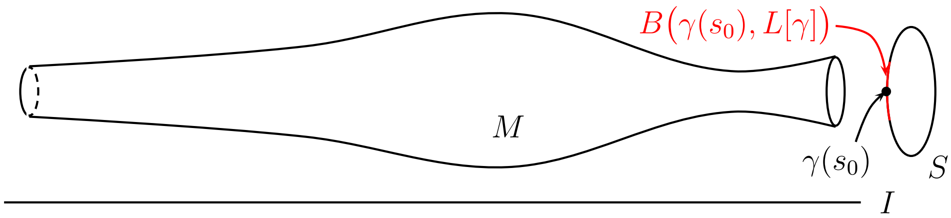
$$t(\theta) = C(\theta - \sinh\theta), \quad f\big(t(\theta) - t_0\big) = C(1 - \cosh\theta).$$

**Remark 1.42.** Let $(M, g)$ a Robertson-Walker spacetime and $c(s) := \big(t(s), \gamma(s)\big)$ a lightlike geodesic. Then $c' = t'\nu + \gamma'$ provides

$$0 = g(c', c') = -(t')^2 + (f \circ t)^2 \cdot \|\gamma'\|^2 \qquad \Longrightarrow \qquad \|\gamma'\| = \frac{|t'|}{f \circ t}.$$

Now let $t' > 0$, then for fixed $s_0 \in \mathbb{R}$, the last equation provides

$$L[\gamma] = \int_{s_0}^{\infty} \|\gamma'(s)\|_S \, \mathrm{d}s = \int_{s_0}^{\infty} \frac{t'(s)}{f\big(t(s)\big)} \, \mathrm{d}s \leq \int_{t(s_0)}^{\infty} \frac{\mathrm{d}t}{f(t)}.$$




For $f$ growing sufficiently fast, i.e. $t^2$ or $e^t$, we have $L[\gamma] < \infty$, so the curve $\gamma$ does not leave the ball with radius $L[\gamma]$ around $\gamma(s_0)$. Therefore, from any spacetime point, there are parts of the universe, which can not be observed, a phenomenon known as the Horizon problem.

## 1.5 The Schwarzschild half-plane

In this section, fix $m > 0$, which we physically interpret as the mass of a rotationally symmetric black hole.

**Definition 1.43.** The Lorentzian manifold

$$M := \mathbb{R} \times \big((0, 2m) \cup (2m, \infty)\big), \qquad g := -h(r)\,\mathrm{d}t \times \mathrm{d}t + \frac{1}{h(r)}\,\mathrm{d}r \otimes \mathrm{d}r,$$

where $h(r) := 1 - \frac{2m}{r}$, is called **Schwarzschild half-plane**.

The non-vanishing Christoffel symbols are

$$\Gamma^1_{00} = \frac{h \cdot h'}{2}, \qquad \Gamma^0_{01} = \Gamma^0_{10} = \frac{h'}{2h}, \qquad \Gamma^1_{11} = -\frac{h'}{2h},$$

and hence, for the sectional curvature, we obtain

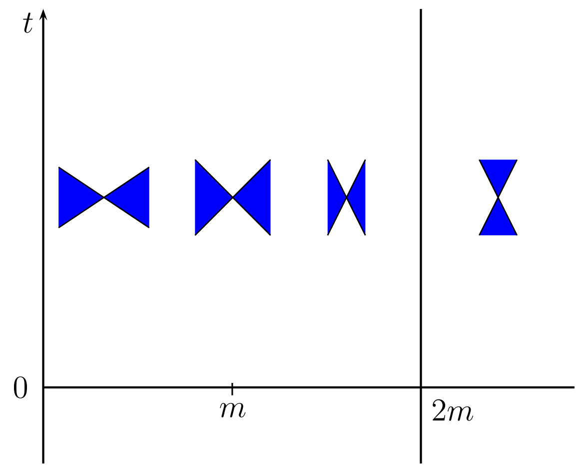
$$K = \frac{R^0_{101}}{g_{11}} = h\big(-\partial_1 \Gamma^0_{10} + \Gamma^0_{01}\Gamma^0_{11} - \Gamma^0_{10}\Gamma^0_{01}\big) = -\frac{h''}{2} = \frac{2m}{r^3}.$$

Let us determine the lightcone of the tangent space in $(t, r)$, i.e. the set of all lightlike vectors. For $X = a\frac{\partial}{\partial t} + b\frac{\partial}{\partial r}$, this is the case if

$$0 = g_{(t,r)}(X, X) = -a^2 \cdot h(r) + \frac{b^2}{h(r)}$$

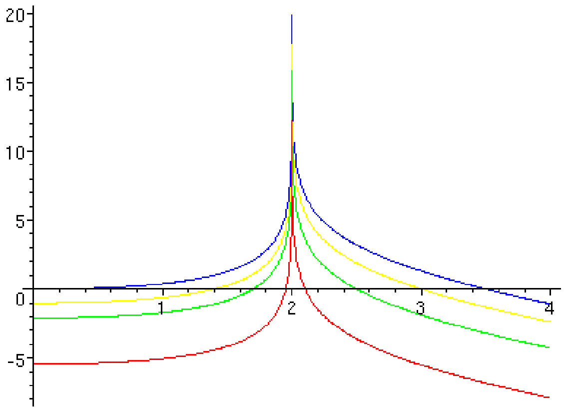
$$\iff \qquad b = \pm h(r) \cdot a.$$

The lightlike curves $c(s) := \big(t(s), r(s)\big)$ of $M$ satisfy $c'(s) = t'(s)\frac{\partial}{\partial t} + r'(s)\frac{\partial}{\partial r}$ with $r' = \pm h(r) \cdot t'$, that is, for fixed $s_0 \in \mathbb{R}$, we have




$$t(s) - t(s_0) = \int_{s_0}^{s} t'(s)\,\mathrm{d}s = \pm \int_{s_0}^{s} \frac{r'(s)\,\mathrm{d}s}{h\big(r(s)\big)} = \pm \int_{r(s_0)}^{r(s)} \frac{r\,\mathrm{d}r}{h(r)}$$

$$= \pm \int_{r(s_0)-2m}^{r(s)-2m} \frac{\rho + 2m}{\rho}\,\mathrm{d}\rho = \pm\left(r(s) - r(s_0) + 2m \log \frac{r(s) - 2m}{r(s_0) - 2m}\right).$$




Due to $\lim_{r \to 0} K = \infty$, $M$ cannot be extended to $r = 0$. Furthermore, there are geodesics, which reach $r = 0$ within "finite time", so $M$ fails to be geodesically complete.
The singularity at $r = 2m$, on the other hand, is removable by some change of coordinates, the so-called "Kruskal-coordinates".

For fixed $t_0 \in \mathbb{R}$, the maps

$$(t, r) \longmapsto (t + t_0, r), \qquad (t, r) \longmapsto (-t, r).$$

are isometries and hence, the subsets $\{t_0\} \times (0, 2m)$ and $\{t_0\} \times (2m, \infty)$ provide a timelike and a spacelike geodesic, respectively.

# 2 Causality

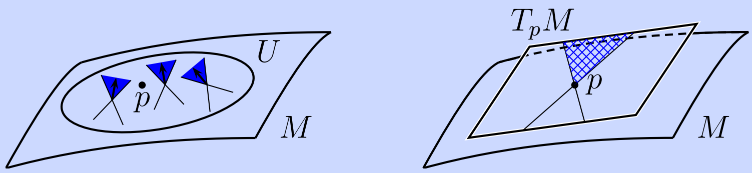
## 2.1 Fundamental Notions

**Definition 2.1.** Let $M$ be a Lorentzian manifold and $\mathscr{P}(TM)$ denote the power set of the tangent bundle. A **time-orientation of** $M$ is a map

$$\zeta\colon\ M \to \mathscr{P}(TM)$$

such that, for all $p \in M$,

- $\zeta(p)$ is one of the connected components $I_\pm(p)$ of $T_pM$ and
- there is a chart $(x, U)$ around $p$ such that $\frac{\partial}{\partial x^0}\big|_q \in \zeta(q)$ for all $q \in U$.




$M$ is called **time-orientable** if it admits a time-orientation $\zeta$, and a pair $(M, \zeta)$ is called **time-oriented Lorentzian manifold**.

**Proposition 2.2.** *Let $M$ be a Lorentzian manifold. Then the following statements are equivalent:*

*(i) $M$ is time-orientable.*

*(ii) $M$ admits a continuous timelike vector field.*

*(iii) $M$ admits a smooth timelike vector field.*

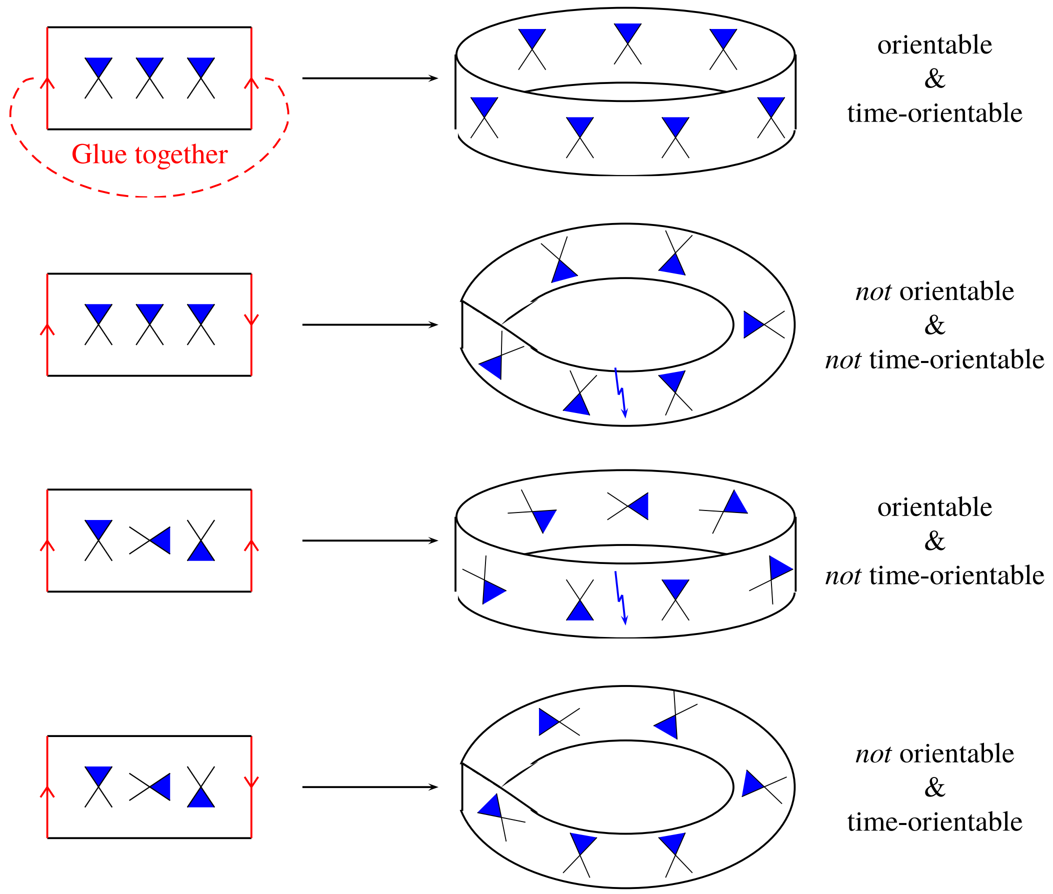
*Proof.* $(iii) \Rightarrow (ii)$: trivial.

$(ii) \Rightarrow (i)$: Let $X$ be a continuous timelike vector field and for all $p$, define $\zeta(p)$ as the connected components that contains $X(p)$. For fixed $p$, choose $(x, U)$ such that $\frac{\partial}{\partial x^0}\big|_q \in \zeta(p)$, and hence $g\big(X, \frac{\partial}{\partial x^0}\big) < 0$ in $p$. By possibly shrinking $U$ and due to continuity of $X$ and $\frac{\partial}{\partial x^0}$, we obtain $g\big(X, \frac{\partial}{\partial x^0}\big) < 0$ on all of $U$ and thus $\frac{\partial}{\partial x^0}\big|_q \in \zeta(q)$ for all $q \in U$.

$(i) \Rightarrow (iii)$: Let $\zeta$ be a time-orientation and $\{p_\alpha\}_\alpha \subset M$ such that $\{U_\alpha\}_\alpha$ yields a locally finite cover of $M$ with corresponding partition of unity $\{\rho_\alpha\}_\alpha$ and $\frac{\partial}{\partial x^0_\alpha}\Big|_q \in \zeta(q)$ for all $q \in U_\alpha$ and $\alpha$. Then the convex linear combination $X := \sum_\alpha \rho_\alpha \frac{\partial}{\partial x^0_\alpha}$ is a well-defined and smooth vector field on all of $M$ and thus $X(p) \in \zeta(p)$ for all $p \in M$ since $\zeta(p)$ is convex, so in particular $X$ is timelike. $\square$

All examples considered in chapter one are time-orientable. Note that, unlike orientablity, time-orientablity not only depends on the topological space $M$ but also on the metric $g$. Moreover, there is no connection between both orientablity properties as the following examples show:

**Example 2.3.**




From now on let $(M, \zeta)$ always be a connected and time-oriented Lorentzian manifold, which we address simply by $M$. A curve $c$ in $M$ is always considered to be continuous and piecewise smooth. Furthermore, we call a causal curve future or past directed if for all $s$, we have $c'(s) \in \overline{\zeta\big(c(s)\big)}$ or $c'(s) \in -\overline{\zeta\big(c(s)\big)}$, respectively.

**Notation 2.4.** For $p, q \in M$, we define the following relations:

$$p \ll q \quad :\Longleftrightarrow \quad \exists \text{ future directed timelike curve from } p \text{ to } q.$$

$$p < q \quad :\Longleftrightarrow \quad \exists \text{ future directed causal curve from } p \text{ to } q.$$

$$p \leq q \quad :\Longleftrightarrow \quad p < q \text{ or } p = q.$$

For $A \subset M$, we define

$$I_+(A) := \{q \in M \mid \exists p \in A\colon \ p \ll q\} \quad \text{the chronological future of } A,$$

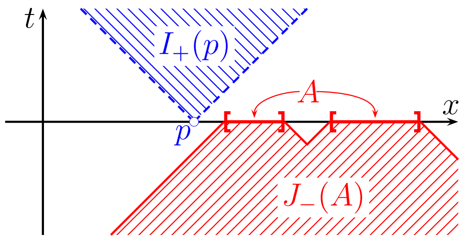
$$J_+(A) := \{q \in M \mid \exists p \in A\colon \ p \leq q\} \quad \text{the causal future of } A.$$

Analogously, one defines the chronological and causal past $I_-(A)$ and $J_-(A)$ of $A$, respectively. Moreover, for $A = \{p\}$, we write $I_\pm(p)$, $J_\pm(p)$.

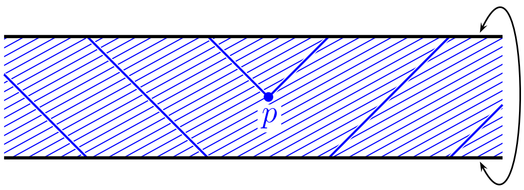
**Remark 2.5.** We have $I_\pm(A) = \bigcup_{p \in A} I_\pm(p)$ and $J_\pm(A) = \bigcup_{p \in A} J_\pm(p)$.

**Example 2.6.**

$M := \mathbb{R}^2_{\text{Mink}}$




$$\begin{aligned} M &:= \mathbb{R}^2_{\text{Mink}}/\mathbb{Z}e_0 \\ &\cong \left(S^1 \times \mathbb{R}, -\mathrm{d}\theta^2 + \mathrm{d}r^2\right) \\ &\Longrightarrow \quad I_\pm(p) = M = J_\pm(p). \end{aligned}$$



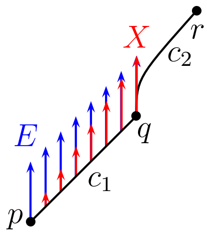

The relation "$\ll$" is transitive, i.e. $p \ll q$ and $q \ll r$ imply $p \ll r$, since one can always connect future directed timelike and causal curves, respectively. Even a stronger version of transitivity holds:

**Proposition 2.7.** *Let $M$ be a connected time-oriented Lorentzian manifold and $p, q, r \in M$. Then the following two statements hold:*

$$p \ll q \text{ and } q \leq r \quad \Longrightarrow \quad p \ll r, \qquad p \leq q \text{ and } q \ll r \quad \Longrightarrow \quad p \ll r.$$

*Proof.* We merely prove the second statement (the proof of the first one is similar). Since the case $p = q$ is trivial, we assume $p < q$. Let $c_1 : [0,1] \to M$ be a causal future directed curve from $p$ to $q$, and $c_2 : [1,2] \to M$ a timelike future directed curve from $q$ to $r$. The idea is to consider the concatination and find a timelike future directed deformation of it. Let $t \mapsto E(t)$ be the parallel vector field along $c_1$ with $E(1) = \dot c_2(1)$ and $X(t) := t \cdot E(t)$. Furthermore, let $c_{1,s}$ denote a variation of $c_1$ with

$$c_{1,s}(0) = p, \quad c_{1,s}(1) = c_2(1+s), \quad \frac{\partial c_{1,s}}{\partial s}\Big|_{s=0}(t) = X(t)$$




By assumption, we have $g\big(\dot c_{1,s}, \dot c_{1,s}\big)\big|_{s=0} \le 0$ and moreover, due to freeness of torsion of the Levi-Civita connection,

$$\begin{aligned}\frac{\partial}{\partial s} g\big(\dot c_{1,s}(t), \dot c_{1,s}(t)\big)\Big|_{s=0} &= 2g\left(\frac{\nabla}{\partial s}\frac{\partial c_{1,s}}{\partial t}(t), \frac{\partial c_{1,s}}{\partial t}(t)\right)\Big|_{s=0} \\ &= 2g\Bigg(\frac{\nabla}{\partial t}\underbrace{\frac{\partial c_{1,s}}{\partial s}(t)}_{=X(t)=tE(t)}, \dot c_1(t)\Bigg)\Big|_{s=0} = 2g\big(E(t), \dot c_1(t)\big) < 0\end{aligned}$$

since $E$ is parallel and both, $E$ and $\dot c_1$, are timelike and future directed. Hence, for $s > 0$ suitably small and all $t \in [0,1]$, we obtain $\frac{\partial}{\partial s} g\big(\dot c_{1,s}(t), \dot c_{1,s}(t)\big) < 0$. It follows that, again for $s > 0$ suitably small, $c_{1,s} : [0,1] \to M$ is timelike and future directed and thus, so is

$$c\colon \quad [0, 2-s] \to M, \qquad t \longmapsto \begin{cases} c_{1,s}(t), & t \in [0,1] \\ c_2(t+s), & t \in [1, 2-s] \end{cases},$$

which connects $p$ and $r$, i.e. $p \ll r$. □

**Corollary 2.8.** *For all $A \subset M$, we have*

$$I_+(A) = I_+\big(I_+(A)\big) = J_+\big(I_+(A)\big) = I_+\big(J_+(A)\big) \qquad J_+\big(J_+(A)\big) = J_+(A).$$

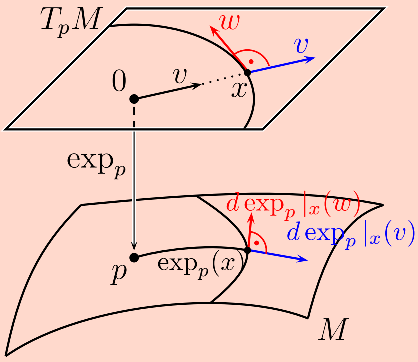
*Proof.* We start with the first statement for which it suffices to show

$$I_+(A) \subset I_+\big(I_+(A)\big) \subset J_+\big(I_+(A)\big) \subset I_+(A).$$

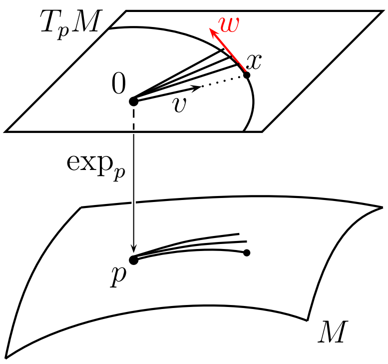
The second inclusion is trivial and the third one follows directly from the Proposition. This remains true if we replace $J_+\big(I_+(A)\big)$ by $I_+\big(J_+(A)\big)$. However, for the first inclusion, we consider $p \in I_+(A)$, i.e. there is some future directed, timelike curve $c\colon [0,1] \to M$ connecting $A$ and $p$. For some $t \in (0,1)$, let $r := c(t) \in I_+(A)$, that is $p \in I_+(r) \subset I_+\big(I_+(A)\big)$.
For the second statement, recall $A \subset J_+(A)$ and furthermore $J_+\big(J_+(A)\big) \subset J_+(A)$ due to transitivity of "$\le$", i.e. $J_+(A) \subset J_+\big(J_+(A)\big) \subset J_+(A)$. □

**Proposition 2.9 (Gauß-Lemma).** *Let $M$ be a Lorentzian manifold, $p \in M$ and $x \in T_pM$ in the domain of* $\exp_p$*. For some $t_0 \in \mathbb{R}$, let furthermore $v := t_0 x \in T_pM$ and $w \in T_pM$, and we obtain*

$$g\big(\mathrm{d}_x \exp_p(v),\, \mathrm{d}_x \exp_p(w)\big) = g(v, w).$$




*Proof.*
Note that the identity, which we would like to prove, is homogeneous in $v$, so without loss of generality, we assume $t_0 = 1$, that is $v = x$. We start by defining




$$\begin{aligned} \psi\colon \quad [0,1] \times (-\varepsilon, \varepsilon) &\longrightarrow M \\ (t, s) &\longmapsto \exp_p\big(t(v + sw)\big). \end{aligned}$$

This map satisfies

$$\frac{\partial\psi}{\partial t}(1,0) = \mathrm{d}_x \exp(v), \qquad \frac{\partial\psi}{\partial s}(1,0) = \mathrm{d}_x \exp(w),$$

i.e. it remains to show $g\Big(\frac{\partial\psi}{\partial t}(1,0), \frac{\partial\psi}{\partial s}(1,0)\Big) = g(v,w)$. For each $s$, the curve $t \mapsto \psi(t,s)$ is a geodesic with velocity vector $v + sw$ at $t = 0$, so it follows

$$\begin{aligned} \frac{\partial}{\partial t} g\left(\frac{\partial\psi}{\partial t}, \frac{\partial\psi}{\partial s}\right) &= g\left(\overbrace{\frac{\nabla}{\partial t}\frac{\partial\psi}{\partial t}}^{=0}, \frac{\partial\psi}{\partial s}\right) + g\left(\frac{\partial\psi}{\partial t}, \frac{\nabla}{\partial t}\frac{\partial\psi}{\partial s}\right) \\ &= g\left(\frac{\partial\psi}{\partial t}, \frac{\nabla}{\partial s}\frac{\partial\psi}{\partial t}\right) = \frac{1}{2}\frac{\partial}{\partial s} \underbrace{g\left(\frac{\partial\psi}{\partial t}, \frac{\partial\psi}{\partial t}\right)}_{=g(v+sw,v+sw)} = g(v,w) + s g(w,w). \end{aligned}$$

For $f(t) := g\Big(\frac{\partial\psi}{\partial t}(t,0), \frac{\partial\psi}{\partial s}(t,0)\Big)$, we obtain $\dot f(t) = g(v,w)$ and $f(0) = 0$, that is

$$f(1) = g\left(\frac{\partial\psi}{\partial t}(1,0), \frac{\partial\psi}{\partial s}(1,0)\right) = g(v,w). \qquad \square$$

For any differentiable curve $c : [0,1] \to M$ and differentiable vector field $X$ on $M$, recall the notations

$$\frac{\nabla(X \circ c)}{\partial t} := \nabla_{\dot{c}(t)} X, \qquad \frac{\nabla \dot{c}}{\mathrm{d}t}(t) := \nabla_{\dot{c}(t)} \dot{c}.$$

**Remark 2.10.** Note that this only holds if one vector is "radial", i.e. a multiple of the vector, where the differential is evaluated. Hence, the Gauß-Lemma allows the exponential map to be understood as a radial isometry.

**Lemma 2.11.** *Let $M$ be a time-oriented Lorentzian manifold, $p \in M$ and $\gamma \colon [0,b] \to T_pM$ a curve running in the domain of $\exp_p$ such that $\gamma(0) = 0$ and $c := \exp_p \circ \gamma$ is timelike and future (past) directed. Then $\gamma(t) \in I_+(I_-)$ for all $t \in (0,b]$.*

*Proof.* a) We merely prove $\gamma(t) \in I_+$ since the other case is completely similar. Let $q(x) := g(x,x)$ the corresponding quadratic form on $T_pM$, i.e. in Minkowski coordinates

$$q(x) = -(x^0)^2 + \sum_{i=1}^{n} (x^i)^2,$$

which are chosen such that $\frac{\partial}{\partial x^0}$ is future directed. Note that for $\xi \in T_x T_p M$,

$$g\big(\mathrm{grad}_x q, \xi\big) = \mathrm{d}_x q(\xi) = -2x^0\xi^0 + 2\sum_{i=1}^{n} x^i \xi^i \qquad \Longrightarrow \qquad \mathrm{grad}_x q = 2x.$$

Then for $v = w = 2x$, the Gauß-Lemma provides

$$g\big(\mathrm{d}_x \exp_p(\mathrm{grad}_x q),\, \mathrm{d}_x \exp_p(\mathrm{grad}_x q)\big) = g\big(\mathrm{grad}_x q, \mathrm{grad}_x q\big) = 4q(x),$$

so $P(x) := \mathrm{d}_x \exp_p(\mathrm{grad}_x q)$ is timelike for all $x \in I \subset T_pM$. Furthermore, it is future directed if $x$ is.
b) Consider the case of $\gamma$ being smooth, so $q\big(\gamma(0)\big) = q(0) = 0$ and $\dot{c}(0) = \mathrm{d}_0 \exp_p\big(\dot{\gamma}(0)\big)$, i.e. $\dot{\gamma}(0)$ is timelike. Hence, for $\varepsilon > 0$ suitably small, we have $\gamma(t) \in I_+$, so in particular $q\big(\gamma(t)\big) < 0$, for all $t \in (0,\varepsilon]$. Therefore, it remains to show that $q\big(\gamma(t)\big) < 0$ for all $t \in (0,b]$. Assume there exists some $t_1 \in (0,b]$ such that $q(\gamma(t_1)) = 0$ and without loss of generality let it be the smallest one. By the mean value theorem there has to be some $t_0 \in (0,t_1)$, where $q \circ \gamma$ assumes an minimum, that is $\gamma(t_0) \in I_+$ and $\frac{\mathrm{d}(q\circ\gamma)}{\mathrm{d}t}(t_0) = 0$. On the other hand, the Gauß-Lemma then yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} q\big(\gamma(t)\big) &= \mathrm{d}_{\gamma(t)} q\big(\dot{\gamma}(t)\big) = g\big(\mathrm{grad}_{\gamma(t)} q, \dot{\gamma}(t)\big) \\ &= g\Big(\mathrm{d}_{\gamma(t)} \exp_p\big(\mathrm{grad}_{\gamma(t)} q\big),\, \mathrm{d}_{\gamma(t)} \exp_p\big((\gamma(t))\big)\Big) = g\big(P(\gamma(t)), \dot{c}(t)\big),\end{aligned}$$

so $\frac{\mathrm{d}(q\circ\gamma)}{\mathrm{d}t}(t_0) < 0$ due to a). It follows that such a $t_1$ does not exist, so $q\big(\gamma(t)\big) < 0$ and hence $\gamma(t) \in I_+$ for all $t \in (0,b]$.

c) Now let $\gamma$ be piecewise smooth and $0 := b_0 < b_1 < \ldots < b_N := b$ a partition of $[0, b]$ such that $\gamma$ is smooth on $(b_i, b_{i+1})$ for all $i$. Due to b), we have $\gamma(t) \in I_+$ for all $t \in (0, b_1]$ and we assume $\gamma\big((0, b_k]\big) \subset I_+$ for some $k = 1, \ldots, N-1$. Due to a), we obtain for the right-sided derivative

$$\lim_{\delta \to 0} \frac{q\big(\gamma(b_k + \delta)\big) - q\big(\gamma(b_k)\big)}{\delta} = g\big(P(\gamma(b_k)), \dot{c}(\gamma(b_k))\big) < 0$$

since by assumption, $\dot{c}(t)$ is timelike and future directed for all $t$. Applying b) leads to $\gamma\big([b_k, b_{k+1}]\big) \subset I_+$ and thus completes the proof. □

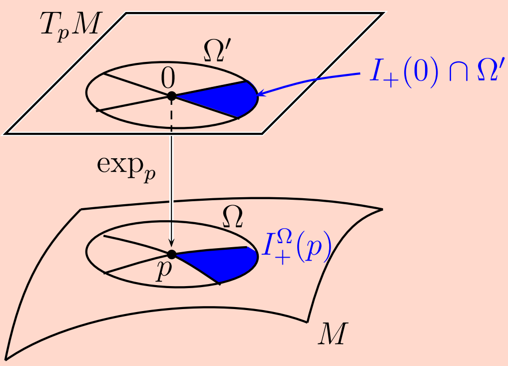
**Notation 2.12.** For $\Omega \subset M$ and any $A \subset \Omega$, we define

$$I_+^\Omega(A) := \big\{ q \in \Omega \,\big|\, \exists p \in A \colon\; p \ll q \text{ in } A \big\}$$

and analogously $I_-^\Omega(A)$ and $J_\pm^\Omega(A)$.

**Corollary 2.13.** *Let $M$ be a time-oriented Lorentzian manifold and $p \in M$. Furthermore, let $\Omega \subset M$ and $\Omega' \subset T_pM$ be open neighborhoods of $p$ and $0$, respectively, such that $\Omega'$ is starshaped with respect to $0$ and $\exp_p \colon \Omega' \to \Omega$ a diffeomorphism. Then we have*

$$I_\pm^\Omega(p) = \exp_p\big(I_\pm(0) \cap \Omega'\big), \qquad J_\pm^\Omega(p) = \exp_p\big(J_\pm(0) \cap \Omega'\big).$$




*Proof.* We only prove the statements for "+" since "−" works analogously. We start with "⊂" in the first equality, so let $q \in I_+^\Omega(p)$, i.e. $p \ll q$. Let $c\colon [0,1] \to \Omega$ be the timelike future directed curve with $c(0) = p$ and $c(1) = q$. Since $\exp_p$ is a diffeomorphism, we obtain a curve $\gamma := \exp_p^{-1} \circ c\colon [0,1] \to \Omega'$, which, due to the Lemma, maps to $I_+(0)$, and thus

$$\exp_p^{-1}(q) = \gamma(1) \in I_+(0) \qquad \Longrightarrow \qquad q \in \exp_p\big(I_+(0) \cap \Omega'\big).$$

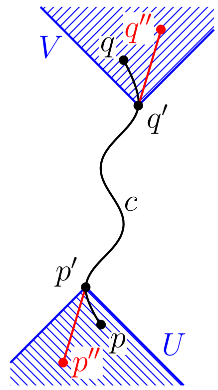
Let $x \in I_+(0) \cap \Omega'$ and consider the ray $t \mapsto tx,\ t > 0$, in $I_+(0) \cap \Omega'$. Hence, $t \mapsto \exp_p(tx)$ is a timelike and future directed geodesic connecting $p$ with $\exp_p(x)$ in $\Omega$, that is $\exp_p(x) \in I_+^\Omega(p)$.

For the second equation, one proves "$\supset$" similarly, so it remains "$\subset$". Let $q \in J_+(p)$ and $(q_i)_{i\in\mathbb{N}} \subset \Omega$ a sequence of points with $q_i \ll q$ and $q_i \to q$. From Proposition 2.7 follows $p \ll q_i$, i.e. $q_i \in I_+^\Omega(p)$ and thus $\exp_p^{-1}(q_i) \in I_+(0) \cap \Omega'$ from the first equation. On the other hand, we obtain $\exp_p^{-1}(q_i) \to \exp_p^{-1}(q)$, i.e. $\exp_p^{-1}(q)$ lies in the closure of $I_+(0) \cap \Omega'$ in $\Omega'$, which is $J_+(0) \cap \Omega'$ and we conclude $q \in \exp_p^{-1}\big(J_+(0) \cap \Omega'\big)$. $\square$

**Proposition 2.14.** *"$\ll$" is an open relation, i.e. $p \ll q$ implies the existence of neighborhoods $U, V$ of $p$ and $q$, respectively, such that $p' \ll q'$ for all $p' \in U$ and $q' \in V$.*

*Proof.*

Let $p \ll q$ with corresponding timelike and future directed curve $c\colon [0,1] \to M$. Set $p' := c(\varepsilon)$ with $\varepsilon > 0$ chosen small enough such that there is an open neighborhood $\Omega$ of $p'$ with $p \in \Omega$ and so that $\exp_{p'}\colon \Omega' \to \Omega$ yields a diffeomorphism for some suitable starshaped neighborhood $\Omega'$ of $0$ in $T_{p'}M$. Then $U := I_-^\Omega(p')$ is an open neighborhood of $p$ in $M$. Analogously, define $V$. Then for all $p'' \in U$ and $q'' \in V$, we have $p'' \ll p' \ll q' \ll q''$ and thus $p'' \ll q''$.



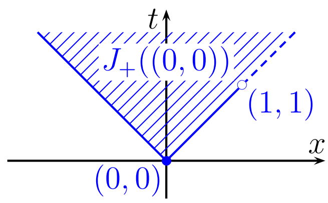

$\square$

**Corollary 2.15.** *For any subset $A \subset M$, the subsets $I_\pm(A) \subset M$ are open.*

*Proof.* The Proposition shows that $I_\pm(A)$ is open in $\Omega$ and so is $I_\pm(A) = \bigcup_{p\in A} I_\pm(p)$. $\square$

**Attention!** $I_\pm(A)$ and $J_\pm(A)$ need not be closed even if $A \subset M$ is.

**Example 2.16.**
Let $M = \mathbb{R}^2 \setminus \{(1,1)\}$ with Minkowski metric. Then $J_+\big((0,0)\big)$ is not closed.




**Proposition 2.17.** *For any $A \subset M$, we have*

$$I_\pm(A) = \mathring{J}_\pm(A), \qquad J_\pm(A) \subset \overline{I_\pm(A)},$$

*and equality holds in the last statement if and only if $J_\pm(A)$ is closed.*

*Proof.* We start with the first statement, for which "$\subset$" is trivial since $I_\pm(A)$ is open and contained in $J_\pm(A)$. For the other direction, let $p \in \mathring{J}_+(A)$ and choose $q \ll p$ such that $q \in \mathring{J}_+(A)$. Hence, there is some $r \in A$ with $r < q \ll p$, which implies $r \ll p$ and thus $p \in I_+(A)$. The proof for "$-$" is similar.
For the second statement, we start with $J_+(p) \subset \overline{I_+(p)}$", so let $q \in J_+(p)$. Since clearly $p \in \overline{I_+(p)}$, we assume $p < q$. Let $c\colon [0,1] \to M$ be the corresponding causal and future directed curve and for all $i \in \mathbb{N}$, set $q_i := c\big(1 - \frac{1}{i}\big)$. Furthermore, choose $r_i$ such that $r_i \to q$, which implies $p < q_i \ll r_i$ and hence $p \ll r_i$, that is $r_i \in I_+(p)$. For any $A \subset M$ it thus follows

$$J_+(A) = \bigcup_{p \in A} J_+(A) \subset \bigcup_{p \in A} \underbrace{\overline{I_+(p)}}_{\subset \overline{I_+(A)}} \subset \overline{I_+(A)}.$$

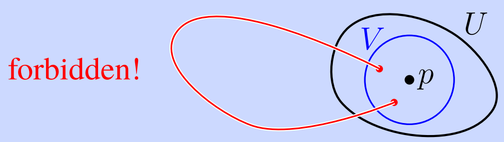
If we have equality, then $J_+(A)$ is obviously closed. Conversely, if $J_+(A)$ is closed, then equality is implied by $\overline{I_+(A)} \subset \overline{J_+(A)} = J_+(A)$. $\square$

**Proposition 2.18.** *Any compact and time-oriented Lorentzian manifold contains a closed timelike curve.*

*Proof.* Let $M$ be a compact and time-oriented Lorentzian manifold, so $\big\{I_+(p)\big\}_{p \in M}$ yields an open cover of $M$. Due to compactness, we find a finite subcover $\big\{I_+(p)\big\}_{p_1,\dots,p_N}$, which, without loss of generality, can be chosen such that no element of this cover contains another one. Hence, if $p_1 \in I_+(p_i)$ for some $i \geq 2$, then $I_+(p_1) \subset I_+(p_i)$, which is a contradiction. Therefore, $p_1$ has to be covered by $I_+(p_1)$, i.e. $p_1 \in I_+(p_1)$, which implies the existence of a timelike and future directed curve that is closed. $\square$

**Definition 2.19.** Let $M$ be a connected and time-oriented Lorentzian manifold. Then $M$ satisfies the

- **chronology condition** if it does not contain any closed timelike curves.
- **causality condition** if it does not contain any closed causal curves.
- **strong causality condition** if for any $p \in M$ and any neighborhood $U$ of $p$, there is a neighborhood $V \subset U$ of $p$ such that any causal curve that starts and ends in $V$ is completely contained in $U$.



**Remark 2.20.** We have

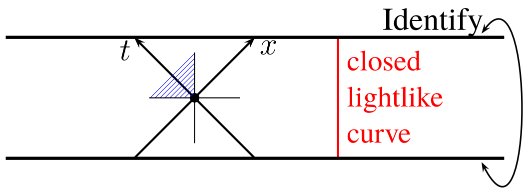
$$\text{strong causality condition} \implies \text{causality condition} \implies \text{chronology condition}.$$

In general, the converse implication do not hold.

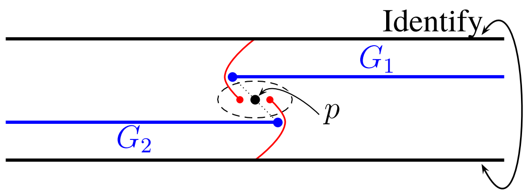
Let $M := \mathbb{R}^2/\big(\mathbb{Z}\cdot(1,1)\big)$ with metric and time-orientation induced by the Minkowski space $\mathbb{R}^2$. This satisfies the choronology condition but not the causality condition since the ray $t \mapsto t\cdot(1,1)$ yields a closed lightlike curve.

**Example 2.21.**
1) Due to Proposition 2.18, any compact and time-oriented Lorentzian manifold does not satisfy any of these conditions.

2) Let $M := \mathbb{R}^2/\big(\mathbb{Z}\cdot(1,1)\big)$ with metric and time-orientation induced by the Minkowski space $\mathbb{R}^2$. This satisfies the choronology condition but not the causality condition since the ray $t \mapsto t\cdot(1,1)$ yields a closed lightlike curve.




3) Let $M := \Big\{\mathbb{R}^2/\big(\mathbb{Z}\cdot(1,0)\big)\Big\}\backslash(G_1\cup G_2)$, where $G_1 := \big\{\big(\frac{1}{8},s\big) \;\big|\; s \geq -\frac{1}{8}\big\}$ and $G_2 := \big\{\big(-\frac{1}{8},s\big) \;\big|\; s \leq \frac{1}{8}\big\}$. Then $M$ satisfies the causality condition but not the strong causality condition.




**Definition 2.22.** The **length of a curve** $c\colon [a,b] \to M$ is defined via

$$L[c] := \int_a^b \sqrt{\big|g\big(\dot c(t),\dot c(t)\big)\big|}\,\mathrm{d}t.$$

**Remark 2.23.** Lightlike curves have length zero.

**Definition 2.24.** For $p,q \in M$, the function

$$\tau(p,q) := \begin{cases} \sup\big\{L[c] \;\big|\; c \text{ is a causal and future directed curve from } p \text{ to } q\big\}, & p < q \\ 0, & p \not< q \end{cases}$$

is called **time difference between** $p$ **and** $q$.

**Example 2.25.**
1) For $(M, g)$ the Minkowski space, we have

$$\tau(p,q) = \sqrt{\left| \langle\langle p-q, p-q \rangle\rangle \right|}.$$

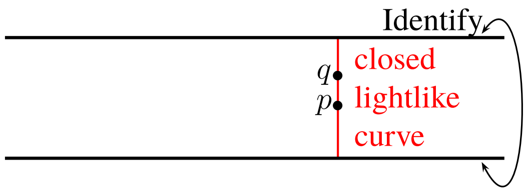
Consider the line segment $t \mapsto tq + (1-t)p$ on $[0,1]$, which clearly yields a causal and future directed curve connecting $p$ and $q$ of length $\sqrt{\left| \langle\langle p-q, p-q \rangle\rangle \right|}$.
For $p - q$ lightlike, we have $\sqrt{\left| \langle\langle p-q, p-q \rangle\rangle \right|} = 0$. On the other hand, all causal curves connecting $p$ and $q$ are necessarily lightlike and hence of length $0$, that is $\tau(p,q) = 0$.
Now let $p - q$ be timelike. After applying some Poincaré transformation and without loss of generality, we can assume that $p = 0$ and $q = (t, 0, \dots, 0)$ for some $T > 0$. Let $c$ be a causal and future directed curve connecting $p$ and $q$, i.e. $\dot{c}^0 > 0$. After some reparametrization, we may assume $c^0(t) = t$, i.e. $c(t) = \big(t, x(t)\big)$ for some curve $x \colon [0,T] \to \mathbb{R}^n$. For the length of $c$, we obtain

$$L[c] = \int_0^T \sqrt{\left| -1 + \|\dot{x}(t)\|^2 \right|}\, \mathrm{d}t = \int_0^T \underbrace{\sqrt{1 - \|\dot{x}(t)\|^2}}_{\leq 1}\, \mathrm{d}t \leq T = \sqrt{\left| \langle\langle p-q, p-q \rangle\rangle \right|}.$$

2) For $M$ the Lorentz cylinder, we have

$$\tau(p,q) = \infty$$

for all $p, q \in M$.




