SINGULARITY THEOREMS - CLASSICAL RESULTS

Objective: Prove Penrose's 1965 singularity theorem, which shows the existence of singularities in certain conditions.

Remarks (1): In general, singularity theorems (Penrose, 1965; Hawking, 1967; Hawking and Penrose, 1970 ...) prove the existence of spacetime singularities independently of the symmetry of spacetime. This is an interesting feature of GR. On the contrary, in classical physics (e.g. classical collapses, Newton's gravitational field, Coulomb's electrostatic field) singularities are unphysical and only arise when unrealistical levels of symmetry are assumed.

Remark (2): We will characterize singularities looking for the "holes" they leave on spacetime, which we will detect through the existence of incomplete geodesics (geodesics of finite afine length in one direction which are nonetheless inextendable). For a discussion on the difficulty of defining singularities, see Wald, Ch. 9.1, or 1410.5226, Sec. 4.1.

0. A REMINDER OF GEODESICS

→ A curve γ is <u>autoparallel</u> if it transports parallelly its own tangent vector ξ^a:

$$\xi^a \nabla_a \xi^b \propto \xi^b$$

Under an affine reparametrization → $\xi^a \nabla_a \xi^b = 0$

→ A spacelike / timelike curve γ is called a <u>geodesic</u> if it extremizes the <u>length functional</u>

$\gamma : S \subset \mathbb{R} \to \mathcal{M}$, $L[\gamma] = \int_S |g_{ab}\xi^a \xi^b| \, ds$ ← length functional

→ Under the <u>Levi-Civita connection</u> autoparallel and geodesic curves are equivalent. Henceforth, we refer to both as <u>geodesics</u>. In particular, a null geodesic w/ tangent vector ξ^a satisfies

$$\xi^a \xi_a = 0, \qquad \xi^a \nabla_a \xi^b \underset{\uparrow}{=} 0$$

assuming affine param.

$\Leftrightarrow \xi_b \xi^b = \text{cst}$ ↓ usually normalized to ± 1

1. NULL HYPERSURFACES

DEF: A <u>null hypersurface</u> is a hypersurface whose normal vector is everywhere null

Rec: Given a hypersurface Σ defined by $f(x) = \text{cst}$, with f a smooth function, the 1-form $(df)_\mu = \partial_\mu f$ is orthogonal to Σ.

E.g.: Schwarzschild spacetime in ingoing EF coordinates

Inverse metric: $\partial_s^2 = \partial_r \partial_v + \partial_v \partial_r + \left(1 - \frac{2M}{r}\right) \partial_r^2 + (\text{angular})$

Consider the family of hypersurfaces $r = \text{cte}$, whose normal 1-form is $n \equiv dr$. The pseudonorm of the normal vector is

$$n \cdot n = g^{rr} = 1 - \frac{2M}{r} \Rightarrow r = 2M \text{ is null hypersurface}$$

The corresponding normal vector (dual to n) is

$$\tilde{n} = \partial_v + \left(1 - \frac{2M}{r}\right) \partial_r \xrightarrow{r \to 2M} \partial_v$$

→ For any hypersurface, the tangent vector is one normal to the normal vector ⇒ in null hypersurfaces, the normal (null) vector is <u>also tangent</u> (see discussion in Wald, p. 42, p. 65).

<u>PROP 1.1</u>: Given a null vector n^a normal to a null hypersurface, its integral curves are null geodesics (called <u>generators</u> of the hypersurface)

$\triangleright$: Let $\mathcal{N}$ be the null hypersurface, defined as $f = cst$ for some function f such that $df|_{\mathcal{N}} \neq 0$.
Let n_a be the 1-form $n \propto df$ normal to $\mathcal{N}$.
Then,

$$n_a n^a = 0 \Rightarrow \nabla_a (n_b n^b)\big|_{\mathcal{N}} \propto n_a \qquad (*)$$

$$\nabla_a n_b \propto \nabla_a (df)_b = \nabla_a \nabla_b f \underset{\substack{\uparrow \\ f \text{ is scalar}}}{=} \nabla_b \nabla_a f = \nabla_b n_a \qquad (**)$$

$$\nabla_a (n_b n^b) = n_b (\nabla_a n^b) + (\nabla_a n_b) n^b = 2 n^b \nabla_a n_b \underset{(**)}{=} 2 n^b \nabla_b n_a \quad (***)$$

$(*), (***) \Rightarrow n^b \nabla_b n_a \big|_{\mathcal{N}} \propto n_a$ ← eq. for autoparallel curve (geodesic) ∎

<u>E.g.:</u> Kruskal spacetime

$$ds^2 = -\frac{32 M^3 e^{-r/2M}}{r} dU\, dV + r^2 d\Omega^2, \qquad r = r(U,V)$$

Consider the family of null hypersurfaces $U = cst$, whose normal 1-form $n = dU$ is <u>everywhere</u> null. Then,

$$n^b \nabla_b n_a \underset{\substack{\uparrow \\ \text{Proof of prop. 1.1}}}{=} \frac{1}{2} \nabla_a (n_b n^b) = 0$$ // (not just $\propto n_a$ because dU is everywhere null)

↳ affinely parametrized geodesic

~~Tangent~~ Normal vector is $\tilde{n} = \frac{r}{16 M^3 e^{-r/2M}} \partial_V$

Recall: $r_* = r + 2M \log\left|\frac{r}{2M} - 1\right|$, $u = t - r_*$, $U = -e^{-u/4M}$

In Kruskal-Szekeres coordinates, $r = 2M \Leftrightarrow U = 0$ or $V = 0$, ergo $U = 0 \Rightarrow r = 2M$. Let $U = 0$, $r = 2M$, we have seen that $\hat{n} \propto \partial_V$. Recall $\hat{n}$ is tangent to affinely parametrized null geodesics. Therefore V is an affine parameter for null geodesics which generate the null hypersurface $U = 0$ (reciprocally, U is affine parameter for null geodesics generating null hypersurface $V = 0$).

2. GEODESIC CONGRUENCES

2.1 Geodesic deviation

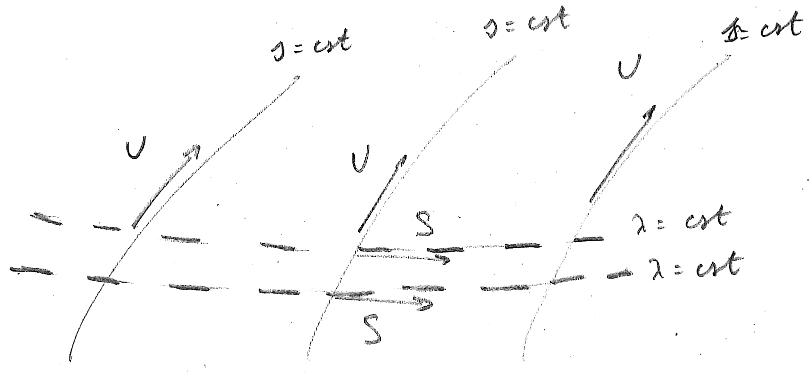
DEF: Let $\mathcal{U} \subset M$ be open (M = full spacetime). A geodesic congruence in $\mathcal{U}$ is a family of geodesics such that one and only one passes through each point of $\mathcal{U}$

Objective: study how gravity (curvature) acts as a "tidal force" deforming families of geodesics. We define a more tractable subset of a congruence

DEF: A 1-parameter family of geodesics (1PFG) is a map $\gamma: (I \subset \mathbb{R}) \times (I' \subset \mathbb{R}) \to M$ such that $\gamma(s = s_0, \lambda)$ is an affinely parametrized geodesic (w/ affine param. λ).

Note (1): The map λ is assumed to be smooth, with smooth inverse and one-to-one. One-to-one in λ means that individual geodesics do not auto-intersect. One-to-one in s means that geodesics do not cut each other (otherwise, they would not belong to a congruence).

Note (2): In the conditions of the definition, a 1PFG formes a 2d-hypersurface Σ in M.



U^a : tangent to curves of constant s (i.e. geodesics)
S^a : " " " " " λ → $\boxed{\text{deviation vector}}$

Consider a chart of Σ: $x^\mu(s, \lambda)$, then

$$S^\mu = \frac{\partial x^\mu}{\partial s}, \quad U^\mu = \frac{\partial x^\mu}{\partial \lambda}$$

Extending this to a neighbourhood of Σ, we get a natural set of coordinates (s, λ, y^1, y^2). This defines the coordinate vectors $S = \partial_s$, ~~$U = \partial_\lambda$~~, which therefore commute: $\partial_s \partial_\lambda = \partial_\lambda \partial_s$, i.e. $U^b \nabla_b S^a = S^b \nabla_b U^a$.