**Proposition 2.26.** *For any time-oriented Lorentzian manifold $M$, we have*

1. $\tau(p,q) > 0 \iff p \ll q$.
2. *If $p \leq q$ and $q \leq r$, then $\tau(p,q) + \tau(q,r) \leq \tau(p,r)$ (inverse triangle inequality).*
3. *The function $\tau \colon M \times M \to R$ is lower semi-continuous.*

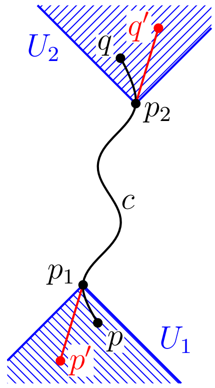
*Proof.* 1) "$\Leftarrow$": If $p \ll q$, then there is a timelike and future directed curve $c$ connecting $p$ and $q$, which therefore is of positive length.
"$\Rightarrow$": Let $\tau(p,q) > 0$, i.e. there is some causal and future directed curve of positive length. Thus, it contains some timelike segment, on which we choose some $p_1, q_1$ such that $p_1 \ll q_1$. It follows $p \leq p_1 \ll q_1 \leq q$, that is $p \ll q$ due to Proposition 2.7.

2) We first consider the case $\tau(p,q), \tau(q,r) < \infty$. For any $\varepsilon > 0$, there are causal and future directed curves $c_1, c_2$ from $p$ to $q$ of length $L[c_1] \geq \tau(p,q) - \varepsilon$ and from $q$ to $r$ of length $L[c_2] \geq \tau(q,r) - \varepsilon$, respectively. It follows

$$\tau(p,r) \leq L[c_1 \cup c_2] = L[c_1] + L[c_2] \geq \tau(p,q) - \varepsilon + \tau(q,r) - \varepsilon$$

and hence the claim since $\varepsilon$ can be chosen arbitrarily small.
Now let $\tau(p,q) = \infty$, i.e. there are arbitrarily long causal and future directed curves from $p$ to $q$. Then concatination with any fixed causal and future directed curve from $q$ to $r$ shows $\tau(p,r)$. The case $\tau(q,r) = \infty$ can be treated similarly.

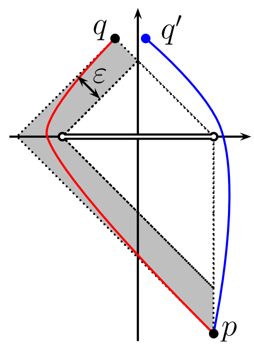
3) Let $p, q \in M$. We distinguish 3 cases:
(i) $\tau(p,q) = 0$: Nothing to show here.
(ii) $0 < \tau(p,q) < \infty$: Choose $\varepsilon > 0$ such that, without loss of generality, $\varepsilon < \frac{\tau(p,q)}{2}$. Let $c\colon [0,1] \to M$ be a causal and future directed curve from $p$ to $q$ with $L[c] \geq \tau(p,q) - \frac{\varepsilon}{2}$. Furthermore, choose $\delta_1, \delta_2 \in (0,1)$ such that $L\left[c|_{[0,\delta_1]}\right], L\left[c|_{[\delta_2,1]}\right] < \frac{\varepsilon}{4}$. For $i = 1,2$, set $p_i := c(\delta_i)$ and $U_i := I_-(p_i)$, which are open neighborhoods of $p$ and $q$, respectively.
Then for any $p' \in U_1$ and $q' \in U_2$, we have




$$\begin{aligned}
\tau(p',q') &\geq \tau(p',p_1) + \tau(p_1,q_1) + \tau(q_1,q') \\
&\geq 0 + L\left[c|_{[\delta_1,\delta_2]}\right] + 0 \\
&\geq L[c] - L\left[c|_{[0,\delta_1]}\right] - L\left[c|_{[\delta_2,1]}\right] \\
&\geq \tau(p,q) - \frac{\varepsilon}{2} - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} \\
&\geq \tau(p,q) - \varepsilon.
\end{aligned}$$

(iii) $\tau(p,q) = \infty$: It follows that there are arbitrarily long causal and future directed curves from $p$ to $q$. The same construction as in (ii) leads to neighborhoods of $p$ and $q$, respectively, such that elements of these neighborhoods are similarly connected by arbitrarily long causal and future directed curves. □

In general, $\tau$ is not continuous, i.e. not upper semi-continuous, as the following example shows:
Let $M := \mathbb{R}^2 \setminus (\{0\} \times [-1,1])$ and arrange $p, q \in M$ as in the picture. Any causal curve connecting both (red curve) has to run entirely in $J_-(q) \cap J_+(p)$ (shaded region), so for $\varepsilon > 0$ small, every causal is almost lightlike, hence short. However, if we shift $q$ to the right, at some point new causal curves (blue) of large length appear.



## 2.2 Curve deformation

**Definition 2.27.** A curve $c\colon [a,b] \to M$, so which we find a function $\alpha\colon [a,b] \to \mathbb{R}$ such that
$$\frac{\nabla}{\mathrm{d}t}\dot{c}(t) = \alpha(t)\dot{c}(t)$$
for all $t \in [a,b]$, is called **pregeodesic**.

Clearly, geodesics are pregeodesics with $\alpha = 0$, and every reparametrization of a geodesic is a pregeodesic since for any geodesic $c$ and $\widetilde{c} := c \circ \phi$, we have
$$\frac{\nabla}{\mathrm{d}t}\dot{\widetilde{c}} = \frac{\nabla}{\mathrm{d}t}\big(\dot{\phi}\cdot(\dot{c}\circ\phi)\big) = \ddot{\phi}\cdot(\dot{c}\circ\phi)\big) + \dot{\phi}^2\cdot\Bigg(\underbrace{\frac{\nabla}{\mathrm{d}t}\dot{c}}_{=0}\circ\phi\Bigg) = \frac{\ddot{\phi}}{\dot{\phi}}\cdot\dot{\widetilde{c}}.$$

Conversely, if $\widetilde{c}$ is a pregeodesic, then setting
$$\phi(t) = \int_a^t \exp\left(\int_a^\tau \alpha(s)\,\mathrm{d}s\right)\mathrm{d}\tau \qquad \Longrightarrow \qquad \alpha(t) = \frac{\ddot{\phi}(t)}{\dot{\phi}(t)}$$
provides a geodesic $c := \widetilde{c} \circ \phi$.

**Remark 2.28.** The proof of Proposition 2.7 already showed: If $c\colon [a,b] \to M$ is a causal curve, $c_s$ a variation of $c$ with $s \in (-\varepsilon,\varepsilon)$ and variational field $X$ such that $g\big(\frac{\nabla X}{\mathrm{d}t},\dot{c}\big) < 0$, then $c_s$ is timelike for $s$ suitably small.

**Lemma 2.29.** *Let $c\colon [a,b] \to M$ be a causal curve, which is not a lightlike pregeodesic. In each neighborhood of c, with respect to the compact-open topology, we find a timelike curve with the same start and endpoint.*

*Proof.* Without loss of generality, we consider $[a,b] = [0,1]$.
a) If there is some $t_0 \in [0,1]$ such that $\dot{c}(t_0)$ is timelike, then $c$ contains a timelike segment and we can deform it to some timelike curve like in the proof of Propsition 2.7.

b) Let $c$ be smooth and lightlike but not a pregeodesic. We obtain $g(\dot{c},\dot{c}) = 0$ and thus $0 = \frac{\mathrm{d}}{\mathrm{d}t}g(\dot{c},\dot{c}) = 2g\big(\frac{\nabla}{\mathrm{d}t}\dot{c},\dot{c}\big)$, i.e. $\frac{\nabla}{\mathrm{d}t}\dot{c}(t)\perp\dot{c}(t)$ for all $t \in [0,1]$. On the other hand, for any lightlike Minkowski vector $v$, we have the decomposition $v^\perp = \mathbb{R}v \oplus E$ for some spacelike subspace $E$. Therefore, if both, $\frac{\nabla}{\mathrm{d}t}\dot{c}(t)$ and $\dot{c}(t)$, were lightlike for all $t$, then $c$ would be a pregeodesic, which is ruled out by assumption. Hence, $\frac{\nabla}{\mathrm{d}t}\dot{c}$ has to be spacelike somewhere, in particular $g\big(\frac{\nabla}{\mathrm{d}t}\dot{c},\frac{\nabla}{\mathrm{d}t}\dot{c}\big)$ does not vanish everywhere.

Choose some timelike $Y_0 \in T_{c(0)}M$ such that $g\big(Y_0, \dot{c}(0)\big) < 0$, and let $Y$ be the parallel vector field along $c$ determined by $Y(0) = Y_0$. It follows that $Y(t)$ is timelike and $g\big(Y(t), \dot{c}(t)\big) < 0$ for all $t \in [0,1]$. If we find some $X(t) := a(t)Y(t) + b(t)\frac{\nabla}{\mathrm{d}t}\dot{c}(t)$ with smooth $a, b$ such that

$$a(0) = b(0) = a(1) = b(1) = 0, \qquad g\left(\frac{\nabla X}{\mathrm{d}t}, \dot{c}\right) < 0,$$

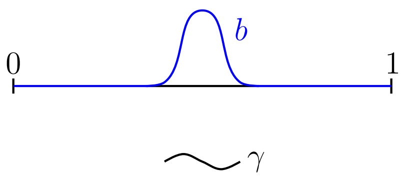
then $X$ determines a variation of $c$ as in the Remark and hence yields the claim. Recall $g\big(\frac{\nabla}{\mathrm{d}t}\dot{c}, \dot{c}\big) = 0$, which leads to

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} g\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \dot{c}\right) = g\left(\frac{\nabla^2}{\mathrm{d}t^2}\dot{c}, \dot{c}\right) + g\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \frac{\nabla}{\mathrm{d}t}\dot{c}\right)$$

$$\begin{aligned} \implies \quad g\left(\frac{\nabla X}{\mathrm{d}t}, \dot{c}\right) &= \dot{a}g(Y, \dot{c}) + ag\left(\overbrace{\frac{\nabla Y}{\mathrm{d}t}}^{=0}, \dot{c}\right) + \dot{b}g\overbrace{\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \dot{c}\right)}^{=0} + bg\left(\frac{\nabla^2}{\mathrm{d}t^2}\dot{c}, \dot{c}\right) \\ &= \dot{a}g(Y, \dot{c}) - bg\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \frac{\nabla}{\mathrm{d}t}\dot{c}\right). \end{aligned}$$

Then $\gamma := \frac{g\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \frac{\nabla}{\mathrm{d}t}\dot{c}\right)}{g(Y, \dot{c})} \geq 0$, which is non-zero somewhere, so we find some smooth $b\colon [0,1] \to \mathbb{R}$ with $b(0) = b(1) = 0$ and

$$\int_0^1 b(t)\gamma(t) = -1.$$




Then $a(t) := \int_0^t (b\gamma + 1)(s)\,\mathrm{d}s$ satisfies $a(0) = a(1) = 0$ and furthermore, we obtain

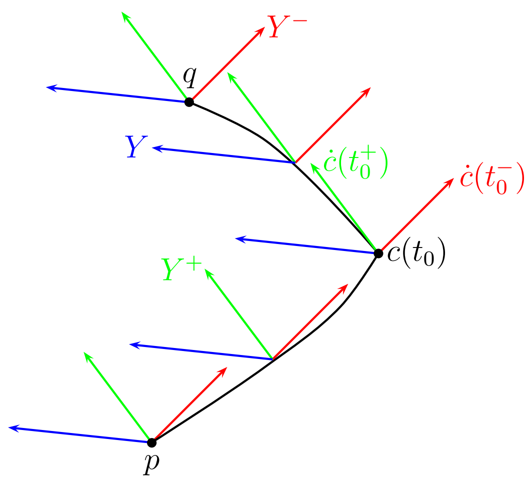
$$g\left(\frac{\nabla X}{\mathrm{d}t}, \dot{c}\right) = \underbrace{(b\gamma + 1)}_{>b\gamma}\underbrace{g(Y, \dot{c})}_{<0} - bg\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \frac{\nabla}{\mathrm{d}t}\dot{c}\right) < bg\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \frac{\nabla}{\mathrm{d}t}\dot{c}\right) - bg\left(\frac{\nabla}{\mathrm{d}t}\dot{c}, \frac{\nabla}{\mathrm{d}t}\dot{c}\right) = 0.$$

c) Finally, let $c$ be piecewise smooth and lightlike. If one of these smooth segments is not a pregeodesic, then it can be deformed to a timelike segment following b) and due to a) the whole curve can be deformed to a timelike one.
Let every segment be a lightlike pregeodesic and let $c$ be not differentiable at $t_0 \in (0,1)$, that is the left- and right-sided derivative $\dot{c}(t_0^{\pm})$ at $t_0$ are lineraly independent. According to a), it suffices to treat the case of one such $t_0$, i.e. $c$ is smooth on $[0, t_0) \cup (t_0, 1]$.

Let $Y^\pm$ be the parallel transport of $\dot c(t_0^\pm)$ along $c$, so $\dot c(t), Y^-(t)$ are linear dependent with the same orientation for all $t \in [0, t_0]$ and so are $\dot c(t), Y^+(t)$ for all $t \in [t_0, 1]$. For $Y := Y^+ - Y^-$, the Cauchy-Schwarz-inequality provides

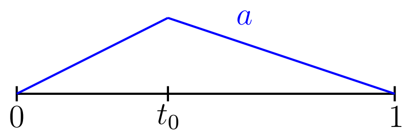
$$g\big(Y(t), \dot c(t)\big) = \begin{cases} g\big(Y^-(t), \dot c(t^-)\big) < 0, & t \in [0, t_0] \\ -g\big(Y^+(t), \dot c(t^+)\big) > 0, & t \in [t_0, 1] \end{cases},$$

since both vectors are lightlike and have the same orientation.




Let $a\colon [0,1] \to \mathbb{R}$ be continuous at $t_0$ and smooth everywhere else such that

$$a(0) = a(1) = 0, \qquad a'(t^\pm) \begin{cases} > 0 & \text{on } [0, t_0] \\ < 0 & \text{on } [t_0, 1] \end{cases}$$

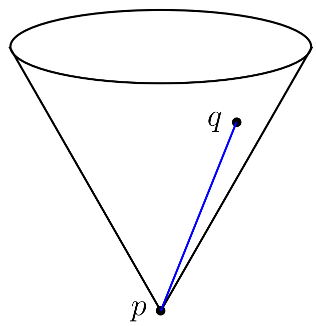



Then $X(t) := a(t)Y(t)$ satisfies

$$X(0) = X(1) = 0, \qquad g\left(\frac{\nabla X}{\mathrm{d}t}, \dot c\right) = a' g(Y, \dot c) < 0$$

and hence the corresponding variation with fixed start and endpoint yields the desired timelike deformation of $c$. □

If $c$ is a lightlike pregeodesic, then, in general, it can not be deformed into a timelike curve with the same start and end point. Consider, for instance, the Minkowski space and $p$ and $q$ connected by some lightlike geodesic as in the picture. Obviously, there is no timelike curve connecting $p$ and $q$ that this geodesic can be deformed into.




**Lemma 2.30.** *Let $c$ be a lightlike geodesic and $c_s$ a variation of it with variational field $X$ such that $X \perp \dot c$ at the start and end point. If there is a sequence $s_i \to 0$ such that $c_{s_i}$ is timelike for all $i$, then we have $X \perp \dot c$ everywhere.*

*Proof.* Let $(s_i)_{i\in\mathbb{N}}$ be such a sequence and without loss of generality, let either all $s_i > 0$ or all $s_i < 0$. Since $c$ is lightlike, we obtain

$$\lim_{i\to\infty} \frac{g(\dot{c}_{s_i}, \dot{c}_{s_i})}{s_i} = \lim_{i\to\infty} \frac{g(\dot{c}_{s_i}, \dot{c}_{s_i}) - g(\dot{c}, \dot{c})}{s_i} = \frac{\partial}{\partial s} g(\dot{c}_s, \dot{c}_s)\Big|_{s=0},$$

which, due to assumption, is either $\leq 0$ or $\geq 0$ for all $t \in [0,1]$, respectively. It follows that

$$\frac{\partial}{\partial s} g(\dot{c}_s, \dot{c}_s)\Big|_{s=0} = 2g\left(\frac{\nabla}{\partial s}\dot{c}_s\Big|_{s=0}, \dot{c}\right) = 2g\left(\frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s} c_s\Big|_{s=0}, \dot{c}\right) = 2g\left(\frac{\nabla X}{\partial t}, \dot{c}\right)$$

is either $\leq 0$ or $\geq 0$ for all $t \in [0,1]$, respectively. On the other hand, we have

$$\int_0^1 g\left(\frac{\nabla X}{\partial t}, \dot{c}\right) \mathrm{d}t = \int_0^1 \left(\frac{\mathrm{d}}{\mathrm{d}t} g(X, \dot{c}) - g\Big(X, \underbrace{\frac{\nabla \dot{c}}{\mathrm{d}t}}_{=0}\Big)\right) \mathrm{d}t = g(X, \dot{c})\big|_0^1 = 0.$$

Due to the constant sign of $g\left(\frac{\nabla X}{\partial t}, \dot{c}\right)$, we obtain $\frac{\mathrm{d}}{\mathrm{d}t} g(X, \dot{c}) = 0$, so $g(X, \dot{c})$ is constant on $[0,1]$, that is $0$. □

If $c_s$ is a variation of $c$ such that every $c_s$ is a geodesic, then the corresponding variational field $J(t) := \frac{\partial c_s(t)}{\partial s}\big|_{s=0}$ satisfies

$$\frac{\nabla^2 J}{\mathrm{d}t^2} = \frac{\nabla}{\partial t}\frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}\Big|_{s=0} = \frac{\nabla}{\partial t}\frac{\nabla}{\partial s}\frac{\partial c_s}{\partial t}\Big|_{s=0} = \frac{\nabla}{\partial s}\underbrace{\frac{\nabla}{\partial t}\frac{\partial c_s}{\partial t}}_{=0}\Big|_{s=0} + R(\dot{c}, J)\dot{c}.$$

**Definition 2.31.** A vector field $J$ along a geodesic $c$ is called **Jacobi field**, if

$$\frac{\nabla^2 J}{\mathrm{d}t^2} + R(J, \dot{c})\dot{c} = 0. \tag{2.1}$$

The Jacobi equation (2.1) is a linear, ordinary differential equation of second order in $J$. Therefore, a jacobi field can be found along all of $c$ and is determined by the data $J(0)$ and $\frac{\nabla J}{\mathrm{d}t}(0)$. The set of jacobi fields constitutes a $2\dim(M)$-dimensional vector space.

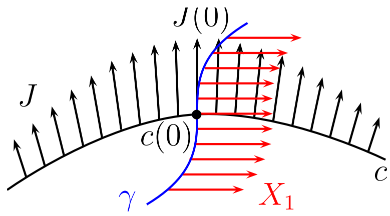
**Lemma 2.32.** *Let $J$ be a smooth vector field along some geodesic $c$. Then we have*

*(i) $J$ is a Jacobi field.*

*(ii) There is a geodesic variation $c_s$ of $c$ with variational field $J$.*

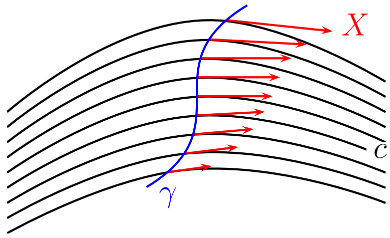
*Proof.* We have already seen $(ii) \Rightarrow (i)$, so it remains to show $(i) \Rightarrow (ii)$.

Let $J$ be a Jacobi-field along $c$. We choose a smooth curve $\gamma$ with $\gamma(0) = c(0)$ and $\dot\gamma(0) = J(0)$. Let $X_1$ and $X_2$ be parallel vector field along $\gamma$ given by $X_1(0) = \dot c(0)$ and $X_2(0) = \frac{\nabla J}{\mathrm{d}t}(0)$. Set




$$X(s) := X_1(s) + sX_2(s), \quad c_s(t) := \exp_{\gamma(s)}\big(tX(s)\big).$$

Then we have




$$c_0(t) = \exp_{\gamma(0)}\big(tX(0)\big) = \exp_{c(0)}\big(t\dot c(0)\big) = c(t),$$

so $c_s$ is a geodesic variation. Therefore, its variational field $\frac{\partial c_s}{\partial s}\big|_{s=0}$ is a Jacobi field, and we complete the proof by showing that it coincides with $J$.
Since Jacobi fields are determined by their initial data, it suffices to check:

$$\frac{\partial c_s}{\partial s}\bigg|_{s=t=0} = \frac{\partial}{\partial s}\exp_{\gamma(s)}(0)\bigg|_{s=0} = \dot\gamma(0) = J(0),$$

$$\frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}\bigg|_{s=t=0} = \frac{\nabla}{\partial s}\frac{\partial}{\partial t}\exp_{\gamma(s)}\big(tX(s)\big)\bigg|_{s=t=0} = \frac{\nabla X}{\partial s}(0) = \underbrace{\frac{\nabla X_1}{\partial s}}_{=0}(0) + X_2(0) = \frac{\nabla J}{\mathrm{d}t}(0). \square$$

**Example 2.33.**

1. Let $c$ be a geodesic and $J(t) := (at+b)\dot c(t)$ for some $a, b \in \mathbb{R}$. Since

   $$\frac{\nabla^2 J}{\mathrm{d}t^2}(t) = \frac{\nabla}{\mathrm{d}t}\Bigg(a\dot c(t) + (at+b)\underbrace{\frac{\nabla \dot c}{\mathrm{d}t}}_{=0}(t)\Bigg) = a\frac{\nabla \dot c}{\mathrm{d}t}(t) = 0 = (at+b)R(\dot c, \dot c)\dot c = R(\dot c, J)\dot c,$$

   $J$ is a Jacobi field with corresponding geodesic variation $c_s(t) = c\big((1+as)t + bs\big)$, i.e. a mere reparametrization of $c$.

2. For $M = \mathbb{R}^n$ the Euclidean or Minkowski space, the Jacobi equation reads $\frac{\nabla^2 J}{\mathrm{d}t^2} = 0$ with solution $J(t) = tX(t) + Y(t)$ for any parallel vector fields $X, Y$. The corresponding geodesic variation of the straight line $c$ is then

   $$c_s(t) = c(t) + s\big(tX(t) + Y(t)\big).$$

3. For $M$ having constant sectional curvature $\kappa$, we obtain

   $$R(X,Y)Z = \kappa\big(g(Y,Z)X - g(X,Z)Y\big)$$

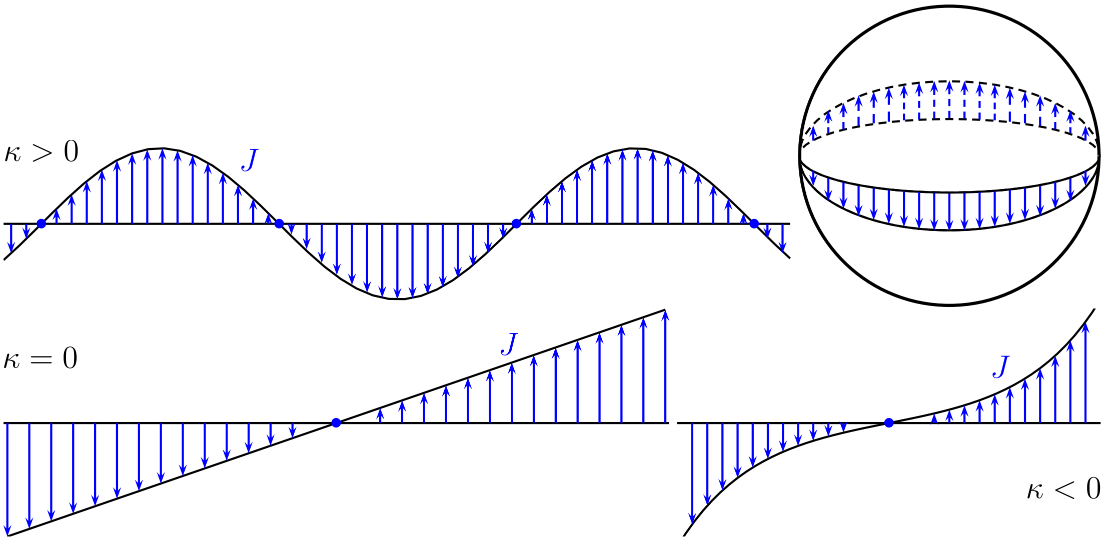
for vector fields $X, Y, Z$. Let $c$ be a geodesic and $X, Y$ parallel vector fields along $c$, which are pointwise orthogonal to $\dot{c}$. For $\delta \in \mathbb{R}$, we introduce the generalized sine and cosine function via

$$\mathfrak{s}_\delta(t) := \begin{cases} \frac{1}{\sqrt{\delta}} \sin\big(\sqrt{\delta} t\big), & \delta > 0, \\ t, & \delta = 0 \\ \frac{1}{\sqrt{|\delta|}} \sinh\big(\sqrt{|\delta|} t\big), & \delta < 0 \end{cases}, \quad \mathfrak{c}_\delta(t) := \begin{cases} \frac{1}{\sqrt{\delta}} \cos\big(\sqrt{\delta} t\big), & \delta > 0, \\ 1, & \delta = 0 \\ \frac{1}{\sqrt{|\delta|}} \cosh\big(\sqrt{|\delta|} t\big), & \delta < 0 \end{cases},$$

which then satisfy $\mathfrak{s}_\delta'' = -\delta \mathfrak{s}_\delta$ and $\mathfrak{c}_\delta'' = -\delta \mathfrak{c}_\delta$. Hence, $J(t) := \mathfrak{s}_{\eta\kappa}(t) X(t) + \mathfrak{c}_{\eta\kappa} Y(t)$, where $\eta := g(\dot{c}, \dot{c})$, fulfills $\frac{\nabla^2 J}{\mathrm{d}t^2} = -\eta\kappa J$. On the other hand, we have

$$R(\dot{c}, J)\dot{c} = \kappa\big(\underbrace{g(J, \dot{c})}_{=0}\, \dot{c} - \underbrace{g(\dot{c}, \dot{c})}_{=\eta}\, J\big) = -\eta\kappa J,$$

so $J$ is a Jacobi field.




Let $P \subset M$ be a semi-Riemannian submanifold and $p \in P$, so for vector fields $X, Y$ on $P$, we have

$$\big(\nabla_X Y\big)(p) = \underbrace{\big(\nabla_X^P Y\big)(p)}_{\in T_pP} + \underbrace{I\!I_p\big(X(p), Y(p)\big)}_{\in N_pP}.$$

Recall the second fundamental form $I\!I_p \colon T_pP \times T_pP \to N_pP$, which is bilinear and symmetric, and we introduce $\widetilde{I\!I}_p \colon T_pP \times N_pP \to T_pP$ via

$$g\big(\widetilde{I\!I}(X, \nu), Y\big) := -g\big(I\!I(X, Y), \nu\big),$$

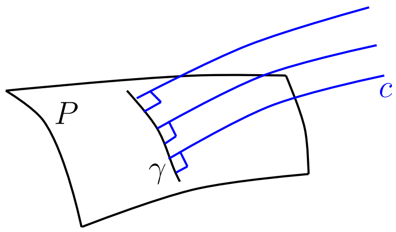
where $X, Y$ are tangential and $\nu$ normal to $P$. This map is clearly bilinear as well, and for some fixed $X$, it is given by $\widetilde{I\!I}(X, \cdot) = -I\!I(X, \cdot)^t \colon N_pP \to T_pP$.

**Lemma 2.34.** *Let $P \subset M$ be a semi-Riemannian submanifold, $p \in P$ and $c$ a geodesic in $M$ with $c(0) = p$ and $\dot{c}(0) \in N_pP$. For $J$ a Jacobi field along $c$, the following statements are equivalent:*

*(i) $J$ is the variational vector field of some geodesic variation $c_s$ of $c$ with $c_s(0) \in P$ and $\dot{c}_s(0) \in N_{c_s(0)}P$ for all $s$.*

*(ii) $J(0) \in T_pP$ and $\tan\big(\frac{\nabla J}{\mathrm{d}t}(0)\big) = \widetilde{I\!I}_p\big(J(0), \dot{c}(0)\big)$, where $\tan\colon T_pM \to T_pP$ is the orthogonal projection.*

*Proof.*
$(i) \Rightarrow (ii)$ : The curve $\gamma(s) := c_s(0)$ runs in $P$, that is $J(0) = \gamma'(0) \in T_pP$. For $X$ some tangential vector field on $P$ and since $g\big(X(\gamma(s)), \dot{c}_s(0)\big) = 0$ for all $s$, we have




$$0 = \frac{\mathrm{d}}{\mathrm{d}s} g\big(X(\gamma(s)), \dot{c}_s(0)\big)\Big|_{s=0} = g\left(\frac{\nabla(X\circ\gamma)}{\mathrm{d}s}(0), \dot{c}(0)\right) + g\left(X(p), \frac{\nabla}{\partial s}\frac{\partial c_s}{\partial t}\Big|_{s=t=0}\right)$$

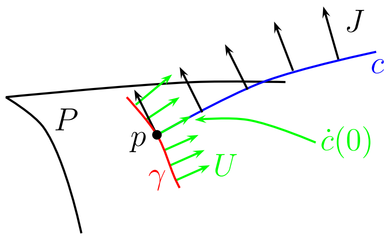
$$= g\Bigg(\underbrace{\frac{\nabla^P(X\circ\gamma)}{\mathrm{d}s}(0)}_{\in T_pP}, \underbrace{\dot{c}(0)}_{\in N_pP}\Bigg) + g\Big(I\!I_p\big(\gamma'(0), X(p)\big), \dot{c}(0)\Big) + g\left(X(p), \frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}\Big|_{s=t=0}\right)$$

$$= g\left(-\widetilde{I\!I}_p\big(J(0), \dot{c}(0)\big) + \frac{\nabla J}{\mathrm{d}t}(0), X(p)\right).$$

This holds for any tangential vector field $X$ and hence provides the identity. $(ii) \Rightarrow (i)$ : Similar to the proof of Lemma 2.32, the geodesic variation is given by $c_s(t) = \exp_{\gamma(s)}\big(tX(s)\big)$ with

$$\gamma(0) = p, \qquad \gamma'(0) = J(0), \qquad X(0) = \dot{c}(0), \qquad \frac{\nabla X}{\mathrm{d}s}(0) = \frac{\nabla J}{\mathrm{d}t}(0),$$

and it remains to show $\gamma(s) \in P$ and $X(s) \in N_{\gamma(s)}P$ for all $s$. Since $J(0) \in T_pP$, we can choose $\gamma$ such that it runs entirely in $P$. Let $U(s) \in N_{\gamma(s)}P$ be the normal-parallel transport of $\dot{c}(0)$ along $\gamma$, i.e. $U(0) = \dot{c}(0)$ and $\mathrm{nor}\big(\frac{\nabla U}{\mathrm{d}s}\big) = 0$, and let $V(s) \in N_{\gamma(s)}P$ denote the corresponding normal-parallel transport of $\mathrm{nor}\big(\frac{\nabla J}{\mathrm{d}t}(0)\big)$ along $\gamma$. Then $X(s) := U(s) + sV(s) \in T_{\gamma(s)}P$ satisfies




$$X(0) = \dot{c}(0), \qquad \mathrm{nor}\left(\frac{\nabla X}{\mathrm{d}s}(0)\right) = \mathrm{nor}\left(\frac{\nabla J}{\mathrm{d}t}(0)\right), \quad \tan\left(\frac{\nabla X}{\mathrm{d}s}(0)\right) = \frac{\nabla U}{\mathrm{d}s}(0). \qquad (2.2)$$

Thus, for any tangent vector field $Y$ along $\gamma$, we obtain

$$\begin{aligned} g\left(\tan\left(\frac{\nabla X}{\mathrm{d}s}(0)\right), Y(0)\right) &= g\left(\frac{\nabla U}{\mathrm{d}s}(0), Y(0)\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\underbrace{g\big(U(s), Y(s)\big)}_{=0}\Big|_{s=0} - g\Bigg(\underbrace{U(0)}_{\in N_pP}, \frac{\nabla Y}{\mathrm{d}s}(0)\Bigg) = -g\big(\underbrace{U(0)}_{=\dot c(0)}, I\!I_p\big(\underbrace{\gamma'(0)}_{=J(0)}, Y(0)\big) \\ &= g\big(\widetilde{I\!I}_p\big(J(0), \dot c(0)\big), Y(0)\big) = g\left(\tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right), Y(0)\right), \end{aligned}$$

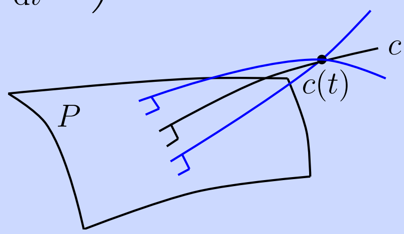
and hence $\tan\big(\frac{\nabla X}{\mathrm{d}s}(0)\big) = \tan\big(\frac{\nabla J}{\mathrm{d}t}(0)\big)$, that is $\frac{\nabla X}{\mathrm{d}s}(0) = \frac{\nabla J}{\mathrm{d}t}(0)$ due to (2.2). □

**Definition 2.35.** Let $P \subset M$ be a semi-Riemannian submanifold, $p \in P$ and $c$ a geodesic in $M$ with $c(0) = p$ and $\dot c(0) \in N_pP$. We say that $P$ has a **focal point along $c$ of order $\mu$ at $t$** if

$$\mu := \left\{ \text{Jacobi fields } J \text{ along } c \;\middle|\; J(0) \in T_pP,\ \tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right) = \widetilde{I\!I}_p\big(J(0), \dot c(0)\big),\ J(t) = 0 \right\}$$

is positive.

We call $t$ the **focal value of $P$ along $c$**.

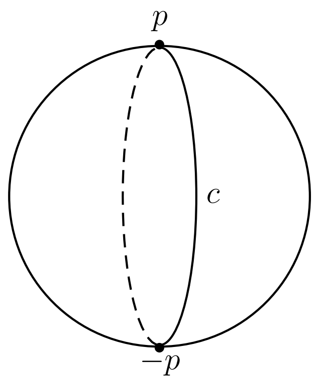



**Example 2.36.**

1. For $P = \{p\}$, we have $\mu = \dim\big\{J(0) = J(t) = 0\big\}$. In this case, we speak of **conjugated points** rather than focal points.
2. Let $M := S^n, P := \{p\}$ and $c$ a geodesic, parametrized by arc length with $c(0) = p$. Then for any parallel vector field $E$ along $c$ with $E \perp \dot c$, we obtain a Jacobi field

$$J(t) := \sin(t) \cdot E(t)$$

with $J(0) = J(k\pi) = 0$ for all $k \in \mathbb{N}$. Hence, along any such geodesic, the point $p$ has a conjugated point at $t = k\pi$, $k \in \mathbb{N}$, of degree $\mu = n - 1$.

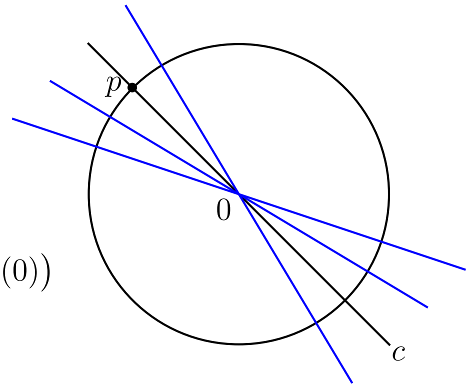


**Remark 2.37.** Let $\dim(M) = n$ and $\dim(P) = m$. The condition $J(0) \in T_pP$ reduces the dimension of the vector space of Jacobi fields along $c$ by $n - m$. Moreover, the condition $\tan\big(\frac{\nabla J}{\mathrm{d}t}(0)\big) = \widetilde{II}_p\big(J(0), \dot{c}(0)\big)$ yields a further reduction by $m$, so the space of Jacobi fields satisfying both conditions is $n$-dimensional. This space contains the Jacobi field $J(t) = t\dot{c}(t)$, which does not vanish for any $t \neq 0$, and therefore

$$\mu \leq n - 1.$$

3. Consider the Euclidean space $\mathbb{R}^{n+1}$ and $P = S^n$. For $p \in P$ set $c(t) := (1-t)p$ and for $E \in T_pP$ consider the vector field $t \mapsto E(t)$ given by parallel transport along $c$. We obtain a Jacobi field

$$J(t) := (1-t)E(t)$$

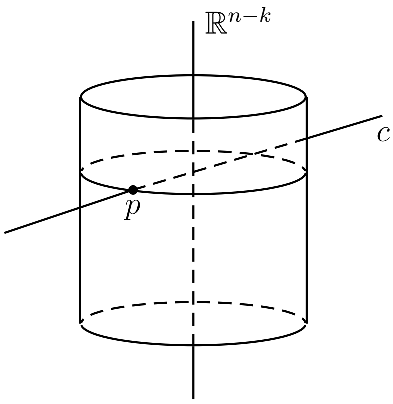



with $J(0) = E$ and $J(1) = 0$ as well as $\tan\big(\frac{\nabla J}{\mathrm{d}t}(0)\big) = -E$. Furthermore, $II(X,Y) = g(X,Y) \cdot \dot{c}(0)$ and thus $\widetilde{II}\big(X, \dot{c}(0)\big) = -X$. It follows that

$$\tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right) = \widetilde{II}\big(E, \dot{c}(0)\big) = \widetilde{II}\big(J(0), \dot{c}(0)\big)$$

and hence, $S^n$ has focal points along $c$ of degree $n$ at $0$, that is $t = 1$.

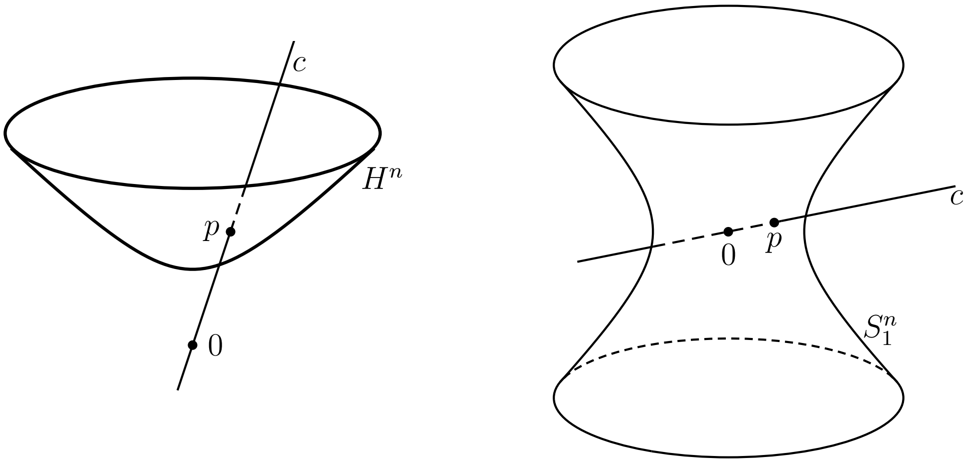
4. Again in $(n+1)$-dimensional Euclidean space, consider the cylinder $P := S^k \times \mathbb{R}^{n-k}$, and for $p = (p_1, p_2) \in P$, let $c(t) := \big((1-t)p_1, p_2\big)$. Let $E \in T_{p_1}S^k$ and define the parallel vector field $E(t)$ and the Jacobi field $J(t)$ as in example 3, which then satisfies




$$J(0) = E, \quad J(1) = 0, \quad \tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right) = -E.$$

Therefore, correspondingly, $P$ has a focal point of order $k$ along $c$ at $t = 1$.

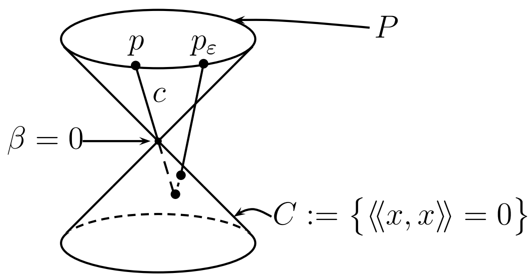
5. In $M = \mathbb{R}^{n+1}_{\text{Mink}}$, consider $P = H^n$ and $P = S^n_1$. For fixed $p \in P$, let $c(t) := (1-t)p$, which is a geodesic in $M$ and due to elementary properties of the submanifolds $H^n, S^n_1$ (see section 1.2), $c$ stands orthogonal to both with respect to the Minkowski metric, i.e. $\dot{c}(0) \in N_pP$. Considering $E(t)$ and $J(t)$ as in example 3 shows that in both cases $P$ has a focal point of degree $n$ along $c$ at $t = 1$.