From here, one can show that the deviation vector satisfies the geodesic deviation equation:

$$\boxed{\underbrace{U^c \nabla_c (U^b \nabla_b S^a)}_{\text{"tidal acceleration" of the geodesics}} = R^a{}_{bcd} U^b U^c S^d}$$

$\boxed{\triangleright}$: Recall the commutation relation $U^b \nabla_b S^a = S^b \nabla_b U^a$ (1)

$$U^c \nabla_c (U^b \nabla_b S^a) \overset{(1)}{=} U^c \nabla_c (S^b \nabla_b U^a) = (U^c \nabla_c S^b)(\nabla_b U^a) + U^c (\nabla_c \nabla_b U^a) S^b$$

$$\overset{(1)}{=} (S^c \nabla_c U^b)(\nabla_b U^a) + U^c S^b \nabla_c \nabla_b U^a$$

$$= (\nabla_b U^a)(S^c \nabla_c U^b) + U^c S^b \nabla_b \nabla_c U^a - U^c S^b R_{cbd}{}^a U^d$$

$$= S^c \nabla_c (\underbrace{U^b \nabla_b U^a}_{=0 \text{ (geod. eq.)}}) - S^b U^c R_{cbd}{}^a U^d$$

$$= R^a{}_{bcd} U^b U^c S^d$$ $\blacksquare$

Symmetries of Riemann tensor

DEF : Given an affinely parametrized geodesic γ w/ tangent vector U^a, a solution of the geodesic deviation equation along γ is called Jacobi field.

→ Aside from the "tidal acceleration" given by the geodesic deviation equation, another relevant quantity for the study of geodesic congruences is

$$B^a{}_b \equiv \nabla_b U^a,$$

which measures the failure of S^a to be parallelly transported along a geodesic (if $B^a{}_b = 0$, then
$U^b \nabla_b S^a = S^b \nabla_b U^a = B^a{}_b S^b = 0$ and S^a
is parallelly transported). Three relevant properties can be proven at this point.

PROP 2.1

(i) $B^a{}_b U^b = 0$; (ii) $B^a{}_b U_a = 0$ for affinely param. geod.

(iii) $U \cdot S = \text{cst}$ along the geodesic w/ tangent vector U

☒:

(i) $B^a{}_b U^b = (\nabla_b U^a) U^b = U^b \nabla_b U^a = 0$ (geod. eq.)

(ii) $B^a{}_b U_a = (\nabla_b U^a) U_a = \frac{1}{2} \nabla_b (U_a U^a) = \frac{1}{2} \nabla_b (\text{cst}) = 0$

(iii) $U^b \nabla_b (U^a S_a) = \underbrace{U^b (\nabla_b U^a)}_{=0 \text{ (geod. eq.)}} S_a + U^a U^b (\nabla_b S_a)$

$$= U^a B_{ab} S^b \overset{(ii)}{=} 0$$ ☒

2.2 The question of the gauge

When constructing geodesics (curves in general), we have the freedom to change the parameter λ of the curves without affecting their length (indeed, length is an integral, so invariability under reparametrizations follows from the theorem of the change of variables). However, vectors related to the curves (tangent, deviation) might be affected by reparametrizations. Let us study how and how to fix this "gauge" freedom in a way as natural as possible.

→ The tangent vector

Consider a curve defined as $x^\mu(\lambda)$, $\lambda \in I \subset \mathbb{R}$. Under a generic reparametrization $\lambda \to \tilde{\lambda}$, we have

$$x^\mu(\lambda) = \tilde{x}^\mu(\tilde{\lambda}) \Rightarrow \frac{\partial x^\mu}{\partial \lambda} = \frac{\partial \tilde{x}^\mu}{\partial \tilde{\lambda}} \frac{d\tilde{\lambda}}{d\lambda} \Rightarrow U^\mu = \tilde{U}^\mu \frac{d\tilde{\lambda}}{d\lambda}$$

Therefore, an affine reparametrization $\tilde{\lambda} = a\lambda + b$, $a, b \in \mathbb{R}$, $a \neq 0$ will reescale the tangent vector by a factor of a. This factor is customarily taken to normalize the tangent vector (for spacelike / timelike curves). Henceforth, we will consider this normalization factor to be fixed

→ The deviation vector

Consider a 1PFG defined as $x^\mu(\lambda, s)$, $\{\lambda, s\} \in I \times I' \subset \mathbb{R}^2$. Assuming an affine parametrization such that the norm of the tangent vector $U^\mu = \frac{\partial x^\mu}{\partial \lambda}$ is fixed, we still have a residual freedom to shift the affine parameter by an s-dependent function, $\lambda \to \tilde{\lambda} = \lambda - a(s)$. As per the previous point, the tangent vector remains fixed. However, the deviation vector transforms as

$$x^\mu(\lambda, s) = \tilde{x}^\mu(\tilde{\lambda}, s)$$

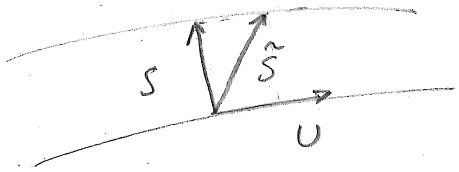
$$S^\mu = \frac{\partial x^\mu}{\partial s} = \frac{d\tilde{x}^\mu}{ds} = \frac{\partial \tilde{x}^\mu}{\partial \tilde{\lambda}} \frac{\partial \tilde{\lambda}}{\partial s} + \frac{\partial \tilde{x}^\mu}{\partial s} = \tilde{U}^\mu \frac{d\tilde{\lambda}}{ds} + \tilde{S}$$

Since $\tilde{U}^\mu = U^\mu$ and $\tilde{\lambda} = \lambda - a(s)$, we reach

$$S^\mu = \tilde{S}^\mu - U^\mu a'(s) \qquad (1)$$

N.B. S^μ and $\tilde{S}^\mu$ point to the same neighbouring geodesic.

Conclusion: reparametrizations of the type $\lambda \to \lambda - a(s)$ shift the deviation vector by a multiple of the tangent vector



→ Fixing the gauge in spacelike / timelike geodesics

From (1), $U \cdot \tilde{S} = U \cdot S + \frac{da}{ds} U^2$, so we can fix $a(s)$ so that $U \cdot S = 0$ at some point on each geodesic. Then, by virtue of PROP 2.1 (iii), $U \cdot S = 0$ everywhere. Intuitively, we remove the component of S colinear w/ U such that U and S are normal

→ Fixing the gauge in null geodesics

The above argument does not work because $a(s)$ does not affect the value of $U \cdot S$. Intuitively, the component of S colinear w/ U cannot be detected through the scalar product because $U \cdot U = 0$. We fix the gauge following a different procedure

1) Construct a vector field N such that

$$N^2 = 0, \quad U \cdot N = -1, \quad U^a \nabla_a N^b = 0 \ (*)$$

Such a field exists

$\boxed{\Sigma}$: Pick a spacelike hypersurface Σ intersecting every geodesic once. Define the vector field N^a on Σ such that $N^2|_\Sigma = 0$ and $N \cdot U|_\Sigma = -1$. Extend N^a by parallel transport along the geodesics. This implies that $U^a \nabla_a N^b = 0$ Furthermore, N^2 and $N \cdot U$ are constant along the geodesics (see below). Therefore, $N^2 = 0$ and $N \cdot U = -1$ everywhere.

$$U^a \nabla^a (N^2) = 2 N_b \underbrace{U^a \nabla^a N^b}_{(*)\ = 0} = 0$$

$$U^a \nabla^a (N^b U_b) = U_b \underbrace{U^a \nabla_a N^b}_{(*)\ = 0} + N^b \underbrace{U^a \nabla_a U^b}_{\substack{=0 \\ (\text{geod. eq.})}} = 0 \qquad \blacksquare$$

Notice that N is not colinear w/ U, otherwise $N \cdot U = 0$

2) Decompose any deviation vector as

$$S^a = \underbrace{\alpha U^a + \hat{S}}_{\text{orthogonal to } U} + \underbrace{\beta N}_{\substack{\text{non-orthogonal to } U \\ \text{parallely transported}}},$$

with $U \cdot \hat{S} = N \cdot \hat{S} = 0 \Rightarrow \hat{S}$ spacelike

Remarks (1): $U \cdot S = -\beta$, therefore βN accounts for the component of $\hat{S}$ non-orthogonal to U. Notice however that the choice of N is non-unique (we will see this is not relevant).

Remarks (2): If we modify S by a multiple of U (by a reparametrization $\lambda \to \lambda - a(s)$, $\hat{S}$ does not change (the change in S is absorbed in α) $\hat{S}$ can be thought as an equivalence class of vector differing by a multiple of U, therefore it belongs to a dimension 2 quotient subspace.

→ The (equivalence class of) deviation vector $\hat{S}$ can be written by projecting S into the dim-2 subspace orthogonal to U and N:

$$\hat{S}^a = P^a_b S^b, \quad P^a_b = \delta^a_b + N^a U_b + U^a N_b$$

Consistently with being such projector,

$$P^a_b U^b = U^a + 0 - U^a = 0$$

$$P^a_b N^b = 0$$

$$P^a_b P^b_c = \delta^a_c + \cancel{N^a U_c} + \cancel{U^a N_c} + N^a U_c - \cancel{N^a U_c} + 0$$
$$+ U^a N_c + 0 - \cancel{U^a N_c} = P^a_c \quad (\text{idempotent})$$

N.B. Since both U and N are parallelly transported along the geodesics, so is P^a_b.

Summary: we have projected S into a 2-d subspace in which, by construction, the gauge redundancy disappears.

But can we still study $\hat{S}^a = P^a_b S^b$ with the same tools described in the previous section? The next proposition proves we can, provided $U \cdot S = 0$ (which is the same condition imposed on timelike / spacelike congruences, and which we will later show that holds for congruences of generators of null hypersurfaces)

PROP 2.2: A deviation vector for which $U \cdot S = 0$ satisfies $U \cdot \nabla \hat{S}^a = \hat{B}^a{}_b \hat{S}^b$, w/ $\hat{B}^a{}_b = P^a_c B^c{}_d P^d_b$

(i.e. the definition of $B^a{}_b$ gets naturally projected into the 2-d subspace)

$\boxtimes$: $U\cdot\nabla \hat{S}^a = U^b \nabla_b (P^a_c S^c) \underset{\substack{\uparrow \\ P \text{ paral. transp.}}}{=} P^a_c U^b \nabla_b S^c = P^a_c B^c{}_d S^d$ (1)

$P^d_e S^e = (\delta^d_e + N^d U_e + U^d N_e) S^e = S^d + N^d \underbrace{U_e S^e}_{U\cdot S = 0} + U^d S^e N_e$ (2)

Since $B^c{}_d U^d = 0$ (prop 2.1 (ii)), $B^c{}_d P^d_e S^e = B^c{}_d S^d$ (3)

(3), (1) $\Rightarrow$ $U\cdot\nabla \hat{S}^a = P^a_c B^c{}_d P^d_e S^e \underset{\substack{\uparrow \\ P \text{ idempotent}}}{=} P^a_c B^c{}_d P^d_e P^e_f S^f$

$= \hat{B}^a{}_e \hat{S}^e$ $\boxtimes$

Remarks: Consider a IPFG contained in the generators of a null hypersurface $\mathcal{N}$. The deviation vector S is tangent to $\mathcal{N}$ (because the geodesics are contained in $\mathcal{N}$), while the tangent vector U is (by construction) normal to $\mathcal{N}$. Hence, $U\cdot S = 0$ ($\Rightarrow \beta = 0$), and the above proposition holds.

2.3 Expansion (and rotation and shear)

Given $\hat{B}^a{}_b$, we can decompose it as

$$\hat{B}^a{}_b = \frac{1}{2}\theta P^a_b + \hat{\sigma}^a{}_b + \hat{\omega}^a{}_b$$

→ $\theta = \hat{B}^a{}_a$ (the trace) is called <u>expansion</u>

→ $\hat{\sigma}_{ab} = \hat{B}_{(ab)} - \frac{1}{2} P^a_b\, \theta$ (traceless symmetric) is called <u>shear</u>

→ $\hat{\omega}_{ab} = \hat{B}_{[ab]}$ (antisymmetric) is called <u>rotation</u>

→ Cf. the timelike case in Wald, p. 217, and the change of dimensionality $2 \to 3$

We will be mostly interested in studying the expansion. Next, we show that its definition is independent of the choice of N in the decomposition of S and illustrate its meaning.

PROP 2.3 Defined as above, $\theta = g^{ab} B_{ab} = \nabla_a U^a$. In particular, it is independent on the choice of N entering the decomposition of S and the definition of P^a_b.

$\boxed{\triangleright}$: $\theta = \hat{B}^a{}_a = g^{ab} \hat{B}_{ab} = g^{ab} P^c_a B_{cd} P^d_b$

$= g^{ab} P^c_a B_{cd} (\delta^d_b + N^d U_b + \cancel{U^d N_b})$ bc. $B_{cd} U^d = 0$

$= g^{ab} (\delta^c_a + N^c U_a + U^c N_a)(B_{cb} + B_{cd} N^d U_b)$

$= g^{ab} (B_{ab} + B_{ad} N^d U_b + N^c U_a B_{cb} + N^c U_a B_{cd} N^d U_b$

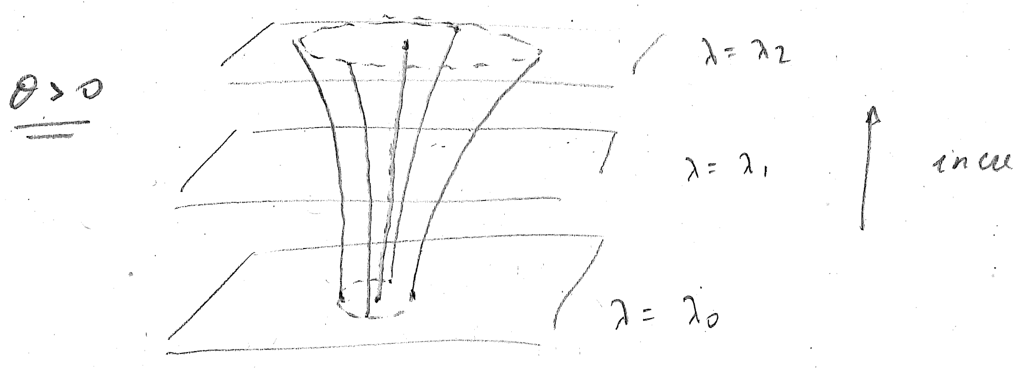
$+ \underbrace{U^c N_a B_{cb}}_{=0 \,:\, U^c B_{cb}} + \underbrace{U^c B_{cd} N^d U_b}_{=0 \,:\, U^c B_{cd}})$

$= g^{ab} B_{ab} + \underbrace{B_{ad} N^d U^a}_{=0 \,:\, B_{ad} U^a} + \underbrace{N^c U^b B_{cb}}_{=0 \,:\, B_{cb} U^b} + \underbrace{N^c U^b B_{cd} N^d U_b}_{=0 \,:\, U^b U_b}$

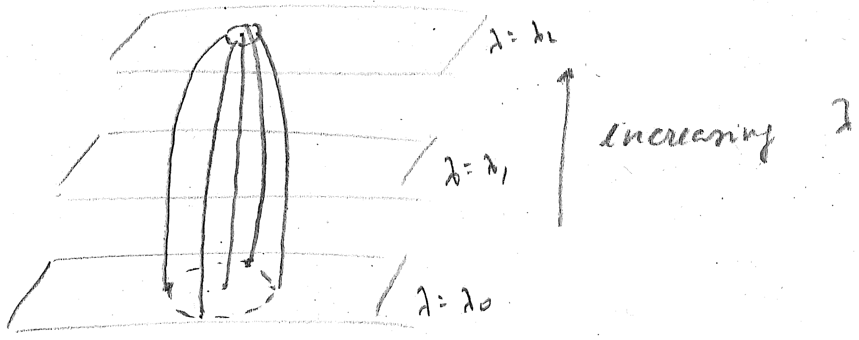
$= g^{ab} B_{ab} = \nabla_a U^a$ ∎

→ Illustration of the geometrical meaning of θ

θ measures the rate of increase of the area element of the surfaces of constant afine parameter in a 1PFG. Therefore, it gives idea of how neighbouring geodesics move apart $(\theta > 0)$ or closer $(\theta < 0)$

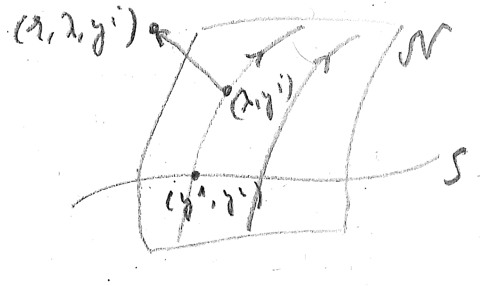


$\theta < 0$



To show this, consider a null hypersurface $\mathcal{N}$ and a 1PFG of generators of $\mathcal{N}$ w/ affine parameter λ. Pick a surface $\Sigma \subset \mathcal{N}$ and use it to define Gaussian null coordinates as follows:

1) Assign coordinates (y^1, y^2) to each point in Σ
2) To each point in $\mathcal{N}$ an affine distance λ from Σ (along the generator that cuts Σ in (y^1, y^2)), assign (y^1, y^2, λ)
3) Define a vector field V^a s.t. $V^a V_a|_{\mathcal{N}} = 0$, $V \cdot \partial_{y^i} = 0$ and $V \cdot U = 0$. Construct its null integral curves. To each point in a neighbourhood of $\mathcal{N}$ an affine distance r from (y^1, y^2, λ) along the integral geodesic of V that cuts $\mathcal{N}$ at (y^1, y^2, λ), assign (y^1, y^2, λ, r).