**Proposition 2.38.** *Let $M$ be a Lorentz manifold, $P \subset M$ a spacelike submanifold and $c\colon [0,b] \to M$ a lightlike geodesic in $M$ with $c(0) = p \in P$ and $\dot c(0) \in N_pP$. Set $q := c(b)$. If $P$ has a focal point along $c$ before $q$, that is for some $t \in (0,b)$, then in every neighborhood of $c$, with respect to the compact-open topology, there is a timelike curve from $P$ to $q$.*

The proof of the Proposition needs some preparation and we start with an example:

**Example 2.39.**
For $M := \mathbb{R}^3_{\mathrm{Mink}}$, consider $P := \{1\} \times S^1$ and point $p := (1,1,0) \in P$ such that the curve $c(t) := (1-t)p$ is a lightlike geodesic. Then $c$ hits $q := (\beta,\beta,0)$ for any $\beta \in \mathbb{R}$, and for $\varepsilon > 0$ small and $p_\varepsilon := \big(1, \cos(\varepsilon), \sin(\varepsilon)\big)$, we calculate




$$\langle\!\langle q - p_\varepsilon, q - p_\varepsilon\rangle\!\rangle$$

$$= \left\langle\!\!\left\langle \begin{pmatrix} \beta - 1 \\ \beta - \cos(\varepsilon) \\ -\sin(\varepsilon)\end{pmatrix}, \begin{pmatrix} \beta - 1 \\ \beta - \cos(\varepsilon) \\ -\sin(\varepsilon)\end{pmatrix} \right\rangle\!\!\right\rangle = -(\beta-1)^2 + \big(\beta - \cos(\varepsilon)\big)^2 + \sin^2(\varepsilon)$$

$$= -\beta^2 + 2\beta - 1 + \beta^2 - 2\beta\cos(\varepsilon) + \cos^2(\varepsilon) + \sin^2(\varepsilon) = 2\beta\big(\underbrace{1 - \cos(\varepsilon)}_{>0}\big).$$

For $\beta = 0$, the connecting curve

$$c_\varepsilon(t) := p_\varepsilon + \frac{t}{b}(q - p_\varepsilon) = \left(1 - t, \cos(\varepsilon) + \frac{t}{b}\big(\beta - \cos(\varepsilon)\big), \sin(\varepsilon)\Big(1 - \frac{t}{b}\Big)\right)$$

is a timelike geodesic, which, on $[0,b]$, converges uniformly to $c$ for $\varepsilon \to 0$.

**Lemma 2.40.** *Let $M$ be a semi-Riemannian manifold, $P \subset M$ a semi-Riemannian submanifold and $c\colon [0,b] \to M$ a geodesic with $c(0) = p \in P$ and $\dot{c}(0) \in N_pP$. Then*

$$\mathcal{T} := \big\{ t \in [0,b] \;\big|\; P \text{ has a focal point along } c \text{ at } t \big\}$$

*is compact and contained in $(0,b]$.*

*Proof.* Let $\mathcal{V}$ denote the set of Jacobi fields $J$ along $c$ with $J(0) \in T_pP$ and $\tan\big(\frac{\nabla J}{\mathrm{d}t}(0)\big) = \widetilde{II}\big(J(0), \dot{c}(0)\big)$. For some Riemannian metric $h$ on $M$ and with

$$\|J\| := \sup_{t\in[0,b]} \big|J(t)\big|_h + \sup_{t\in[0,b]} \left|\frac{\nabla J}{\mathrm{d}t}(t)\right|_h,$$

it becomes an $n$-dimensional normed vector space, where $n := \dim(M)$.

We show that $\mathcal{T}$ is closed. Consider a sequence $\{t_i\}_{i\in I}$ in $\mathcal{T}$ converging to some $t \in (0,b]$ and corresponding $J_i \in \mathcal{V}\backslash\{0\}$ with $J_i(t_i) = 0$, i.e. $\{J_i\}_{i\in I} \subset \mathcal{T}$. Without loss of generality, e assume $\|J_i\| = 1$. Let $J \in \mathcal{V}$ denote the limit of some subsequence, which therefore fulfills $\|J\| = 1$, and thus, in particular, $J \neq 0$.
Let $\Pi_t^s\colon T_{c(s)}M \to T_{c(t)}M$ denote the parallel transport along $c$ with respect to the Levi-Civita-connection given by the original semi-Riemannian metric on $M$. Then we have

$$\begin{aligned} \big|J(t)\big|_h &\leq \big|J(t) - \Pi_t^{t_i} J(t_i) + \Pi_t^{t_i} J(t_i) - \Pi_t^{t_i} \underbrace{J_i(t_i)}_{=0}\big|_h \\ &\leq \underbrace{\big|J(t) - \Pi_t^{t_i} J(t_i) + \Pi_t^{t_i} J(t_i)\big|_h}_{\to 0} + C\|J - J_i\| \end{aligned}$$

for some constant $C > 0$ and by continuity of $J$. It follows that $J(t) = 0$, that is $t \in \mathcal{T}$.
Now assume $t = 0$, i.e. $J(0) = 0$ and

$$\tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right) = \widetilde{II}\big(J(0), \dot{c}(0)\big) = 0.$$

It remains to show $\operatorname{nor}\big(\frac{\nabla J}{\mathrm{d}t}(0)\big) = 0$, which then yields $J = 0$ and therefore contradicts $\|J\| = 1$. Note that

$$\left|\operatorname{nor}\left(\frac{\Pi_0^{t_i} J(t_i)}{t_i}\right)\right|_h = \left|\operatorname{nor}\left(\frac{\Pi_0^{t_i} J(t_i) - J(0)}{t_i - 0}\right)\right|_h \longrightarrow \left|\operatorname{nor}\left(\frac{\nabla J}{\mathrm{d}t}(0)\right)\right|_h,$$

and on the other hand

$$\left|\operatorname{nor}\left(\frac{\Pi_0^{t_i} J(t_i)}{t_i}\right)\right|_h = \frac{1}{t_i}\left|\operatorname{nor}\left(\int_0^\tau \Pi_0^\tau\left(\frac{\nabla J}{\mathrm{d}\tau}(\tau) - \frac{\nabla J_i}{\mathrm{d}\tau}(\tau)\right)\mathrm{d}\tau - \underbrace{J_i(0)}_{\in T_pP}\right)\right|_h$$
$$\leq \frac{1}{t_i}\int_0^{t_i}\left|\Pi_0^\tau\left(\frac{\nabla(J-J_i)}{\mathrm{d}\tau}(\tau)\right|_h \mathrm{d}\tau$$
$$\leq \frac{C}{t_i}\int_0^{t_i}\|J-J_i\|\,\mathrm{d}\tau = C\|J-J_i\| \longrightarrow 0$$

for some constant $C > 0$. □

**Corollary 2.41.** *Let $M, P, p$ as in the Lemma. If $P$ has a focal point along $c$, then there is a well-defined first focal value* $\min \mathcal{T} > 0$.

**Remark 2.42.** Furthermore, one can show

| | | |
|---|---|---|
| $M$ Riemannian | $\Longrightarrow$ | $\mathcal{T}$ discrete |
| $M$ Lorentzian and $c$ causal | $\Longrightarrow$ | $\mathcal{T}$ discrete |
| $M$ Lorentzian and $c$ spacelike | $\Longrightarrow$ | Every compact subset of $(0,b]$ is the set of focal points for some choice of $M$ and $c$. |

*Proof of Proposition 2.38.* Let $t_0 > 0$ be the first focal value of $P$ along $c$, so there is a Jacobi field $J \neq 0$ along $c$ with

$$J(0) \in T_pP, \qquad \tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right) = \widetilde{II}\big(J(0), \dot{c}(0)\big), \qquad J(t_0) = 0.$$

Note that since $t_0$ is the first focal value, we have $J(t) \neq 0$ for all $t_0 \in (0, t_0)$.

**Claim a):** There is some $\delta \in (0, b - t_0)$ such that $J = f \cdot U$ on $[0, t_0 + \delta]$, where $U$ is a spacelike unit normal field along $c$ and $f$ a smooth function with $f\big|_{(0,t_0)} > 0$ and $f\big|_{(t_0,t_0+\delta)} < 0$. Since $J$ is a Jacobi field and $c$ a geodesic, we have

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle J, \dot{c}\rangle = \left\langle \frac{\nabla^2 J}{\mathrm{d}t^2}, \dot{c}\right\rangle = \big\langle R(\dot{c}, J)\dot{c}, \dot{c}\big\rangle = 0,$$

so $\langle J, \dot{c}\rangle$ is a function of the form $\alpha t + \beta$, $\alpha, \beta \in \mathbb{R}$, which, by assumption, vanishes for $t = 0$ and $t = t_0$. Therefore, $\langle J, \dot{c}\rangle = 0$, that is $J \perp \dot{c}$ on all of $[0, b]$.

This does not rule out that $J$ is tangential to $c$ somewhere since $c$ was assumed to be lightlike. If there was some $t_1 \in (0, t_0)$ such that $J(t_1) = a\dot{c}(t_1)$, then $\widetilde{J}(t) := J(t) - \frac{at}{t_1}\dot{c}(t)$ would provide a Jacobi field along $c$ with $\widetilde{J}(0) = J(0) \in T_pP$ and $\frac{\nabla \widetilde{J}}{\mathrm{d}t} = \frac{\nabla J}{\mathrm{d}t} - \frac{a}{t_1}\dot{c}$, which therefore satisfies

$$\tan\left(\frac{\nabla \widetilde{J}}{\mathrm{d}t}(0)\right) = \tan\left(\frac{\nabla J}{\mathrm{d}t}(0)\right) = \widetilde{I\!I}\big(J(0), \dot{c}(0)\big) = \widetilde{I\!I}\big(\widetilde{J}(0), \dot{c}(0)\big).$$

Furthermore, $\widetilde{J}(t_1) = 0$, so $t_1 < t_0$ would be a focal value, which contradicts the assumption on $t_0$. Thus, $J$ is nowhere tangential to $\dot{c}$ and hence spacelike on $(0, t_0)$.
We write $J := \varphi \cdot Y$, where $Y$ is a smooth vector field along $c$ and

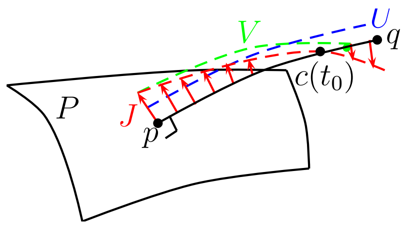
$$\varphi(t) := \begin{cases} t(t_0 - t), & J(0) = 0 \\ t_0 - t, & J(0) \neq 0 \end{cases}.$$

By the recent results, $Y$ has to be spacelike on $(0, t_0)$ and clearly $J(0) \neq 0$ implies $Y(0) \neq 0$. This is also true if $J(0) = 0$ because then $Y(0) = \frac{1}{t_0}\frac{\nabla J}{\mathrm{d}t}(0) \neq 0$, and similarly, $Y(t_0) = \frac{1}{\varphi'(t_0)}\frac{\nabla J}{\mathrm{d}t}(t_0) \neq 0$ since $\varphi'(t_0) \neq 0$.
Moreover, $Y(0), Y(t_0)$ are spacelike. For $J(0) \neq 0$, we obviously have $Y(0) = \frac{1}{t_0}J(0)$, which is spacelike. For $J(0) = 0$, $\langle J, \dot{c}\rangle = 0$ yields

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\langle J, \dot{c}\rangle = \left\langle \frac{\nabla J}{\mathrm{d}t}, \dot{c}\right\rangle \qquad \Longrightarrow \qquad \frac{\nabla J}{\mathrm{d}t} \perp \dot{c}.$$

If $Y(0)$ was tangential to $\dot{c}$, i.e. $Y(0) = a\dot{c}(0)$, we would have $\frac{\nabla J}{\mathrm{d}t}(0) = \varphi'(0)Y(0) = at_0\dot{c}(0)$ and thus, due to the initial conditions, $J(t) = at_0 \cdot \dot{c}(t)$, which contradicts $J(t_0) = 0$. Analogously $Y(t_0) = a\dot{c}(t_0)$ implies $\frac{\nabla J}{\mathrm{d}t}(t_0) = a\varphi'(t_0) \cdot \dot{c}(t_0)$ and hence $J(t) = a\varphi'(t_0) \cdot t\dot{c}(t)$, which also does not vanish for $t = t_0$.
It follows that $Y$ is a nowhere vanishing smooth and spacelike vector field on $[0, t_0]$ and therefore also on $[0, t_0 + \delta]$ for some $\delta > 0$. Setting $U := \frac{Y}{|Y|}$ and $f := \varphi \cdot |Y|$ thus proves Claim a).

**Claim b):** There is some $\delta \in (0, b - t_0)$ and a vector field $V$ along $c$ with $V(0) = J(0)$ and $V(t_0 + \delta) = 0$ such that $V \perp \dot{c}$ on $[0, t_0 + \delta]$ and $\left\langle \frac{\nabla^2 V}{\mathrm{d}t^2} + R(V, \dot{c})\dot{c}, V\right\rangle > 0$ on $(0, t_0 + \delta)$.
For $f, U$ as in the proof of claim a), we write $V := (f + h)U = J + hU$ for some $h$ yet to be determined. Then (2.1) provides




$$\begin{aligned} \frac{\nabla^2 V}{\mathrm{d}t^2} + R(V, \dot{c})\dot{c} &= \frac{\nabla^2 J}{\mathrm{d}t^2} + h''U + 2h'\frac{\nabla U}{\mathrm{d}t} + h\frac{\nabla^2 U}{\mathrm{d}t^2} + R(J, \dot{c})\dot{c} + hR(U, \dot{c})\dot{c} \\ &= h''U + 2h'\frac{\nabla U}{\mathrm{d}t} + h\left(\frac{\nabla^2 U}{\mathrm{d}t^2} + R(U, \dot{c})\dot{c}\right) \end{aligned}$$

and thus

$$\left\langle \frac{\nabla^2 V}{\mathrm{d}t^2} + R(V,\dot c)\dot c, V\right\rangle = (f+h)\Big(h'' \underbrace{\langle U,U\rangle}_{=1} + 2h' \underbrace{\left\langle \frac{\nabla U}{\mathrm{d}t}, U\right\rangle}_{=0} + h \underbrace{\left\langle \frac{\nabla^2 U}{\mathrm{d}t^2} + R(U,\dot c)\dot c, U\right\rangle}_{=:l}\Big)$$

$$= (f+h)(h''+hl).$$

Pick some $\alpha > 0$ such that $l \geq -\alpha^2$ on $[0, t_0+\delta]$ and set $h(t) := \beta\big(e^{\alpha t}-1\big)$ with $\beta := -\frac{f(t_0+\delta)}{h(t_0+\delta)} > 0$, i.e. $f+h > 0$ on $(0,t_0]$ and $(f+h)(t_0+\delta) = 0$. Without loss of generality, we assume that $f+h$ has no positive zeros before $t_0+\delta$, that is $f+h>0$ on $(0,t_0+\delta)$. It follows that

$$\left\langle \frac{\nabla^2 V}{\mathrm{d}t^2} + R(V,\dot c)\dot c, V\right\rangle = \underbrace{(f+h)}_{>0}\underbrace{(h''+hl)}_{>0} > 0$$

since $h''+hl = \alpha^2 h + \alpha^2\beta + hl \geq \alpha^2\beta > 0$.

**Claim c):** There is a smooth vector field $X$ along $c$ such that on $[t_0, t_0+\delta]$, we have

$$X(0) = I\!I\big(J(0),J(0)\big), \qquad X(t_0+\delta) = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \langle X, \dot c\rangle\right) \leq 0.$$

We directly calculate

$$\begin{aligned}\big\langle I\!I\big(J(0),J(0)\big), \dot c(0)\big\rangle &= -\big\langle \widetilde{I\!I}\big(J(0),\dot c(0)\big), J(0)\big\rangle \\ &= -\left\langle \tan\!\left(\frac{\nabla J}{\mathrm{d}t}(0)\right), \underbrace{J(0)}_{\in T_pP}\right\rangle = -\left\langle \frac{\nabla J}{\mathrm{d}t}(0), J(0)\right\rangle\end{aligned}$$

as well as

$$\begin{aligned}\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle &= \left\langle J + b\big(e^{\alpha t}-1\big)U, \frac{\nabla J}{\mathrm{d}t} + \alpha\beta e^{\alpha t}U + b\big(e^{\alpha t}-1\big)\frac{\nabla U}{\mathrm{d}t}\right\rangle \\ &= \left\langle J, \frac{\nabla J}{\mathrm{d}t}\right\rangle + \big\langle V, \alpha\beta e^{\alpha t}U\big\rangle + b\big(e^{\alpha t}-1\big)\left\langle U, \frac{\nabla J}{\mathrm{d}t}\right\rangle.\end{aligned}$$

This particularly implies $\big\langle V(0), \frac{\nabla V}{\mathrm{d}t}(0)\big\rangle = -\big\langle I\!I\big(J(0),J(0)\big), \dot c(0)\big\rangle + \alpha\beta\|J(0)\|$.
We assume $\big\langle I\!I\big(J(0),J(0)\big), \dot c(0)\big\rangle =: -a \neq 0$, so there is some $L_0 \in N_pP$ such that $I\!I\big(J(0),J(0)\big) = aL_0$ and $\big\langle L_0, \dot c(0)\big\rangle = -1$. Let $L$ denote the parallel vector field along $c$ determined by $L(0) = L_0$, for which therefore $\langle L, \dot c\rangle = -1$ holds. Then

$$X(t) := \left(\left\langle V(t), \frac{\nabla V}{\mathrm{d}t}(t)\right\rangle + \frac{\alpha\beta\|J(0)\|}{t_0+\delta}(t - t_0 - \delta)\right)L(t)$$

yields a vector field along $c$, which satisfies

$$X(0) = \Big(a + \alpha\beta\|J(0)\| - \frac{\alpha\beta\|J(0)\|}{t_0+\delta}(t_0+\delta)\Big)L(0) = aL_0 = \mathit{II}\big(J(0),J(0)\big)$$

$$X(t_0+\delta) = \left\langle \underbrace{V(t_0+\delta)}_{=0}, \frac{\nabla V}{\mathrm{d}t}(t_0+\delta)\right\rangle \cdot L(t_0+\delta) = 0.$$

Furthermore, $\langle L, \dot{c}\rangle = -1$ ensures

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \langle X, \dot{c}\rangle\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle - \left(\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \frac{\alpha\beta\|J(0)\|}{t_0+\delta}(t-t_0-\delta)\right)\right]\\ &= -\frac{\alpha\beta\|J(0)\|}{t_0+\delta} \leq 0.\end{aligned}$$

Now consider the case $\big\langle \mathit{II}\big(J(0),J(0)\big), \dot{c}(0)\big\rangle = 0$, i.e. $\big\langle V(0), \frac{\nabla V}{\mathrm{d}t}(0)\big\rangle = \alpha\beta\|J(0)\|$. Here we choose two parallel vector fields $L$ and $Z$ along $c$ with $\langle L,\dot{c}\rangle = -1$ and $Z(0) := \mathit{II}\big(J(0),J(0)\big)$ such that $\langle Z, \dot{c}\rangle = 0$. Correspondingly, we set

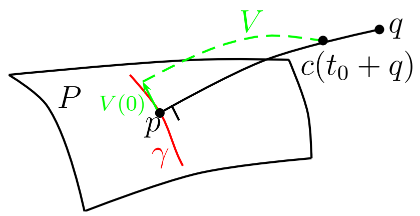
$$X(t) := \left(\left\langle V(t), \frac{\nabla V}{\mathrm{d}t}(t)\right\rangle + \frac{\alpha\beta\|J(0)\|}{t_0+\delta}(t-t_0-\delta)\right)L(t) + \left(1 - \frac{t}{t_0+\delta}\right)Z(t)$$

and analogously to the previous case, we obtain

$$X(0) = 0\cdot L(0) + Z_0 = \mathit{II}\big(J(0),J(0)\big), \qquad X(t_0+\delta) = 0\cdot L(0) + 0\cdot Z_0 = 0,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \langle X, \dot{c}\rangle\right) = \frac{\alpha\beta\|J(0)\|}{t_0+\delta}\overbrace{\langle L, \dot{c}\rangle}^{=-1} - \frac{\overbrace{\langle Z, \dot{c}\rangle}^{=0}}{t_0+\delta} \leq 0.$$

The existence of such $V$ and $X$ now allows the construction of the claimed timelike curve. Let $\gamma\colon (-\varepsilon,\varepsilon) \to P$ be a smooth curve with $\gamma(0) = p$ and




$$\dot{\gamma}(0) = J(0) = V(0), \qquad \frac{\nabla^P\dot{\gamma}}{\mathrm{d}t}(0) = 0.$$

For instance, choose $\gamma(s) := \exp_p^P\big(sJ(0)\big)$. This satisfies

$$\frac{\nabla\dot{\gamma}}{\mathrm{d}t}(0) = \underbrace{\frac{\nabla^P\dot{\gamma}}{\mathrm{d}t}(0)}_{=0} + \mathit{II}\big(J(0),J(0)\big) = X(0).$$

Now choose a variation $c_s$ of $c$ with

$$c_s(0) = \gamma(s), \qquad c_s(t_0+\delta) = c(t_0+\delta), \qquad \left.\frac{\partial \dot{c}_s}{\partial s}\right|_{s=0} = V, \qquad \left.\frac{\nabla}{\partial s}\frac{\partial \dot{c}_s}{\partial s}\right|_{s=0} = X.$$

Since $c$ is lightlike, we have $\langle \dot c_s, \dot c_s\rangle\big|_{s=0} = 0$, and moreover

$$\frac{1}{2}\frac{\partial}{\partial s}\langle \dot c_s, \dot c_s\rangle\big|_{s=0} = \left\langle \frac{\nabla \dot c_s}{\partial s}, \dot c_s\right\rangle = \left\langle \frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}\bigg|_{s=0}, \dot c\right\rangle$$
$$= \left\langle \frac{\nabla V}{\mathrm{d}t}, \dot c\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\langle V, \dot c\rangle}_{=0} - \left\langle V, \underbrace{\frac{\nabla \dot c}{\mathrm{d}t}}_{=0}\right\rangle = 0.$$

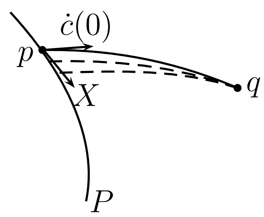
For the second derivative, we obtain

$$\frac{1}{2}\frac{\partial^2}{\partial s^2}\langle \dot c_s, \dot c_s\rangle\bigg|_{s=0} = \frac{\partial}{\partial s}\left\langle \frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}, \dot c_s\right\rangle\bigg|_{s=0} = \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \left\langle \frac{\nabla}{\partial s}\frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}\bigg|_{s=0}, \dot c\right\rangle$$
$$= \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \underbrace{\left\langle \frac{\nabla}{\partial t}\frac{\nabla}{\partial s}\frac{\partial c_s}{\partial s}\bigg|_{s=0}, \dot c\right\rangle}_{=\left\langle \frac{\nabla X}{\mathrm{d}t}, \dot c\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle X, \dot c\rangle} + \big\langle R(V,\dot c)V, \dot c\big\rangle$$
$$= \frac{\mathrm{d}}{\mathrm{d}t}\left(\left\langle V, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \langle X, \dot c\rangle\right) - \left\langle V, \frac{\nabla^2 V}{\mathrm{d}t^2}\right\rangle - \big\langle R(V,\dot c)\dot c, V\big\rangle$$
$$\overset{\text{c)}}{\leq} -\left\langle \frac{\nabla^2 V}{\mathrm{d}t^2} + R(V,\dot c)\dot c, V\right\rangle \overset{\text{b)}}{<} 0.$$

It follows $\langle \dot c_s, \dot c_s\rangle < 0$, i.e. $c_s$ is timelike on $[0, t_0+\delta]$ for $s \neq 0$ sufficiently small. □

**Lemma 2.43.** *Let $P \subset M$ be a spacelike submanifold and $c\colon [0,b] \to M$ a lightlike geodesic with $c(0) =: p \in P$, but $\dot c(0) \notin N_pP$.*
*Then in every neighborhood of $c$, we find a timelike from $P$ to $q := c(b)$.*

*Proof.*
By assumption, there is some $X_0 \in T_pP$ such that $\langle X_0, \dot c(0)\rangle \neq 0$, and without loss of generality, we assume $\langle X_0, \dot c(0)\rangle > 0$. Let $X$ be the corresponding parallel vector field along $c$ determined by $X(0) = X_0$. Then the vector field $V(t) := \left(1 - \frac{t}{b}\right)X(t)$ fulfills $V(0) = X_0$ and $V(b) = 0$. Furthermore, choose a variation $c_s$ of $c$ with $\frac{\partial c_s}{\partial s}\big|_{s=0} = V$, $c_s(0) \in P$ and $c_s(b) = q$ for all $s$.




Since $c$ is lightlike, we have $\big\langle \dot c_s, \dot c_s\big\rangle\big|_{s=0} = 0$ and moreover

$$\frac{\partial}{\partial s}\langle \dot c_s(t), \dot c_s(t)\rangle = 2\left\langle \frac{\nabla V}{\mathrm{d}t}(t), \dot c(t)\right\rangle = -\frac{2}{b}\big\langle X(t), \dot c(t)\big\rangle = -\frac{2}{b}\langle X_0, \dot c(0)\rangle < 0.$$

Thus, for $s > 0$ sufficiently small, we have $\big\langle \dot c_s, \dot c_s\big\rangle < 0$, i.e. $c_s$ is timelike. □

Therefore, we just proved

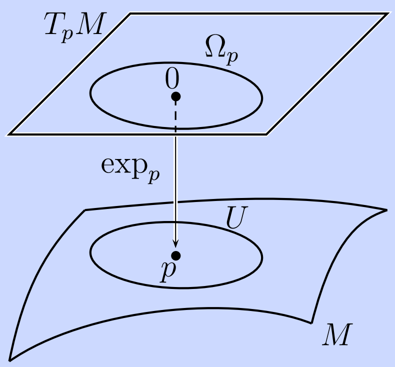
**Theorem 2.44.** *Let $M$ be a Lorentzian manifold, $P \subset M$ a spacelike submanifold and $c\colon [0,b] \to M$ a causal curve with $c(0) =: p \in P$.*
*Then in every neighborhood of $c$, we find a timelike curve from $P$ to $q := c(b)$, unless $c$ is, up to parametrization, a lightlike geodesic with $\dot{c}(0) \in N_pP$ such that $P$ has no focal values $t_0 < b$ along $c$.*

**Remark 2.45.** Compare this to the corresponding statement of Riemannian geometry:

*Let $M$ be a Riemannian manifold, $P \subset M$ a submanifold and $c\colon [0,b] \to M$ a curve with $c(0) =: p \in P$.*
*Then in every neighborhood of $c$, we find a shorter curve from $P$ to $q := c(b)$, unless $c$ is, up to parametrization, a geodesic with $\dot{c}(0) \in N_pP$ such that $P$ has no focal values $t_0 < b$ along $c$.*

## 2.3 Convex sets

**Definition 2.46.** Let $M$ be a semi-Riemannian manifold. An open subset $U \subset M$ is called **convex**, if for each $p \in U$, there is an open subset $\Omega_p \subset T_pM$, which is starshaped with respect to $0$, such that $\exp_p\colon \Omega_p \to U$ is a diffeomorphism.

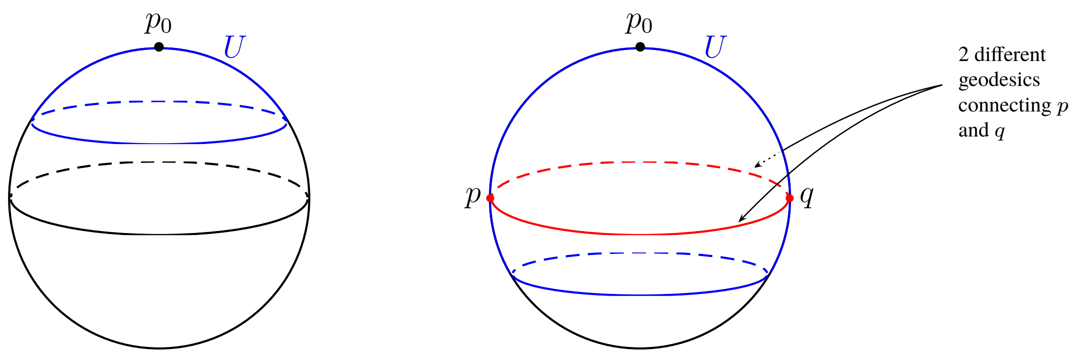



In particular, for every $q \neq p$ in $U$, there is, up to reparametrization, exactly one geodesic from $p$ to $q$ that runs entirely in $U$.

**Example 2.47.**

1. Let $M = \mathbb{R}^n$ Euclidean or Minkowski space. Then $U$ is convex if and only if for all $p, q \in U$ the segment of the straight line connecting both is contained in $U$.

2. Let $M = S^n$ the standard sphere and $U := B_r(p_0)$ the ball of radius $r$ centered at $p_0$. For $r \leq \frac{\pi}{2}$, $U$ is convex, but not for $r > \frac{\pi}{2}$.




Let $M$ be a semi-Riemannian manifold and $D \subset TM$ the maximal domain of $\exp$, which is open. We set

$$E\colon \quad D \longrightarrow M \times M, \qquad X \longmapsto \big(\pi(X), \exp(X)\big),$$

where $\pi\colon TM \to M$ denotes the footpoint map.
For fixed $p \in M$, we write $\mathscr{D}_p \subset T_pM$ for the maximal domain of $\exp_p$.

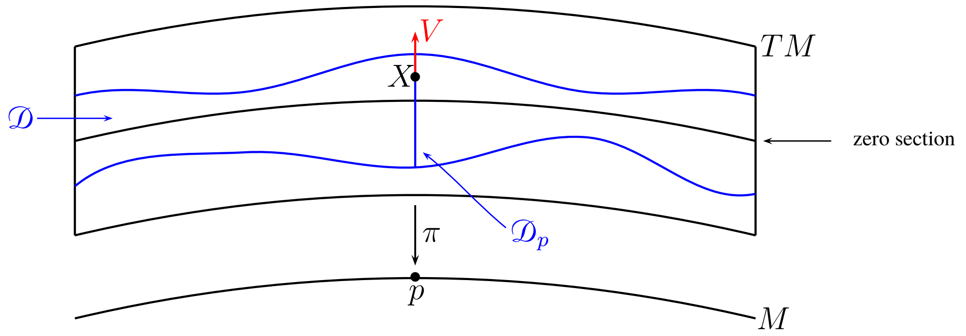
**Lemma 2.48.** *Let $M$ be a semi-Riemannian manifold, $p \in M$ and $X \in T_pM$. Furthermore, let $\exp_p\colon \mathscr{D}_p \to M$ be non-singular at $X$, i.e. $\mathrm{d}\exp_p\big|_X$ is invertible. Then also $E$ is non-singular at $X$.*

*Proof.* We just have to show injectivity of $\mathrm{d}E\big|_X$, so let $V \in \ker \mathrm{d}E\big|_X \subset T_XTM$. For $\pi_i\colon M \times M \to M$, $i = 1, 2$, the projection on the $i$. factor, we obtain

$$\mathrm{d}\pi\big|_X(V) = \mathrm{d}(\pi_1 \circ E)\big|_X(V) = \mathrm{d}\pi_1\big|_{E(X)}\big(\underbrace{\mathrm{d}E\big|_X(V)}_{=0}\big) = 0,$$

so $V$ has no component in "horizontal" direction, that is $V \in T_XT_pM \cong T_pM$. Therefore, we can apply $\mathrm{d}\exp_p\big|_X$ and employ its injectivity:

$$\mathrm{d}\exp_p\big|_X(V) = \mathrm{d}(\pi_2 \circ E)\big|_X(V) = \mathrm{d}\pi_2\big|_{E(X)}\big(\mathrm{d}E\big|_X(V)\big) = 0 \qquad \Longrightarrow \qquad V = 0.$$




$\square$

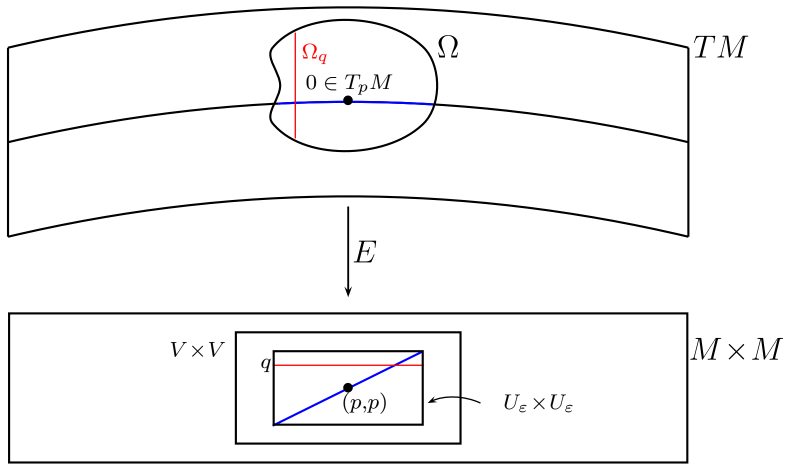
**Corollary 2.49.** *If* $\exp_p$ *is non-singular at* $X \in T_pM$*, then there is a neighborhood of* $X$ *in* $TM$ *that is mapped diffeomorphically to some neighborhood of* $\big(p, \exp_p(X)\big)$ *in* $M \times M$ *under* $E$*. In particular, there is a neighborhood of* $0 \in T_pM$ *in* $TM$ *that is mapped diffeomorphically to some neighborhood of* $(p,p) \in M \times M$.

**Proposition 2.50.** *Let* $M$ *be a semi-Riemannian manifold and* $p \in M$*. Then every neighborhood of* $p$ *contains a convex neighborhood of* $p$.

*Proof.* a) Let $V$ be some neighborhood of $p$, on which, without loss of generality, Riemannian normal coordinates are well-defined, i.e. there is a subset of $T_pM$, which is mapped diffeomorphically to $V$ under $\exp_p$. Furthermore, let $(e_1, \dots, e_n)$ be an orthonormal basis of $T_pM$. The normal coordinates $(x^1, \dots, x^n)$ of any $q \in V$ are characterized by $\exp_p\big(\sum_{i=1}^n x^i e_i\big) = q$. For $\varepsilon > 0$, set

$$U_\varepsilon := \left\{ \exp_p\left(\sum_{i=1}^n x^i e_i\right) \,\middle|\, \sum_{i=1}^n \left(x^i\right)^2 < \varepsilon \right\},$$

so for $\varepsilon$ suitably small, we have $U_\varepsilon \subset V$. Thus, due to Corollary 2.49, $E = (\pi, \exp)$ maps some neighborhood $\Omega \subset TM$ of $0 \in T_pM$ diffeomorphically to $U_\varepsilon \times U_\varepsilon$, that is $\Omega = E^{-1}\big(U_\varepsilon \times U_\varepsilon\big)$. Hence, for each $q \in U_\varepsilon$, $\exp_q$ maps $\Omega_q := \Omega \cap T_qM$ diffeomorphically to $\{q\} \times U_\varepsilon \cong U_\varepsilon$, i.e. it remains to show that $\Omega_q$ is starshaped with respect to $0 \in T_qM$.



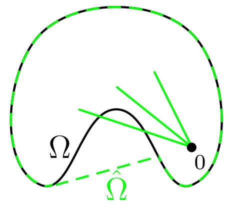

b) For any $q \in U_\varepsilon$, consider the symmetric matrix $B$ determined by the components

$$b_{ij}(q) := \delta_{ij} - \sum_{k=1}^n \Gamma_{ij}^k\big(x(q)\big)x^k.$$

Then $B(p)$ is the identity matrix and hence positive definite. For $\varepsilon$ small enough, we can assume this also for $B(q)$ for all $q \in U_\varepsilon$.

c) Let $\widehat{\Omega}$ denote the "starshaped hull" of $\Omega$, which means

$$\widehat{\Omega} := \{tv \mid v \in \Omega,\ t \in [0,1]\}.$$




Again, for $\varepsilon$ chosen small enough, we can assume

$$E(\widehat{\Omega}) \subset V \times V.$$

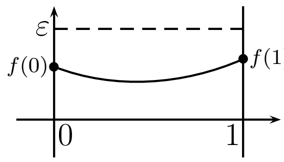
Let $v \in \Omega_q$ and thus $tv \in \widehat{\Omega} \cap T_qM$ for all $t \in [0,1]$. Therefore, Riemannian normal coordinates of $\exp_q(tv)$ exist for all $t$, for which we write $x^1(t), \ldots, x^n(t)$. Consider

$$f(t) := \sum_{k=1}^{n} \left(x^k(t)\right)^2 \qquad \Longrightarrow \qquad \dot{f} = 2\sum_{k=1}^{n} x^k \dot{x}^k, \quad \ddot{f} = 2\sum_{k=1}^{n} \left(\left(\dot{x}^k\right)^2 + x^k \ddot{x}^k\right).$$

The geodesic equation and b) provide

$$\ddot{f} = 2\sum_{k=1}^{n} \left(\left(\dot{x}^k\right)^2 - x^k \sum_{i,j=1}^{n} \Gamma_{ij}^k \dot{x}^i \dot{x}^j\right) = 2\sum_{i,j=1}^{n} b_{ij} \dot{x}^i \dot{x}^j > 0.$$

It follows that $f$ is concave, so since $\exp_q(0) = q \in U_\varepsilon$ and $\exp_q(v) \in \exp_q(\Omega_q) = U_\varepsilon$, we have $f(0), f(1) < \varepsilon$ and thus $f(t) < \varepsilon$ for all $t \in [0,1]$. Therefore $\exp_q(tv) \in U_\varepsilon$ and thus $tv \in \Omega_q$ for all $t \in [0,1]$, i.e. $\Omega_q$ is starshaped with respect to $0$.



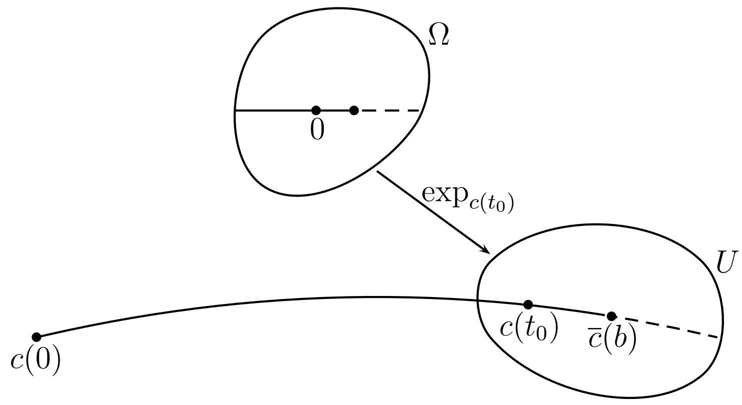

□

**Lemma 2.51.** *Let $M$ be a semi-Riemannian manifold and $c\colon [0,b) \to M$ a geodesic. Then the following statements are equivalent:*

*(i) $c$ can be continuously extended to $b$.*

*(ii) $c$ can be extended as a geodesic to $[0, b+\varepsilon)$ for some $\varepsilon > 0$.*

*Proof.* We merely prove (i)⇒ (ii) since the other direction is trivial. Let $\bar{c}$ be the continuous extension of $c$ to $[0,b]$ and let $U$ be a convex neighborhood of $\bar{c}(b)$. Choose some $t_0 \in [0,b)$ such that $c(t_0) \in U$. Due to convexity, there is an open subset $\Omega \subset T_{c(t_0)}M$ such that $\exp_{c(t_0)}\colon \Omega \to U$ is a diffeomorphism.




The curve $\exp_{c(t_0)}^{-1} \circ c$ is a line segment in $\Omega$, which is continuously extendible to $\exp_{c(t_0)}^{-1}\big(\bar{c}(b)\big)$. Now we extend this line segment to $\Omega$, so its image under $\exp_{c(t_0)}$ provides the desired extension as a geodesic. $\square$

**Remark 2.52.** Let $U \subset M$ be convex and define the map

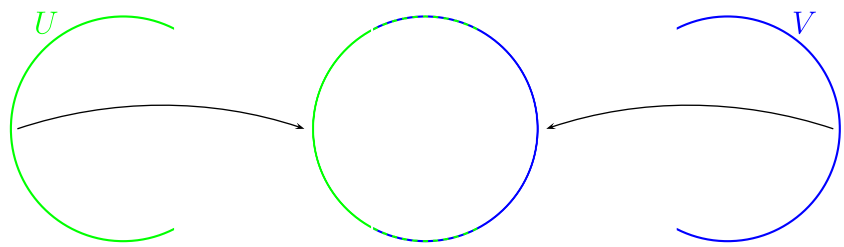
$$\Delta\colon \quad U \times U \longrightarrow TM, \qquad (p,q) \longmapsto \dot{c}_{pq}(0), \tag{2.3}$$

where $c_{p,q}\colon [0,1] \to U$ denotes the geodesic with $c(0) = p$ and $c(1) = q$. Then we obtain

$$E\big(\Delta(p,q)\big) = E\big(\dot{c}_{pq}(0)\big) = \big(p, \exp_{c(0)}\big(\dot{c}_{pq}(0)\big)\big) = (p,q),$$

so $\Delta$ is the inverse of the local diffeomorphism $E$ and thus in particular smooth.