In $\mathcal{N}$ (i.e. at $r=0$), $g_{r\lambda}|_{\mathcal{N}} = \partial_r \cdot U = 1$ and $g_{ri}|_{\mathcal{N}} = V \cdot \partial_{y^i} = 0$. This is actually valid for all r because $g_{r\mu,r} = 0$ (which follows from $V^\mu \nabla_\mu V^\nu = 0$)

Also, $\underset{(r=0)}{g_{\lambda\lambda}|_{\mathcal{N}}} = U_\mu U^\mu = 0$ and $\underset{(r=0)}{g_{\lambda i}|_{\mathcal{N}}} = U^\mu \underbrace{(\partial_{y^i})_\mu}_{\text{tangent to } \mathcal{N} \Rightarrow \text{ orthogonal to } U} = 0$

$\Rightarrow$ $g_{\lambda\lambda} = rF$, $g_{\lambda i} = r h_i$
for smooth F, h_i

Therefore, the metric in this set of coordinates is

$$ds^2 = dr\,d\lambda + d\lambda\,dr + rF\,d\lambda^2 + r h_i\,(d\lambda\,dy^i + dy^i\,d\lambda) + h_{ij}\,dy^i\,dy^j$$

$$\Rightarrow \left.ds^2\right|_{\mathcal{N}\,(r\to 0)} = dr\,d\lambda + d\lambda\,dr + h_{ij}\,dy^i\,dy^j \qquad (1)$$

On a surface of constant λ within $\mathcal{N}$,

$\left.ds^2\right|_{\mathcal{N},\,\lambda=\text{cst}} = h_{ij}\,dy^i\,dy^j \Rightarrow h \equiv \det(h_{ij})$ is the area element of surfaces of constant λ within $\mathcal{N}$.

Now, $\left.U^\mu\right|_{\mathcal{N}} = (0,1,0,0) \overset{(1)}{\Longrightarrow} \left.U_\mu\right|_{\mathcal{N}} = (1,0,0,0)$, and from PROP 2.1, $U_\mu B^\mu{}_\nu = U^\mu B^\nu{}_\mu = 0$, therefore $\left.B^\lambda{}_\mu\right|_{\mathcal{N}} = \left.B^\mu{}_\lambda\right|_{\mathcal{N}} = 0$. PROP 2.3 prescribes to compute θ as $\theta = B^\mu{}_\mu$, so $\left.\theta\right|_{\mathcal{N}} = B^i{}_i = \nabla_i U^i = \partial_i U^i + \Gamma^i_{i\mu} U^\mu$

$$= \underline{\underline{\Gamma^i_{i\lambda}}} \overset{(1)}{=} \frac{1}{2} h^{ij} h_{ij,\lambda} = \frac{\partial \log\sqrt{h}}{\partial\lambda}$$

Jacobi's identity

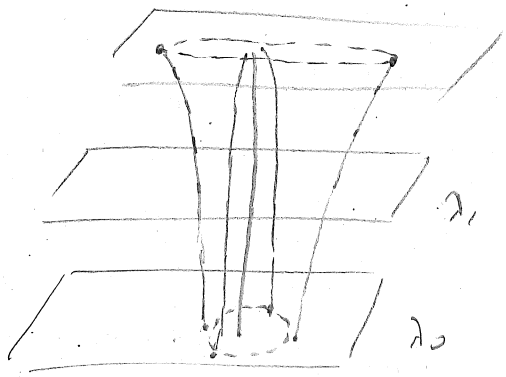
$\Rightarrow \theta\sqrt{h} = \dfrac{\partial\sqrt{h}}{\partial\lambda} \rightarrow \theta$ gives the rate of increase of the area element in surfaces of constant affine parameter as one moves along the geodesics. ∎

Two final comments on the shear and the rotation

1) The rotation of a congruence of geodesics containing the generators of a null hypersurface $\mathcal{N}$ is zero on $\mathcal{N}$ ($\left.\hat{\omega}_{ab}\right|_{\mathcal{N}} = 0$ for a IPFG of generators of $\mathcal{N}$)

→ See proof in p. 52 of Reall's notes, the proof is easier if one uses the wedge product.

2) Like with the expansion, it is easy to give a geometric interpretation of the shear as the "distorsion" of an area element (more precisely, the rate of change of the semiaxis of an originally circular ellipse).



3. SOME TOPICS RELATED TO θ

3.1 Raychaudhuri's equation

Raychaudhuri's equation determines the evolution of θ along the geodesics of a congruence. For null geodesics, it is

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \hat{\sigma}^{ab}\hat{\sigma}_{ab} + \hat{\omega}^{ab}\hat{\omega}_{ab} - R_{ab}U^a U^b$$

(Remark: For timelike geodesics, remove hats and $\frac{1}{2} \to \frac{1}{3}$, see Wald p. 218)

$\boxtimes$: $\displaystyle \frac{d\theta}{d\lambda} = U\cdot\nabla(B^a{}_b P^b_a) \underset{\substack{\uparrow \\ P \text{ trans. par.}}}{=} P^b_a\, U\cdot\nabla B^a{}_b = P^b_a U^c \nabla_c \nabla_b U^a$

$$= P^b_a\, \underbrace{U^c(\nabla_b \nabla_c U^a}_{\longrightarrow} + R^a{}_{dcb}U^b) \quad \text{(used def. of Riemann)}$$

$$= P^b_a\left[\nabla_b(\underbrace{U^c\nabla_c U^a}_{=0 \text{ (geod. eq.)}}) - (\nabla_b U^c)(\nabla_c U^a)\right] + P^b_a R^a{}_{dcb} U^c U^d$$

$$= -P^b_a B^c{}_b B^a{}_c - \underbrace{P^b_a R^a{}_{dcb} U^c U^d}_{=(*)}$$

$$(*) = P^b_a\, g^{ae}\, R_{edcb}\, U^c U^d = g^{ae}\, U^c U^d \big(\delta^b_a + \underbrace{\cancel{N^b U_c + U^c N_b}}_{\text{sym. in } bc}\big) \underbrace{R_{edbc}}_{\text{antisym}}$$

$$= g^{ae}\, U^c U^d\, \delta^b_a\, R_{edcb} = R^b{}_{dcb}\, U^c U^d = - R^b{}_{dbc} U^c U^d$$

$$= - R_{cd}\, U^c U^d \qquad \text{(Ricci is symmetric)}$$

Therefore,

$$\frac{d\theta}{d\lambda} = - P^b_a\, B^c{}_b\, B^a{}_c - R_{cd}\, U^c U^d \qquad (**)$$

Also, notice that

$$\hat{B}^c{}_a\, \hat{B}^a{}_c = P^c_\ell\, B^\ell{}_m \underbrace{P^m_a\, P^a_r}_{P^m_r} B^r{}_s\, P^s_c = B^\ell{}_m\, B^r{}_s\, P^m_r\, P^s_\ell$$

$$= B^\ell{}_m\, B^r{}_s\, P^m_r \underbrace{(\delta^s_\ell + U^s N_\ell \cancel{+ N^s U_\ell})}_{\text{Prop 2.1}} = \underbrace{B^\ell{}_m\, P^m_r\, B^r{}_\ell}_{\text{1st term in RHS above } (**)}$$

Therefore

$$\frac{d\theta}{d\lambda} = - \hat{B}^c{}_a\, \hat{B}^a{}_c - R_{cd}\, U^c U^d \qquad (***)$$