**Warning!** Convexity of $U, V \subset M$ does not imply that $U \cap V$ is convex. For instance, on $M := S^1$, take the subsets




Here $U$ and $V$ are convex but $U \cap V$ is not even connected.

On the other hand, this implication holds if additionally their union is contained in another convex subset:

**Lemma 2.53.** *Let $U, V, W \subset M$ be convex and $U \cup V \subset W$. Then $U \cap V$ is convex.*

*Proof.* Let $p \in U \cap V$ and $\Omega^W := \exp_p^{-1}(W)$, i.e. $\Omega^W \subset T_pM$ is open and starshaped with respect to $0$. Analogously, we define $\Omega^V$ and $\Omega^U$.
For $v \in \Omega^U$, let $q := \exp_p(v) \in U$, so the unique geodesic connecting $p$ and $q$ in $W$ is $t \mapsto \exp_p(tv)$. Due to convexity, this coincides with the corresponding unique geodesic in $U$, i.e. $tv \in \Omega^U$ for all $t \in [0,1]$. For the same reasons, we also have $tv \in \Omega^V$ for all $t \in [0,1]$, so $\Omega^U$ and $\Omega^V$ are starshaped with respect to $0$ and hence so is $\Omega^U \cap \Omega^V$, which implies convexity of $U \cap V$. $\square$

**Definition 2.54.** An open cover $\mathcal{U} := \{U_\alpha\}_\alpha$ of some semi-Riemannian manifold $M$ is called **convex cover of** $M$ if all countable intersections of its elements are convex sets, i.e. $U_{\alpha_1} \cap \ldots, U_{\alpha_k}$ is convex for all $\alpha_j$ and $k \in \mathbb{N}$.

**Proposition 2.55.** *Let $M$ be a semi-Riemannian manifold and $\mathcal{U}$ an open cover of $M$. Then there is a convex refinement, i.e. there is a convex cover $\mathcal{K} := \{K_\beta\}_\beta$, and for all $\beta$, there is some $\alpha$ such that $K_\beta \subset U_\alpha$.*

*Proof.* Define

$$\mathcal{U}_1 := \{U \subset M \mid U \text{ is convex and } U \subset U_\alpha \text{ for some } \alpha\},$$

which is a cover of $M$ due to Proposition 2.50. Let $\mathcal{U}_2$ be a refinement of $\mathcal{U}_1$ such that for all $U, V \in \mathcal{U}_2$ with $U \cap V \neq \emptyset$, we find some $K \in \mathcal{U}_1$ containing $U \cup V$ (recall that the topology of $M$ is countably generated). Then

$$\mathcal{K} := \{U \subset M \mid U \text{ is convex and contained in some element of } \mathcal{U}_2\}$$

yields another cover of $M$, from which we show that it is convex.

For any $k \in \mathbb{N}$, let $U_1, \ldots, U_k \in \mathcal{K}$ with $U_1 \cap \ldots \cap U_k \neq \emptyset$. Then

$$\begin{aligned}
U_1 \cup U_2 \subset \underbrace{W_1}_{\in \mathcal{U}_2} \cup \underbrace{W_2}_{\in \mathcal{U}_2} \subset \underbrace{K}_{\in \mathcal{U}_1} \quad &\Longrightarrow \quad U_1 \cap U_2 \text{ is convex,} \\
\big(U_1 \cap U_2\big) \cup U_3 \subset U_1 \cup U_3 \subset \underbrace{\widetilde{K}}_{\in \mathcal{U}_1} \quad &\Longrightarrow \quad U_1 \cap U_2 \cap U_3 \text{ is convex,} \\
&\qquad \vdots \\
&\qquad U_1 \cap \ldots \cap U_k \text{ is convex.}
\end{aligned}$$

□

**Lemma 2.56.** *Let $M$ be a convex and time-oriented Lorentzian manifold and $p, q \in M$. Then we have*

*(i) For $p \neq q$, we have*

$$q \in J_+(p) \ \big(I_+(p)\big) \quad \Longleftrightarrow \quad \Delta(p,q) \in T_pM \text{ is causal (timelike) and future directed.}$$

*(ii) $J_+(p) = \overline{I_+(p)}$*

*(iii) The relation "$\leq$" is closed, i.e. $p_n \to p, q_n \to q$ and $p_n \leq q_n$ for all $n$ imply $p \leq q$.*

*(iv) Every causal curve $c\colon [0,b) \to M$, which runs in some compact subset of $M$, is continuously extendible to $b$.*

*Proof.* For $\Omega := \exp_p^{-1}(M)$, Corollary 2.13 yields

$$J_+(p) = \exp_p\big(J_+(0) \cap \Omega\big), \qquad I_+(p) = \exp_p\big(I_+(0) \cap \Omega\big),$$

so the claims (i) - (iii) follow from the respective statements on Minkowski space.
Let $\{t_i\}_{i \in \mathbb{N}}$ be a monotoneously increasing sequence with $\lim\limits_{i\to\infty} t_i = b$. Since it is contained in a compact set, $\big(c(t_i)\big)_{i \in \mathbb{N}}$ has at least one accumulation point, and we show that there is exactly one, which then extends $c$. Let $p, q$ be accumulation points, i.e. there is a monotoneously increasing subsequence $\{s_i\}_{i\in\mathbb{N}} \subset \{t_i\}_{i\in\mathbb{N}}$ such that

$$c\big(s_{2i}\big) \longrightarrow p, \qquad c\big(s_{2i+1}\big) \longrightarrow q$$

if $i \to \infty$. Without loss of generality, we assume $c$ to be future directed, so for all $i$ we have

$$c\big(s_{2i}\big) \leq c\big(t_{2i+1}\big) \leq c\big(s_{2i+2}\big)$$

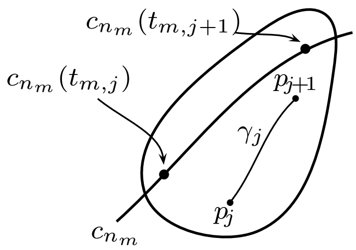
and hence $p \leq q \leq p$ due to (iii). Therefore, (i) implies that $\Delta(p,q)$ is future and past directed, that is $\Delta(p,q) = 0$, which provides $p = q$. □

## 2.4 Quasi-limits

**Definition 2.57.** Let $M$ be a time-oriented Lorentzian manifold, $\mathcal{K}$ a convex cover of $M$ and $(c_n)_{n\in\mathbb{N}}$ a sequence of causal and future directed curves.
A **limit sequence for** $(c_n)_n$ **relative to** $\mathcal{K}$ is a (finite or infinite) sequence of points $p_0 < p_1 < \ldots$ such that

1. For each $p_j$, there is a subsequence $(c_{n_m})_m$, and for each $m = m(j)$, there are parameter values $t_{m,0} < t_{m,1} < \ldots$ such that for all $j$, we have
   1a. $\lim\limits_{m\to\infty} c_{n_m}(t_{m,j}) = p_j$
   1b. $p_j, p_{j+1}$ and, for all $m \geq m(j)$, the segments $c_{n_m}\big([t_{m,j}, t_{m,j+1}]\big)$ are contained in one element of $\mathcal{K}$.
2. If $(p_j)_j$ is infinite, it is nonconvergent. If $\{p_j\}_j$ is finite, it has more than one point and no strictly longer sequence satisfies 1.

**Remark 2.58.** Let $K_j \in \mathcal{K}$ contain $p_j, p_{j+1}$. By assumption, we have $c_{n_m}(t_{m,j}) < c_{n_m}(t_{m,j+1})$, and since "$\leq$" is a closed relation in convex sets, we obtain $p_j \leq p_{j+1}$. On the other hand, $p_j < p_{j+1}$ in $M$ implies $p_j < p_{j+1}$ also in $K_j$, so the unique geodesic $\gamma_j$ connecting $p_j$ and $p_{j+1}$ is causal. The broken geodesic

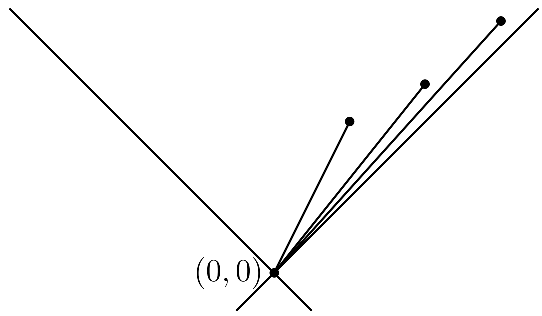



$$\gamma := \gamma_0 \cup \gamma_1 \cup \gamma_2 \cup \ldots$$

is called **quasi-limit of the sequence** $(c_n)_n$ **related to** $\mathcal{K}$.

**Example 2.59.** In all following examples, let $M := \mathbb{R}^2_{\text{Mink}}$.

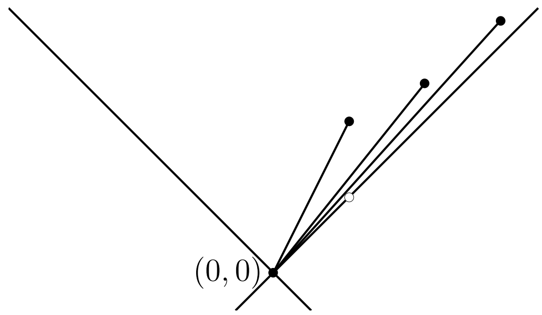
1.) Let $C_n$, $n \in \mathbb{N}$, denote the line segment connecting $(0,0)$ with $\big(n + \frac{1}{n}, n\big)$. Every limit sequence lies on the lightlike geodesic $\big\{(s,s) \bigm| s \geq 0\big\}$, so up to parametrization, $s \mapsto (s,s)$ is the unique quasi-limit of the sequence of curves $(c_n)_n$.



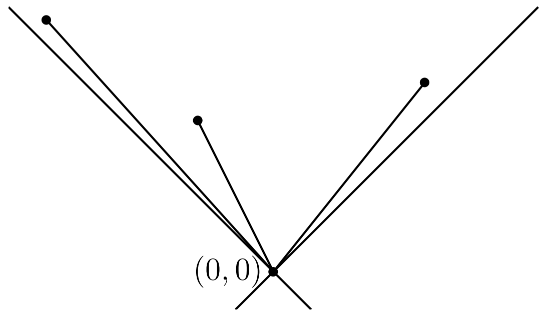
2.) Let $C_n$ be as in the first example. Every accumulation point of $\big(c_n(t_n)\big)_n$ lies on the lightlike geodesic $\big\{(s,s) \;\big|\; s \geq 0\big\}$, but for a limit sequence $p_1 < p_2 < \ldots$, we cannot have

$$p_j = (s_j, s_j), \quad s_j < 1$$
$$p_{j+1} = (s_{j+1}, s_{j+1}), \quad s_{j+1} > 1$$

since those points would not be contained in the same convex set. For example, $p_j := \big(1 - \frac{1}{j}, 1 - \frac{1}{j}\big)$ yields a limit sequence and $s \mapsto (s,s)$ on $[0,1]$ a quasi-limit.



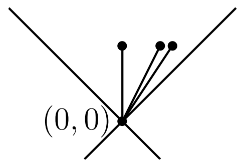

3.) For $C_n,\ n \in \mathbb{N}$, the line segment connecting $(0,0)$ and $\big(n + \frac{1}{n}, (-1)^n n\big)$, we have two quasi-limits $s \mapsto (s, \pm s),\ s \geq 0$.




4.) For $C_n,\ n \in \mathbb{N}$, the line segment connecting $(0,0)$ and $\big(1, 1 - \frac{1}{n}\big)$, the points

$$p_1 := (0,0) < p_2 := (1,1)$$

yield a limit sequence.




**Proposition 2.60.** *Let $M$ be a time-oriented Lorentzian manifold with convex cover $\mathcal{K}$ and $c_n \colon I_n \to M, n \in \mathbb{N}$, causal and future directed curves with $\lim\limits_{n\to\infty} c_n(0) = p \in M$, where $I_n = [0, b_n]$ if $b_n < \infty$, and $I_n = [0, b_n)$ if $b_n \leq \infty$. Then the following statements are equivalent:*

*(i) The sequence $(c_n)_n$ has a quasi-limit relative to $\mathcal{K}$ with $p_0 = p$.*

*(ii) There is a neighborhood $U$ of $p$ such that infinitely many $c_n$ are not entirely contained in $U$, i.e. $c_n \nrightarrow p$.*

*Proof.* (i) $\Longrightarrow$ (ii): Let $p_0 := p < p_1 < p_2 < \ldots$ denote the limit sequence and we choose disjoint neighborhoods $U, U_1$ of $p, p_1$. Due to $c_{n_m}(t_{m,1}) \to p_1$, we have $c_{n_m}(t_{m,1}) \in U_1$ for almost all $m$, and hence $c_{n_m}(t_{m,1}) \notin U$, so $U$ does the job.
(ii) $\Longrightarrow$ (i): a) Let $\mathcal{U}$ be a locally finite refinement of $\mathcal{K}$ such that for all $U \in \mathcal{U}$, the closure $\overline{U}$ is compact and contained in some $K \in \mathcal{K}$. We assume that $U_0 \in \mathcal{U}$ is a neighborhood of $p$

such that infinitely many $c_n$ are not entirely contained in $U$.
Let $\left(c_n^{(1)}\right)_n$ the subsequence of curves, which leave $U_0$, and set

$$t_{n,1} := \inf\left\{t > 0 \,\middle|\, c_n^{(1)}(t) \notin U_0\right\},$$

that is $c_n^{(1)}\left(t_{n,1}\right) \in \partial U_0$. Due to compactness, there is a convergent subsequence, and we define $p_1$ as the corresponding limit. Then $c_n^{(1)}(0) < c_n^{(1)}\left(t_{n,1}\right)$ implies

$$p = \lim_{n\to\infty} c_n^{(1)}(0) \leq \lim_{n\to\infty} c_n^{(1)}\left(t_{n,1}\right) = p_1$$

and $p_0 < p_1$ follows from $p_0 \in U_0$ and $p_1 \in \partial U_0$, that is $p_0 \neq p_1$.
We repeat this procedure as often as possible, where we respect the following convention: If several $U \in \mathcal{U}$ contain the point $p_j$, we define $U_j$ as the so far least chosen one. Then property 1 in Definition 2.57 holds by construction.
b) We proceed with property 2 and start with the case that the constructed sequence $p < p_1 < p_2 < \ldots$ is infinite. Assume that $p_j$ converges to some $q \in M$ and let $V \in \mathcal{U}$ be a neighborhood of $q$. Then almost all $p_j$ are contained in $V$ and hence, almost all corresponding neighborhoods $U_j$ hit $V$. Therefore, due to local finiteness, at least one $U_j$ must have been chosen infinitely many times in the construction. On the other hand, note that $V$ itself would also serve as such an $U_j$. Then the convention, we followed in the construction, demands that $V$ must have been chosen infinitely many times as well, i.e. $p_j \in \partial V$ for infinitely many $j$, which contradicts $p_j \to q$.
c) Now assume the construction to break down after finitely many steps, i.e. produces $k$ points $p = p_0 < p_1 < \ldots < p_k$. Hence, there is a subsequence of curves $\left(c_n^{(k+1)}\right)_n$, which entirely run in $U_k \in \mathcal{U}$, and compactness of $\overline{U}_k$ as well as Lemma 2.56 provide a continuous extension of $c_n^{(k+1)}$ to $b_n$, if it was not already in the first place. Furthermore, after maybe switching over to some subsequence, we have $c_n^{(k+1)}(b_n) \to q \in \overline{U}_k$.
c1) Assume $q = p_k$ and that the finite sequence is extendible by some $p_{k+1} > p_k$ such that $p = p_0 < \ldots p_k < p_{k+1}$ has property 1. Then on $U_k$ would have

$$\underbrace{c_n^{(k+1)}\left(t_{n,k+1}\right)}_{\to p_{k+1}} < \underbrace{c_n^{(k)}\left(b_n\right)}_{\to q} \qquad \Longrightarrow \qquad p_k < p_{k+1} \leq q = p_k,$$

which implies $p_k = p_{k+1}$, since $U_k$ is contained in some convex set. Therefore, $p_{k+1}$ does not yield an extension, so $p = p_0 < \ldots < p_k$ is the limit sequence.
c2) If $p_k \neq q$, i.e. $q > p_k$, $p = p_0 < p_1 < \ldots < p_k < p_{k+1} = q$ has property 1 by construction and 2 by c1), so it yields a limit sequence. □

**Remark 2.61.** Due to property 2 in Definition 2.57, any quasi-limit $\gamma$ is future-inextendible, i.e. if it is parametrised on $[a, b)$, there is no continuous extension to $b$ (otherwise the sequence $(p_j)_j$ would converge to $\gamma(b)$).

## 2.5 Cauchy hypersurfaces

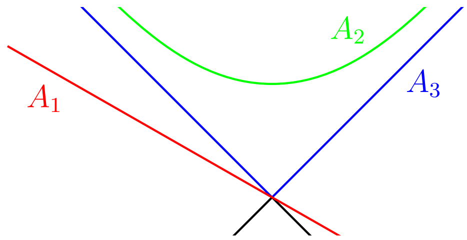
**Definition 2.62.** Let $M$ be a connected and time-oriented Lorentzian manifold.
A subset $A \subset M$ is called **achronal** if there no $p, q \in A$ such that $p \ll q$. In other words, every timelike curve hits $A$ at most once.
A subset $A \subset M$ is called **acausal** if there no $p, q \in A$ such that $p \leq q$. In other words, every causal curve hits $A$ at most once.

**Remark 2.63.**

1. Acausality implies achronality but not cive verca.
2. Subsets of achronal (acausal) subsets are achronal (acausal).
3. The closure of an achronal set $A$ is achronal: If there were $p, q \in \overline{A}$ with $p \ll q$ then choose sequences $(p_n)_n, (q_n)_n$ in $A$ converging to $p$ and $q$, then $p_n \ll q_n$ for $n$ large enough according to Proposition 2.7, which is a contradiction.

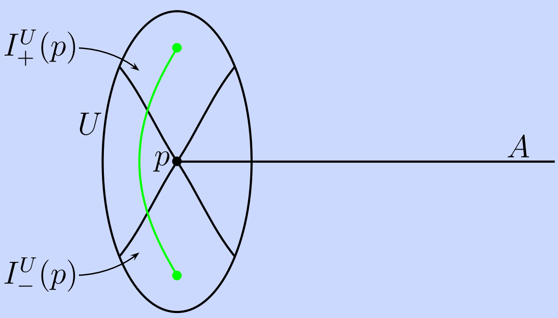
**Example 2.64.**

Let $M = \mathbb{R}^n_{\text{Mink}}$. The subset $A_1$ (spacelike hypersurface), $A_2$ (the $(n-1)$-dimensional hyperbolic space) and $A_3$ (future light cone) are achronal subsets. $A_1$ and $A_2$ are even acausal.




**Definition 2.65.** The **edge of an achronal subset** $A$ is defined as the subset

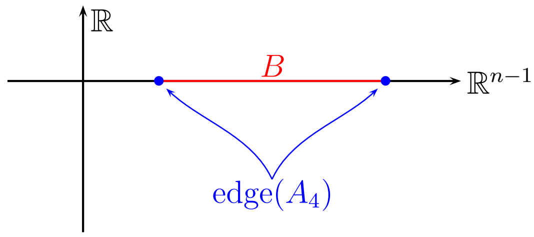
$$\mathrm{edge}(A) := \left\{ p \in \overline{A} \;\middle|\; \begin{array}{l}\text{for all open neighborhoods } U \text{ of } p \text{, there is a timelike} \\ \text{curve in } U \text{ from } I^U_-(p) \text{ to } I^U_+(p)\text{, which does not hit } A\end{array} \right\}.$$



**Example 2.66.** Let $M$ be $n$-dimensional Minkowski space.

1. For the subsets in Remark 2.63, we have $\operatorname{edge}(A_i) = \emptyset$ for $i = 1, 2, 3$.

2. Let $A_4 := \{0\} \times B$ for some $B \subset \mathbb{R}^{n-1}$, then $\operatorname{edge}(A_4) = \{0\} \times \partial B$:
   For $b \in \mathring{B}$ with some open neighborhood $U_0 \subset B$ (with respect to $\mathbb{R}^{n-1}$), let $p := (0, b)$ and $U \subset B$ the double cone over $U_0$, that is




$$U = I_+(U_0) \cup U_0 \cup I_-(U_0).$$

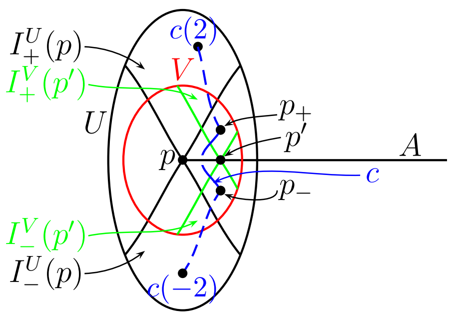
   Every timelike curve from $I_-^U(p)$ to $I_+^U(p)$ has to meet $\{0\} \times B$ and thus $p \notin \operatorname{edge}(A_4)$.
   On the other hand, $b \in \partial B$ implies $V \backslash B \neq \emptyset$ for any neighborhood $V$ of $b$ with respect to $\mathbb{R}^{n-1}$. Hence, for each open neighborhood $U$ of $p$ (with respect to $\mathbb{R}^n$), we find some $p' := (b', 0) \in \{0\} \times \big(V \backslash B \cap U\big)$ and a timelike curve connecting $I_-^U(p')$ and $I_+^U(p')$ running through $p'$. Since $\{0\} \times \mathbb{R}^{n-1}$ is achronal, it is hit by this curve at most once, which is in $p'$, so it does not meet $A_4$, i.e. $p \in \operatorname{edge}(A_4)$.

**Remark 2.67.** If $A$ is achronal, we have $\overline{A} \backslash A \subset \operatorname{edge}(A)$:
For $p \in \overline{A} \backslash A$ and any open neighborhood $U$ of $p$, there is a timelike curve connecting $I_-^U(p)$ and $I_+^U(p)$ and running through $p$. Since $\overline{A}$ is achronal, there is no other intersection point of $\overline{A}$ with this curve, in particular not with $A$. Therefore, $p \notin A$ implies that the curve does not meet $A$ at all, that is $p \in \operatorname{edge}(A)$.

**Lemma 2.68.** *For every achronal subset $A \subset M$, the subset $\operatorname{edge}(A)$ is closed.*

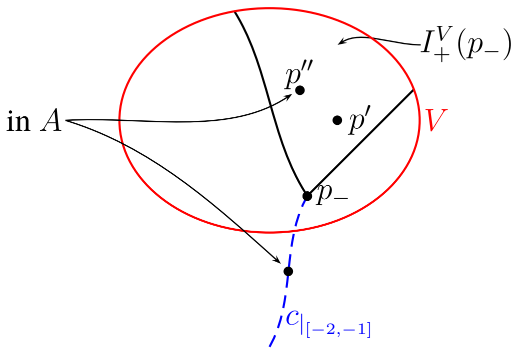
*Proof.* We show $\operatorname{edge}(A) = \overline{\operatorname{edge}(A)}$, so let $p \in \overline{\operatorname{edge}(A)}$ and $U$ a neighborhood of $p \in M$. Furthermore, let $V \subset U$ be an open neighborhood of $p$ contained in $I_+^U\big(I_-^U(p)\big) \cap I_-^U\big(I_+^U(p)\big)$. Since $p \in \overline{\operatorname{edge}(A)}$, there is some $p' \in V \cap \operatorname{edge}(A)$. It follows that there is a timelike curve $c\colon [-1, 1] \to V$ with

$$p_\pm := c(\pm 1) \in I_\pm^V(p'),$$




which does not meet $A$. Note that $p_\pm \in V \subset I_+^U\big(I_-^U(p)\big) \cap I_-^U\big(I_+^U(p)\big)$ implies that we can extend $c$ to some timelike and future directed curve $c\colon [-2, 1] \to U$ such that $c(-2) \in I_-^U(p)$. For similar reasons, we can moreover extend it a to timelike and future directed curve $c\colon [-2, 2] \to U$ such that $c(2) \in I_+^U(p)$.

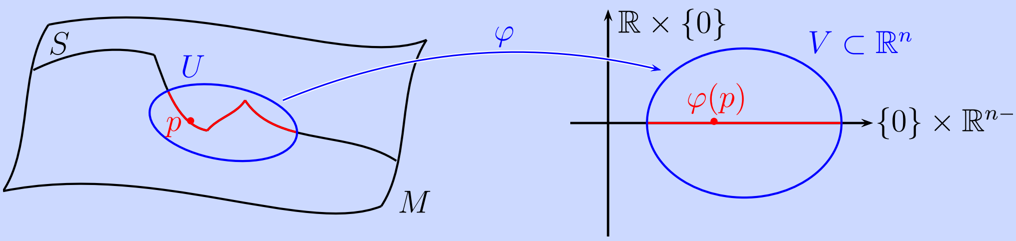
We show that this yields a timelike curve in $U$ connecting $I_+^U(p), I_-^U(p)$, which does not hit $A$, either. Assume it does. Since $I_+^V(p_-)$ is a neighborhood of $p'$ and $p' \in \mathrm{edge}(A) \cap \overline{A}$, there is some $p'' \in A \cap I_+^V(p_-)$. Then any concatination of $c\big|_{[-2,-1]}$ with some timelike and future directed curve from $p_-$ to $p''$ would meet $A$ twice, which yields a contradiction since $A$ was assumed to be achronal.




□

**Definition 2.69.** A subset $S$ of an $n$-dimensional differentiable manifold $M$ is called **topological hypersurface**, if for each $p \in S$, there is an open neighborhood $U$ of $p$ in $M$ and a homeomorphism $\varphi\colon U \to V$, with some open $V \subset \mathbb{R}^n$ open, such that

$$\varphi(U \cap S) = V \cap \big(\{0\} \times \mathbb{R}^{n-1}\big).$$




**Example 2.70.** The subset $S := C_+(0) \subset \mathbb{R}^n$ is a topological hypersurface with, for instance,

$$\varphi\colon\ \mathbb{R}^n \to \mathbb{R}^n, \qquad \big(x^0, \underbrace{x^1, \dots, x^{n-1}}_{=:\hat{x}}\big) \longmapsto \big(x^0 - \|\hat{x}\|, \hat{x}\big).$$

**Proposition 2.71.** *Let $A \subset M$ be achronal. Then the following statements are equivalent:*

*(i)* $A \cap \mathrm{edge}(A) = \emptyset$.

*(ii) $A$ is a topological hypersurface.*

*Proof.* (ii) $\Rightarrow$ (i): Let $A$ be a topological hypersurface with $U, V, \varphi$ as in Definition 2.69 for some $p \in A$. Without loss of generality, let $U$ be connected, so $U \backslash A \approx V \backslash \{x^0 = 0\}$ implies that $U \backslash A$ has two connected components. The subsets $I_\pm^U(p)$ are open and connected and, due to achronality, have empty intersection with $A$. Moreover, every timelike curve through $p$ meets both, $I_-^U(p)$ and $I_+^U(p)$, and, again by applying the homeomorphism, also both connected components. Hence, $I_\pm^U(p)$ are contained in different connected components and since $U$ is separated by $A$, every continuous timelike curve connecting $I_-^U(p)$ and $I_+^U(p)$ has to meet $A \cap U$, that is $p \notin \mathrm{edge}(A)$.

(i) ⇒ (ii): For $p \in A$, let $\widetilde{U}$ be a neighborhood such that every timelike curve connecting $I_-^{\widetilde{U}}(p)$ and $I_+^{\widetilde{U}}(p)$ meets $A$. Without loss of generality, we assume $\widetilde{U}$ to be a coordinate neighborhood, i.e. there is a diffeomorphism $\xi\colon \widetilde{U} \to \xi(\widetilde{U}) \subset \mathbb{R}^n$, and $\frac{\partial}{\partial x^0}$ to be timelike and future directed. Then $\widetilde{U}$ contains a smaller neighborhood $U$ of $p$ such that

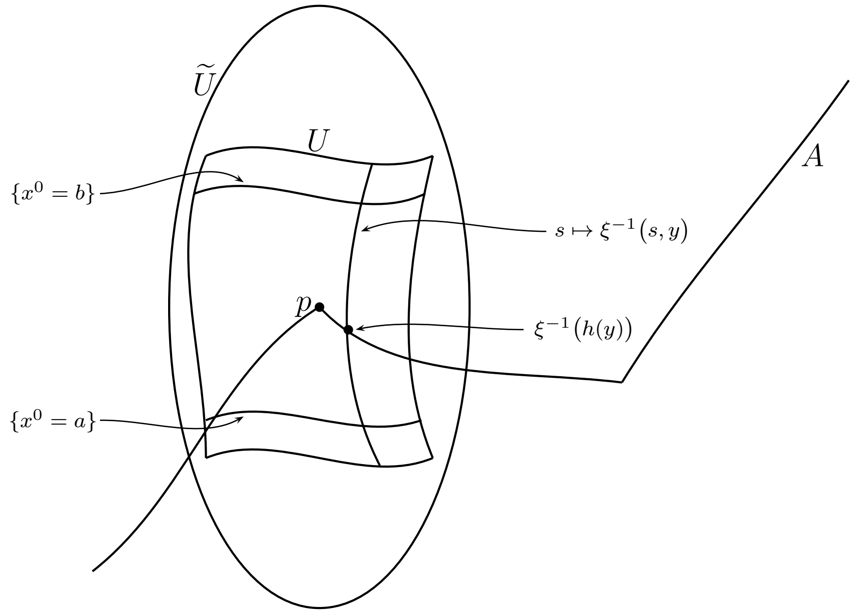
1. $\xi(U) = (a-\delta, b+\delta) \times N =: V$ for some $a, b \in \mathbb{R},\ \delta > 0$ and $N \subset \mathbb{R}^{n-1}$ open,
2. $\{x \in \widetilde{U} \mid x^0 = a\} \subset I_-^{\widetilde{U}}(p)$ and $\{x \in \widetilde{U} \mid x^0 = b\} \subset I_+^{\widetilde{U}}(p)$.

For fixed $y \in N \subset \mathbb{R}^{n-1}$, the curve

$$[a,b] \longrightarrow U, \qquad s \longmapsto \xi^{-1}(s,y),$$

is timelike and meets $A$ by assumption on $\widetilde{U}$. Since $A$ is achronal, this determines a map $h\colon N \to (a,b)$ by demanding $\xi^{-1}\big(h(y),y\big) \in A$ and we obtain

$$U \cap A = \xi^{-1}\Big(\Big\{\big(h(y),y\big) \;\Big|\; y \in N)\Big\}\Big).$$




Assuming continuity of $h$, the map

$$\varphi\colon\ U \longrightarrow V, \qquad p \longmapsto \Big(\xi(p)^0 - h\big(\widehat{\xi(p)}\big), \widehat{\xi(p)}\Big)$$

yields a homeomorphism with $\varphi(p) = (0,y)$ if $p \in U \cap A$ since then $p = \xi^{-1}\big(h(y),y\big)$ for some $y \in N$. If $h$ was not continuous, there would exist a sequence $(y_n)_{n\in\mathbb{N}}$ and $y \in N$ such that $y_n \to y$ but $h(y_n) \nrightarrow h(y)$. Let $q := \xi^{-1}\big(h(y),y\big)$. Since $h(N)$ is contained in the compact set $[a,b]$, there is a subsequence $(y_m)_m$ such that $h(y_m)$ converges to some $r \neq h(y)$, and hence $\xi^{-1}(r,y) \in I_-^U(q) \cup I_+^U(q) =: I^U(q)$. Therefore, $I^U(q)$ is an open neighborhood of $\xi^{-1}(r,y)$, i.e. for $m$ large enough, we have $\xi^{-1}\big(h(y_m),y_m\big) \in I^U(q)$, which contradicts achronality of $A$. □

**Corollary 2.72.** *Let $A \subset M$ be achronal. Then the following statements are equivalent:*

*(i)* $\mathrm{edge}(A) = \emptyset$.

*(ii)* *$A$ is a closed topological hypersurface.*

*Proof.* (i) $\Rightarrow$ (ii): Due to Proposition 2.71, $A$ is a topological hypersurface. It is moreover closed since $\overline{A}\backslash A \subset \mathrm{edge}(A)$, that is $A = \overline{A}$.
(ii) $\Rightarrow$ (i): Proposition 2.71 ensures $A \cap \mathrm{edge} A = \emptyset$. On the other hand, by definition, $\mathrm{edge}(A)$ is a subset of $\overline{A}$, so $A = \overline{A}$ provides the claim. $\square$

**Definition 2.73.** A subset $B$ of a time-oriented Lorentzian manifold $M$ is called **future set** or **past set** if $I_+(B) \subset B$ or $I_-(B) \subset B$, respectively.

**Example 2.74.** For $M = \mathbb{R}^n_{\text{Mink}}$, an example for a future set would be

$$\left\{ \left(x^0, \widehat{x}\right) \in \mathbb{R} \times \mathbb{R}^{n-1} \;\middle|\; x^0 - \|\widehat{x}\| \geq 0 \right\}.$$

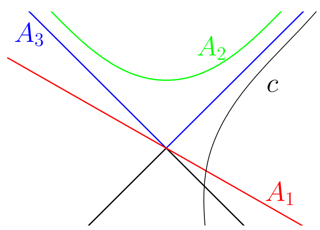
**Remark 2.75.** If $B$ is a future set, then $M\backslash B$ is a past set.

**Corollary 2.76.** *Let $M$ be a connected and time-oriented Lorentzian manifold and $B \subset M$ a future set with $B \notin \{\emptyset, M\}$. Then $\partial B$ is an achronal and closed topological hypersurface.*

*Proof.* Regarding Corollary 2.72, we just have to show that $\partial B$ is achronal and $\mathrm{edge}(\partial B) = \emptyset$. Let $p \in \partial B$ and $q \in I_+(p)$, so $I_-(q)$ is an open neighborhood of $p$, which moreover meets $B$ and thus $q \in I_+(B)$. Since $I_+(B)$ is open and $B$ a future set, we have $I_+(B) \subset \mathring{B}$ and in particular $I_+(p) \subset \mathring{B}$. Analogously, we obtain $I_-(p) \subset \mathrm{int}\big(M\backslash B\big)$.
It follows that $I_+(\partial B) \cap \partial B$, so $\partial B$ is achronal since any timelike curve running through $\partial B$ fails to meet it again. Moreover, every timelike curve connecting $I_-(p)$ and $I_+(p)$ has to hit $\partial B$, which shows $\mathrm{edge}(\partial B) = \emptyset$. $\square$

**Definition 2.77.** A subset $S \subset M$ of a time-oriented Lorentzian manifold is called **Cauchy hypersurface of** $M$ if any inextendible timelike curve hits it exactly once.

**Example 2.78.**
Let $M = \mathbb{R}^n_{\text{Mink}}$ and $A_1, A_2, A_3$ the subsets of Example 2.64. Although all of them are achronal, only $A_1$ is also a Cauchy hypersurface since, for instance, the inextendible timelike curve $c$ does not hit $A_2$ or $A_3$.

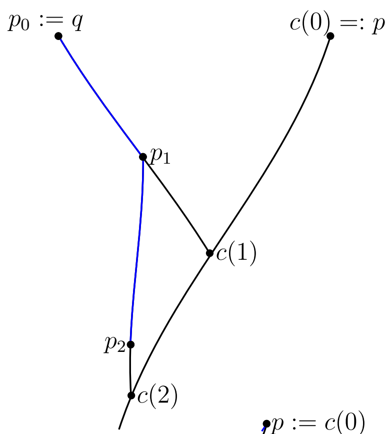


**Lemma 2.79.** *Let $M$ be a connected Lorentzian manifold, $A \subset M$ closed and $c\colon [0,b) \to M\backslash A$ a past directed causal curve with $c(0) =: p$, which is past-inextendible in $M$.*

*(i) For all $q \in I_+^{M\backslash A}(p)$, there is a timelike past directed curve $\widetilde{c}\colon [0,b) \to M\backslash A$ with $\widetilde{c}(0) = q$, which is past-inextendible in $M$.*

*(ii) There is a timelike past directed curve $\widetilde{c}\colon [0,b) \to M\backslash A$ with $\widetilde{c}(0) = p$, which is past-inextendible in $M$, unless $c$ is a lightlike pregeodesic without conjugated points.*

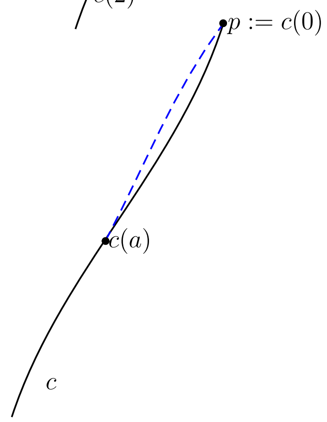
*Proof.* Without loss of generality, let $b = \infty$ and $\big(c(n)\big)_n$ non-convergent. Choose some metric $d$ on $M$, which induces the given topology on $M$.




(i): Let $p_0 := q \gg p$, where "$\gg$" is with respect to $M\backslash A$, so $c(1) \ll p$ since $c(1) \leq c(0) \ll p_0$. On the corresponding timelike connecting curve, choose some $p_1$ such that $0 < d\big(p_1, c(1)\big) < 1$.
Inductively, we find $p_k$ with $c(k) \ll p_k \ll p_{k-1}$ and $d\big(p_k, c(k)\big) < \frac{1}{k}$, which provides a timelike past directed curve $\widetilde{c}$ through all $p_k$ and starting at $p_0 = q$. Furthermore, it is past-inextendible: If $\widetilde{c}$ was extendible by some $p_\infty$, then $p_k \to p_\infty$ and thus

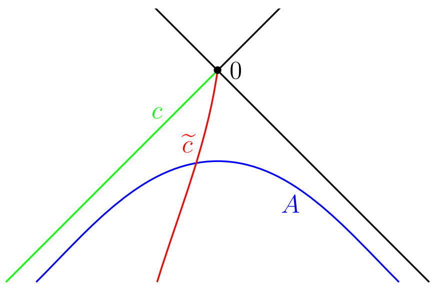
$$d\big(p_\infty, c(k)\big) \leq \underbrace{d\big(p_\infty, p_k\big)}_{\to 0} + \underbrace{d\big(p_k, c(k)\big)}_{<1/k} \longrightarrow 0.$$

This leads to $c(k) \to p_\infty$ and hence a contradiction.




(ii): Let $c$ not be a lightlike pregeodesic without conjugated points, so neither is $c\big|_{[0,a]}$ for some $a > 0$. According to Theorem 2.44, there is a timelike curve from $c(0)$ to $c(a)$ in $M\backslash A$.
It follows that $p \in I_+^{M\backslash A}\big(c(a)\big)$, so we can apply the proof of (i) by replacing $c$ by $c\big|_{[0,a]}$, $p$ by $c(a)$ and $q$ by $p$. □

**Remark 2.80.**
The assumption on $c$ in (ii) can not be dropped as the following example shows: Let $M = \mathbb{R}^n_{\text{Mink}}$ and $A = -H^{n-1}$.
Lightlike past directed lines, which start at $0$, do not hit $A$, but every timelike, past directed and past-inextendible curve $\widetilde{c}$, which starts at $0$, does.



**Proposition 2.81.** *Let $M$ be a connected and time-oriented Lorentzian manifold and $S \subset M$ a Cauchy hypersurface of $M$. Then we have*

*(i) $S$ is achronal.*

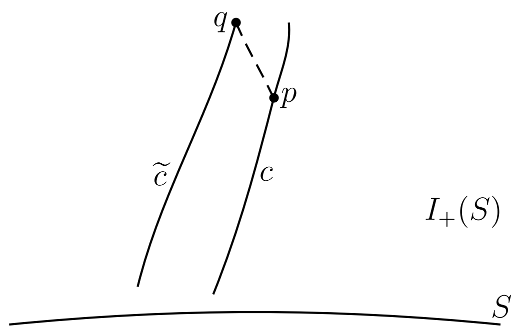
*(ii) $S$ is a closed topological hypersurface.*

*(iii) Every inextendible causal curve hits $S$.*

*Proof.* (i): If there was a timelike curve hitting $S$ at least twice, each maximal extension as a timelike curve would do either, which yields a contradiction.
(ii): Through any $p \in M$, we find an inextendible timelike curve, which therefore has to hit $S$, so we clearly have $M = I_-(S) \cup S \cup I_+(S)$. This is a disjoint union because if $I_\pm \cap S$ or $I_-(S) \cap I_+(S)$ were not empty, this would imply the existence of a timelike curve meeting $S$ twice. It follows that the subsets $I_\pm(S) \cup S$ are closed as complements of the open sets $I_\mp(S)$, so they contain $\overline{I_\pm(S)}$, respectively, and hence,

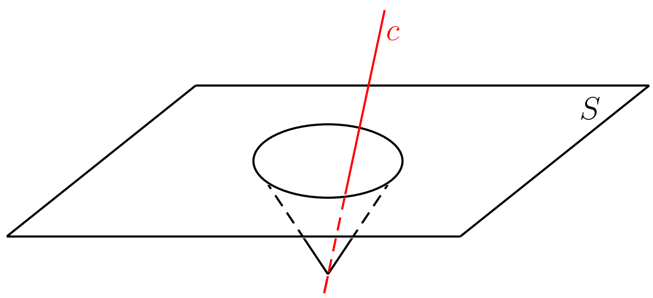
$$\partial I_\pm(S) = \overline{I_\pm(S)} \cap M \backslash I_\pm(S) \subset \big(I_+(S) \cup S\big) \cap \big(I_-(S) \cup S\big) = S.$$

This implies $\partial I_-(S) = S = \partial I_+(S)$ since $S \subset \partial I_\pm(S)$ holds for any subset $S$ of a Lorentzian manifold. Thus, every timelike curve from $I_-(S)$ to $I_+(S)$ has to meet $S$, that is $\mathrm{edge}(S) = \emptyset$ and the claim follows from Corollary 2.72.
(iii): Assume $c$ to be a causal inextendible curve in $M$, which does not meet $S$ and, without loss of generality, runs entirely in $I_+(S)$. For some point $p$ on that curve, let $q \in I_+^{M\backslash S}(p)$, so due to Lemma 2.79, there is a past-inextendible curve (with respect to $M$) in $M\backslash S$, which starts in $q$ and runs entirely in $I_+(p)$ as well. Then the maximal future extension of $c$ in $M$ also runs entirely in $I_+(S)$ and therefore would yield an inextendible timelike, which does not meet $S$ and thus contradicts the assumption on $S$.