The first term reads

$$\hat{B}^c{}_a\, \hat{B}^a{}_c = \hat{B}^c{}_a \left(\tfrac{1}{2}\theta\, P^a_c + \hat{\sigma}^a{}_c + \hat{\omega}^a{}_c \right)$$

1) $\frac{1}{2}\theta\, P^a_c\, \hat{B}^c{}_a = \frac{1}{2}\theta\, P^a_c\, P^c_b\, B^b{}_d\, P^d_a = \frac{1}{2}\theta\, P^d_b\, B^b{}_d$

$$= \tfrac{1}{2}\theta \underbrace{(\delta^d_b + \cancel{U^d N_b + N^d U_b})}_{\text{Prop. 2.1}} B^b{}_d = \tfrac{1}{2}\theta\, B^b{}_b = \tfrac{1}{2}\theta^2$$

2) $B^c{}_a\, \hat{\sigma}^a{}_c = \hat{\sigma}^a{}_c \left(\frac{1}{2} P^c_a\, \theta + \hat{\sigma}^c{}_a + \hat{\omega}^c{}_a \right)$

$$= \tfrac{1}{2}\theta \underbrace{\hat{\sigma}^a{}_a}_{=0 \text{ (traceless)}} + \hat{\sigma}^a{}_c\, \hat{\sigma}^c{}_a + \hat{\sigma}^a{}_c\, \hat{\omega}^c{}_a$$

$$= \hat{\sigma}^{ac}\, \hat{\sigma}_{ac} - \hat{\sigma}^{ac}\, \hat{\omega}_{ac} \qquad (\sigma \text{ sym; } \omega \text{ antisym})$$

3) $B^c{}_a\, \hat{\omega}^a{}_c = \underbrace{0}_{\text{traceless}} + \hat{\omega}^a{}_c\, \hat{\sigma}^c{}_a + \hat{\omega}^a{}_c\, \hat{\omega}^c{}_a$

$$= \hat{\omega}^{ac}\, \hat{\sigma}_{ac} - \hat{\omega}^{ac}\, \hat{\omega}_{ac}$$

Plugging back in (***)

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \hat{\sigma}^{ac}\hat{\sigma}_{ac} + \hat{\omega}^{ac}\hat{\omega}_{ac} - R_{ab}U^aU^b$$

→ Recall the <u>Null Energy condition</u> (NEC):

$$\boxed{\text{For all } V^a \text{ null, } T_{ab}V^aV^b \geq 0 \quad \text{(NEC)}}$$

Quick reminder: <u>energy conditions</u>

They allow to rewrite conditions on the Ricci tensor as conditions on the energy-momentum tensor (passing by the Einstein equations). The most important ones are:

- Weak energy condition: $T_{ab}V^aV^b \geq 0 \;\; \forall V^a$ causal
- Null energy condition: $T_{ab}V^aV^b \geq 0 \;\; \forall V^a$ null
- Dominant energy condition: WEC and $-T^a{}_bV^b$ is a future-directed causal vector (or zero) for all future-directed timelike vectors.
- Strong energy condition: $(T_{ab} - \frac{1}{2}g_{ab}T^c{}_c)V^aV^b \geq 0$ $\forall V^a$ causal

Remark (1): DEC ⇒ WEC ⇒ NEC
SEC ⇒ NEC
<u>but</u> SEC ⇏ WEC

Remark (2): A positive cosmological constant respects DEC (and therefore WEC, NEC) but <u>does not</u> respect SEC.

Remark (3): For Penrose's theorem (1965), only the NEC is relevant; other energy conditions are relevant for other singularity theorems.

Provided the NEC is fulfilled, and following Raychaudhuri's equation, we can prove the following lemma.

<u>LEMMA</u> : Provided the NEC holds, the generators of a null hypersurface satisfy

$$\frac{d\theta}{d\lambda} \leq -\frac{1}{2}\theta^2$$

$\boxtimes$: From the last comment in section 2.3, $\hat{\omega}=0$. As explained in section 2.3, $\hat{S}$ are spacelike, so the metric restricted to the 2-d subspace in which $\hat{S}$ are is positive-definite, hence $\hat{\sigma}^{ab}\hat{\sigma}_{ab} \geq 0$. From Einstein equations,

$$\left(R_{ab} - \frac{1}{2}R\, g_{ab}\right)U^a U^b = 8\pi T_{ab}\, U^a U^b,$$

which for null U^a reduces to $R_{ab}U^aU^b = 8\pi T_{ab}U^aU^b$. Using the NEC, $R_{ab}U^aU^b \geq 0$. Plugging all the above in the RHS of Raychaudhuri's equation,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - (\text{something} \geq 0) \Rightarrow \frac{d\theta}{d\lambda} \leq -\frac{1}{2}\theta^2 \quad \blacksquare$$

From this lemma follows a very important corollary.

<u>COR. 2.4</u>. Assume the NEC. If $\theta = \theta_0 < 0$ at some point p on a generator γ of a null hypersurface, then $\theta \to -\infty$ along γ within affine parameter distance $2/|\theta_0|$ <u>provided γ extends that far</u>.

$\boxtimes$: WLOG, let $\lambda = 0$ at p. From the LEMMA above (which holds thanks to the NEC),

$$-\frac{1}{\theta^2}\frac{d\theta}{d\lambda} \geq 1/2 \Rightarrow \frac{d\theta^{-1}}{d\lambda} \geq 1/2$$

Integrating w.r.t. λ, $\theta^{-1} - \theta_0^{-1} \geq \lambda/2$. Solving the inequality for θ,

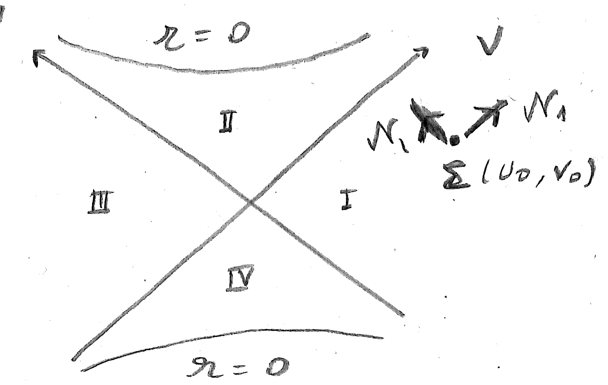
$$\theta \leq \frac{\theta_0}{1 + \lambda\theta_0/2} \xrightarrow[\lambda \to 2/|\theta_0|]{} -\infty \quad (\text{bc. } \theta_0 < 0) \ \blacksquare$$

3.2 Trapped surfaces

Given a 2d, orientable spacelike surface Σ, for every point p there are two future directed null vectors $U^a_{(1)}$, $U^a_{(2)}$. These fields define two families of null geodesics starting on Σ and orthogonal to it, when themselves are the generators of two null hypersurfaces N_1, N_2. By analogy with simple situations, they are called "ingoing" and "outgoing" null geodesics. Intuitively, ingoing hypersurfaces have $\theta_1 < 0$ (geodesics focus) and outgoing hypersurfaces have $\theta_2 > 0$ (geodesics disperse). However, in certain situations, both families of geodesics behave equally. If $\theta_i < 0$, $i = 1,2$, we say Σ is a "trapped surface" (more formally defined below).

<u>DEF</u>: A compact, orientable, spacelike, 2d-surface Σ is <u>trapped</u> if both families of null geodesics orthogonal to it have $\theta < 0$ everywhere on Σ.

E.g. Kruskal spacetime (cf. example after PROP 1.1)



A point (U_0, V_0) in the above diagram is a 2-sphere (spatial surface). The null hypersurfaces orthogonal to Σ, N_1 and N_2 are of the form $U=$ cst, $V=$ cst. The corresponding tangent vectors (with signed fixed to be future oriented ones)

$$U_{(1)} \propto r e^{r/2M} \partial_V, \quad U_{(2)} \propto r e^{r/2M} \partial_U$$

The expansion is

$$\theta_1 = \nabla_a U_{(1)}{}^a = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g}\, U_{(1)}{}^\mu) = 2 e^{r/2M} \partial_V r$$

$$ds^2 = -\frac{32M^3 e^{-r/2M}}{r} dU\, dV + r^2 d\Omega^2$$

Recall $UV = -e^{r/2M}\left(\frac{r}{2M} - 1\right)$. Taking ∂_V on both sides, $U = -\frac{r e^{r/2M}}{4M^2} \partial_V r \Rightarrow \partial_V r = -\frac{4M^2 U}{r e^{r/2M}}$

so $\boxed{\theta_1 = -\frac{8M^2}{r} U}$

Similarly, $\boxed{\theta_2 = -\frac{8M^2}{r} V}$

Region I: $\theta_1 > 0$ (outgoing expand), $\theta_2 < 0$ (ingoing focus)
(intuitive, this is exterior solution)

Region IV: $\theta_1 < 0$, $\theta_2 > 0$
(same as region I w/ a flipped sign bc. of the isometry relation between exterior solutions)

Region II ("BH"): $\theta_1 < 0$, $\theta_2 < 0$ $\rightarrow$ every (v_0, s_0) 2-sphere w/ $v_0, s_0 \in$ Region II is a trapped surface (all null geod. focus)

Region III ("WH"): $\theta_1 > 0$, $\theta_2 > 0$; the opposite to region II, all geodesics diverge.