□

**Remark 2.82.** In general, a Cauchy hypersurface is not hit exactly once by any *causal* curve. Let, for instance, $M = \mathbb{R}^n_{\mathrm{Mink}}$ and consider $S$ and $c$ as in the picture. Here the intersection is a whole lime segment.



**Theorem 2.83.** *Let $M$ be a connected time-oriented Lorentzian manifold, $X$ a smooth timelike vector field on $M$ and $S \subset M$ a Cauchy hypersurface. Then the map $\rho\colon M \to S$, which assigns to each $p \in M$ the unique intersection point of the corresponding integral curve of $X$ with $S$, is well-defined, continuous and open, and we have $\rho\big|_S = \mathrm{id}_S$. In particular, $S$ is connected.*

*Proof.* a) Let $c\colon (a,b) \to M,\ -\infty \leq a < 0 < b \leq \infty$, denote the maximal integral curve of $X$ and $p := c(0)$. It follows that $c$ is inextendible since if $c$ was continuously extendible by $q$ to $b$, the integral curve of $X$ through $q$ would extend $c$ as a integral curve, which contradicts maximality. Therefore, $\rho(p)$ is the unique intersection point of $c$ with $S$, i.e. $\rho$ is well-defined.
b) Consider the flow $\Psi\colon \mathcal{D} \times M \times \mathbb{R} \to M$ of $X$, where $\mathcal{D}$ denotes the maximal domain of $\Psi$. Since $S \subset M$ is a topological hypersurface, so is $S \times \mathbb{R}$ of $M \times \mathbb{R}$ and $\mathcal{D}(S) := (S \times \mathbb{R}) \cap \mathcal{D}$ of $\mathcal{D}$. Note that $\psi := \Psi\big|_{\mathcal{D}(S)}$ is continuous as a restriction of a continuous map and bijective by assumption on $S$ and $X$. Since $\mathcal{D}(S)$ and $M$ are topological hypersurfaces of the same dimension, the map $\psi\colon \mathcal{D}(S) \to M$ is a homeomorphism by Brouwer's Theorem ("Every continuous and injective map between topological manifolds of the same dimension is open").
It follows that $\rho = \pi \circ \psi^{-1}$ is open and continuous since the projection $\pi\colon M \times \mathbb{R} \to M$ is.
c) For $p \in S$, the unique intersection point of the corresponding integral curve of $X$ with $S$ is $p$, i.e. $\rho(p) = p$.
Since $\rho$ is continuous and surjective, i.e. $\rho(M) = S$, and $M$ is connected, $S$ is connected as well. □

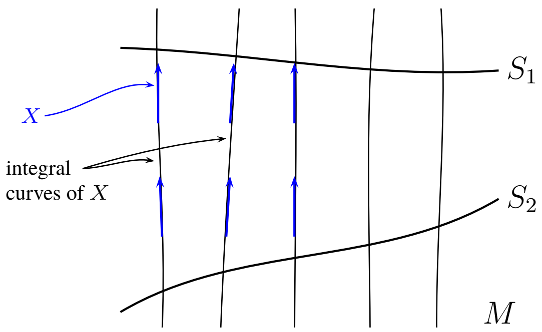
**Corollary 2.84.** *Any two Cauchy hypersurfaces $S_1, S_2$ of $M$ are homeomorphic.*

*Proof.*
Let $X$ be a smooth timelike vector field $M$ and $\rho_i\colon M \to S_i,\ i = 1, 2$, maps defined as in Theorem 2.83. Then the maps

$$\rho_1\big|_{S_2}\colon\quad S_2 \longrightarrow S_1$$

$$\rho_2\big|_{S_1}\colon\quad S_1 \longrightarrow S_2$$

are inverses of each other and hence homeomorphisms.




□

## 2.6 Globally hyperbolic subsets

Let $M$ always be a connected and time-oriented Lorentzian manifold.

**Definition 2.85.** A subset $\Omega \subset M$ is called **globally hyperbolic** if

1. The strong causality condition (Definition 2.19) holds on $\Omega$.
2. For all $p, q \in \Omega$, the **causal diamonds**

$$J(p,q) := J_+(p) \cap J_-(q)$$

are compact and contained in $\Omega$.

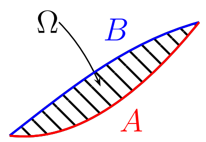
**Example 2.86.** Let $M = \mathbb{R}^n_{\text{Mink}}$. For arbitrary subsets $A, B \subset M$, consider

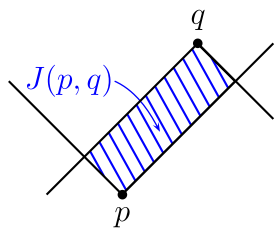
$$\Omega := J_+(A) \cap J_-(B).$$

Clearly, property 1 holds and $J(p,q)$ is compact for all $p, q \in \Omega$. For all $p, q \in \Omega$, we moreover have

$$\begin{aligned} J(p,q) \subset J_+(p) &\subset J_+(\Omega) \\ &\subset J_+\big(J_+(A)\big) = J_+(A), \end{aligned}$$

and analogously $J(p,q) \subset J_-(B)$, that is $J(p,q) \subset \Omega$.




**Lemma 2.87.** *Let $K \subset M$ be compact and let the strong causality condition hold on $K$. Furthermore, let $c\colon [0,b) \to M$, $0 < b \leq \infty$, a future-inextendible causal curve starting in $K$, i.e. $c(0) \in K$. Then there exists some $t_0 \in (0,b)$ such that $c(t) \notin K$ for all $t \in [t_0, b)$.*

*Proof.* Assume otherwise and let $(s_i)_{i\in\mathbb{N}} \subset (0,b)$ be a sequence converging to $b$ with $s_i < s_{i+1}$ and $c(s_i) \in K$ for all $i$, which implies $c(s_i) \to p \in K$ for some subsequence. Since $c$ is future-inextendible, there has to be a sequence $(t_i)_{i\in\mathbb{N}} \subset (0,b)$ converging to $b$ with $t_i < t_{i+1}$ for all $i$ but $c(t_i) \nrightarrow p$. Therefore, by passing on to some subsequence, we find a neighborhood $U$ of $p$ such that $c(t_i) \notin U$ for all $i$ and $s_1 < t_1 < s_2 < t_2 < \ldots$.
On the other hand, we have $c(s_i), c(s_{i+1}) \in V$ for all neighborhoods $V \subset U$ of $p$ and $i$ large enough, so the strong causality condition demands $c\big([s_i, s_{i+1}]\big) \subset U$ and consequently $c(t_i) \in U$, which yields a contradiction. □

**Lemma 2.88.** *Let $K \subset M$ be compact and let the strong causality condition hold on $K$. Furthermore, let $c_n\colon [0,1] \to K$ denote future directed causal curves with $c_n(0) \to p$, $c(1) \to q$ and $p \neq q$.*
*Then there exists a causal future directed broken geodesic $\gamma$ from $p$ to $q$ and a subsequence $\big(c_{n_m}\big)_m$ such that*

$$\lim_{m\to\infty} L\big[c_{n_m}\big] \leq L[\gamma].$$

*Proof.* Due to Proposition 2.60, $(c_n)_n$ has a limit sequence $p =: p_0 < p_1 < p_2 < \ldots$ and we start by showing that is is finite.
Assume it was infinite, so the corresponding quasi-limit would be a causal future-inextendible curve starting in $p$, which, by Lemma 2.87, leaves $K$ without return, i.e. $p_i \notin K$ for all $i$ large enough. Since $c_{n_m}\big(t_{m,i}\big) \to p_i$, we would have $c_{n_m}\big(t_{m,i}\big) \notin K$ for alle $m, i$ large enough, but $c$ was assumed to run in $K$.
We obtain the limit sequence $p =: p_0 < p_1 < \ldots < p_N := q$. The corresponding quasi-limit therefore is a causal future directed broken geodesic from $p$ to $q$. The points $p_i, p_{i+1}$ as well as the segments $c_{n_m}\big([t_{m,i}, t_{m,i+1}]\big)$ are contained in a convex set, which depends on $i$ but not on $m$. Then for instance Gauß' Lemma implies

$$L\Big[c_{n_m}\big|_{[t_{m,i},t_{m,i+1}]}\Big] \leq \big|\Delta\big(p_{m,i}, p_{m,i+1}\big)\big|,$$

where $p_{m,j} := c_{n_m}\big(t_{m,j}\big)$, $|\cdot| := \sqrt{|\langle\cdot,\cdot\rangle|}$ and $\Delta$ the map defined in (2.3). It follows that

$$L\big[c_{n_m}\big] \leq \sum_{i=0}^{N-1} \big|\Delta\big(p_{m,i}, p_{m,i+1}\big)\big|.$$

For $m \to \infty$, the right hand side converges to $\big|\Delta\big(p_i, p_{i+1}\big)\big| = L[\gamma]$. After passing to some subsequence, $\big(L\big[c_{n_m}\big]\big)_m$ converges as well and we obtain

$$\lim_{m\to\infty} L\big[c_{n_m}\big] \leq L[\gamma]. \qquad \square$$

**Lemma 2.89.** *For $p < q$ in $M$, let $J(p,q)$ be compact and let the strong causality condition hold on $J(p,q)$. Then there exists a causal geodesic from $p$ to $q$ of length $\tau(p,q)$.*

*Proof.* Let $c_n\colon [0,1] \to M$ denote causal future directed curves with $c_n(0) = p, c_n(1) = q$ and $L[c_n] \to \tau(p,q)$. The strong causality condition yields $c_n\big([0,1]\big) \subset J(p,q)$ for all $n$, so Lemma 2.88 ensures the existence of a causal future directed broken geodesic $\gamma$ from $p$ to $q$ with

$$\tau(p,q) = \lim_{m\to\infty} L\big[c_{n_m}\big] \leq L[\gamma] \leq \tau(p,q)$$

and hence $L[\gamma] = \tau(p,q)$. If $\gamma$ actually was not smooth in some $t_0$, i.e. not an (unbroken) geodesic, it is well-known that we would then find a variation with fixed endpoints with non-zero first variation of arc length. By some similar procedure as in the proof of Lemma 2.29, we would find a longer causal curve from $p$ to $q$ but by definition of $\tau$, $L[\gamma] = \tau(p,q)$ implies that $\gamma$ already is the longest curve. □

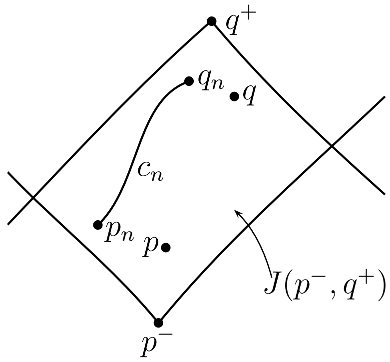
**Remark 2.90.** In the Riemannian geometry, this Lemma corresponds to the statement contained in the Theorem of Hopf-Rinow that on a complete Riemannian manifold, any two points can be connected by some shortest geodesic.

**Proposition 2.91.** *Let $\Omega \subset M$ be a globally hyperbolic subset. Then $\tau$ is continuous and finite on $\Omega \times \Omega$.*

*Proof.* Note that $\tau < \infty$ on $\Omega \times \Omega$ follows directly from Lemma 2.89. Furthermore, $\tau$ is always lower semi-continuous, so we just have to check upper semi-continuity. Suppose there was some $(p,q) \in \Omega \times \Omega$, where $\tau$ fails to be upper semi-continuous, i.e. we find $\delta > 0$ as well as sequences $(p_n)_{n\in\mathbb{N}}$, $(q_n)_{n\in\mathbb{N}}$ converging to $p, q$ such that

$$\tau(p_n, q_n) \geq \tau(p,q) + \delta$$

for all $n$. Choose causal future directed curves $c_n\colon [0,1] \to M$ with $c_n(0) = p_n, c_n(1) = q_n$ and $L[c_n] \geq \tau(p_n, q_n) - \frac{1}{n}$. Since $\Omega$ is open, we therefore find $p^-, q^+ \in \Omega$ such that $p^- \ll p, q^+ \ll q$. Since $I_+(p^-), I_-(q^+)$ are open neighborhoods of $p$ and $q$, respectively, we have $p_n \in I_+(p^-)$ and $q_n \in I_-(q^+)$ for $n$ large enough, and moreover




$$c_n\big([0,1]\big) \subset I_+(p^-) \cap I_-(q^+) \subset J(p^-, q^+). \tag{2.4}$$

Due to global hyperbolicity of $\Omega$, $J(p^-, q^+)$ is compact and satisfies the strong causality condition, so Lemma 2.88 implies the existence of a broken geodesic $\gamma$ from $p$ to $q$ and a subsequence $\big(c_{n_m}\big)_m$ such that

$$\lim_{m\to\infty} L\big[c_{n_m}\big] \leq L[\gamma] \leq \tau(p,q).$$

It follows that

$$L\big[c_{n_m}\big] \geq \tau\big(p_{n_m}, q_{n_m}\big) - \frac{1}{n_m} \geq \tau(p,q) + \delta - \frac{1}{n_m},$$

which, for $m \to \infty$, yields $\tau(p,q) \geq \tau(p,q) + \delta$ and hence a contradiction. □

**Proposition 2.92.** *Let $\Omega \subset M$ be an open and globally hyperbolic subset. Then "$\leq$" is a closed relation on $\Omega$.*

*Proof.* Let $(p_n)_{n\in\mathbb{N}}, (q_n)_{n\in\mathbb{N}} \subset \Omega$ converging to some $p, q \in \Omega$, respectively, such that $p_n \leq q_n$ for all $n$. We have to show that then $p \leq q$. Since "$=$" is a closed relation, the statement is trivial if $p_n = q_n$ for infinitely many $n$, so without loss of generality, we assume $p_n < q_n$ for all $n$ (otherwise pass on to a suitable subsequence).
Let $c_n\colon [0,1] \to M$ be the corresponding causal future directed curves with $c_n(0) = p_n$ and $c_n(1) = q_n$. Like in the proof of Proposition 2.91, we choose $p^-, q^+ \in \Omega$ such that (2.4) holds, so we find a causal broken geodesic from $p$ to $q$ and hence, we have $p \leq q$. □

## 2.7 Cauchy developments and Cauchy horizones

Let $M$ always be a connected and time-oriented Lorentzian manifold.

**Definition 2.93.** For $A \subset M$ achronal, the set

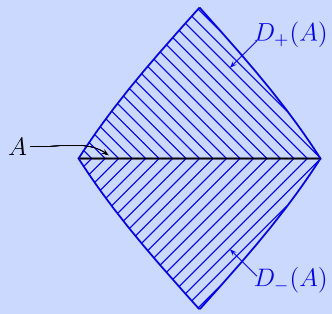
$$D(A) := \left\{p \in M \,\middle|\, \text{Every past-inextendible causal curve through } p \text{ meets } A\right\}$$

is called **future Cauchy development of** $A$.

Analogously one defines the **past Cauchy development of** $A$ and we call

$$D(A) := D_+(A) \cup D_-(A)$$

**Cauchy development of** $A$.




**Remark 2.94.**

a) $A \subset D_\pm(A) \subset A \cup I_\pm(A) \subset J_\pm(A)$.

b) $D_\pm(A) \cap I_\mp(A) = \emptyset$ since $A$ is achronal.

c) Fom a) and b) follows

$$A \subset D_+(A) \cap D_-(A) \subset D_+(A) \cap \big(A \cup I_-(A)\big) = D_+(A) \cap A = A,$$

and thus $A = D_+(A) \cap D_-(A)$.

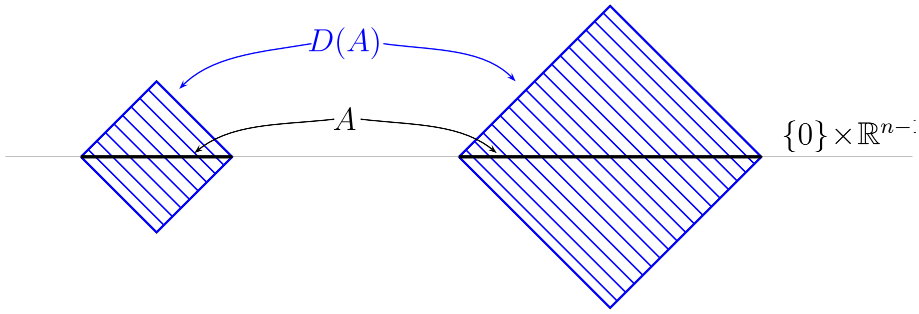
d) $D(A) \cap I_\pm(A) = D_\pm(A)\backslash A$.

**Example 2.95.** 1. For $M = \mathbb{R}^n_{\text{Mink}}$ and $A := \{0\} \times \mathbb{R}^{n-1}_{\text{Mink}}$, we have

$$D_\pm(A) = J_\pm(A) = A \cup I_\pm(A).$$

2. Let $M$ be any Lorentzian manifold with some Cauchy hypersurface $S$. By definition, every inextendible timelike curve meets $S$, that is $I_\pm(S) \cup S \subset D_\pm(S)$ and a) yields equality. On the other hand, we found the (disjoint) decomposition $M = I_-(S) \sqcup S \sqcup I_+(S)$, so b) implies equality and thus $M = D(S)$.

3. For $M = \mathbb{R}^n_{\text{Mink}}$ and some $B \subset M$, let $A := \{0\} \times B$. Then $D(A)$ is the double cone over $B$.

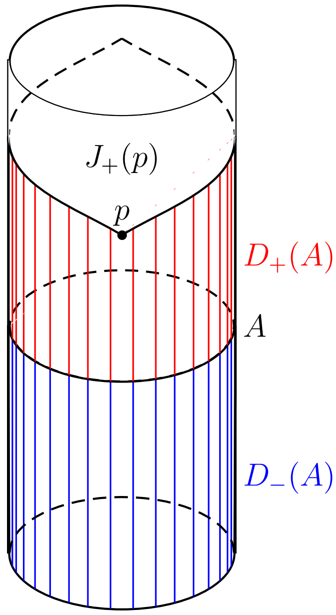



4. For $(M, g) = \left(\mathbb{R} \times S^1, -\mathrm{d}t^2 + \mathrm{d}\theta^2\right)$ and $A = \{0\} \times S^1$, we clearly have

$$D_\pm(A) = J_\pm(A).$$

For $p \in I_+(A)$ and $\widetilde{M} := M \backslash \{p\}$, we still have $D_-(A) = J_-(A)$, but $D_+(A)$ is given by the union of $A$ and the open region between $S$ and both future directed null-geodesics emanating from $p$, i.e.

$$D_+(A) = J_+(A) \backslash J_+(p).$$

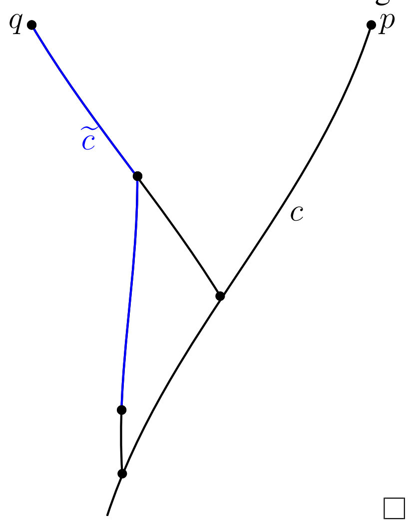



**Lemma 2.96.** *Let $A \subset M$ be an achronal set.*

*(i) Every causal past directed curve, which starts in $D_+(A)$ and leaves $D_+(A)$, meets $A$.*

*(ii) Every past or future-inextendible causal curve through $p \in \mathring{D}(A)$ meets $I_-(A)$ or $I_+(A)$, respectively.*

*Proof.* (i): Let $c\colon [0,b] \to M$ be a causal and past directed curve with $c(0) \in D_+(A)$ and $c(b) \notin D_+(A)$. Hence, we find some past-inextendible causal curve $\gamma$, which starts in $c(b)$ but does not hit $A$. Then the concatenation $c \cup \gamma$ yields a past-inextendible causal curve through $c(0) \in D_+(A)$ and thus meets $A$, i.e. $c$ does.



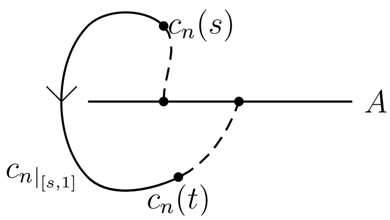

(ii): We just prove the first case, since both can be treated similarly. Remark 2.94 a) yields

$$D(A) \subset A \cup I_+(A) \cup I_-(A).$$

Let $c$ be a past-inextendible causal curve starting in $p \in \mathring{D}(A)$, so for $p \in I_-(A)$, the claim is trivial. Let $p \in A \cup I_+(A)$ and choose $q \in I_+(p) \cap D(A)$. The proof of Lemma 2.79 (i) with $A = \emptyset$ shows that there is a past-inextendible timelike curve $\widetilde{c}$ starting in $q$ such that $c$ meets $I\big(\widetilde{c}(s)\big)$ for all $s$. Remark 2.94 d) implies $q \in D_+(A)$, so $\widetilde{c}$ meets $A$ in some $\widetilde{c}(s)$ and therefore $c$ meets $I_-(A)$. $\square$

**Theorem 2.97.** *Let $A \subset M$ be achronal. Then $\mathring{D}(A)$ is globally hyperbolic.*

*Proof.* a) We start by showing that $\mathring{D}(A)$ has the causality property.
Assume there was a causal loop $c$ through some point in $\mathring{D}(A)$, i.e. due to Lemma 2.96 (ii), we find points $q_\pm \in I_\pm(A)$ on $c$. Hence, there are $q'_\pm \in A$ such that $q_\pm \in I_\pm(q'_\pm)$, that is $q'_+ \ll q_+ \le q_- \ll q'_-$. This would imply the existence of a timelike curve meeting both $q'_\pm$, which contradicts achronality of $A$.
b) We show that $\mathring{D}(A)$ has the strong causality property.
Assume that it does not hold at some $p \in \mathring{D}(A)$, i.e. there is a sequence of causal future directed curves $c_n\colon [0,1] \to M$, $n \in \mathbb{N}$, with $\lim\limits_{n\to\infty} c_n(0) = p = \lim\limits_{n\to\infty} c_n(1)$ and a neighborhood $U$ of $p$ such that for all $n$, $c_n$ does not entirely run in $U$. Due to Proposition 2.60, there is a limit sequence $p =: p_0 < p_1 < \ldots$ of $(c_n)_{n\in\mathbb{N}}$. If it is finite, then $p_N = p$, that is $p < p$ and hence a contradiction to the causality condition.
Therefore, suppose the limit sequence is infinite and the corresponding quasi-limit $\gamma$ future-inextendible. According to Lemma 2.96 (ii), it meets $I_+(A)$ and does not leave it, that is $p_i \in I_+(A)$ for some element $p_i$ of the limit sequence. Possibly passing on to some subsequence and after a reparametrization, there exists $s \in (0,1)$ such that $c_n(s) \to p_i$. In particular, we have $c_n(s) \in I_+(A)$ for $n$ large enough.
Applying Proposition 2.60 to $\big(c_n|_{[s,1]}\big)_n$ provides a limit sequence $p =: q_0 > q_1 > \ldots$. If it was finite, we would have $q_{N'} = p$ and again $p < p_i = q_{N'} < q_0 = p$, which contradicts the causality condition. Therefore, we consider an infinite limit sequence with a past-inextendible causal curve $\widehat{\gamma}$ starting in $p \in \mathring{D}(A)$ as the corresponding quasi-limit. From Lemma 2.96 (ii) follows that $\widehat{\gamma}$ hits $I_-(A)$, so we would find $n \in \mathbb{N}$ and $t \in [0,1]$ such that $c_n(t) \in I_-(A)$, which contradicts achronality of $A$.



c) We show that $J(p,q)$ is compact for all $p, q \in \mathring{D}(A)$.
From $p \not\leq q$ follows $J(p,q) = \emptyset$, so here the claim is trivial. Furthermore, $p = q$ implies $J(p,q) = \{p\}$, since for any $r \in J(p,p)\backslash\{p\}$, we would have $p < r < p$, which contradicts achronality of $A$.
Thus, we consider $p < q$ and we show that every sequence $(x_n)_{n\in\mathbb{N}} \subset J(p,q)$ has a subsequence, which converges in $J(p,q)$. Let $c_n \colon [0,1] \to M$ causal future directed curves from $p$ to $q$ through $x_n$. Furthermore, let $\mathcal{K}$ be a cover of $M$ by open and convex subsets $U$ such that the closures $\overline{U}$ are compact and contained in some open and convex set and such that, due to Proposition 2.60, we find a limit sequence $p =: p_0 < p_1 < \ldots$ of $(c_n)_{n\in\mathbb{N}}$ relative to $\mathcal{K}$. We show that we can always find a finite one:
c1) Assume that all such limit sequences were infinite. Adopting the approach of b), we find a subsequence of reparametrised $c_n$ such that $c_n(s) \to p_i \in I_+(A)$ for some fixed $s$. Again the sequence $\left(c_n\big|_{[s,1]}\right)_n$ provides a limit sequence $q =: q_0 > q_1 > \ldots$, which has to be infinite since otherwise we would have found a finite limit sequence

$$p = p_0 < p_1 < \ldots < p_i = q_N < \ldots < q_0 = q$$

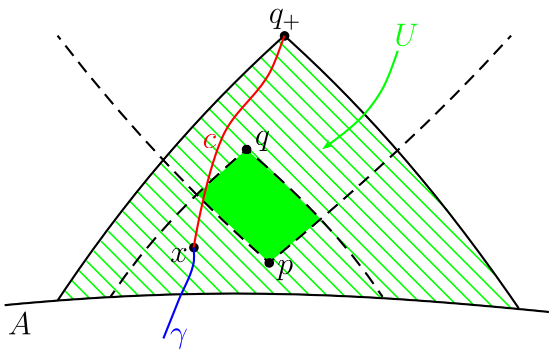
of $(c_n)_{n\in\mathbb{N}}$. The corresponding quasi-limit is past-inextendible and as in b), it hits $I_-(A)$, which contradicts achronality of $A$.
c2) Let $p =: p_0 < p_1 < \ldots < p_N := q$ be a finite limit sequence of $(c_n)_{n\in\mathbb{N}}$ relative to $\mathcal{K}$. Passing on a subsequence of $(c_n)_{n\in\mathbb{N}}$ yields $x_n \in c_n\big([s_{n,i}, s_{n,i+1}]\big)$ for all $n$ and fixed $i$, so all $x_n$ are contained in some fixed $U \in \mathcal{K}$. Due to compactness of $\overline{U}$, a subsequence of $(x_n)_{n\in\mathbb{N}}$ converges to some $x \in \overline{U} \subset V$, where $V$ is open and convex. Now Lemma 2.56 implies $x \in J(p,q)$ since

$$c_n(s_{n,i}) \leq x_n \leq c_n(s_{n,i+1}) \quad \Longrightarrow \quad p_i \leq x \leq p_{i+1} \quad \Longrightarrow \quad p \leq x \leq q.$$

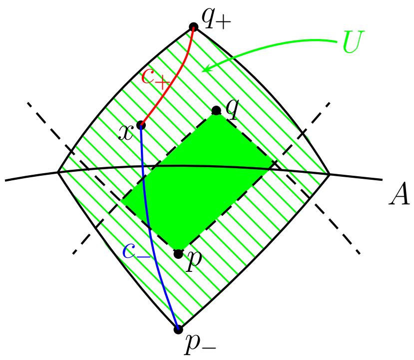
d) It remains to show $J(p,q) \subset \mathring{D}(A)$ for all $p, q \in \mathring{D}(A)$.
Clearly, we only have to consider $p < q$ and we start with the case $p, q \in I_+(A)$. Choose $q_+ \in I_+(q) \cap D(A) \subset I_+(A) \cap D_+(A)$ and set $U := I_+(A) \cap I_-(q_+)$. Since

$$J(p,q) \subset J_+\big(I_+(A)\big) \cap J_-\big(I_-(q_+)\big) = I_+(A) \cap I_-(q_+) = U,$$

this yields an open neighborhood of $J(p,q)$ and we show that it is contained in $D(A)$. For $x \in U$, let $c$ be a timelike future directed curve from $x$ to $q_+$, which fails to meet $A$ due to achronality. Hence, for any past-inextendible causal curve $\gamma$ starting in $x$, the concatenation $c \cup \gamma$ yields a past-inextendible causal curve, which starts in $q_+$ and therefore meets $A$. It follows that $\gamma$ meets A and thus $x \in D_+(A)$.



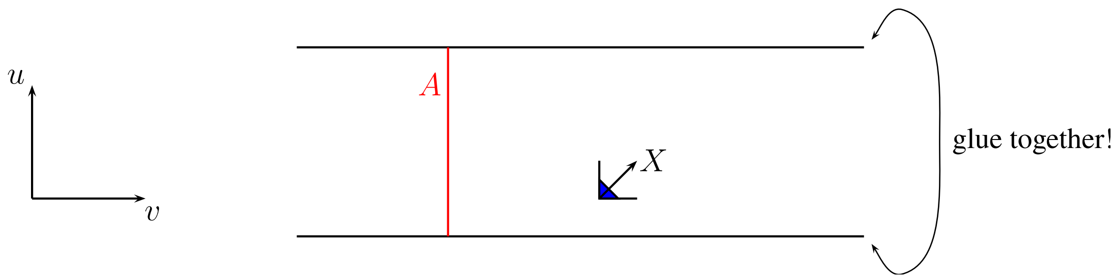
Now consider the case $p \in I_-(A), q \in I_+(A)$. Choose $p_- \in I_-(p) \cap D(A)$ and $q_+ \in I_+(q) \cap D(A)$, so $U := I_+(p_-) \cap I_-(q_+)$ is again a neighborhood of $J(p,q)$ and we show $U \subset D(A)$. Let $x \in U$. Since for $x \in A$, the claim directly follows from $A \subset D(A)$, we assume $x \notin A$. Let $c_-, c_+$ be timelike future directed curves from $p_-$ to $x$ and $x$ to $q_+$, respectively. Due to achronality of $A$, at least one of both curves does not meet $A$.




□

On page 421, [O'Neill1983] claims that for achronal set $A$, the causality condition holds on all of $D(A)$. The following example demonstrates that this is *not* the case.

**Example 2.98.** Let $M := S^1 \times \mathbb{R}$ with coordinates $(u,v)$ and $g := -\,\mathrm{d}u \otimes \mathrm{d}v - \mathrm{d}v \otimes \mathrm{d}u$. The time-orientation is determined by $X := \frac{\partial}{\partial u} + \frac{\partial}{\partial v}$ and we consider the subset $A := S^1 \times \{0\}$.

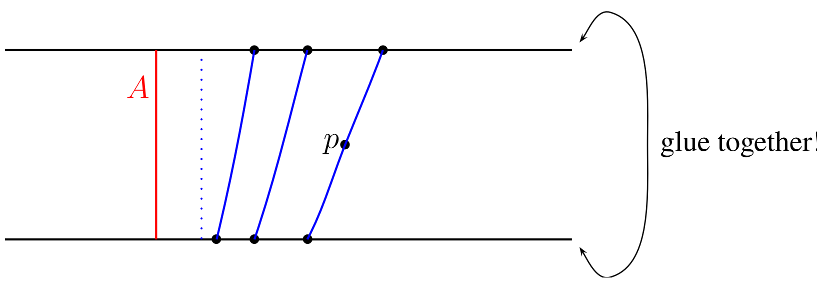



The subset $A$ is achronal: For any timelike curve $s \mapsto \big(u(s), v(s)\big)$, we obtain

$$0 > g\left(u' \frac{\partial}{\partial u} + v \frac{\partial}{\partial v}, u' \frac{\partial}{\partial u} + v \frac{\partial}{\partial v}\right) = -2u'v'.$$

Note that $v' \neq 0$, which implies either $v'(s) > 0$ for all $s$ or $v'(s) < 0$ for all $s$. Hence, this curve meets $A$ at most once.

On the other hand, $A$ fails to be acausal: For each $p \in M \backslash A$, there is a inextendible timelike curve through $p$, which does not meet $A$. It follows that $D(A) = A$ and $A$ does not satisfy the causality condition since $A$ itself yields a causal loop.



**Corollary 2.99.** *Let $M$ be a Lorentzian manifold, which exhibits a Cauchy hypersurface. Then $M$ is globally hyperbolic.*

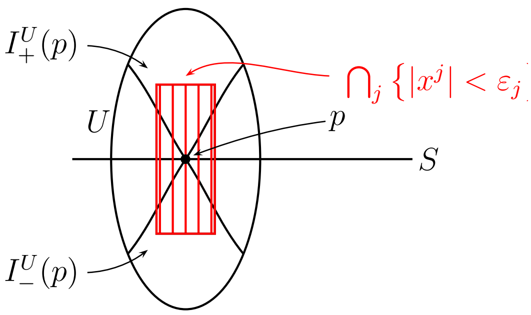
*Proof.* Let $S \subset M$ be a Cauchy hypersurface of $M$, so $\mathring{D}(S)$ is globally hyperbolic by Theorem 2.97. On the other hand, we have $D(S) = M = \mathring{M} = \mathring{D}(S)$. □

**Lemma 2.100.** *Every spacelike achronal (smooth) hypersurface is acausal.*

*Proof.* Suppose there was a causal future directed curve $c\colon [0,1] \to M$ with $c(0), c(1) \in S$, where $S$ is a spacelike achronal hypersurface. By Theorem 2.44, there is a timelike curve from $S$ to $c(1)$, which contradicts achronality of $A$, unless $c$ is a lightlike pregeodesic without focal points before $c(1)$ with $\dot{c}(0) \perp S$. Since $S$ is spacelike, $N_{c(0)}S$ is timelike, so $\dot{c}(0)$ and thus $c$ is not lightlike, i.e. $c$ fails to be causal. □

**Proposition 2.101.** *Let $S \subset M$ be an acausal topological hypersurface. Then $D(S)$ is open and globally hyperbolic.*

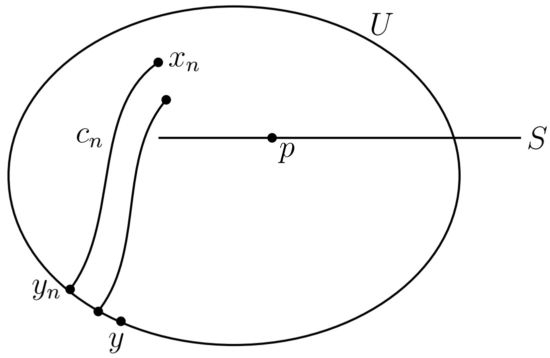
*Proof.* Recall that due to acausality of $S$, the union $I := I_-(S) \cup S \cup I_+(S)$ is a disjoint one, since if $I_-(S) \cap S$ or $I_+(S) \cup I_-(S)$ were not empty, we would find timelike curves hitting $S$ at least twice.
a) We show that $I \subset M$ is open in $M$, i.e. every $p \in S$ is contained in $\mathring{I}$:
By Proposition 2.71, we have $S \cap \mathrm{edge}(S) = \emptyset$ and thus $p \notin \mathrm{edge}(S)$. Hence, there exists a neighborhood $U$ of $p$ such that all timelike curves in $U$ from $I^U_-(p)$ to $I^U_+(p)$ meet $S$. Let $x^0, \dots, x^{n-1}$ be Riemannian normal coordinates around $p$ with timelike $x^0$ and $|x^j| < \varepsilon_j$ for some fixed $\varepsilon_j > 0$. Choosing the $\varepsilon_j$ suitably small ensures $\{x^0 = \pm\varepsilon_0\} \subset U$. Then the $x^0$-coordinate lines meet $S$ and therefore run entirely in $I$, so $\bigcap_{j=0}^{n-1} \{|x^j| < \varepsilon_j\}$ yields an open neighborhood of $I$.




b) Next we show that $S$ is actually contained in $\mathring{D}(S)$:
Suppose $p \in S \backslash \mathring{D}(S)$ and let $U$ be an open neighborhood of $p$ such that $\overline{U}$ is compact and $U \subset V \cap I$ for some convex subset $V$. Since $p \notin \mathring{D}(S)$, we find a sequence $(x_n)_{n\in\mathbb{N}}$ in $M \backslash D(S)$ converging to $p$. Without loss of generality, for each $n$, let $x_n \in I_+(S) \cap U$, and since $x_n \notin D_+(S)$, there is a past-inextendible timelike curve $c_n$, which starts in $x_n$ and does not meet $S$. Since $\overline{U}$ is contained in the convex set $V$, everys $c_n$ hits the boundary $\partial U$

(Lemma 2.56 (iv)) and we call the first intersection point $y_n$. Then we have $y_n \leq x_n$ and due to compactness of $\partial U$, we find a subsequence converging to some $y \in \partial U$. By Lemma 2.56, the relation "$\leq$" is closed, that is $y \leq p$ and even $y < p$ since $y \neq p$. In particular, $y \in I$.
b1) If $y \in I_+(S)$, we would find $q \in S$ such that $q \ll y$ and thus $q \ll p$, which contradicts achronality of $S$.

b2) $y \in S$ contradicts $y < p$ since $S$ is achronal.




b3) If $y \in I_-(S)$, there would exist some $n$ such that $y_n = c_n(t_n) \in I_-(S)$. On the other hand, by definition of $y_n$, we have

$$c_n\big([0,t_n]\big) \subset \overline{U} \subset I = I_-(S) \sqcup S \sqcup I_+(S).$$

Recall that $c_n$ does not meet $S$, so $c_n\big([0,t_n]$ has to be contained in $I_-(S)$, which contradicts the assumption $c_n(0) = x_n \in I_+(S)$.
c) We show that $D(S)$ is open if $S \subset M$ is closed:
For $S$ closed, it suffices to show that $D_+(S)\backslash S = I_+(S) \cap D(S)$ is open since it would directly follow that $D_-(S)\backslash S$ is open and therefore, $D(S)$ can be written as the union of 3 open sets:

$$D(S) = \big(D_+(S)\backslash S\big) \cup S \cup \big(D_-(S)\backslash S\big) \subset \underbrace{\big(D_+(S)\backslash S\big) \cup \mathring{D}(S) \cup \big(D_-(S)\backslash S\big)}_{\text{open}} \subset D(S).$$

Assume $p \in D_+(S)\backslash S$ was not an inner point. Then we find a sequence $(x_n)_{n\in\mathbb{N}} \not\subset D_+(S)\backslash S$ converging to $p$ such that for each $n$, there is a past-inextendible causal curve $c_n \colon [0,b) \to M$ starting in $x_n$, which does not meet $S$ (except maybe in $x_n$). Due to b) and since $S$ is closed, $\big\{M\backslash S, \mathring{D}(S)\big\}$ yields an open cover of $M$. By Proposition 2.55, there is a refinement $\mathcal{K}$ by open and convex sets, which therefore either are not contained in $\mathring{D}(S)$ or do not meet $S$. Let $\gamma$ be a quasi-limit of $(c_n)_{n\in\mathbb{N}}$ relative to $\mathcal{K}$ starting in $p$, which is a past-inextendible causal curve. Due to $p \in D_+(S)$, $\gamma$ meets $S$ in some unique point $\gamma(s)$ and for $p = p_0 > p_1 > \dots$ the corresponding limit sequence, let $i$ be the index such that

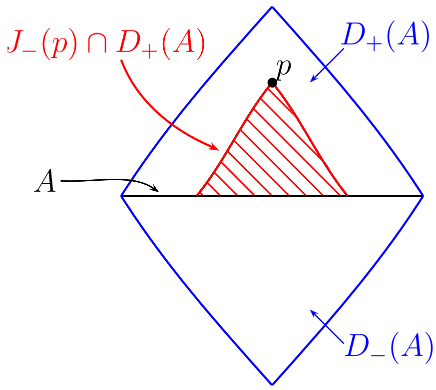
$$p_i > \gamma(s) \geq p_{i+1}.$$

Hence, the element of $\mathcal{K}$, which contains the corresponding segment of $\gamma$, meets $S$ (in $\gamma(s)$) and is thus contained in $\mathring{D}(S)$ by choice of $\mathcal{K}$. Acausality of $S$ implies $p_i \notin S$, that is $p_i \in D_+(S)\backslash S = I_+(S) \cap D(S)$. It is even contained in the open set $I_+(S) \cap \mathring{D}(S) \subset D_+(S)$, so for $n$ large enough, $c_n$ has to meet $D_+(S)$. Consequently, as a past-inextendible causal curve, it also meets $S$, which yields a contradiction.
d) Finally assume that $S$ is not closed. Note that due to acausality, the Cauchy developments of different connected components of $S$ are pairwise disjoint, so without loss of generality, assume $S$ to be connected. Clearly, $S$ is closed in $I$, so replacing $M$ by its connected open submanifold $I$ in c) shows that $D(S)$ is open in $I$. On the other hand, the Cauchy development of $S$ in $I$ and in $M$ coincide, i.e. $D(S)$ is an open subset of $M$ and thus globally hyperbolic by Theorem 2.97. $\square$

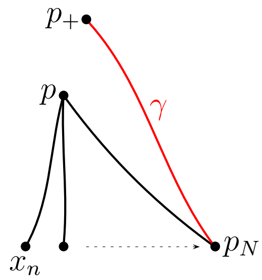
**Remark 2.102.** Proposition 2.101 becomes wrong if one replaces "acausal" by "achronal" as is demonstrated in Example 2.98.

**Lemma 2.103.** *For each achronal subset $A \subset M$ and $p \in \mathring{D}(S) \backslash I_-(A)$, the intersection $J_-(p) \cap D_+(A)$ is compact.*

*Proof.*
Let $(x_n)_{n\in\mathbb{N}} \subset J_-(p) \cap D_+(A)$ a sequence with no subsequence converging to $p$, and for each $n$, choose a causal past directed curve $c_n$ from $p$ to $x$. Note that if such a sequence does not, there is nothing to proof. By Proposition 2.60, there is a limit sequence $p =: p_0 > p_1 > \dots$.




a) Assume that the limit sequence is finite, i.e. $p > p_1 > \dots > p_N$ and we find a subsequence of $(x_n)_{n\in\mathbb{N}}$ converging to $p_N$, so it remains to show $p_N \in D_+(A)$. Let $p_+ \in I_+(p) \cap D_+(A)$, that is $p_+ \gg p \geq p_N$ and thus $p_+ \gg p_N$, so there is a timelike past directed curve $\gamma$ from $p_+$ to $p_N$. If $\gamma$ does not meet $A$, we directly have $p_N \in D_+(A)$. If $\gamma$ meets $A$, then $p_N \in A \cap I_-(A)$, but $p_N \in I_-(A)$ would imply also $x_n \in I_-(A)$ for $n$ large enough, which contradicts $x_n \in D_+(A)$.




b) If the limit sequence is infinite, the quasi-limit $\gamma$ is a past-inextendible causal curve curve, which starts in $p$ and meets $I_-(A)$ by Lemma 2.96. Therefore, $p_i \in I_-(A)$ for $i$ large enough and consequently, for $n$ large enough, $x_n \in I_-(A)$, which again contradicts $x_n \in D_+(A)$. □

**Definition 2.104.** For all $p \in M$ and $A \subset M$, set

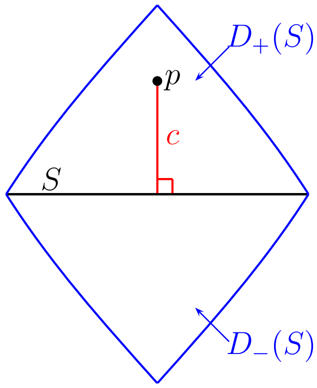
$$\tau(A, p) := \sup_{q \in A} \tau(p, q).$$

**Theorem 2.105.** *Let $S \subset M$ be a closed, achronal, spacelike and smooth hypersurface and $p \in D(S)$ such that there is a geodesic $c$ from $S$ to $p$ of length $\tau(S, p)$. Then $c$ is orthogonal to $S$, it has no focal point before $p$, and it is timelike, unless $p \in S$.*

*Proof.* Due to Lemma 2.100, $S$ is acausal, and by Proposition 2.101, $D(S)$ is open and globally hyperbolic. Without loss of generality, we only consider $p \in D_+(S)$. Furthermore, Lemma 2.103 implies compactness of $J_-(p) \cap D_+(S)$ and hence

$$J_-(p) \cap S = J_-(p) \cap D_+(S) \cap S$$

is compact as well since $S$ is closed. Proposition 2.91 ensures continuity of $\tau$ on $J_-(p) \cap S$, so the maximum of $\tau(\cdot, p)$ on $J_-(p) \cap S$ is attained at some $q$. By Lemma 2.89, there is a causal geodesic $c$ of length $\tau(q, p)$ connecting $q$ and $p$. If $c$ was not orthogonal to $S$ or if $c$ had a focal point before $p$, this curve could be deformed into a longer timelike curve from $S$ to $p$, which contradicts the maximality of the length of $c$.




□

**Definition 2.106.** Let $A \subset M$ be achronal. Then we call

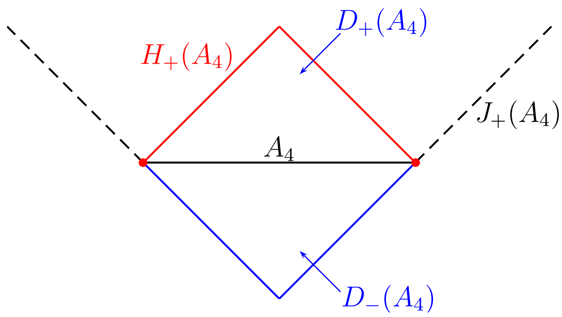
$$H_+(A) := \overline{D_+(A)} \backslash I_-\big(D_+(A)\big) = \big\{p \in \overline{D_+(A)} \;\big|\; I_+(p) \cap D_+(A) = \emptyset\big\}$$

**future Cauchy horizon of** $A$. Analogously, one defines $H_-(A)$, the **past Cauchy horizon of** $A$. The **Cauchy horizon of** $A$ is given by

$$H(A) := H_+(A) \cup H_-(A).$$

**Example 2.107.** Let $M := \mathbb{R}^n_{\mathrm{Mink}}$.

1. For $A_1 := \{0\} \times \mathbb{R}^{n-1}$, we obtain $D_\pm(A_1) = J_\pm(A_1)$, and consequently $H(A) = \emptyset$.
2. For $A_2 := H^{n-1}$ the $(n-1)$-dimensional hyperbolic space, we have $D(A_2) = D_+(A_2) = I_+(0)$ but $\overline{D_-(A_2)} = C_+(0)$, so $H_+(A_2) = \emptyset$ and $H_-(A_2) = C_+(0)$.
3. For $A_3 := C_+(0)$, we have $D(A_3) = D_+(A_3) = J_+(0)$ and $H_\pm(A_3) = H_\pm(A_2)$.
4. For $\dim(M) = 2$ we consider $A_4 := \{0\} \times (-1, 1)$ and obtain $H_+(A_4)$ as in the picture. Particularly note that $H_+(A_4) \not\subset J_+(A_4)$ since $(0, \pm 1) \in H_+(A_4) \backslash J_+(A_4)$.



**Lemma 2.108.** *For all achronal $A \subset M$, we have*

*(i)* $H_\pm(A)$ *is closed.*

*(ii)* $H_\pm(A)$ *is achronal.*

*(iii) If $A$ is closed, then*

$$\overline{D_+(A)} = \big\{p \in M \,\big|\, \text{every past-inextendible timelike curve through } p \text{ meets } A\big\}.$$

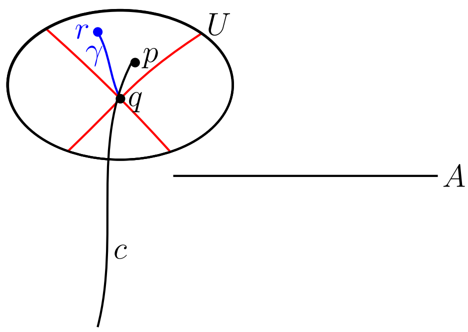
*(iv) If $A$ is closed, then*

$$\partial D_\pm(A_\pm) = A \cup H_\pm(A).$$

*Proof.* (i): Since $\overline{D_\pm(A)}$ is closed and $I_\pm\big(D_\pm(A)\big)$ is open, this follows by definition.
(ii): Since $I_+\big(H_+(A)\big)$ is open and $I_+\big(H_+(A)\big) \cap D_+(A)$ empty by definition, we obtain $I_+\big(H_+(A)\big) \cap \overline{D_+(A)} = \emptyset$ and therefore $I_+\big(H_+(A)\big) \cap H_+(A) = \emptyset$. Note that this implies achronality of $H_+(A)$.
(iii): We introduce the short-hand notation

$$X := \big\{p \in M \,\big|\, \text{every past-inextendible timelike curve through } p \text{ meets } A\big\}$$

and start with "$\subset$": If there was a $p \in \overline{D_+(A)} \backslash X$, we would find a past-inextendible timelike curve $c\colon [0,b) \to M$, which starts in $p$ but does not hit $A$. In particular $p \notin A$, and since $A$ is closed, there is an open and convex neighborhood $U$ of $p$ such that $U \cap A = \emptyset$.




Choose $\varepsilon > 0$ such that $q := c(\varepsilon) \in U$ and thus $p \in I_+^U(q)$. Since $I_+^U(q)$ is an open neighborhood of $p$ and $p \in \overline{D_+(A)}$, there is some $r \in I_+^U(q) \cap D_+(A)$ with $\gamma$ the corresponding timelike and past directed curve from $r$ to $q$. Due to convexity, $\gamma$ runs entirely in $U$, so it does not meet $A$. On the other hand, the concatenation $\gamma \cup c\big|_{[\varepsilon,b)}$ yields a past-inextendible timelike curve, which starts in $r \in D_+(A)$, so it has to meet $A$. Contradiction!
We proceed with "$\supset$": Let $p \in \overline{D_+(A)}$ and choose $q \in I_-^{M\backslash\overline{D_+(A)}}(p)$, so in particular $q \notin \overline{D_+(A)}$. Therefore, we find a past-inextendible causal curve in $M$, which starts in $p$ but does not meet $A$. Lemma 2.79 implies the existence of a past-inextendible timelike curve, which starts in $r \in D_+(A)$, i.e. it has to meet $A$. Contradiction!
(iv): We start with $A \subset D_+(A)$: We already know that $A$ is contained in $D_+(A)$, and if there was some $p \in A \cap \mathring{D}_+(A)$, we could choose $q \in \mathring{D}_+(A) \cap I_-(p)$, which implies the existence of a past-inextendible timelike $c$ starting in $q$. Since $q \in D_+(A)$, $c$ must hit $A$ in some $r \in A$, i.e. $r \ll p$. On the other hand, $p, r \in A$, which yields a contradiction.

Next we show $H_+(A) \subset \partial D_+(A)$: By definition, we have $H_+(A) \subset \overline{D_+(A)}$. If there was any $p \in H_+(A) \cap \mathring{D}_+(A)$, the intersection $I_+(p) \cap D_+(A)$ would not be empty, which contradicts $p \in H_+(A)$.
It remains to show $\partial D_+(A) \subset A \cup H_+(A)$: Assume there was any $p \in \partial D_+(A) \backslash \big(A \cup H_+(A)\big)$. Then, in particular, $p \in \overline{D_+(A)} \backslash A$, so (iii) implies $p \in I_+(A)$. On the other hand, $p \in \overline{D_+(A)} \backslash H_+(A)$, so there exists a $q \in I_+(p) \cap D_+(A)$, and $I_+(A) \cap I_-(q)$ is an open neighborhood of $p$. We complete the proof by showing that this neighborhood is contained in $D_+(A)$. Then $p$ would have to be an inner point, which contradicts $p \in \partial D_+(A)$.
Let $r \in I_+(A) \cap I_-(q)$ and $c$ a past-inextendible causal curve starting in $r$. Furthermore, let $\gamma$ be a timelike and past directed curve from $q$ to $r$, which necessarily stays in $I_+(A)$ due to $r \in I_+(A)$ and therefore fails to meet $A$ since achronality of $A$ demands $A \cap I_+(A) = \emptyset$. On the other hand, $q \in D_+(A)$ implies that $\gamma \cup c$ hits $A$, so $c$ has to hit $A$ and hence $r \in D_+(A)$.□

**Proposition 2.109.** *For any closed and acausal topological hypersurface $S \subset M$, we have*

*(i)* $H_+(S) = I_+(S) \cap \partial D_+(S) = \overline{D_+(S)} \backslash D_+(S)$.

*(ii)* $H_+(S) \cap S = \emptyset$.

*(iii)* *$H_+(S)$ is a closed and achronal topological hypersurface.*

*(iv)* *In each point of $H_+(S)$ starts a past-inextendible lightlike geodesic without any conjugated points, which runs entirely in $H_+(S)$.*

*Proof.* (i): We already know that $H_+(S) \subset \overline{D_+(S)} \subset S \cup I_+(S)$, where the last inclusion follows from Lemma 2.108 (iii). If there was any $p \in H_+(S) \cap D_+(S)$, $I_+(p)$ would hit $D(S)$ since according to Proposition 2.101, $D(S)$ is open, but due to achronality of $S$, we have $I_+(p) \cap D_-(S) = \emptyset$. Therefore, $I_+(p)$ has to meet $D_+(S)$, which contradicts $p \in H_+(S)$, and thus $H_+(S) \cap D_+(S) = \emptyset$, that is $H_+(S) \subset \partial D_+(S)$. Moreover, from $S \subset D_+(S)$ follows that also $H_+(S) \cap S = \emptyset$, so the inclusion, we started with, implies $H_+(S) \subset I_+(S)$. On the other hand, Lemma 2.108 (iv) provides

$$I_+(S) \cap \partial D_+(S) = \big(S \cup H_+(S)\big) \cap I_+(S) = H_+(S) \cap I_+(S) = H_+(S).$$

It remains to show $\overline{D_+(S)} \backslash D_+(S) \subset H_+(S)$ (we already proved the converse inclusion). Hence, for all $p \in \overline{D_+(S)} \backslash D_+(S)$, we show $I_+(p) \cap D_+(S) = \emptyset$. Let $q \in I_+(p)$ and $\gamma$ a timelike and past directed curve from $q$ to $p$. Since $p \notin S \cup I_-(S)$ and $p \notin D_+(S)$, $\gamma$ does not meet $S$ and there is a past-inextendible causal curve $c$ starting in $p$, which does not meet $S$ either. Therefore, $\gamma \cup c$ is past-inextendible causal curve, which starts in $q$ but does not meet $S$, that is $q \notin D_+(S)$.
(ii): Follows directly from (i) and achronality of $S$:

$$H_+(S) \cap S = \partial D_+(S) \cap \underbrace{I_+(S) \cap S}_{=\emptyset} = \emptyset.$$

(iii): Consider the past set $B := D_+(S) \cup I_-(S)$ (Definition 2.73), so Corollary 2.76 ensures that $\partial B$ is a topological hypersurface. On the other hand, (i) and $I_-(S) \cap I_+(S) = \emptyset$ imply

$$H_+(S) = \partial D_+(S) \cap I_+(S) = \partial B \cap I_+(S),$$

so $H_+(S)$ is an open subset of a topological hypersurface and thus a topological hypersurface on its own right. Achronality and closedness of $H_+(S)$ follows from Lemma 2.108.
(iv): For $p \in H_+(S)$, (i) ensures the existence of a past-inextendible causal curve $c$, which starts in $p$ and does not meet $S$. Due to Lemma 2.108, such a curve fails to be timelike, so $c$ can not be deformed to a timelike curve starting in $p$ and avoiding $S$. By Lemma 2.79 (ii), $c$ has to be a lightlike (pre-)geodesic without any conjugated points, and it remains to show that it does not leave $H_+(S)$.
If $c$ hit $D_+(S)$, it would hit $S$ as well, which yields a contradiction. If $c(s) \notin \overline{D_+(S)}$ for some parameter value $s$, we would find a past-inextendible timelike curve $\gamma$, which starts in $c(s)$ and does not meet $S$. Applying Lemma 2.79 (ii) to $c\big|_{[0,s]\cup\gamma}$ provides a timelike, which starts in $p$ and does not hit $S$, which contradicts $p \in \overline{D_+(S)}$. □

**Corollary 2.110.** *For any non-empty, closed and acausal topological hypersurface $S \subset M$, we have*

*(i) $S$ is a Cauchy hypersurface if and only if $H(S) = \emptyset$.*

*(ii) $S$ is a Cauchy hypersurface if every non-inextendible lightlike geodesic meets $S$.*

*Proof.* (i): From Proposition 2.101 we know that $D(S)$ is open. Moreover, $S = \overline{D_+(S)} \cap \overline{D_-(S)}$ since any non-inextendible timelike curve through some $p \in \big(\overline{D_+(S)} \cap \overline{D_-(S)}\big)\backslash S$ would have to meet $S$ in the future and the past of $p \notin S$, which contradicts achronality of $S$. It follows that

$$\begin{aligned}\partial D(S) &= \overline{D(S)}\backslash D(S) = \big(\overline{D_+(S)} \cup \overline{D_-(S)}\big)\backslash D(S) = \big(\overline{D_+(S)}\backslash D(S)\big) \cup \big(\overline{D_-(S)}\backslash D(S)\big)\\ &= \big(\overline{D_+(S)}\backslash D_+(S)\big) \cup \big(\overline{D_+(S)}\backslash D_-(S)\big) = H_+(S) \cup H_-(S) = H(S),\end{aligned}$$

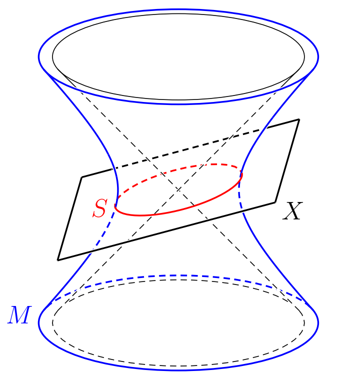
where the fourth equality is due to Proposition 2.109 (i). Recall that we $M$ is always assumed to be connected, which leads to

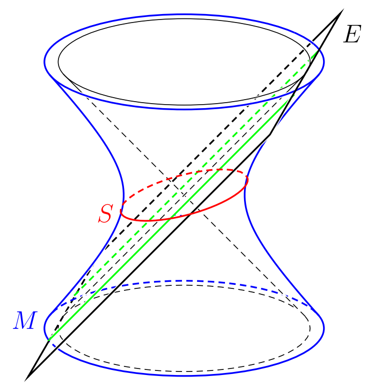
$$H(S) = \emptyset \quad \Longleftrightarrow \quad \partial D(S) = \emptyset \quad \Longleftrightarrow \quad D(S) = M,$$

and the last statement is obviously equivalent to $S$ being a Cauchy hypersurface.
(ii): Regarding (i), we show that for all $p \in H(S)$, there is a past-inextendible lightlike geodesic through $p$, which does not meet $S$. Without loss of generality, assume $p \in H_+(S)$. According to Proposition 2.109 (iv), there is a past-inextendible lightlike geodesic $c$, which runs entirely in $H_+(S)$ and does not hit $S$ since $H_+(S) \cap S = \emptyset$ due to Proposition 2.109 (ii). If the maximal extension of $c$ to a future-inextendible geodesic met $S$ in some $q \in S$, then $q \geq p$. On the other hand, $p \in H_+(S) \subset I_+(S)$ implies $q \in I_+(S) \cap S$, which contradicts achronality of $S$. □

**Example 2.111.**

1. Let $M = S_1^n(r)$ be the de-Sitter space and $S := M \cap X$, where $X \subset \mathbb{R}^{n+1}$ a spacelike hyperplane. It follows that $S$ is an acausal hypersurface.
   Moreover, due to the results of section 1.2, every lightlike geodesic is of the form $M \cap E$ for some degenerate hyperplane $E$. The intersection $X \cap E$ therefore is a spacelike straight line and thus hits $M$. It follows that $S$ is a Cauchy hypersurface and in particular, according to Corollary 2.99, $M$ is globally hyperbolic.

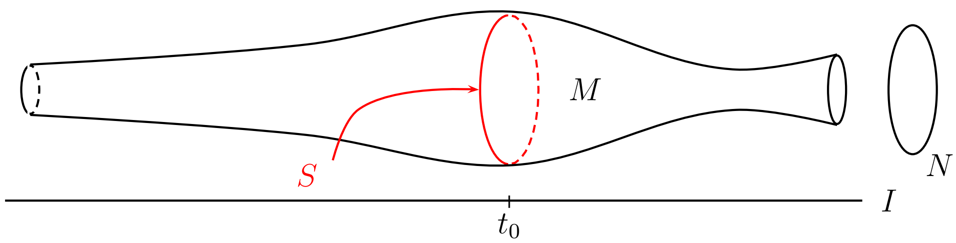



2. Let $(M, g)$ be a Robertson-Walker spacetime, i.e.

$$M = I \times N, \qquad g = -\,\mathrm{d}t \otimes \mathrm{d}t + f(t)^2 g_N$$

for some complete Riemannian manifold $(N, g_N)$ and $f \in C^\infty(I, \mathbb{R}_+)$, and consider $S := \{t_0\} \times N$ for some fixed $t_0 \in I$.




For any causal curve $c(s) = \big(t(s), \gamma(s)\big)$, we obtain

$$0 \geq g\big(c'(s), c'(s)\big) = -\big(t'(s)\big)^2 + f\big(t(s)\big)^2 \cdot g_N\big(\gamma'(s), \gamma'(s)\big),$$

and thus $|t'| \geq g \cdot \|\gamma'\|_N$. Note that $t'(s) = 0$ would imply $\gamma'(s) = 0$ and consequently $c'(s) = 0$, which is not causal. Therefore, $t' > 0$ on all of $I$ or $t' < 0$ on all of $I$, so $S$ is acausal since $t(s) = t_0$ for at most one $s$.
Now let $c = (t, \gamma)\colon (a, b) \to M$ be an inextendible lightlike geodesic and without loss of generality, let $t' > 0$. Choose $s_0 \in (a, b)$ and set $\delta := \max_J f$, where

$$J := \begin{cases} \big[t_0, t(s_0)\big], & t_0 \leq t(s_0) \\ \big[t(s_0), t_0\big], & t_0 \geq t(s_0) \end{cases} .$$

Due to Corollary 1.37, the function $t'(f \circ t) =: \eta > 0$ is constant and thus $t' = \frac{\eta}{f \circ t} \geq \frac{\eta}{\delta} > 0$ on $J$. For $\tau(s) := t(s_0) + \frac{\eta}{\delta}(s - s_0)$, we obtain $t(s_0) = \tau(s_0)$ and moreover, $t(s) \geq \tau(s)$ if $s \geq s_0$ and $t(s) \leq \tau(s)$ if $s \leq s_0$. Since $\tau(s_1) = t_0$ for $s_1 := s_0 + \frac{\delta}{\eta}\big(t_0 - t(s_0)\big)$ and $\gamma$ is a pregeodesic in the complete Riemannian manifold $N$, the equation $t(s) = t_0$ has a solution in the compact interval spanned by $s_0$ and $s_1$ as long as $t \in J$. It follows that $S$ is a Cauchy hypersurface and hence, $M$ is globally hyperbolic.

## 2.8 Hawking's singularity theorem

**Reminder:** For $M$ a semi-Riemannian manifold and $S \subset M$ a $p$-dimensional submanifold, for all $x \in S$, the mean curvature vector $H(x)$ is defined by

$$H(x) := \frac{1}{p} \sum_{j=1}^{p} \varepsilon_j II(e_j, e_j),$$

where $e_1, \ldots, e_p$ denotes a generalized orthonormal basis of $T_xS$ and $\varepsilon_j := g(e_j, e_j)$.

**Theorem 2.112 (Hawking's singularity theorem).** *Let $M$ be an $n$-dimensional connected and time-oriented Lorentzian manifold with*

$$\mathrm{ric}(X, X) \geq 0$$

*for all timelike $X \in TM$. Let $S \subset M$ be a Cauchy hypersurface with mean curvature vector field $H$ and future directed unit normal field $\nu$. Assume that there exists some $\beta > 0$ such that*

$$\langle H, \nu \rangle \geq \beta.$$

*Then the length of every timelike and future directed curve starting in $S$ is bounded by $\frac{1}{\beta}$.*

**Physical interpretation:** The Lorentzian manifold $M$ models the spacetime and the Cauchy hypersurface $S$ the present spacelike universe. Einstein's field equations in dimension $4$ read

$$8\pi T = \mathrm{ric} - \frac{1}{2}\mathrm{scal} \cdot g,$$

where $T$ is the energy-momentum-tensor. This implies $8\pi \mathrm{tr}_g(T) = \mathrm{scal} - \frac{1}{2} \cdot 4\mathrm{scal} = -\mathrm{scal}$, so the field equations can be reformulated via

$$8\pi T = \mathrm{ric} + 4\pi \mathrm{tr}_g(T) \cdot g,$$

which leads to

$$\mathrm{ric}(X, X) \geq 0 \qquad \Longleftrightarrow \qquad T(X, X) \geq \frac{1}{2}\mathrm{tr}_g(T) g(X, X)$$

for all timelike vectors $X$. This inequality is known as the *strong energy inequality*, where $T(X,X)$ is interpreted as the energy density measured by an observer, whose world line has the tangent vector $X$. Furthermore, the condition $\langle H,\nu\rangle \geq \beta$ stands for a spacelike universe, which contracts at a rate at least $\beta$, so Hawking's theorem states that the time, such a universe exists, is at most $\frac{1}{\beta}$, which therefore stands for the time, a big crunch singularity would occur at the latest.
For the proof, we need the following proposition:

**Proposition 2.113.** *Let $M$ be a Lorentzian manifold, $c\colon [a,b] \to M$ a timelike geodesic and $c_s$ a smooth variation of $c$ by timelike curves with variational field $V := \frac{\partial c_s}{\partial s}\big|_{s=0}$ and acceleration field $A := \frac{\nabla}{\partial s}\frac{\partial c_s}{\partial s}\big|_{s=0}$. Then we have*

1. $\frac{\mathrm{d}}{\mathrm{d}s}L[c_s]\big|_{s=0} = \left\langle V(a), \frac{\dot{c}(a)}{|\dot{c}(a)|}\right\rangle - \left\langle V(b), \frac{\dot{c}(b)}{|\dot{c}(b)|}\right\rangle$,
2. $\frac{\mathrm{d}^2}{\mathrm{d}s^2}L[c_s]\Big|_{s=0} = \left\langle A(a), \frac{\dot{c}(a)}{|\dot{c}(a)|}\right\rangle - \left\langle A(b), \frac{\dot{c}(b)}{|\dot{c}(b)|}\right\rangle$

$$-\int_a^b \frac{1}{|\dot{c}|}\left(\left\langle R(V,\dot{c})V,\dot{c}\right\rangle + \left\langle \frac{\nabla V}{\partial t}, \frac{\nabla V}{\partial t}\right\rangle + \left\langle \frac{\nabla V}{\partial t}, \frac{\dot{c}}{|\dot{c}|}\right\rangle^2\right)\mathrm{d}t.$$

*Proof.* Let $c\colon [a,b] \to M$ be a timelike geodesic and $c_s$ a smooth variation of $c$ with variational field $V$ and acceleration field $A$.
(1) For $V_s = \frac{\partial c_s}{\partial s}$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}L[c_s] = \frac{\mathrm{d}}{\mathrm{d}s}\int_a^b \sqrt{-\langle \dot{c}_s, \dot{c}_s\rangle}\,\mathrm{d}t = \int_a^b \frac{-2\left\langle \frac{\nabla \dot{c}_s}{\partial s}, \dot{c}_s\right\rangle}{2\sqrt{-\langle \dot{c}_s,\dot{c}_s\rangle}}\,\mathrm{d}t$$

$$= -\int_a^b \left\langle \frac{\nabla}{\partial s}\frac{\partial c_s}{\partial t}, \frac{\dot{c}_s}{|\dot{c}_s|}\right\rangle \mathrm{d}t = -\int_a^b \left\langle \frac{\nabla}{\partial t}\frac{\partial c_s}{\partial s}, \frac{\dot{c}_s}{|\dot{c}_s|}\right\rangle \mathrm{d}t = -\int_a^b \left\langle \frac{\nabla V_s}{\partial t}, \frac{\dot{c}_s}{|\dot{c}_s|}\right\rangle \mathrm{d}t,$$

which, for $s=0$, provides the claim:

$$\frac{\mathrm{d}}{\mathrm{d}s}L[c_s]\Big|_{s=0} = -\int_a^b \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\dot{c}}{|\dot{c}|}\right\rangle \mathrm{d}t$$

$$= -\int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}t}\left\langle V, \frac{\dot{c}}{|\dot{c}|}\right\rangle - \left\langle V, \underbrace{\frac{\nabla}{\partial t}\frac{\dot{c}}{|\dot{c}|}}\right\rangle_{=0}\right)\mathrm{d}t = -\left\langle V, \frac{\dot{c}}{|\dot{c}|}\right\rangle\Big|_a^b.$$

(2) The second claim follows directly from direct calculation of the second variation:

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}L[c_s]\bigg|_{s=0} = -\frac{\mathrm{d}}{\mathrm{d}s}\int_a^b \left\langle \frac{\nabla V_s}{\partial t}, \frac{\dot c_s}{|\dot c_s|}\right\rangle \mathrm{d}t\bigg|_{s=0}$$

$$= -\int_a^b \left( \left\langle \frac{\nabla}{\partial s}\frac{\nabla V_s}{\partial t}\bigg|_{s=0}, \frac{\dot c}{|\dot c|}\right\rangle + \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla}{\partial s}\bigg|_{s=0}\frac{\dot c_s}{|\dot c_s|}\right\rangle \right)\mathrm{d}t$$

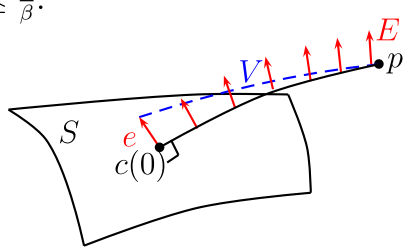
$$= -\int_a^b \left( \left\langle R(V,\dot c)V + \frac{\nabla}{\partial t}\underbrace{\frac{\nabla V_s}{\partial s}\bigg|_0}_{=A}, \frac{\dot c}{|\dot c|}\right\rangle + \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{1}{|\dot c|^2}\left(|\dot c|\frac{\nabla \dot c_s}{\partial s}\bigg|_0 - \frac{\partial |\dot c_s|}{\partial s}\bigg|_0 \dot c\right)\right\rangle \right)\mathrm{d}t$$

$$= -\int_a^b \left( \left\langle R(V,\dot c)V, \frac{\dot c}{|\dot c|}\right\rangle + \frac{\mathrm{d}}{\mathrm{d}t}\left\langle A, \frac{\dot c}{|\dot c|}\right\rangle - \left\langle A, \underbrace{\frac{\nabla}{\mathrm{d}t}\frac{\dot c}{|\dot c|}}_{=0}\right\rangle + \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{1}{|\dot c|} + \frac{\left\langle \frac{\nabla V}{\mathrm{d}t}, \dot c\right\rangle \dot c}{|\dot c|^3}\right\rangle \right)\mathrm{d}t$$

$$= -\left\langle A, \frac{\dot c}{|\dot c|}\right\rangle\bigg|_a^b - \int_a^b \left( \left\langle R(V,\dot c)V, \frac{\dot c}{|\dot c|}\right\rangle + \frac{1}{|\dot c|}\left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t}\right\rangle + \frac{1}{|\dot c|^3}\left\langle \frac{\nabla V}{\mathrm{d}t}, \dot c\right\rangle^2 \mathrm{d}t. \quad \square$$

*Proof of Theorem 2.112.* Let $\gamma$ be a future directeed timelike curve from $S$ to some $p$. It follows that $p$ is an element of $I_+(S) = D_+(S)\backslash S$, so due to Theorem 2.105, there is a timelike geodesic $c\colon [0,b] \to M$ with $c(0) \in S, \dot c(0) \perp S, c(b) = p$ and $L[c] = \tau(S,p)$. Moreover, without loss of generality, we assume $c$ to be parametrized with respect to proper time, i.e. $|\dot c| = 1$ and thus $L[c] = b$. Therefore, it remains to show $b \leq \frac{1}{\beta}$.

Let $e \in T_{c(0)}S$ be a unit vector and $E$ the spacelike, parallel unit vector field along $c$ given by $E(0) = e$. Furthermore, let $c_s$ be a variation of $c$ with variational field $V(t) = \left(1 - \frac{t}{b}\right)E(t), c_s(0) \in S$ and $c_s(0) \in S$. Since $c$ is the longest connection of $S$ with $p$, by Proposition 2.113, we have




$$0 \geq \frac{\mathrm{d}^2}{\mathrm{d}s^2}L[c_s]\bigg|_{s=0} = \langle A(0), \dot c(0)\rangle - 0$$

$$- \int_0^b \left[ \left(1 - \frac{t}{b}\right)^2 \langle R(E,\dot c)E, \dot c\rangle + \left\langle -\frac{1}{b}E, -\frac{1}{b}E\right\rangle + \overbrace{\left\langle -\frac{1}{b}E, \dot c\right\rangle^2}^{=0} \right]\mathrm{d}t$$

$$= \left\langle \frac{\nabla}{\partial s}\frac{\partial c_s}{\partial s}(0), \dot c(0)\right\rangle - \int_0^b \left[ \left(1 - \frac{t}{b}\right)^2 \langle R(E,\dot c)E, \dot c\rangle + \frac{1}{b^2}\right]\mathrm{d}t$$

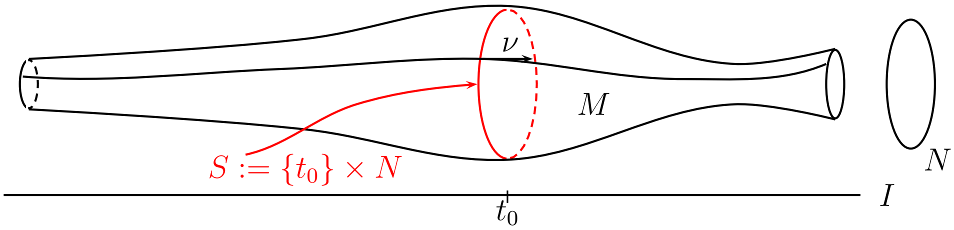
$$= \langle II(V,V), \nu\rangle + \int_0^b \left[ \left(1 - \frac{t}{b}\right)^2 \langle R(\dot c, E)E, \dot c\rangle \;\mathrm{d}t - \frac{1}{b}.$$

To an orthonormal basis $e_1, \dots, e_{n-1}$ of $T_{c(0)}S$, we obtain corresponding $E_1, \dots, E_{n-1}$, and summation yields

$$0 \geq \langle (n-1)H, \nu \rangle + \int_0^b \left(1 - \frac{t}{b}\right)^2 \underbrace{\mathrm{ric}(\dot c, \dot c)}_{\geq 0} - \frac{n-1}{b} \geq (n-1)\beta + 0 - \frac{n-1}{b},$$

that is $b \leq \frac{1}{\beta}$. □

**Example 2.114.** 1) Let $M$ be an $(n+1)$-dimensional Robertson-Walker spacetime.

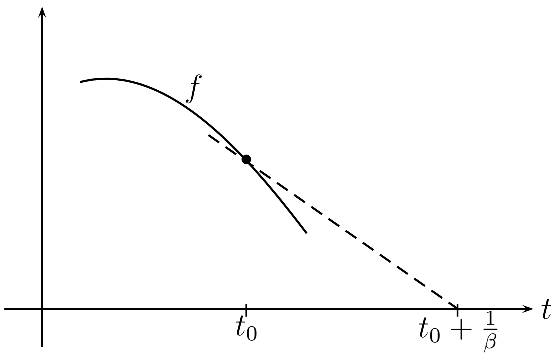



We already proved that $\mathrm{ric}(\nu, \nu) = -n\frac{f''}{f}$, where $\nu := \frac{\partial}{\partial t}$. For the proof of Hawking's singularity theorem, we employed the assumption $\mathrm{ric}(X, X) \geq 0$ merely for $x := \dot c(t)$, where stands $c$ stands for a geodesic that starts in $S$ and orthogonally to it.

Here, we actually have $X = \nu$, so the assumption $\mathrm{ric}(\nu, \nu) \geq 0$ is equivalent to $f'' \leq 0$, i.e. $f$ being concave. The shape operator of $S := \{t_0\} \times N$ in $M$ with respect to $\nu$ is given by

$$W = \frac{f'(t_0)}{f(t_0)} \cdot \mathrm{id},$$

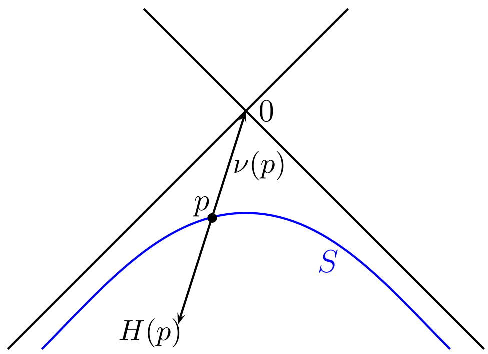
and hence, $H = \frac{f'(t_0)}{f(t_0)}\nu$ and $\langle H, \nu \rangle \geq \beta$ if and only if $\frac{f'(t_0)}{f(t_0)} \leq -\beta$.




2) Let $M$ be $n$-dimensional Minkowski space and thus $\mathrm{ric} = 0$. Consider $S := -H^{n-1}(r)$, where $H^{n-1}(r) := \left\{x \in M \mid \langle x, x \rangle = -r^2,\ x^0 > 0\right\}$. Then for any $p \in S$, we have $H(p) = -\frac{p}{r^2}$ and $\nu = -\frac{p}{r}$, and therefore

$$\big\langle H(p), \nu(p) \big\rangle = -\frac{1}{r^3} \underbrace{\langle p, p \rangle}_{=-r^2} = \frac{1}{r} =: \beta.$$

On the other hand, we know that the maximal timelike geodesics that start in $S$ have infinite length! The reason for that is that $S$ is not a Cauchy hypersurface in $\mathbb{R}^n_{\mathrm{Mink}}$, but it is in $D(S) = I_-(0)$, where the maximal timelike geodesics that start in $S$ indeed have length $r = \frac{1}{\beta}$. In this example, the estimate in Hawking's singularity theorem is sharp.



## 2.9 Penrose's singularity theorem

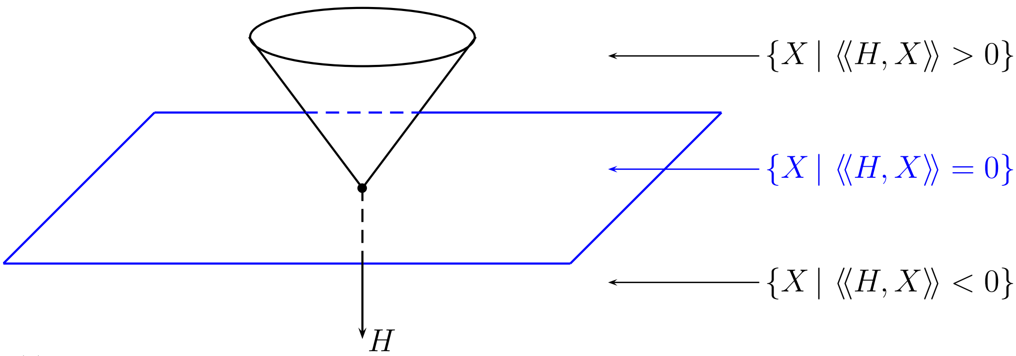
**Lemma 2.115.** *Let $H \in \mathbb{R}^n$ and $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ the Minkowski product. Then the following statements are equivalent:*

*(i) For all future directed and lightlike $X \in \mathbb{R}^n$, we have $\langle\!\langle H, X \rangle\!\rangle > 0$.*

*(ii) For all future directed and causal $X \in \mathbb{R}^n$, we have $\langle\!\langle H, X \rangle\!\rangle > 0$.*

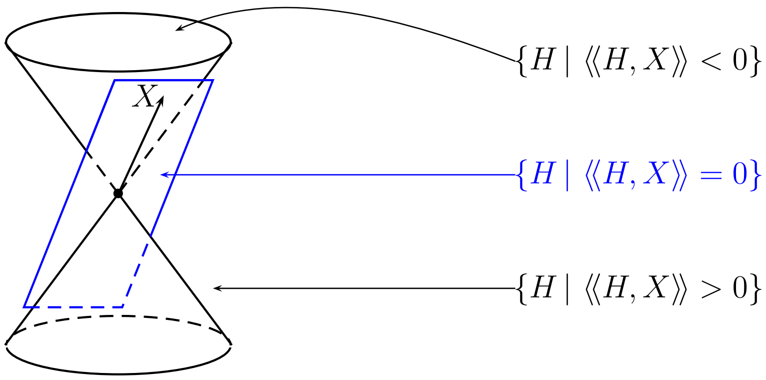
*(iii) $H$ is past directed and timelike.*

*Proof.* $(iii) \Rightarrow (ii)$: After applying some some time-orientation-preserving Lorentz transformation, without loss of generality, we may assume $H = -ce_0$ for some $c > 0$.




$(ii) \Rightarrow (i)$: trivial.