3.3. Conjugate points

DEF: Points p, q on a geodesic are conjugate if there exists a Jacobi field (cf Sec 2.1) which is non-zero but vanishes at p and q.

The presence of conjugate points in a congruence (i.e. locally vanishing deviation vectors) can be associated with the existence of caustics (focusing of geodesics):

THM 3.1: Consider a null geodesic congruence which includes all of the null geodesics through p^*. If $\theta \rightarrow -\infty$ at a point q on some null geodesic γ through p, then q is conjugate to p along γ.

* In a congruence, only one geodesic passes through each point, so p is not strictly part of the congruence

$\boxtimes$: See p. 224 Wald.

The concept of conjugate points can be extended to define the conjugate point to a 2d spacelike surface Σ. Let U^a be one of the two future-directed null vector fields normal to Σ and γ a null geodesic along one of the corresponding null hypersurfaces N. We say a point $p \in \gamma$

is conjugate to Σ if there exists a non-zero Jacobi field tangent to Σ that vanishes at p. (see p. 230 Wald)

The results so far are enough to prove the singularity theorem we are interested in. But there is a lot more to study about them, in particular related to the infinitesimal variation of the length of curves. Below are some important results and references

More on conjugate points

<u>Thm:</u> Let γ be a smooth timelike curve. The necessary and sufficient condition for γ to maximize the proper time between p and q is that γ is a <u>geodesic</u> with <u>no conjugate points</u> between p and q.

$\boxed{\Sigma}$: See thm. 9.3.3 (Wald) and prop. 4.5.8 (Hawking and Ellis)

Null curves have zero length, so the above does not apply. Yet, conjugate points control whether null curves can be deformed into timelike curves

<u>Thm</u>: Let μ be a smooth causal curve, $p, q \in \mu$. There does not exist a smooth 1-param family of causal curves λ_α connecting p, q w/ $\lambda_0 = \mu$, $\lambda_{\alpha \neq 0}$ timelike iff μ is a null geodesic w/ no point conjugate to p along μ between p and q.

$\boxed{\Sigma}$: See thm. 9.3.8 (Wald) and props. 4.5.11 and 4.5.12 (Hawking and Ellis).

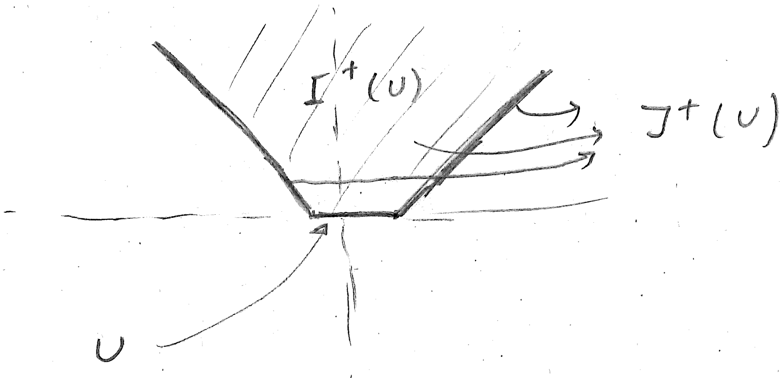
4. CAUSAL STRUCTURE

DEF: Let (M, g) be a time-orientable spacetime, $U \subset M$.

1) Chronological future $I^+(U)$: set of points in M reachable from U by a future-directed timelike curve
2) Causal future $J^+(U)$: set of points in M reachable from U by a future-directed causal curve

Chronological past $I^-(U)$ and causal past $J^-(U)$ are defined analogously.

In Minkowski, the topology of these sets is very simple:



$\boxed{\text{⚠ In Minkowski!}}$

$I^+(U)$ is made of timelike geodesics
$\dot{I}^+(U) = \overline{I^+(U)} \setminus I^+(U)$ is made of null geodesics.

For a more generic spacetime, this is only true locally

THM 4.1: Given $p \in M$, there exists a convex normal neighbourhood of p, $U \subset M$. For any $q, r \in U$, there exists a unique geodesic $\gamma \subset U$ connecting q, r.
$I^+(p) \cap U$ is made of timelike geodesics
$\dot{I}^+(p) \cap U$ is made of null geodesics
Like in Minkowski!

$\boxtimes$: See prop. 4.5.1 in Hawking and Ellis

Remark (1): The theorem is stronger than just saying "Spacetime is locally flat". We are not just saying the topology of M is locally that of $\mathbb{R}^4$, but that is causal structure is precisely that of Minkowski

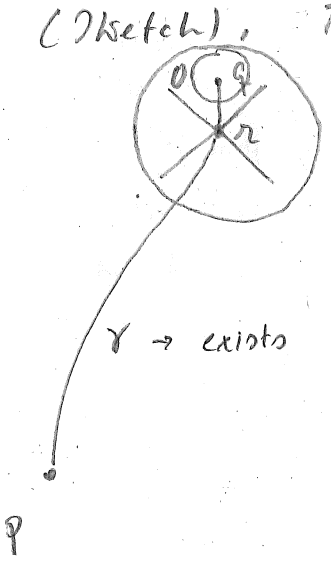
Remark (2): Locally, every Lorentzian metric looks like Minkowski, but this is not what the theorem talks about. The metric tells us the algebraic structure of the tangent bundle $T_p \notin M$, while this theorem tells us the causal structure of $U_p \subset M$.

COR. 4.2: If $q \in J^+(p) \setminus I^+(p)$, there exists a null geodesic from p to q.

$\boxtimes$: We know a causal curve exists from p to q. Such a curve is the continuous image of a compact set (e.g. $I = [0,1] \subset \mathbb{R}$), therefore compact. Therefore, we can cover it with a finite amount of convex normal neighbourhoods and apply THM 4.1 for each of them. (Rigorous proof → Prop. 4.5.10 Hawking-Ellis) $\blacksquare$

→ $I^+(p)$ is an open set

$\boxtimes$: (Sketch). Take $q \in I^+(p)$. A small neighbourhood O of q is also in $I^+(p)$. Indeed,



U → here, Minkowski holds (THM 4.1)

γ → exists bc $q \in I^+(p)$

$O \subset I^+(r) \Rightarrow O \subset I^+(p)$ and O is an open set which contains q $\blacksquare$

Some extra results that can be proven from here:

1) $I^+(I^+(S)) = I^+(S)$

2) $I^+(S) = I^+(\bar{S})$

3) If a null geodesic joins p and q, then $q \in \overline{I^+(p)}$

4) $\overline{J^+(S)} = \overline{I^+(S)}$

5) $I^+(S) = \text{int}\,[\overline{I^+(S)}] = \text{int}\,[J^+(S)]$

6) $\dot{I}^+(S) = \dot{J}^+(S)$

See Wald, Ch. 8, and Hawking-Ellis, Ch. 6

<u>DEF</u>: $S \subset M$ is <u>achronal</u> if no two points in S are joined by a timelike curve.

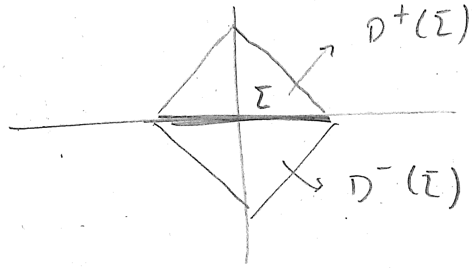
<u>DEF</u>: Let $p \in M$ sit on a curve $\gamma \subset M$, we say p is a <u>future endpoint</u> of γ if for all neighbourhood U of p, there exists λ_0 such that $\gamma(\lambda)$ stays in U for $\lambda > \lambda_0$ (i.e. the curve "does not go beyond p".