$(i) \Rightarrow (iii)$:




$$\Longrightarrow \quad \bigcap_{X \in C_+(0) \setminus \{0\}} \big\{ H \bigm| \langle H, X \rangle > 0 \big\} = I_-(0). \qquad \square$$

**Definition 2.116.** A connected time-oriented Lorentzian manifold is called **timelike future-complete** if for all future directed timelike vectors $X \in TM$, the geodesic $t \mapsto \exp_{\pi(X)}(tX)$ is defined on all of $[0, \infty)$. Similarly, one defines **timelike past-complete** as well as **lightlike future- and past-complete** Lorentzian manifolds, respectively.

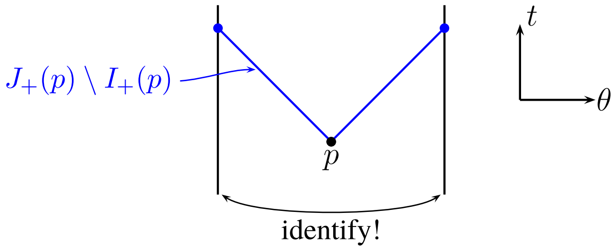
**Example 2.117.** Let $M := D\big(-H^{n-1}\big) = I_-(0) \subset \mathbb{R}^n_{\text{Mink}}$. The Lorentzian manifold $M$ is not timelike or lightlike future-complete but it is timelike and lightlike past-complete.

**Definition 2.118.** A closed and achronal subset $A \subset M$ is called **future-trapped** or **past-trapped** if $J_+(A)\backslash I_+(A)$ or $J_-(A)\backslash I_-(A)$ is compact, respectively.

**Example 2.119.** For

$$(M,g) := \big(\mathbb{R}\times S^1, -\mathrm{d}t^2 + \mathrm{d}\theta^2\big),$$

the subset $A := \{p\}$ is both, future- and past-trapped.




**Remark 2.120.**

1. Recall that achronality implies $A \cap I_+(A) = \emptyset$, that is $A \subset J_+(A)\backslash I_+(A)$, and hence, future-trapped subsets have to be compact.

2. For arbitrary $A \subset M$, the subset $J_+(A)\backslash I_+(A)$ is achronal since

$$p \in I_+\big(J_+(A)\backslash I_+(A)\big) \subset I_+\big(J_+(A)\big) = I_+(A) \qquad \Longrightarrow \qquad p \notin J_+(A)\backslash I_+(A),$$

   and therefore $I_+\big(J_+(A)\backslash I_+(A)\big) \cap \big(J_+(A)\backslash I_+(A)\big) = \emptyset$.

**Lemma 2.121.** *Let $M$ be an $n$-dimensional Lorentzian manifold, $p \in M$, $\ell \in T_pM$ lightlike and $e_1, \dots, e_{n-2}$ spacelike and orthonormal with $e_j \perp \ell$ for all $j$. Then we have*

$$\operatorname{ric}(\ell,\ell) = \sum_{j=1}^{n-2} \big\langle R(\ell, e_j)e_j, \ell\big\rangle.$$

*Proof.* Consider some spacelike $e_{n-1}$ and some timelike $e_n$ such that $e_{n+1} + e_n$ is a multiple of $\ell$ and $e_1, \dots, e_n$ is a generalized orthonormal basis of $T_pM$. By definition, we obtain

$$\operatorname{ric}(\ell,\ell) = \sum_{j=1}^{n-1} \big\langle R(\ell, e_j)e_j, \ell\big\rangle - \big\langle R(\ell, e_n)e_n, \ell\big\rangle,$$

so we have to show $\big\langle R(\ell, e_{n-1})e_{n-1}, \ell\big\rangle = \big\langle R(\ell, e_n)e_n, \ell\big\rangle$. Note that $e_{n-1} + e_n$ being a multiple of $\ell$ implies

$$\big\langle R(\ell, e_{n-1} + e_n)e_{n-1}, \ell\big\rangle = 0, \qquad \big\langle R(\ell, e_{n-1} + e_n)e_n, \ell\big\rangle = 0,$$

and the claim follows from subtracting the second equation from the first one. $\square$

**Proposition 2.122.** *Let $M$ be a Lorentzian manifold, $P \subset M$ a spacelike submanifold of codimension 2 and mean curvature vector field $H$. Furthermore, let $c\colon [0,b] \to M$ be a lightlike geodesic that starts in some $p \in P$ such that $\dot c(0) \in N_pP$. Moreover, we assume*

*(i)* $\operatorname{ric}\big(\dot c(t), \dot c(t)\big) > 0$ *for all* $t \in [0,b]$,

*(ii)* $\big\langle H(p), \dot c(0)\big\rangle \geq \frac{1}{b}$.

*Then $c$ has a focal point in* $(0,b]$.

*Proof.* Assume that $c$ has no focal point in $(0,b]$.
a) Let $e_1,\dots,e_{n-2}$ be an orthonormal basis of $T_pP$, to which we consider Jacobi fields $J_i$ along $c$ determined by the initial values $J_i(0) = e_i$ and $\frac{\nabla J_i}{\mathrm{d}t}(0) = \widetilde{I\!I}\big(e_i, \dot c(0)\big)$. In addition, let $J_0(t) := t\cdot \dot c(t)$ be the Jacobi field given by $J_0(0) = 0$ and $\frac{\nabla J}{\mathrm{d}t}(0) = \dot c(0)$.

b) We show that $J_0(t),\dots,J_{n-2}(t)$ constitutes a basis of $c(t)^\perp$ for all $t \in [0,b]$. Note that $J_0(t) \perp \dot c(t)^\perp$ since $c$ is lightlike. For all $i = 1,\dots,n-2$, we have

$$\begin{cases} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\big\langle J_i, \dot c\big\rangle & = \big\langle \frac{\nabla^2}{\mathrm{d}t^2} J_i, \dot c\big\rangle = \big\langle R(\dot c, J_i)\dot c, \dot c\big\rangle = 0, \\ \big\langle J_i(0), \dot c(0)\big\rangle & = \langle e_i, \dot c(0)\rangle = 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}\big\langle J_i(t), \dot c(t)\big\rangle\big|_{t=0} & = \big\langle \frac{\nabla J_i}{\mathrm{d}t}(0), \dot c(0)\big\rangle = \Big\langle \widetilde{I\!I}\big(e_i, \dot c(0)\big), \dot c(0)\Big\rangle = 0, \end{cases}$$

and thus $\big\langle J_i, \dot c\big\rangle = 0$ due to well-posedness of the initial value problem, that is $J_i(t) \in \dot c(t)^\perp$ for all $t$. It remains to show linear independence of $J_0(t),\dots,J_{n-2}(t)$. Assume it was not for some $t \in (0,b]$, i.e. $\sum_{i=0}^{n-2} \alpha_i J_i(t) = 0$ for $\alpha_0,\dots,\alpha_{n-2} \in \mathbb{R}$ not all equal to zero. This provides a non-trivial Jacobi field $J := \sum_{i=0}^{n-2} \alpha_i J_i$ satisfying

$$J(0) = \sum_{i=0}^{n-2} \alpha_i e_i \in T_pP, \qquad J(t) = 0,$$

$$\begin{aligned} \tan\Big(\frac{\nabla J}{\mathrm{d}t}(0)\Big) &= \tan\Big(\sum_{i=1}^{n-2} \alpha_i \widetilde{I\!I}\big(e_i, \dot c(0)\big) + \alpha_0 \dot c(0)\Big) \\ &= \tan\Big(\sum_{i=1}^{n-2} \alpha_i \widetilde{I\!I}\big(e_i, \dot c(0)\big)\Big) = \tan\Big(\widetilde{I\!I}\big(J(0), \dot c(0)\big)\Big), \end{aligned}$$

so $t$ would be a focal point, which contradicts the assumption.

c) We show $\big\langle \frac{\nabla J_i}{\mathrm{d}t}, J_j\big\rangle = \big\langle J_i, \frac{\nabla J_j}{\mathrm{d}t}\big\rangle$ for all $i,j$. The symmetries of $R$ provide

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\bigg(\Big\langle \frac{\nabla J_i}{\mathrm{d}t}, J_j\Big\rangle - \Big\langle J_i, \frac{\nabla J_j}{\mathrm{d}t}\Big\rangle\bigg) &= \Big\langle \frac{\nabla^2 J_i}{\mathrm{d}t^2}, J_j\Big\rangle - \Big\langle J_i, \frac{\nabla^2 J_j}{\mathrm{d}t^2}\Big\rangle \\ &= \big\langle R(\dot c, J_i)\dot c, J_j\big\rangle - \big\langle J_i, R(\dot c, J_j)\dot c\big\rangle = 0, \end{aligned}$$

and hence, $\left\langle \frac{\nabla J_i}{\mathrm{d}t}, J_j \right\rangle - \left\langle J_i, \frac{\nabla J_j}{\mathrm{d}t} \right\rangle$ is constant. Furthermore, by symmetry of $II$, we obtain

$$\left\langle \frac{\nabla J_i}{\mathrm{d}t}(0), J_j(0) \right\rangle - \left\langle J_i(0), \frac{\nabla J_j}{\mathrm{d}t}(0) \right\rangle = \left\langle \widetilde{II}\big(e_i, \dot{c}(0), e_j\right\rangle - \left\langle e_i, \widetilde{II}\big(e_j, \dot{c}(0)\big)\right\rangle$$
$$= -\left\langle II(e_i, e_j), \dot{c}(0) \right\rangle + \left\langle e_j, e_i, \dot{c}(0) \right\rangle = 0,$$

for all $i, j \geq 1$, and moreover,

$$\left\langle \frac{\nabla J_0}{\mathrm{d}t}(0), J_j(0) \right\rangle - \left\langle J_0(0), \frac{\nabla J_j}{\mathrm{d}t}(0) \right\rangle = \langle \dot{c}(0), e_j \rangle - \left\langle 0, \frac{\nabla J_j}{\mathrm{d}t}(0) \right\rangle = 0$$

for all $j = 0, \ldots, n-2$, which proves the claim.

d) Let $V$ be a smooth vector field along $c$ with $V(t) \perp \dot{c}(t)$ for all $t$ with $V(0) \in T_p P$ and $V(b) = 0$. We show that

$$\int_0^b \left( \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t} \right\rangle - \left\langle R\big(\dot{c}, V\big)V, \dot{c} \right\rangle \right) \mathrm{d}t - \left\langle \dot{c}(0), II\big(V(0), V(0)\big) \right\rangle \geq 0$$

and equality if and only if $V$ is tangential to $c$.
Due to b), there are smooth functions $f_i \colon (0, b] \to \mathbb{R}$ such that $V(t) = \sum_{i=0}^{n-2} f_i(t) J_i(t)$ for all $t \in (0, b]$, and we introduce the vector fields $X := \sum_{i=0}^{n-2} \dot{f}_i \cdot J_i$ and $Y := \sum_{i=0}^{n-2} f_i \cdot \frac{\nabla J_i}{\mathrm{d}t}$. With them, we have $\frac{\nabla V}{\mathrm{d}t} = X + Y$, and therefore,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \langle V, Y \rangle &= \left\langle \frac{\nabla V}{\mathrm{d}t}, Y \right\rangle + \left\langle V, \frac{\nabla Y}{\mathrm{d}t} \right\rangle = \langle X + Y, Y \rangle + \left\langle V, \frac{\nabla Y}{\mathrm{d}t} \right\rangle \\
&= \langle X + Y, Y \rangle + \left\langle V, \sum_{i=0}^{n-2} \dot{f}_i \cdot \frac{\nabla J_i}{\mathrm{d}t} \right\rangle + \left\langle V, \sum_{i=0}^{n-2} f_i \cdot \frac{\nabla^2 J_i}{\mathrm{d}t^2} \right\rangle \\
&= \langle X + Y, Y \rangle + \sum_{i,j=0}^{n-2} \left\langle f_j \cdot J_j, \dot{f}_i \cdot \frac{\nabla J_i}{\mathrm{d}t} \right\rangle + \left\langle V, \sum_{i=0}^{n-2} f_i \cdot R\big(\dot{c}, J_i\big)\dot{c} \right\rangle \\
&= \langle X + Y, Y \rangle + \sum_{i,j=0}^{n-2} f_j \dot{f}_i \left\langle J_j, \frac{\nabla J_i}{\mathrm{d}t} \right\rangle + \left\langle V, R(\dot{c}, V)\dot{c} \right\rangle \\
&\overset{\text{c)}}{=} \langle X + Y, Y \rangle + \langle X, Y \rangle - \left\langle R(\dot{c}, V)\dot{c}, V \right\rangle \\
&= \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t} \right\rangle - \langle X, X \rangle - \left\langle R(\dot{c}, V)\dot{c}, V \right\rangle.
\end{aligned}$$

For small $\varepsilon > 0$, integration yields

$$\int_\varepsilon^b \langle X, X \rangle \,\mathrm{d}t = \int_\varepsilon^b \left( \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t} \right\rangle - \left\langle R(\dot{c}, V)V, \dot{c} \right\rangle - \frac{\mathrm{d}}{\mathrm{d}t} \left\langle V, Y \right\rangle \right) \mathrm{d}t$$

$$= \int_\varepsilon^b \left( \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t} \right\rangle - \langle R(\dot c, V)V, \dot c\rangle \right) \mathrm{d}t - \langle V(\varepsilon), Y(\varepsilon)\rangle.$$

Note that in addition to $J_0(t), \ldots, J_{n-2}(t)$, we find a further basis $\frac{1}{t}J_0(t) = \dot c(t), \ldots, J_{n-2}(t)$ of $\dot c(t)^\perp$, which, however, is also the case for $t = 0$. Hence, $tf_0(t), f_1(t), \ldots, f_{n-2}(t)$ and continuously extendible to $t = 0$, and furthermore, $V(0) = \sum_{i=1} f_i(0)J_i(0) = \sum_{i=1} f_i(0)e_i$ since $V(0) \in T_pP$. It follows that

$$\begin{aligned}
\langle V(\varepsilon), Y(\varepsilon)\rangle &= \left\langle V(\varepsilon), f_0(\varepsilon)\dot c(\varepsilon) + \sum_{i=1}^{n-2} f_i(\varepsilon)\cdot \frac{\nabla J_i}{\mathrm{d}t}(\varepsilon) \right\rangle = \left\langle V(\varepsilon), \sum_{i=1}^{n-2} f_i(\varepsilon)\cdot \frac{\nabla J_i}{\mathrm{d}t}(\varepsilon) \right\rangle \\
&\xrightarrow{\varepsilon\downarrow 0} \left\langle V(0), \sum_{i=1}^{n-2} f_i(0)\cdot \frac{\nabla J_i}{\mathrm{d}t}(0) \right\rangle = \left\langle V(0), \sum_{i=1}^{n-2} f_i(0)\cdot \widetilde{I\!I}\big(e_i, \dot c(0)\big) \right\rangle \\
&= \left\langle V(0), \widetilde{I\!I}\big(V(0), \dot c(0)\big) \right\rangle = -\left\langle I\!I\big(V(0), V(0)\big), \dot c(0)\right\rangle.
\end{aligned}$$

We just proved that

$$\int_0^b \left( \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t} \right\rangle - \langle R(\dot c, V)V, \dot c\rangle \right) \mathrm{d}t + \left\langle I\!I\big(V(0), V(0)\big), \dot c(0)\right\rangle = \lim_{\varepsilon\to 0} \int_\varepsilon^b \langle X, X\rangle \,\mathrm{d}t.$$

Since $X(t)$ is defined as a linear combination of $J_1(t), \ldots, J_{n-2}(t)$, we have $X(t)\perp \dot c(t)$, i.e. $X(t)$ fails to be timelike. Therefore, $\langle X(t), X(t)\rangle \geq 0$ for all $t \in (0, b]$, which is the desired inequality. Clearly, equality holds if and only if $\langle X, X\rangle = 0$, i.e. $X(t)$ is lightlike for all $t$ and consequently, $\dot f_1 = \ldots = \dot f_{n-2} = 0$. Because $f_i(b) = 0$ for all $i$, this is equivalent to $f_1 = \ldots = f_{n-2} = 0$, which leads to $V(t) = f_0(t)J_0(t) = tf_0(t)\dot c(t)$, so $V$ is tangential to $c$.

e) For some unit vector $e \in T_pP$, let $E$ denote the corresponding parallel vector field along $c$ with $E(0) = e$ and consider $V(t) := \big(1 - \frac{t}{b}\big)E(t)$. Then $V$ satisfies the conditions in d) but is not tangential on $c$, and hence,

$$\begin{aligned}
0 &< \int_0^b \left( \left\langle \frac{\nabla V}{\mathrm{d}t}, \frac{\nabla V}{\mathrm{d}t} \right\rangle - \langle R(\dot c, V)V, \dot c\rangle \right) \mathrm{d}t - \left\langle \dot c(0), I\!I\big(V(0), V(0)\big)\right\rangle \\
&= \int_0^b \left( \frac{1}{b^2} - \left(1 - \frac{t}{b}\right)^2 \langle R(\dot c, E)E, \dot c\rangle \right) \mathrm{d}t - \langle \dot c(0), I\!I(e, e)\rangle \\
&= \frac{1}{b} - \int_0^b \left(1 - \frac{t}{b}\right)^2 \langle R(\dot c, E)E, \dot c\rangle \,\mathrm{d}t - \langle \dot c(0), I\!I(e, e)\rangle.
\end{aligned}$$

Considering $e := e_i$ and summing over all $i$, Lemma 2.121 provides the desired contradiction:

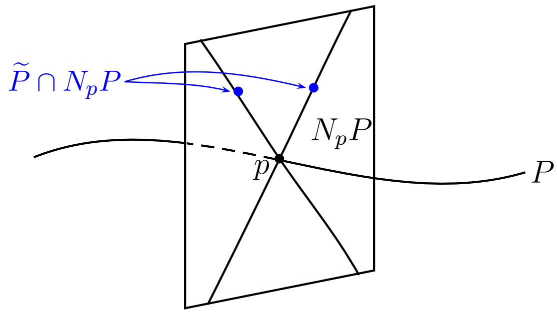
$$0 < \frac{n-2}{b} - \int_0^b \left(1 - \frac{t}{b}\right)^2 \underbrace{\mathrm{ric}(\dot c, \dot c)}_{\geq 0} \,\mathrm{d}t - (n-2) \underbrace{\langle \dot c(0), H(p)\rangle}_{\geq \frac{1}{b}} \leq 0. \qquad \square$$

**Proposition 2.123.** *Let $M$ be a connected, time-oriented and lightlike future-complete Lorentzian manifold with $\operatorname{ric}(X, X) \geq 0$ for all lightlike $X \in TM$. Furthermore, let $P \subset M$ be a compact, achronal and spacelike submanifold of codimension 2. Then $P$ is future-trapped if the mean curvature vector field $H$ of $P$ is timelike past directed.*

*Proof.* a) For some arbitrary Riemannian metric $h$ on $M$, set

$$\widetilde{P} := \big\{ X \in TM \bigm| X \text{ is lightlike future directed and } h(X, X) = 1 \big\}.$$

The footpoint map $\pi \colon NP \to P$ turns $\widetilde{P}$ into a two-fold cover of $P$. Moreover, $\widetilde{P}$ is compact.




b) For $X \in \widetilde{P}$, we have $\langle\!\langle H\big(\pi(X)\big), X \rangle\!\rangle > 0$ by Lemma 2.115, and due to compactness, we find some $b > 0$ such that $\langle\!\langle H\big(\pi(X)\big), X \rangle\!\rangle > \frac{1}{b}$ for all $X \in \widetilde{P}$.Since $M$ is lightlike future-complete, $t \mapsto c_X(t) := \exp(tX)$ is well-defined on $[0, \infty)$, i.e. on $[0, b]$, in particular. Now Proposition 2.122 ensures that $c_X$ has a focal point in $(0, b]$.

c) Let $q \in J_+(P) \backslash I_+(P)$. By Theorem 2.44, there is a lightlike and future directed geodesic $c$ from $P$ to $q$ without any focal points before $q$ and such that $\dot{c}(0) \in NP$. Clearly, we have $c = c_X$ and thus, $c_X(t) = q$ for some $X \in \widetilde{P}$ and $t \in [0, b]$. Therefore, $J_+(P) \backslash I_+(P)$ for the compact subset $K := \big\{ tX \bigm| 0 \leq t \leq b,\ X \in \widetilde{P} \big\}$, so $\exp(K)$ is compact as well.

d) For $(q_n)_{n\in\mathbb{N}} \subset J_+(P) \backslash I_+(P)$ and after maybe restricting to some subsequence, we have $q_n \to q_\infty \in \exp_( K) \subset J_+(P)$. Suppose $q_\infty \in I_+(P)$, that is $q_n \in I_+(P)$ for all $n$ large enough since $I_+(P)$ is open, which contradicts the assumption on $(q_n)_{n\in\mathbb{N}}$. Hence, $q_\infty \in J_+(P) \backslash I_+(P)$, i.e. $J_+(P) \backslash I_+(P)$ is compact. $\square$

**Lemma 2.124.** *Let $M$ be a globally hyperbolic Lorentzian manifold and $K \subset M$ compact. Then $J_\pm(K)$ is closed.*

*Proof.* For $(p_i)_{i\in\mathbb{N}} \subset J_+(K)$ with $p_i \to p \in M$, we show $p \in J_+(K)$. Choose $(q_i)_{i\in\mathbb{N}} \subset K$ with $q_i \leq p_i$, so after maybe passing to some subsequence, $(q_i)_{i\in\mathbb{N}}$ converges to some $q \in K$. Recall that $\leq$ is a closed relation by Proposition 2.92, so $q \leq p$, and hence, $p \in J_+(q) \subset J_+(K)$. The proof for $J_-(K)$ is similar. $\square$

**Theorem 2.125 (Penrose's singularity theorem).** *Let $M$ be a conneceted and time-oriented Lorentzian manifold with* $\mathrm{ric}(X,X) \geq 0$ *for all lightlike $X \in TM$. Furthermore, assume that there is a non-compact Cauchy hypersurface $S \subset M$ and a non-empty, compact, spacelike and achronal submanifold $P \subset M$ of codimension 2 with past directed and timelike mean curvature vector field. Then $M$ fails to be lightlike future complete.*

*Proof.* Assume that $M$ was lightlike future complete.

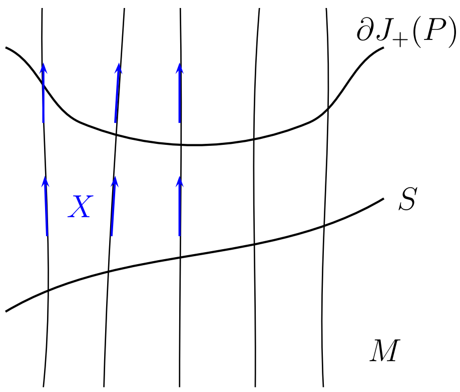
a) Since $M$ has a Cauchy hypersurface, it is globally hyperbolic by Corollary 2.99. Moreover, $J_+(P)$ is closed due to compactness of $P$ by Lemma 2.124, and thus

$$J_+(P)\backslash I_+(P) = \overline{J_+(P)}\backslash \mathring{J}_+(P) = \partial J_+(P).$$

Recall that $J_+(P)$ is a future set, so $\partial J_+(P)$ is closed topological hypersurface by Corollary 2.76. Moreover, due to Proposition 2.123, it is compact, so $\partial J_+(P)$ represents an achronal and compact topological hypersurface of $M$.

b) If $\partial J_+(P)$ was empty, we would have $J_+(P) = I_+(P)$, which is open and closed at the same time, and furthermore non-empty since $P \subset J_+(P)$. This implies $M = I_+(P)$ because $M$ is connected, so particularly $P \subset I_+(P)$, which contradicts achronality of $P$.

c) Let $\rho \colon \partial J_+(P) \to S$ the map given by the flow of some smooth and time-oriented vector field $X$ just like in Theorem 2.83, which is well-defined since $S$ is a Cauchy hypersurface. Furthermore, it is injective due to achronality of $\partial J_+(P)$, and thus a continuous and injective map between topological manifolds of the same dimension. By Brouwer's theorem (see for instance [Vick1973]), $\rho\big(\partial J_+(P)\big)$ is open. On the other hand, $\partial J_+(P)$ is compact, and therefore, so is its image under $\rho$. It follows that $\rho\big(\partial J_+(P)\big) = S$ since $M$ and hence $S$ is connected, which contradicts compactness of $S$. □




**Example 2.126.** 1) *Exterior Schwarzschild model:*
For fixed $m > 0$ and $h(r) := 1 - \frac{2m}{r}$, consider the spacetime given by

$$M := \mathbb{R} \times (2m, \infty) \times S^2, \qquad g := -h(r)\,\mathrm{d}t \otimes \mathrm{d}t + \frac{1}{h(r)}\mathrm{d}r \otimes \mathrm{d}r + r^2 g_{S^2}, \tag{2.5}$$

where $g_{S^2}$ stands for the standard metric on $S^2$. Some direct calculation shows $\mathrm{ric} = 0$.
Now let $S := \{0\} \times (2m, \infty) \times S^2 \subset M$, which is a closed and spacelike hypersurface with unit normal field $\frac{1}{\sqrt{h(r)}}\frac{\partial}{\partial t}$. It is non-compact but totally geodesic as the fixed point of the isometry $(t, r, \gamma) \mapsto (-t, r, \gamma)$.

For $c(s) =: \big(t(s), r(s), \gamma(s)\big)$ some causal curve in $M$, we have

$$0 \geq g(\dot{c}, \dot{c}) = -h(r)\cdot \dot{t}^2 + \frac{\dot{r}^2}{h(r)} + r^2 \|\dot{\gamma}\|^2_{S^2} \qquad \Longrightarrow \qquad \dot{t}^2 \geq \frac{\dot{r}^2}{h(r)^2} + \frac{r^2}{h(r)}\|\dot{\gamma}\|^2_{S^2}.$$

Therefore, any zero point $s_0$ of $\dot{t}$ would directly be a zero point of $\dot{r}$ and $\dot{\gamma}$, which would lead to $\dot{c}(s_0) = 0$ and thus to a contradiction to causality of $c$. It follows that $\dot{t}(s) > 0$ for all $s$ or $\dot{t} < 0$ for all $s$, i.e. $c$ hits $S$ at most once, so $S$ is acausal.
Now let $c(s) =: \big(t(s), r(s), \gamma(s)\big)$ denote some future-inextendible lightlike geodesic $(a, b) \to M$. It is not hard to see that as long as $r$ stays in some compact interval $[r_1, r_2] \subset (2m, \infty)$, we can solve the geodesic equation, i.e. the case $a > -\infty$ or $b < \infty$ only occurs if for $s \to a$ or $s \to b$, we would obtain $r(s) \to 2m$ or $r(s) \to \infty$, respectively.
From $\dot{t}^2 \geq \frac{\dot{r}^2}{h(r)^2}$ follows

$$t(s_1) - t(s_0) = \int_{s_0}^{s_1} \dot{t}(s)\,\mathrm{d}s \geq \int_{s_0}^{s_1} \frac{\dot{r}(s)}{h\big(r(s)\big)}\,\mathrm{d}s \geq \int_{r(s_1)}^{r(s_0)} \frac{\mathrm{d}r}{h(r)} \geq \Big[r + 2m\log(r-2m)\Big]_{r(s_0)}^{r(s_1)},$$

where, without loss of generality, we assumed $\dot{r} \geq 0$. Note that $r + 2m\log(r-2m) \to \infty$ and $r + 2m\log(r-2m) \to -\infty$ for $r \to \infty$ and $r \to 2m$, respectively, and recall that $\frac{\partial}{\partial t}$ is a Killing, that is $\frac{\mathrm{d}}{\mathrm{d}t} g\big(\dot{c}, \frac{\partial}{\partial t}\big) = g\big(\dot{c}, \nabla_{\dot{c}}\frac{\partial}{\partial t}\big) = 0$. Thus, if $r$ stays in the the compact interval $[r_0, r_1]$ and hence $(a, b) = (-\infty, \infty)$, we extract the constant $-E := g\big(\dot{c}, \frac{\partial}{\partial t}\big) = -h(r)\dot{t}$. In particular, this implies

$$\dot{t} = \frac{E}{h(r)} \geq \min_{r\in[r_0,r_1]} \frac{E}{h(r)} =: \tau > 0 \qquad \Longrightarrow \qquad t(s_1) - t(s_0) \geq \tau(s_1 - s_0).$$

In any case, the $t$-component runs over all of $\mathbb{R}$, so $c$ hits $S$ and hence, $S$ is a Cauchy hypersurface by Corollary 2.110.
Now consider $P := \{0\} \times \{r_0\} \times S^2$, which is a non-empty, compact and spacelike submanifoldof codimension 2 contained in $S$. Since $S \subset M$ is totally geodesic, the mean curvature vector field of $P$ in $M$ coincides with the one of $P$ in $S$, i.e. spacelike, such that we can not apply Theorem 2.125.

2) *Interior Schwarzschild model:*
In (2.5), replace $(2m, \infty)$ by $(0, 2m)$, so $h < 0$, and hence, $\frac{\partial}{\partial t}$ is spacelike and $\frac{\partial}{\partial r}$ is timelike. Let $S := \mathbb{R} \times \{r_0\} \times S^2$ for some $r_0 \in (0, 2m)$. A similar discussion as in 1) shows that $S$ is a spacelike, non-compact Cauchy hypersurface and $T := \{0\} \times (0, 2m) \times S^2$ is totally geodesic. Therefore, the mean curvature field of $P := \{0\} \times \{r_0\} \times S^2$ in $M$ is the same as in $T$. Because $HG$ is normal, we have $H = c(r_0)\frac{\partial}{\partial r}$, where $c(r_0)$ does not depend on the point in $P$ since the group $\mathrm{SO}_3$ acts isometrically on $M$ and transitively on $P$. One calculates $c(r_0) = -\frac{1}{r_0}$, so the mean curvature vector field is timelike and, provided the right choice of time-orientation, past directed. Finally, Theorem 2.125 ensures the existence of lightlike geodesics, whose maximal domain is not all of $\mathbb{R}$, which are interpreted as the worldlines of photons falling into the black hole.

## 2.10 Structure of globally hyperbolic Lorentzian manifolds

**Reminder:** A Lorentzian manifold $M$ is globally hyperbolic if $J(p,q) := J_+(p) \cap J_-(q)$ is compact for all $p, q \in M$ and the strong causality condition holds (Definition 2.85). We already know that the existence of a Cauchy hypersurface implies global hyperbolicity (Corollary 2.99). Let $M$ always be a connected and time-oriented Lorentzian manifold. Furthermore, let $f\colon M \to \mathbb{R}$ some smooth function with $f > 0$ and $\int_M f \,\mathrm{dvol} = 1$.

**Remark 2.127.**

1. Here, $\mathrm{dvol}$ denotes the Riemannian volume element. For integrable functions $\varphi$ with support contained in some chart $x\colon U \to V$ and coordinates $x^1, \ldots, x^n$, we have

$$\int\limits_M \varphi \,\mathrm{dvol} = \int\limits_M \varphi(x) \sqrt{\left|\det\big(g_{ij}(x)\big)\right|}\, \mathrm{d}x^1 \ldots \mathrm{d}x^n.$$

2. There exist some $f \in C^\infty(M)$ with $f > 0$ and $\int_M f \,\mathrm{dvol} = 1$. In the case of finite volume $\mathrm{vol}(M) := \int_M \mathrm{dvol} < \infty$, it is trivially given by the constant function $f := \frac{1}{\mathrm{vol}(M)}$. Otherwise, choose a partition of unity $(\rho_j)_{j \in \mathbb{N}}$ and constants $c_j := \big(2^j \int_M \rho_j \,\mathrm{dvol}\big)^{-1} > 0$. Then $f := \sum_{j=1}^\infty c_j \rho_j$ is clearly positive and moreover satisfies

$$\int\limits_M f \,\mathrm{dvol} = \sum_{j=1}^\infty c_j \int\limits_M \rho_j \,\mathrm{dvol} = \sum_{j=1}^\infty 2^{-j} = 1.$$

We introduce the functions

$$v_\pm\colon \quad M \longrightarrow [0,1], \qquad v_\pm(p) := \int\limits_{I_\pm(p)} f \,\mathrm{dvol}.$$

**Lemma 2.128.** *For any future directed and timelike curve $c\colon (a,b) \to M$, $v_+ \circ c$ is monotoneously decreasing and $v_- \circ c$ is monotoneously increasing.*
*For $M$ moreover satisfying the chronology condition, even strict monotonicity holds.*

*Proof.* Clearly, we have $c(t_1) \ll c(t_2)$ if $t_1 < t_2$, so in particular,

$$I_+\big(c(t_2)\big) \subset I_+\big(c(t_1)\big) \quad \overset{f>0}{\Longrightarrow} \quad \int\limits_{I_+(c(t_2))} f \,\mathrm{dvol} \le \int\limits_{I_+(c(t_1))} f \,\mathrm{dvol},$$

that is $v_+\big(c(t_2)\big) \le v_+\big(c(t_1)\big)$. Analogously, we obtain monotonicity of $v_- \circ c$.
Assume $M$ to satisfy the chronology condition and consider the open set $I_-\big(c(t_2)\big) \cap I_+\big(c(t_1)\big)$, which is non-empty since it contains $c(t)$ for all $t \in (t_1, t_2)$. Furthermore, it is disjoint with

$I_+\big(c(t_2)\big)$ because any element of $I_-\big(c(t_2)\big)\cap I_+\big(c(t_2)\big)$ would imply the existence of a timelike closed curve, which contradicts the chronology condition. It follows that

$$v_+\big(c(t_2)\big) = \int\limits_{I_+(c(t_2))} f\,\mathrm{dvol} < \int\limits_{I_+(c(t_2))\cup\big(I_-(c(t_2))\cap I_+(c(t_1))\big)} f\,\mathrm{dvol} \le \int\limits_{I_+(c(t_1))} f\,\mathrm{dvol} = v_+\big(c(t_1)\big). \qquad \square$$

**Remark 2.129.** In general, $v_\pm$ fail to be continuous.

**Example 2.130.** Let

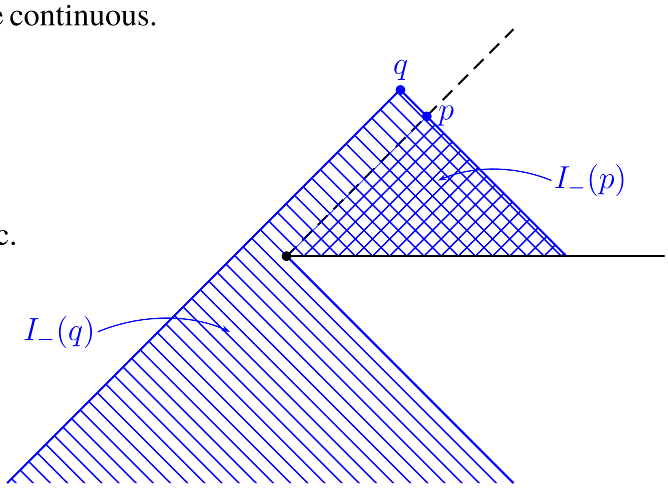
$$M := \mathbb{R}^2 \backslash \big\{(0,x) \,\big|\, x \ge 0\big\}$$

be equipped with the Minkowski metric. Then $v_-$ is discontinuous along

$$\big\{(t,x) \,\big|\, t = x > 0\big\}$$

and $v_+$ is discontinuous along

$$\big\{(t,x) \,\big|\, t = -x < 0\big\}.$$




**Lemma 2.131.** *The functions $v_\pm$ are lower semi-continuous.*

*Proof.* Let $(p_i)_{i\in\mathbb{N}} \subset M$ be a sequence converging to some $p \in M$, for which we show

$$\liminf_{i\to\infty} v_\pm(p_i) \ge v_\pm(p).$$

We carry out the proof only for $v_+$ since it works similarly for $v_-$. For any $q \in I_+(p)$, we have $p \in I_-(q)$ and thus, for $i$ large enough, $p_i \in I_-(q)$ and $I_+(q) \subset I_+(p_i)$. It follows that $v_+(p_i) \ge \int_{I_+(q)} f\,\mathrm{dvol}$ and therefore

$$\liminf_{i\to\infty} v_+(p_i) \ge \int\limits_{I_+(q)} f\,\mathrm{dvol}. \tag{2.6}$$

Choose some future directed timelike curve $c\colon [0,1] \to M$ with $c(0) = p$ and put $q_n := c\big(\frac{1}{n}\big)$. Then we directly have $I_+(q_n) \subset I_+(q_{n+1})$ and consequently

$$\bigcup_{n=1}^{\infty} I_+(q_n) \subset I_+(p).$$

Conversely, for any $z \in I_+(p)$, that is $p \in I_-(z)$ and $q_n \in I_-(z)$ for $n$ large enough, we find $z \in I_+(q_n)$, so also the converse inclusion holds, i.e. equality. From that, we deduce the claim:

$$\liminf_{i\to\infty} v_+(p_i) \overset{(2.6)}{\ge} \lim_{n\to\infty} \int\limits_{I_+(q_n)} f\,\mathrm{dvol} = \int\limits_{I_+(p)} f\,\mathrm{dvol} = v_+(p). \qquad \square$$

**Lemma 2.132.** *If $M$ is globally hyperbolic, then $v_\pm$ are continuous.*

*Proof.* Let $(p_i)_{i\in\mathbb{N}} \subset M$ be a sequence converging to some $p \in M$, for which we now show

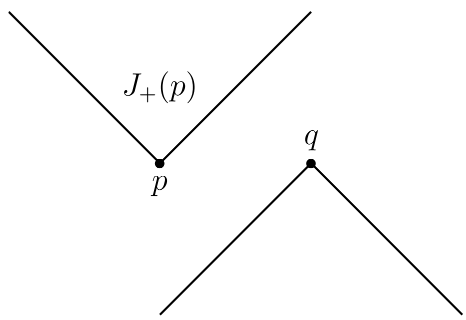
$$\limsup_{i\to\infty} v_\pm(p_i) \leq v_\pm(p).$$

Again, we give the proof only for $v_+$ since everything works analogously for $v_-$.

a) For $q \in M\backslash J_+(p)$, we show that $I_-(q) \cap I_+(p_i) = \emptyset$ if $i$ is large enough.
Assume that there exist $r_i \in I_-(q) \cap I_+(p_i)$, i.e. $q \ll r_i \ll p_i$, for infinitely many $i$. Since $p_i \to p$ and "$\geq$" is closed in the globally hyperbolic case, we obtain $q \geq p$ and hence $q \in J_+(p)$, which contradicts the assumption.

b) Due to Lemma 2.124, $J_+(p)$ is closed, so

$$M\backslash J_+(P) = \bigcup_{q\in M\backslash J_+(p)} I_-(q).$$




Since the topology of $M$ has a countable basis, there is a countable and dense subset $\{x_n\}_{n\in\mathbb{N}} \subset M\backslash J_+(p)$. To every $x_n$, choose some $q_n \in M\backslash J_+(p)$ such that $x_n \in I_-(q_n)$, which leads to

$$M\backslash J_+(p) = \bigcup_{n=1}^{\infty} I_-(q_n).$$

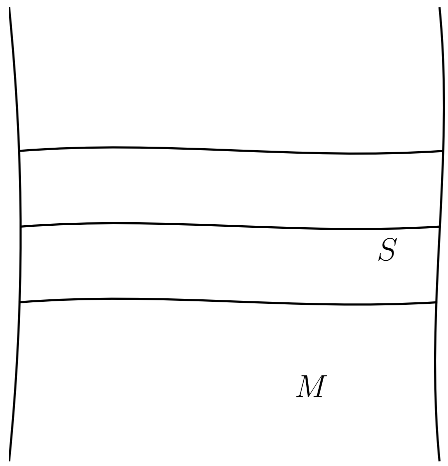
For the open sets $X_N := \bigcup_{n=1}^{N} I_-(q_n)$, we directly obtain $X_N \subset X_{N+1}$ and $X_N \cap I_+(p_i) = \emptyset$ for $i$ large enough by a). For every $N$, we therefore have

$$\limsup_{i\to\infty} \int_{I_+(p_i)\cup X_N} f\,\mathrm{dvol} \leq \int_M f\,\mathrm{dvol} = 1 \quad\Longrightarrow\quad \limsup_{i\to\infty} v_+(p_i) \leq 1 - \int_{X_N} f\,\mathrm{dvol}. \tag{2.7}$$

Recall that $\lim\limits_{N\to\infty} X_N = M\backslash J_+(p)$ and $\partial I_+(p)$ is a zero set, so taking the limit $N\to\infty$ provides

$$\limsup_{i\to\infty} v_+(p_i) \leq 1 - \int_{M\backslash J_+(p)} f\,\mathrm{dvol} = \int_{J_+(p)} f\,\mathrm{dvol} = \int_{I_+(p)} f\,\mathrm{dvol} = v_+(p). \qquad \square$$

**Theorem 2.133 (Geroch 1970).** *Let $M$ be a globally hyperbolic Lorentzian manifold. Then there is a Cauchy hypersurface $S \subset M$ and a homeomorphism $\mathbb{R}\times S \to M$, under which any $\{t\}\times S$ is mapped to a Cauchy hypersurface.*




*Proof.* a) We start by proving $\lim_{t\to b} v_+\big(c(t)\big) = 0$ for any future-inextendible timelike curve $c\colon [a,b) \to M$. Analogously, one shows $\lim_{t\to b} v_-\big(c(t)\big) = 0$ if $c$ is past-inextendible instead.
By assumption, for any $q \in M$, the causal diamond $J\big(c(a), q\big)$ is compact, so according to Lemma 2.87, we find some $t_0 \in [a,b)$ such that $c(t) \notin J\big(c(a), q\big)$ for all $t \geq t_0$. It follows that $c(t) \notin J_-(q)$ and consequently, $I_+\big(c(t)\big) \cap I_-(q) = \emptyset$ for all $t \geq t_0$.
Choose $(q_n)_{n\in\mathbb{N}} \subset M$ such that $M = \bigcup_{n=1}^\infty I_-(q_n)$, and set $X_N := \bigcup_{n=1}^N I_-(q_n)$. To each $N$, we find some $t_N \in [a,b)$ such that $I_+\big(c(t)\big) \cap X_N = \emptyset$ for all $t \geq t_N$. For those $t$, similarly to (2.7), we obtain

$$v_+\big(c(t)\big) = \int_{I_+(c(t))} f \,\mathrm{dvol} \leq 1 - \int_{X_N} f\,\mathrm{dvol},$$

which directly leads to

$$\limsup_{t\to b} v_+\big(c(t)\big) \leq 1 - \int_{X_N} f\,\mathrm{dvol} \qquad \overset{N\to\infty}{\Longrightarrow} \qquad \limsup_{t\to b} v_+\big(c(t)\big) \leq 1 - \int_M f\,\mathrm{dvol} = 0.$$

b) Next we show that $S(v_0) := \left\{q \in M \;\middle|\; \frac{v_-(q)}{v_+(q)} = v_0\right\}$ is a Cauchy hypersurface for all $v_0 > 0$. Clearly, it is achronal since $\frac{v_-}{v_+} \circ c$ is strictly monotonic for each timelike curve $c$. Let $c\colon (a,b) \to M$ non-extendible, future directed and timelike abd choose some $t_0 \in (a,b)$. For all $t \geq t_0$, we obtain

$$\frac{v_-\big(c(t)\big)}{v_+\big(c(t)\big)} \geq \frac{v_-\big(c(t_0)\big)}{v_+\big(c(t)\big)} \xrightarrow{t\uparrow b} \infty$$

as well as

$$\frac{v_-\big(c(t)\big)}{v_+\big(c(t)\big)} \leq \frac{v_-\big(c(t)\big)}{v_+\big(c(t_0)\big)} \xrightarrow{t\downarrow a} 0$$

for all $t \leq t_0$. It follows that $\frac{v_-}{v_+} \circ c$ is strictly monotoneously increasing, which therefore maps $(a,b)$ bijectively to $(0,\infty)$, so for each $v_0$, there is exactly one $t \in (a,b)$ such that $\frac{v_-(c(t))}{v_+(c(t))} = v_0$. This implies $c(t) \in S$, i.e. $c$ hits $S$.

c) Now let $\rho_{v_1,v_2}\colon S(v_1) \to S(v_2)$ denote the homeomorphism induced by the flow of some smooth and timelike vector field on $M$. Set

$$\phi\colon \quad M \longrightarrow \mathbb{R} \times S(1), \qquad q \longmapsto \left(\log \frac{v_-(q)}{v_+(q)}, \rho_{\frac{v_-(q)}{v_+(q)},q}(q)\right).$$

This map is continuous and maps the Cauchy hypersurface $S(v)$ bijectively to $\{\log(v)\} \times S(1)$. In particular, the map itself is a bijection and therefore a homoemorphism. □

**Theorem 2.134 (Bernal-Sánchez 2004).** *Any globally hyperbolic Lorentzian manifold $(M, g)$ is isometric to*

$$\left(\mathbb{R} \times S, -\beta\, \mathrm{d}\tau^2 + g_\tau\right),$$

*where $\beta\colon \mathbb{R} \times S \to \mathbb{R}$ is smooth and positive and $g_\tau$ a smooth family of Riemannian metric on $S$. Furthermore, $\{\tau_0\} \times S$ is a smooth and spacelike Cauchy hypersurface for all $\tau_0$.*

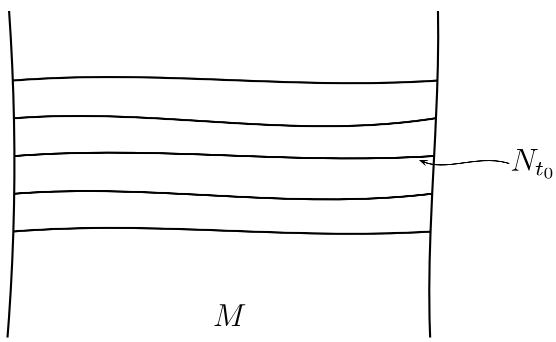
The theorem of Bernal and Sánchez provides the last piece for the main result of this section:

**Theorem 2.135.** *Let $M$ be a connected and time-oriented Lorentzian manifold. Then the following statements are equivalent:*

1. *$M$ is globally hyperbolic.*
2. *$M$ has a (topological) Cauchy hypersurface.*
3. *$M$ has a smooth, spacelike Cauchy hypersurface.*
4. *$M$ is isometric to $(\mathbb{R} \times S, -\beta\, \mathrm{d}\tau^2 + g_\tau)$ as in Theorem 2.134.*

*Proof.* $(4) \Rightarrow (3)$ and $(3) \Rightarrow (2)$ are trivial. $(2) \Rightarrow (1)$ follows from Corollary 2.99 and finally, $(1) \Rightarrow (4)$ is a consequence of Theorem 2.134. □

We dedicate the rest of the section to the proof of the theorem of Bernal and Sánchez. Let $M$ be globally hyperbolic and $t := \log \frac{v_-}{v_+}\colon M \to \mathbb{R}$ the continuous and surjective function due to Geroch, which is strictly monotoneously increasing along each timelike curve and such that $N_{t_0} := t^{-1}(t_0)$ are Cauchy hypersurfaces for each $t_0 \in \mathbb{R}$. Then we obtain




$$J_+(N_{t_0}) = t^{-1}\big([t_0, \infty)\big), \qquad J_-(N_{t_0}) = t^{-1}\big((-\infty, t_0]\big),$$
$$I_+(N_{t_0}) = t^{-1}\big((t_0, \infty)\big), \qquad I_-(N_{t_0}) = t^{-1}\big((-\infty, t_0)\big).$$

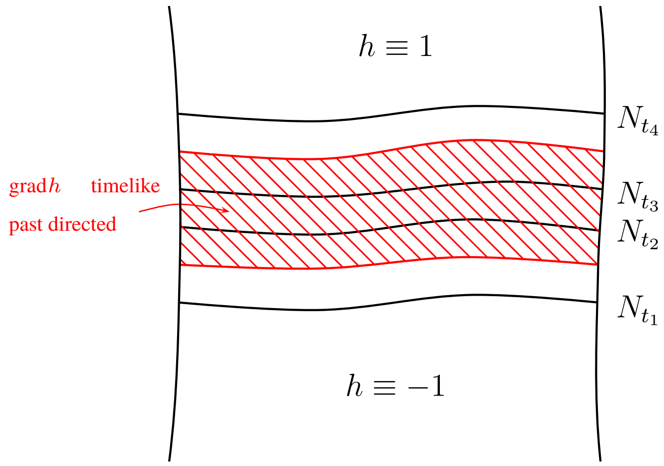
**Lemma 2.136.** *For all $t_1 < t_2 < t_3 < t_4$, there is a smooth function $h\colon M \to \mathbb{R}$ such that*

*(i)* $-1 \le h \le 1$.

*(ii)* *If $\operatorname{grad}_p h \ne 0$, it is timelike and past directed at $p$.*

*(iii)* $h \equiv -1$ *on* $t^{-1}\big((-\infty, t_1)\big)$ *and* $h \equiv 1$ *on* $t^{-1}\big((t_4, \infty)\big)$.

*(iv)* $\operatorname{grad} h \ne 0$ *on* $t^{-1}\big((t_2, t_3)\big)$.




We start with the proof of the theorem of Bernal and Sánchez and prove Lemma 2.136 afterwards.

*Proof of Theorem 2.134.* a) To each $k \in \mathbb{Z}$ and $t_1 := k-2, t_2 := k-1, t_3 := k+1, t_4 := k+2$, Lemma 2.136 provides some $h_k\colon M \to \mathbb{R}$, and we set $\tau := h_0 + \sum_{k=1}^{\infty} \big(h_{-k} + h_k\big)$.

| $t$ | | $-k-2$ | | $-k+2$ | | $0$ | | $k-2$ | | $k+2$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $h_{-k} + h_k$ | $-2$ | | $\in [-2, 0]$ | | | $0$ | | | $\in [0, 2]$ | | $2$ |

For each compact $K \subset M$, there is some $k_0$ such that $K \subset t^{-1}\big((-k_0 + 2, k_0 - 2)\big)$ and hence, $h_{-k} + h_k = 0$ on $K$ for all $k \ge k_0$. It follows that $\tau$ is a well-defined and smooth function.

b) We show that $\operatorname{grad}\tau$ is timelike and past directed on $M$.
For all $p \in M$, we either have $\operatorname{grad}_p h_k = 0$ or $\operatorname{grad}_p h_k$ is timelike and past directed. Let $p \in t^{-1}\big((k-1, k+1)\big)$, i.e. $\operatorname{grad}_p h_k$ is timelike and past directed. In particular, the level sets $S_{\tau_0} := \tau^{-1}(\tau_0)$ are spacelike hypersurfaces.

c) We show that $S_{\tau_0}$ are actually Cauchy hypersurfaces.
Let $c\colon (a,b) \to M$ be an inextendible, timelike and future directed curve, so by b), we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\tau\big(c(s)\big) = \Big\langle \mathrm{grad}_{c(s)}\tau, \dot{c}(s)\Big\rangle > 0.$$

Therefore, $\tau \circ c$ is strictly monotoneously increasing, that is, every value $\tau_0$ is assumed at most once and thus, $c$ meets $S_{\tau_0}$ at most once. For some suitably fixed $t_0$ and $p \in t^{-1}\big([t_0,\infty)\big)$, we obtain

$$\tau(p) = h_0(p) + \sum_{k=1}^{\infty}\big(h_{-k}(p) + h_k(p)\big) \geq -1 + \sum_{k=1}^{\lfloor t_0-2\rfloor} 2 = -1 + 2\lfloor t_0 - 2\rfloor = 2\lfloor t_0\rfloor - 5$$

by the properties of $h_{-k}+h_k$. Choose $\tau_+$ such that $2\lfloor t_0\rfloor - 5 > \tau_+ > \tau_0$. Since $N_{t_0}$ is a Cauchy hypersurface, there is some $s_0 \in (a,b)$ such that $c(s_0) \in N_{t_0}$, and thus, $\tau\big(c(s_0)\big) \geq \tau_+ > \tau_0$. One shows analogously that $\tau \circ c$ also takes values $\leq \tau_0$, so indeed, $c$ meets $S_{\tau_0}$.

d) Consider the diffeomorphism $\phi\colon \mathbb{R} \times S_0 \to M$ induced by $\tau$ and the flow along $\mathrm{grad}\,\tau$. The corresponding pulled back metric along $\phi$ on $\mathbb{R} \times S_0$ is of the form $-\beta\,\mathrm{d}\tau^2 + g_\tau$ since the level sets $S_{\tau_0}$ are spacelike and $\mathrm{grad}\,\tau \perp S_\tau$ because for any differentiable function, its gradient vector field is always orthogonal to its level sets. □

Before proving Lemma 2.136, we need some further technical Lemma and Whitney's famous embedding theorem:

**Theorem 2.137 (Whitney).** *Every $n$-dimensional differentiable manifold can be embedded in $\mathbb{R}^{2n+1}$ as a closed submanifold.*

*Proof.* (see, for instance, [Whitney1936]). □

**Corollary 2.138.** *Every differentiable manifold can be given a complete Riemannian metric.*

*Proof.* Let $M$ be an $n$-dimensional, differentiable manifold and $\iota\colon M \hookrightarrow \mathbb{R}^{2n+1}$ the differentiable embedding given by Theorem 2.137 such that $\iota(M) \subset \mathbb{R}^{2n+1}$ is a closed submanifold. For $g_0$ the Euclidean standard metric on $\mathbb{R}^{2n+1}$, we obtain a complete Riemannian manifold $(M, \iota^* g_0)$. In order to see this, it suffices to show that every sequence in $\iota(M)$ that is a Cauchy sequence with respect to $\iota^* g_0$, converges. Let $(p_i)_{i\in\mathbb{N}} \subset \iota(M)$ be a Cauchy sequence with respect to $g_0$. Due to completeness, it converges in $\mathbb{R}^{2n+1}$ to some $p$, and since $\iota(M)$ is closed, we have $p \in \iota(M)$, so the claim follows from the Hopf-Rinow-theorem. □

**Lemma 2.139.** *Let $d_R$ be the Riemannian distance function given by some complete Riemannian auxiliary metric $g_R$ on $M$. Let $N \subset M$ be a closed subset covered by a family $\mathcal{W} := \{W_\alpha \subset M \mid \alpha \in I\}$ of open subsets of $M$. Furthermore, assume that every $W_\alpha$ is contained in an open subset $\mathcal{C}_\alpha$ with $\operatorname{diam}(\mathcal{C}_\alpha) := \sup\{d_R(p,q) \mid p, q \in \mathcal{C}_\alpha\} < 1$. Then there is a countable, locally finite subfamily $\mathcal{W}' = \{W_j\}_{j\in\mathbb{N}} \subset \mathcal{W}$ covering $N$. The corresponding subfamily $\{C_j\}_{j\in\mathbb{N}} \subset \mathcal{C}$ is locally finite as well.*

*Proof.* Let $p \in M$ and $B_p(r)$ the open ball centered at $p$ of radius $r > 0$. Since $(M, g_R)$ is complete, the closure $\overline{B_p(r)}$ is compact and so are the subsets

$$K_m := \overline{B_p(m)} \backslash B_p(m-1), \qquad N := K_m \cap N$$

for all $m \in \mathbb{N}$. Note that $M \subset \bigcup_{m\in\mathbb{N}} K_m$ and consequently, $N \subset \bigcup_{m\in\mathbb{N}} N_m$. In fact, due to compactness, $N_m$ is covered already by finitely many subsets $W_{1,m}, \ldots, W_{k_m,m} \in \mathcal{W}$, so

$$\mathcal{W}' := \{W_{j,m} \mid m \in \mathbb{N}, j = 1, \ldots, k_m\},$$

is countable, and therefore so is also the corresponding subfamily of $\mathcal{C}$. Moreover, $\mathcal{W}'$ covers $N$ and it is locally finite since $W_{j,m} \cap W_{j',m'} = \emptyset$ for $|m - m'| \geq 3$. □

*Proof of Lemma 2.136.* We perform the construction in several steps. For that, let $M$ be always globally hyperbolic and $t := \log \frac{v_-}{v_+} \colon M \to \mathbb{R}$ Geroch's continuous and surjective 'time function':

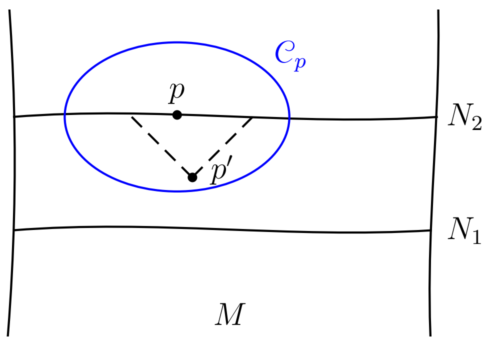
a) Let $t_1 < t_2$, and we write $N_i := N_{t_i}, i = 1, 2$. For $p \in N_2$ and $\mathcal{C}_p \subset I_+(N_1)$ a convex neighborhood of $p$, we show that there is a smooth function $H_p \colon M \to [0, \infty)$ such that

(i) $H_p(p) = 1$.

(ii) $\operatorname{supp}(H_p)$ is compact and contained in $\mathcal{C}_p \cap I_+(N_1)$.

(iii) For $q \in J_-(N_2)$, the gradient $\operatorname{grad}_q H_p$ is either zero or timelike and past directed.

For $\tau$ the time difference function (Definition 2.24) on $\mathcal{C}_p$ and $p' \in I_-(p) \cap I_+(N_1)$ such that $J_+(p') \cap J_-(N_2) \subset \mathcal{C}_p$, we define $H_p$ on $I_-(N_2)$ via




$$H_p(q) := \begin{cases} e^{\tau(p',p)^{-2} - \tau(p',q)^{-2}}, & q \in I_-(N_2) \cap \mathcal{C}_p, \\ 0, & q \in I_-(N_2) \backslash \mathcal{C}_p. \end{cases}$$

and suitably extend it to all of $M$ with compact support in $\mathcal{C}_p$. By construction, $H_p(p) = e^0 = 1$ and $H_p(q) = 0$ if $q \notin I_+(p')$ since then $\tau(p', q) = 0$, so (i) and (ii) are satisfied.

For (iii), consider $f\colon q \mapsto \frac{1}{2}\tau(p',q)^2$ on $\mathcal{C}_p$, i.e. $f = 0$ outside of $I_+(p')$. For $q \in I_+(p')$, let $\gamma\colon [0,1] \to \mathcal{C}_p$ be the unique timelike, future directed geodesic from $p'$ to $q$, so for all $t, t' \in [0,1]$ with $t' < t$ and $c := \sqrt{-g\big(\dot\gamma(1),\dot\gamma(1)\big)}$, we obtain

$$\tau\big(\gamma(t'),\gamma(t)\big) = \int_{t'}^{t} |\dot\gamma(s)|\,\mathrm{d}s = c(t-t').$$

It follows that going forward/backward along $\gamma$ means going in the direction of maximal increase/decrease of $q \mapsto \tau(p',q)$, i.e. $\operatorname{grad}_{\gamma(t)}\tau$ is proportional to $\dot\gamma(t)$. In particular, we have $\operatorname{grad}_q f = \alpha \cdot \dot\gamma(1)$ for some $\alpha \neq 0$, which is

$$\alpha = \alpha \cdot \frac{g\big(\dot\gamma(1),\dot\gamma(1)\big)}{-c^2} = -\frac{g\big(\operatorname{grad}_q f,\dot\gamma(1)\big)}{c^2} = -\frac{1}{c^2}\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{f\big(\gamma(t)\big)}_{=-\frac{1}{2}c^2t^2}\Bigg|_{t=1} = 1.$$

Therefore, $\operatorname{grad}_q f$ is timelike and future directed. Note that, where it does not vanish, we have

$$H_p(q) = \exp\Big(\tau(p',p)^{-2} + \frac{1}{2}f(q)^{-1}\Big) \qquad \Longrightarrow \qquad \operatorname{grad}_q H_p = -\frac{1}{2}\operatorname{grad}_q f \cdot H_p(q),$$

which proves (iii) since $H_p(q) \geq 0$.

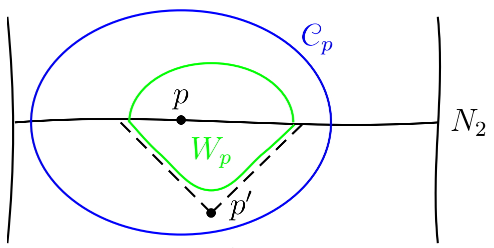
b) We show that there is a smooth function $H\colon M \to [0,\infty)$ such that

(i) $H = 0$ on $J_-(N_1)$.

(ii) $H > \frac{1}{2}$ on $N_2$.

(iii) $\operatorname{grad} H$ is timelike past directed on $V := H^{-1}\big((0,\frac{1}{2})\big) \cap I_-(N_2)$.

Let $d_R$ be the auxiliary metric from Lemma 2.139. For $p \in N_2$, let $\mathcal{C}_p$ be a convex neighborhood of $p$ with $\operatorname{diam}(\mathcal{C}_p) < 1$ and $H_p$ as defined in a). Then




$$H_p^{-1}\left(\left(\frac{1}{2},\infty\right)\right) =: W_p \subset \mathcal{C}_p,$$

since $\operatorname{supp}(H_p) \subset \mathcal{C}_p$, and $\mathcal{W} := \{W_p\}_{p\in N_2}$ satisfies the assumptions of Lemma 2.139 with $N := N_2$. Therefore, we find a countable and locally finite subfamily $\mathcal{W}' = \{W_{p_j}\}_{j\in\mathbb{N}} \subset \mathcal{W}$, which covers $N_2$. Let $W_j := W_{p_j}$ and $H_j := H_{p_j}$ the corresponding functions with compact support in $\mathcal{C}_j := \mathcal{C}_{p_j}$, and we set

$$H := \sum_{j\in\mathbb{N}} H_j. \tag{2.8}$$

Due to local finiteness of $\{\mathcal{C}_j\}_{j\in\mathbb{N}}$ due to Lemma 2.139 and hence also $\big\{\operatorname{supp}(H_j)\big\}_{j\in\mathbb{N}}$, this is a well-defined and smooth function $M \to [0,\infty)$. Now (i) and (iii) follow from the respective properties of the $H_j$ by a), and moreover, for all $j \in \mathbb{N}$ and $p \in W_j$, we have $H_j(p) > \frac{1}{2}$, so (ii) follows from $N_2 \subset \bigcup_j W_j$.

c) For $t_1 < t_0 < t_2$ with Cauchy hypersurface $N := N_{t_0}$, we show that there is an open set $U$ with $J_-(N) \subset U \subset I_-(N_{t_2})$ and a function $H^+\colon M \to [0,\infty)$ with

(i) $\mathrm{supp}(H^+) \subset I_+(N_{t_1})$

(ii) For $p \in U$ such that $H^+(p) > 0$, $\mathrm{grad}_p H^+$ is timelike past directed.

(iii) We have $H^+$ and thus $\mathrm{grad}\, H^+$ timelike past directed on $J_+(N) \cap U$.

We construct $U$ as follows: Choose convex neighborhoods, which cover $N$ and are completely contained in $I_-(N_{t_2})$ (see, for instance, the construction of $H$ in the proof of b)). Define $U$ as their union united with $I_-(N)$ and set $N_2 := N_{t_0}, N_1 := N_{t_1}$. Then (2.8) provides such a function $H^+$.

d) For $t_0 < t_2$, let $N := N_{t_0}$ be the corresponding Cauchy hypersurface and $U \subset I_-(N_{t_2})$ an open neighborhood of $J_-(N)$. We show that there is a function $H^-\colon M \to (-\infty, 0]$ with

(i) $\mathrm{supp}(H^-) \subset U$

(ii) For $p \in U$, we have either $\mathrm{grad}_p H^- = 0$ or $\mathrm{grad}_p H^-$ is timelike past directed.

(iii) $H^- \equiv -1$ on $J_-(N)$.

The proof is similar to the one for b). We adopt all quantities introduced so far but add "$\sim$" if they refer to the converse time orientation. For instance, $\widetilde{\tau}$ denotes the time difference with respect to the converse time orientation, $\widetilde{I_-}(N) = I_+(N)$ etc. Let $N_2 := N_{t_0}$ and $N_1 := N_{t_2}$, and we cover $N$ by convex neighborhoods $C_p$ as in the proof of b), which are completely contained in $U$. Similarly, define $H$ via (2.8) with $H_j$ replaced by $-\widetilde{H}_j$, which are constructed analogously to a). This satisfies (i) and (ii). Note that $H(p) < -\frac{1}{2}$ for all $p \in N$ due to b) (ii) and the sign, and define $H^- := \Phi \circ H$, where $\Phi\colon \mathbb{R} \to \mathbb{R}$ denotes a smooth function with

$$\Phi\big|_{\left(-\infty,-\frac{1}{2}\right)} = -1, \qquad \Phi'\big|_{\left[-\frac{1}{2},0\right]} > 0, \qquad \Phi(0) = 0.$$

These properties ensure that (i) and (ii) still hold. Note that a timelike gradient points in the direction of maximal decrease of a function, so $H^-$ is decreasing in past direction by (ii) and thus, (iii) follows from $H^-\big|_N < -\frac{1}{2}$.

e) For some $t_0 \in \mathbb{R}$, let $N := N_{t_0}$. We show that there is a function $\widehat{H} := \widehat{H}_{t_0}\colon M \to \mathbb{R}$ such that (i)-(iii) of Lemma 2.136 are satisfied as well as $N \subset V_{t_0} := \mathrm{int}\big(\mathrm{supp}(\mathrm{grad}\, \widehat{H})\big)$.
Let $U$ and $H^\pm$ as in c) and d), and note that $(H^+ - H^-)\big|_U > 0$. We set

$$\widehat{H}_{t_0}(p) := \frac{2H^+}{H^+ - H^-} - 1$$

which is a smooth function with $\widehat{H}_{t_0}\big|_{M\setminus U} \equiv 1$ due to $\mathrm{supp}(H^-) \subset U$. Furthermore,

$$\mathrm{grad}\, \widehat{H}_{t_0} = 2 \cdot \frac{H^+ \cdot \mathrm{grad}\, H^- - H^- \cdot \mathrm{grad}\, H^+}{(H^+ - H^-)^2}$$

vanishes or is timelike past directed, so the properties of $H^\pm$ imply that $\widehat{H}_{t_0}$ does the job.

f) For $t_1 < t_0 < t_2 < t_3 < t_4$, consider the compact subset $K \subset t^{-1}\big([t_1, t_2]\big)$. We show that there is a function $\widetilde{H}\colon M \to \mathbb{R}$, which satisfies the assumptions on $\widehat{H}$ in e) and additionally, $K \subset V = \operatorname{int}\big(\operatorname{supp}(\operatorname{grad}\widehat{H})\big)$.
For each $N_{t_0}$ with $t_0 \in [t_1, t_2]$, according to e), choose the corresponding $\widehat{H}_{t_0}$, so $K$ can be covered by the corresponding open sets $V_{t_0}$. By compactness, there are finitely many $V_{t_{01},\dots,V_{t_{0,m}}}$ already cover $K$, so the function

$$\widetilde{H} := \frac{1}{m}\sum_{i=1}^{m}\widehat{H}_{t_{0,i}}$$

does the job.

g) Let $(v_j)_{j\in\mathbb{N}} \subset \mathbb{R}^n_{\text{Mink}}$ be a sequence of timelike vectors from the same cone $I_+$ or $I_-$. We show that if $v := \sum_{j=1}^{\infty} v_j$ is convergent, $v$ is timelike as well and contained in the same cone. Since the causal cone is closed, $\sum_{j=2}^{\infty}$ is causal, so the claim follows from the fact that the sum $v_1 + \sum_{j=2}^{\infty}$ of a timelike and a causal vector, which lie in the same cone, is timelike. h) Finally, we show the claim of Lemma 2.136:
Let $t_1 < t_2 < t_3 < t_4$ and $\{G_j\}_{j\in\mathbb{N}}$ be an exhaustion of $M$ be relatively compact subsets, i.e. open sets such that

$$\overline{G_j} \text{ is compact}, \qquad \overline{G_j} \subset G_{j+1}, \qquad M = \bigcup_{j=1}^{\infty} G_j.$$

Furthermore, let $K_j := \overline{G_j} \cap J_+(N_{t_2}) \cap J_-(N_{t_3})$, and for each $K_j$, consider the function $\widetilde{H}_j := \widetilde{H}$ with $\widetilde{H}$ given as in f) for $K := K_j$ with $K_j \subset V_j := \operatorname{int}\big(\operatorname{grad}\widetilde{H}_j\big)$. Due to local normal convergence, the naive ansatz

$$h := \sum_{j=1}^{\infty}\frac{1}{2^j}\cdot\widetilde{H}_j$$

defines a continuous $\mathbb{R}$-valued function, but it is neither clear whether it is also smooth nor if partially differentiation and summation commute. We ensure these properties by slight adaptions:
Choose a countable and locally finite atlas $\mathcal{A} := \{W_j\}_{j\in\mathbb{N}}$ orf $M$ such that every chart $(W, x^1, \dots, x^n) \in \mathcal{A}$ is relatively compact and the restriction of some larger chart containsing $\overline{W}$. Every $G_j$ overlaps with only finitely many $W_{j,1}, \dots, W_{j,k_j}$, and since $D_j := \overline{\bigcup_{i=1}^{k_j} W_{j,i}}$ is compact, there is some $c_j > 1$ with $\widetilde{H}_j < c_j$ on $D_j$ such that

$$\forall s < j,\ q \in D_j,\ l_1, \dots, l_s \in \{1, \dots, n\}: \qquad \left|\frac{\partial^s \widetilde{H}_j}{\partial x_{l_1}\partial x_{l_2}\dots\partial x_{l_s}}(q)\right| < c_j.$$

We define

$$H^* := \sum_j \frac{1}{2^j c_j}\cdot\widetilde{H}_j, \tag{2.9}$$

which defines a $C^s$-function on $M$ for all $s$: Let $j_0 \in \mathbb{N}$ and $W \in \mathcal{A}$ with $p \in G_{j_0} \cap W$. For all $j > \max\{j_0, s\}$, the summand $\frac{1}{2^j c_j} \cdot \widetilde{H}_j$ and all its partial derivatives up to order $s$ are bounded by $\frac{1}{2^j}$ on $G_{j_0} \cap W$. It follows that the series (2.9) and the series' of the corresponding partial derivatives converge uniformly in a neighborhood of $p$, so the partial derivatives of the series are given by the series of the respective derivatives of the summands. Thus, $H^*$ satisfies all demanded properties of $h$ in Lemma 2.136 except (i). Instead, we have

$$H^*\big(J_-(N_{t_1})\big) \equiv c_- < 0 \qquad \text{and} \qquad H^*\big(J_-(N_{t_4})\big) \equiv c_+ > 0.$$

For any smooth $\psi\colon \mathbb{R} \to \mathbb{R}$ with $\psi' > 0$ and $\psi(c_\pm) = \pm 1$, $h := \psi \circ H^*$ does the job. □

We close this section and these lecture notes with an overview about rather recent results and improvements concerning the characterization of globally hyperbolic Lorentzian manifolds.

**Definition 2.140.** Let $M$ a globally hyperbolic Lorentzian manifold. A **Cauchy time-function** is a smooth function $t\colon M \to \mathbb{R}$ with timelike past directed gradient at each point such that $t^{-1}(\{s\}) \subset M$ is a Cauchy hypersurface for all $s \in \mathbb{R}$.

Note that Cauchy time-functions are strictly monotoneously increasing along any causal future directed curve, and that Geroch's theorem 2.133, more precisely its proof, provides the existence of such a function for every globally hyperbolic Lorentzian manifold. By some further result (see Theorem 1.2 of [Bernal-Sánchez2006]), also the, in some sense, converse statement holds:

**Theorem 2.141 (Bernal-Sánchez 2005).** *Let $M$ be a globally hyperbolic Lorentzian manifold and $S \subset M$ be a spacelike smooth Cauchy hypersurface. Then there exists a Cauchy time-function $t$ such that $S = t^{-1}(\{0\})$.*

It is moreover possible to relax the causality assumptions as it turns out that, provided compactness of the causal diamonds, causality implies strong causality. For this, we introduce further properties of spacetimes in the "causal hierarchy of spacetimes":

**Definition 2.142.** A connected and time-oriented Lorentzian manifold $M$ is called

- **reflecting** if $p \in \overline{J_+(q)} \;\Leftrightarrow\; q \in \overline{J_-(p)}$ for all $p, q \in M$.
- **non-totally vicious** if there is some $p \in M$, through which no timelike loop passes.
- **future (past) distinguishing** if $I_{+(-)}(p) = I_{+(-)}(q)$ implies $p = q$.
  If it is both, we just call it **distinguishing**.
- **causally simple** if it is distinguishing and $J_\pm(p) \subset M$ are closed subsets for all $p \in M$.

With the terms introduced so far, this hierarchy reads:

$$\begin{aligned} &\text{globally hyperbolic} \;\Rightarrow\; \text{causally simple} \;\Rightarrow\; \text{strongly causal} \;\Rightarrow\; \text{distinguishing} \\ &\quad\Rightarrow\; \text{causal} \;\Rightarrow\; \text{chronological} \;\Rightarrow\; \text{non-totally vicious.} \end{aligned} \tag{2.10}$$

A complete list and proofs can be found, for instance, [Minguzzi-Sánchez2008]. In fact, for causal simplicity, it suffices to demand causality instead of being distinguishing:

**Proposition 2.143.** *Let $J_\pm(p) \subset M$ be closed subsets for all $p \in M$. Then $M$ is causal if and only if it is distinguishing.*

*Proof.* We only show the implication that does not follow from (2.10). Assume $I_+(p) = I_-(q)$ for some $p, q \in M$ and let $\{q_j\}_{j\in\mathbb{N}} \subset M$ be a sequence with $q_j << q$ for all $j$. Closedness of $J_+(p)$ implies $q \in J_+(p)$ because

$$q \in \overline{I_+(q))} = \overline{I_+(p)} = J_+(p)$$

due to Proposition 2.17. Analogously, one shows $p \in J_+(q)$, i.e. $p \ll q \ll p$, that is $p = q$ due to causality. □

**Lemma 2.144.** *Let $M$ be a connected and time-oriented Lorentzian manifold. Then the following implications hold:*

$$\forall p, q: \quad J(p,q) \text{ is compact} \quad \Longrightarrow \quad \forall p: \quad J_\pm(p) \text{ is closed} \quad \Longrightarrow \quad M \text{ is reflecting.}$$

*Proof.* We start with the first arrow: Assume $J_+(p)$ was not closed for some $p \in M$, so there is some $r \in \overline{J_+(p)} \backslash J_+(p)$ and we choose $q \in I_+(r)$. Due to Proposition 2.17, we find a sequence $\{r_j\}_{j\in\mathbb{N}} \subset I_+(p)$ with $r_n \to r$. Since $r \in I_-(q)$, which is an open subset, we have $r_j \ll q$ for all $j$ large enough. Hence, there is a subsequence contained in $J(p,q)$, which converges to $r \notin J(p,q)$ and thus contradicts compactness of $J(p,q)$.
For the second arrow, recall that $p \in I_+(q) \;\Leftrightarrow\; q \in I_-(p)$ always holds for all $p, q \in M$ by definition of $I_\pm$, and hence, the claim follows directly from Proposition 2.17. □

**Theorem 2.145 (Bernal-Sánchez 2006).** *Let $M$ be a connected and time-oriented Lorentzian manifold such that $J(p,q)$ is compact for all $p, q \in M$. Then the causal condition and the strong causal condition are equivalent.*

*Proof.* By Proposition 2.143 and Lemma 2.144, compactness of the causal diamonds and causality imply causal simplicity and thus strong causality due to the causal hierarchy (2.10).□

Therefore, apart from the condition on the $J_\pm$-subsets, for global hyperbolicity and causal simplicity it suffices to demand causality, respectively. In fact, in some circumstances, global hyperbolicity hold even without such a condition:

**Proposition 2.146.** *A reflecting and non-totally vicious Lorentzian manifold $M$ is chronological.*

*Proof.* Suppose it was not, i.e. there is some $p \in M$ with $p \ll p$, and let $q \in M$ not be passed by a timelike loop, that is $q \notin I_+(q) \cup I_-(q)$. It follows that either $I_+(p) \neq M$ or $I_-(p) \neq M$ since otherwise, we would have $q \in I_\pm(p)$ and thus $q \ll p \ll q$, which contradicts the assumption on $q$. Without loss of generality, we assume $I_+(p) \neq M$, so let $r \in \partial I_+(p)$, that is $p \in \partial I_-(r)$ by reflectivity. Due to $p \ll p$, $I_+(p)$ is an open neighborhood of $p$, which implies $r \in I_+(p)$, a contradiction. □

**Theorem 2.147 (Hounnonkpe-Minguzzi 2019).** *Let $M$ be a connected, non-compact and time-oriented Lorentzian manifold of dimension $\geq 3$ such that $J(p,q)$ is compact for all $p, q \in M$. Then $M$ is globally hyperbolic.*

We only sketch the proof and refer to [Hounnonkpe-Minguzzi2019] for a complete argumentation with references.

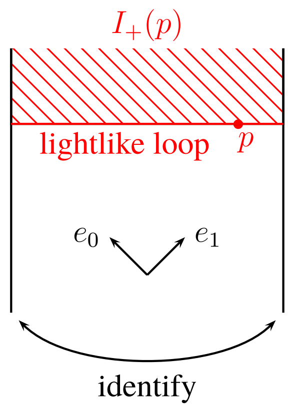
*Sketch of the proof.* a) Assume $M = I_\pm(p)$ for all $p \in M$, so particularly $M = J_\pm(p)$ and thus $M = J(p,q)$ for all $p, q \in M$, which would imply compactness of $M$. Hence, $M$ is non-totally vicious and also reflecting and thus chronological by Lemma 2.144 and Proposition 2.146.

b) Next assume that $M$ was not causal, so let $c\colon I \to M$ be a causal curve with $a, b \in I,\ a < b$ such that $c(a) = c(b)$. However, due to chronology, there cannot be any two points $p \ll q$ on $c$ since then $p \ll q \leq p$, which would imply $p \ll p$. By Lemma 2.29, $c$ is a geodesic (up to parametrization), and it particularly follows that, once parametrized, $c$ cannot develop "corners", that is $\dot{c}(a) \propto \dot{c}(b)$.
Let $p, r \in M$ be points on $c$ and thus $r \in J_\pm(p)$ as $c$ is a causal loop, which directly leads to $J_\pm(p) = J_\pm(r)$. Since $c$ is lightlike, we have $\partial I_+(c) = \partial I_+(p)$, which is non-empty as it contains $c$ and yields an achronal subset due to the properties of $c$. Since $J_+(p)$ is a future set, $\partial I_+(p)$ is an achronal topological hypersurface by Corollary 2.76 and furthermore Lipschitz.

Note that lightlike curves have a domain consisting entirely of critical points, so be the Morse-Sard theorem its image cannot fill a manifold of dimension $\geq 2$. Since $\dim M \geq 3$, we therefore find $q \in \partial I_+(p) \backslash c$, i.e. $q \in J_+(p)$ due to closedness and the curve connecting $p$ and $q$ has to be a lightlike geodesic since otherwise $q \in I_+(p)$. Recall $\partial I_+(c) = \partial I_+(p)$, so for all $r \in c$, there is a lightlike curve connecting $r$ and $q$, which cannot be given by $c$. Therefore, there is only a piecewise lightlike curve connecting $r, p$ and $q$ with a corner at $p$, which can be deformed into a timelike one, that is $q \in I_+(r) = I_+(p)$ due to closedness of $c$. This is a contradiction to the assumption on $q$, so $c$ fails to exist. $\square$

Clearly, we have to demand non-compactness due to Proposition 2.18. The dimensional condition comes from the Morse-Sard theorem and indeed, there are counterexamples in dimension 2:

**Example 2.148.** Let $G_j := \left\{\binom{0}{j} + \binom{1}{1}\right\} \subset \mathbb{R}^2$ and consider the cylinder $M := \mathbb{R}^2 / \bigcup_{j \in \mathbb{N}} G_j$ together with the induced Minkowski metric, i.e. oriented in lightlike direction. Due to continuity of the quotient map, the causal diamonds are compact since they are in $\mathbb{R}^2_{\text{Mink}}$, but it is not hard to find a closed lightlike curve.
The picture shows that for each $p \in M$, the boundary of $I_+(p)$ is indeed represented by only one lightlike loop, so an argumentation as in the proof of Theorem 2.147 does not apply.



# Bibliography


[Bernal-Sánchez2005] A. N. Bernal and M. Sánchez, *Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes*. Comm. Math. Phys. 257 (2005), 43–50.

[Bernal-Sánchez2006] A. N. Bernal and M. Sánchez, *Further results on the smoothability of Cauchy hypersurfaces and Cauchy time functions*. Lett. Math. Phys. 77 (2006), 183–197.

[Bernal-Sánchez2007] A. N. Bernal and M. Sánchez, *Globally hyperbolic spacetimes can be defined as 'causal' instead of 'strongly causal'*. Class. Quantum Grav. 24 (2007), 745–749.

[Geroch1970] R. Geroch, *Domain of dependence.* J. Math. Phys. 11 (1970), 437-449.

[Hounnonkpe-Minguzzi2019] R. A. Hounnonkpe and E. Minguzzi, *Globally hyperbolic spacetimes can be defined without the causal condition*. Class. Quantum Grav. 26 (2019), 197001.

[Milnor1965] J. Milnor: *Topology from the differential viewpoint.* University Press of Virginia, Charlottesville 1965.

[Minguzzi-Sánchez2008] E. Minguzzi, M. Sánchez, *The causal hierarchy of spacetimes*. H. Baum and D. Alekseevsky (eds.), vol. Recent developments in pseudo-Riemannian geometry, ESI Lect. Math. Phys., (Eur. Math. Soc. Publ. House, Zurich, 2008), p. 299-358.

[O'Neill1983] B. O'Neill: *Semi-Riemannian geometry.* Academic Press, New York 1983.

[Sakai1996] T. Sakai: *Riemannian geometry.* AMS, Providence 1996.

[Vick1973] J.W. Vick: *Homology Theory.* Academic Press, New York 1973.

[Warner1983] F. W. Warner: *Foundations of Differential Manifolds and Lie Groups.* Springer-Verlag, New York-Berlin-Heidelberg 1983.

[Whitney1936] H. Whitney: *Differentiable manifolds.* Ann. Math. 37 (1936), 1-36

# Index