<u>DEF</u>: A curve is <u>inextendable to the future</u> if it has no future endpoint.

<u>N.B.</u> These definitions are at the core of singularity theorems, as they allow to define <u>incomplete curves</u>. A curve is incomplete when it has <u>no future (past) endpoint</u>, and yet, its <u>affine parameter remains finite</u>. For these to detect real singularities, we have to consider spacetimes in which points have not been artificially removed: inextendable spacetimes (i.e. spacetimes which are not isometric to a proper subset of another spacetime).

DEF: Let $\Sigma \subset M$ be closed and achronal. We define its future domain of dependence as the set of points $p \in M$ such that every past-directed causal inextendible curve through p crosses Σ (i.e. the part of spacetime which is fully determined causally by Σ). We denote it $D^+(\Sigma)$. A past domain of dependence is defined analogously.

In Minkowski



The domain of dependence is defined as $D(\Sigma) = D^+(\Sigma) \cup D^-(\Sigma)$.

DEF: A closed, achronal set Σ for which $D(\Sigma) = M$ is called Cauchy surface. A spacetime which has a Cauchy surface is called globally hyperbolic.

Finally, we state three theorems that characterize the behaviour of $\dot{J}^+(U)$, and which will be necessary to prove the singularity theorem.

THM 4.3: Let $U \subset M$. Then $\dot{J}^+(U)$ is an achronal 3d submanifold of M.

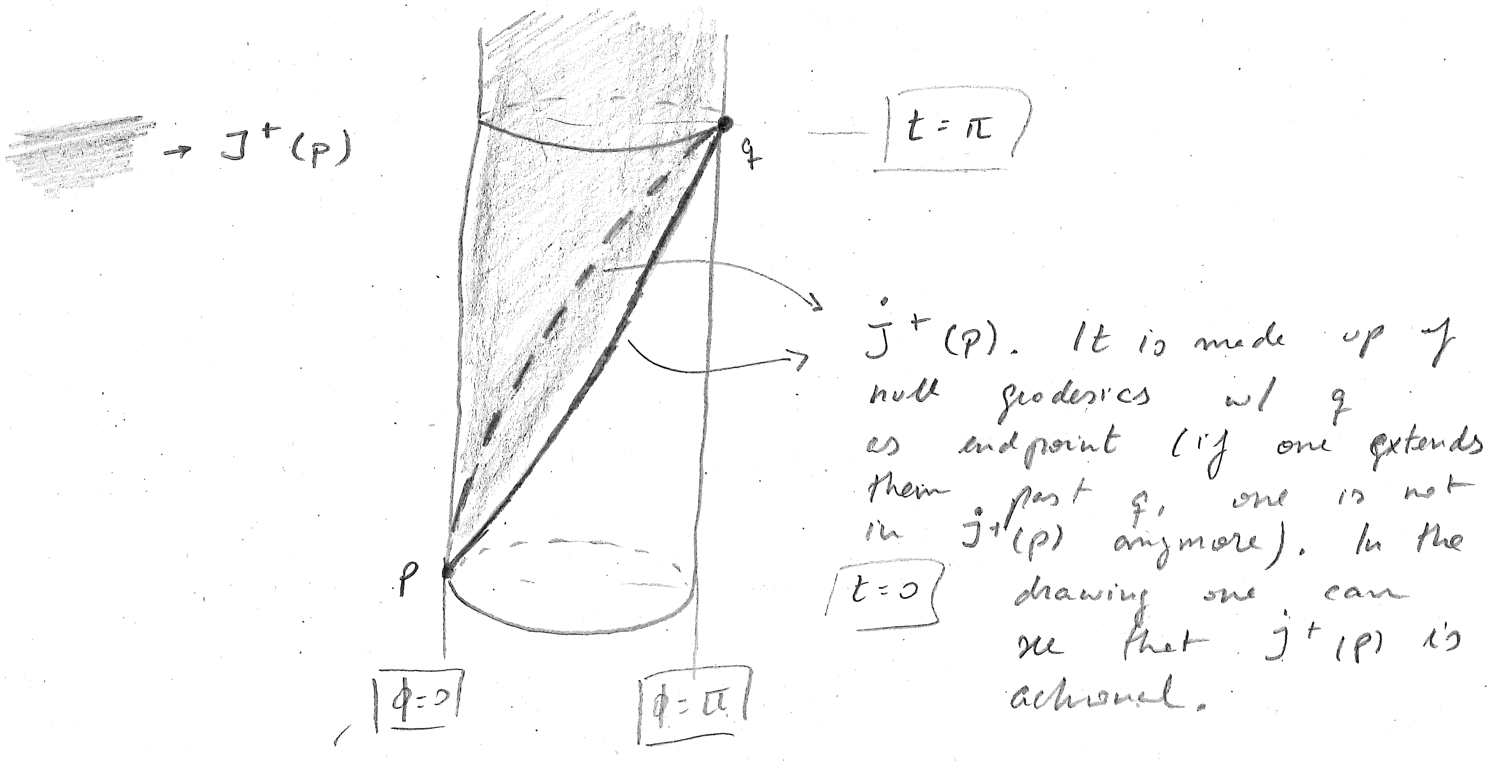
$\boxtimes$: (Sketch, see details in thm 8.1.3 in Wald)

To prove the achronality, take $q \in \dot{I}(U)$. One can see that $I^+(q) \subset I^+(U)$. Indeed, if $p \in I^+(q)$, then $q \in I^-(p)$, and since $I^-(p)$ is open and

$q \in \dot{I}^+(U)$, $\mathcal{O} \cap I^+(U) \neq \emptyset$ for some neighbourhood of q $\mathcal{O} \subset I^-(p)$, which implies $p \in I^+[\mathcal{O} \cap I^+(U)] \subset I^+(U)$. Analogously, one sees that $I^-(q) \subset M \setminus I^+(U)$. If $\dot{I}^+(U)$ were not achronal, there would exist $q, r \in \dot{I}(U)$ with $r \in I^+(q) \subset I^+(U) \Rightarrow I^+(U) \cap \dot{I}^+(U) \neq \emptyset$ which contradicts $I^+(U)$ being open. ($I^+(U)$ open $\Leftrightarrow I^+(U) = \mathrm{int}[I^+(U)]$)

To prove the manifold structure, we construct a homeomorphism (locally) between $\dot{I}^+(U)$ and $\mathbb{R}^3$. This is done by introducing Riemannian normal coordinates in a neighbourhood of each $q \in \dot{I}^+(U)$. The integral curves of the timelike coordinate vector ∂_{x_0} cut $\dot{I}^+(U)$ at one point <u>and only one</u> (because of the aforementioned achronality). This naturally induces a one-to-one map between each $q \in \dot{I}^+(U)$ and the coordinates (x^1, x^2, x^3) defining the curve that cuts $\dot{I}^+(U)$ at q. ∎

E.g. Einstein static universe in 2d w/ $M = \mathbb{R} \times S^1$



$$ds^2 = -dt^2 + d\phi^2$$

THM 4.4: Let $U \subset M$ closed. Then every $p \in \dot{J}^+(U)$ with $p \notin U$ lies on a null geodesic λ lying entirely in $\dot{J}^+(U)$ and such that λ is either past-inextendible or has a past endpoint on U

$\boxtimes$: (Sketch; see Thm. 8.1.6 in Wald for details)

We use the following lemma (Lemma 8.1.5 in Wald, see details also in Lemma 6.2.1 in Hawking-Ellis).

> Lemma: Let $\{\lambda_n\}$ be a sequence of future inextendible causal curves w/ a limit point p. Then, there exists a future inextendible causal curve λ passing through p which is a limit curve of the sequence $\{\lambda_n\}$

Take $M \setminus U$, which is a manifold because U is closed. Choose a sequence of points $\{q_n\} \subset I^+(U)$ such that $q_n \xrightarrow[n]{} p$. For each q_n, take a curve λ_n connecting it to U; each λ_n is therefore inextendible in $M \setminus U$. By the previous lemma a past inextendible causal curve λ passes through p. By construction, $\lambda \subset \overline{I^+(U)}$. Moreover, since $p \in \dot{I}^+(U)$, λ must be a null geodesic lying in $\dot{I}^+(U)$ (recall COR 4.2). Finally, since λ is past inextendible in $M \setminus U$, it is either past inextendible in M or has and endpoint in U $\boxtimes$

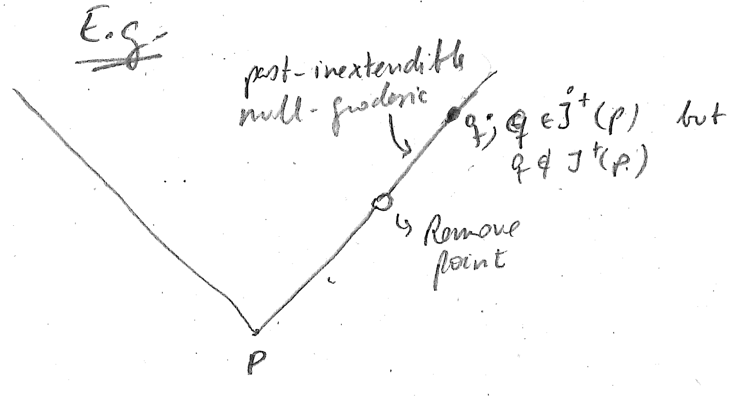
THM 4.4 can be strengthened in the case of global hyperbolicity.

<u>THM 4.5</u>: Let Σ be a 2d orientable, compact, spacelike submanifold of a <u>globally hyperbolic</u> spacetime. Every $p \in \dot{J}^+(\Sigma)$ lies on a future-directed null geodesic starting from Σ, orthogonal to Σ and with no conjugate point to Σ between Σ and p.

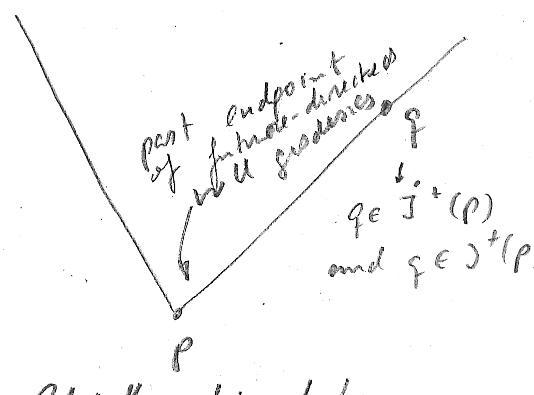
$\boxed{\Sigma}$: The proof can be found in thm 9.3.11 in Wald. Here, we just note to remarks.

1. The role played by the global hyperbolicity hypothesis is to remove the possibility of having a past inextendible null geodesic. In the globally hyperbolic case, the null geodesic must reach Σ
2. The absence of conjugate points along the null geodesic is related to the last theorem in Sec. 3.3. If there were such conjugate points, the curve joining Σ and p could be deformed into a timelike curve, which would be in contradiction w/ $p \in \dot{J}^+(\Sigma)$ (since $I^+(\Sigma) \cap \dot{J}^+(\Sigma) = \emptyset$).

<u>E.g.</u>



Non-globally hiperbolic



Globally hiperbolic

5. PENROSE SINGULARITY THEOREM

Let (M,g) be a globally hyperbolic spacetime with a non-compact Cauchy surface Σ. Assume that the null energy condition is satisfied and that M contains a trapped surface T. Let $\theta_0 < 0$ be the minimum value of θ on T for both sets of null geodesics orthogonal to T. Then, at least one of these geodesics is future inextendible and has affine length no greater than $2/|\theta_0|$ (i.e. at least one geodesic is incomplete).

$\boxed{D}$: We reason by reductio ad absurdum. Assume that all future inextendible null geodesics orthogonal to T have affine length greater than $2/|\theta_0|$. We will reach a contradiction in three steps.

1) $\dot{J}^+(T)$ is a compact set

From COR 2.4 along any of these null geodesics, we will reach $\theta \to -\infty$. From THM 3.1, a conjugate point to T is found along these geodesics (at most, at affine distance $2/|\theta_0|$).

Let $p \in \dot{J}^+(T) \setminus T$. From THM 4.5, p lies on a future directed null geodesic γ starting from T which is orthogonal to T and has no point conjugate to T between T and p, which implies p cannot lie beyond the point on γ conjugate to T, therefore it cannot lie at an affine distance from T larger than $2/|\theta_0|$. We conclude $\dot{J}^+(T)$ is a subset of a compact set. (see more formal proof in the box below).

$$f: A = \underbrace{T}_{\text{compact}} \times \underbrace{[0, 2/t_{0\cdot 1}]}_{\text{compact}} \longrightarrow f(A) \subset M$$

$\overbrace{\phantom{T \times [0, 2/t_{0\cdot1}]}}^{\text{compact}}$

$$(p, \lambda) \longmapsto f(p, \lambda)$$

Point along the null geodesic that exits Σ at p at an affine distance λ

f is continuous in p ("Cauchy stability") and λ, therefore $f(A) = \hat{A}$ is compact.

$\dot{J}^+(T) \subset \hat{A}$ compact

Since $\dot{J}^+(T)$ is closed and it is a subset of a compact set, then $\dot{J}^+(T)$ is itself compact.

2) $\dot{J}^+(T)$ is homeomorphic to Σ.

Pick a timelike vector field T^a and construct its integral curves. By definition of Cauchy surface, these curves will intersect Σ precisely once. By the achronality of $\dot{J}^+(T)$ (THM 4.3), these curves intersect $\dot{J}^+(T)$ at most once. Therefore, mapping points of $\dot{J}^+(T)$ into Σ through these timelike integral curves, we construct a continuous, one-to-one map $\alpha: \dot{J}^+(T) \to \Sigma$. Restricting the arrival to $\alpha[\dot{J}^+(T)]$, then $\alpha: \dot{J}^+(T) \to \alpha[\dot{J}^+(T)] \subset \Sigma$ is a homeomorphism.

Since $\dot{J}^+(T)$ is compact, so is $\alpha[\dot{J}^+(T)]$ (we only need continuity of α for this).

Since $\alpha[\dot{J}^+(T)] \subset \Sigma$, Σ Haussdorf (this is a consequence of assuming M Haussdorf), then $\alpha[\dot{J}^+(T)]$ is closed in the topology of Σ.

At the same time, $\dot{J}^+(T)$ is a 3d manifold (THM 4.3), so it is locally homeomorphic to $\mathbb{R}^3$, which implies $\alpha[\dot{J}^+(T)]$ is open (again, in the 3d topology of Σ). Since Σ is connected (this is a consequence of assuming M is connected), and $\alpha[\dot{J}^+(T)] \subset \Sigma$ is both open and close in the topology of Σ, either $\alpha[\dot{J}^+(T)] = \phi$ or $\alpha[\dot{J}^+(T)] = \Sigma$. The former cannot be by construction. Therefore, $\alpha[\dot{J}^+(T)] = \Sigma$.

3) The contradiction

$\dot{J}^+(T)$ is compact (1), therefore so is $\Sigma = \alpha[\dot{J}^+(T)]$ (2), but by hypothesis Σ is not compact #

$\Rightarrow$ $\boxed{\text{At least one inextendable null geodesic has finite affine length.}}$

■

<u>N.B.</u> The hypothesis of the Cauchy surface not being compact is a reasonable one, as this is the case for every globally-hyperbolic, asymptotically Minkowski (at spatial infinity) spacetime, see 1901.03928, p. 55.

APPENDIX : Very quick reminder of topology

→ Given a set A, we denote $\text{int}(A)$ its interior, $\mathring{A}$ its boundary, $\bar{A}$ its closure. In general, $\text{int}\, A \cap \mathring{A} = \emptyset$ and $\bar{A} = \text{int}\, A \cup \mathring{A}$

→ A set A is open if for every $p \in A$, a neighbourhood of p is contained in A. A set is closed if its complementary is open. $\begin{cases} \text{int}(A) \to \text{open} \\ \mathring{A} \to \text{closed} \end{cases}$

→ A set A is compact if it can be covered by a finite number of open subsets

→ A space A is Haussdorf if two points $p, q \in A$, $p \neq q$, can be isolated in disjoint neighbourhoods, i.e. "it is possible to tell different points appart"

→ A space A is connected if it cannot be subdivided into disjoint non-empty subsets, i.e "it cannot be split in separate pieces"

N.B. In general, spacetime is assumed to be connected and Haussdorf

→ Closed subsets of compact sets are also compact

→ Compact subsets of closed sets in Haussdorf spaces are also closed.

→ A continuous map with continuous inverse which is both one-to-one and onto is a homeomorphism.

→ Continuous image of a compact set is compact

→ Homeomorphic image of open (closed) set is open (closed)

→ Connected spaces have only two subsets that are both open and close: the empty and the total.

→ A n-dim manifold is locally homeomorphic to $\mathbb{R}^n$.