Singularity theorems

Literature:
- Senovilla, Garfinkle: "The 1965 Penrose sing. theorem"
- Pun Wai Tong: "Causality, Conjugate points and Singularity thms in spacetime"

1 Introduction

The Einstein equation:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi T_{\mu\nu}$$

Some of the earliest solutions (e.g. Schwarzschild and Friedman spacetimes) exhibit singular behaviour

Schwarzschild: $ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2\left(d\theta^2 + \sin^2\theta \, d\varphi^2\right)$

- spherically symmetric and static vacuum solution $(R_{\mu\nu} = 0)$

$\rightarrow$ $R = 0$, $R_{\mu\nu} R^{\mu\nu} =: S = 0$ by definition

but the Kretschmann scalar $K := R_{\mu\nu\rho\lambda} R^{\mu\nu\rho\lambda} \sim r^{-6}$ and so it diverges as $r \to 0$, which implies that singular behaviour is coordinate independent, unlike the $r = 2M$ singularity which is removed by a suitable coordinate transformation

$\rightarrow$ we wish to understand singular metrics and define them in a coordinate independent way

Motivation: we want to develop a systematic way of dealing with global causal structure via curves in spacetime

2 Causal structure

Def Chronological future $I^+(p)$ of a point $p \in M$ (manifold) is a set of points $q \in M$ such that there exists a timelike curve $\lambda : [0,1] \to M$ that is future directed and $\lambda(0) = p$, $\lambda(1) = q$. Analogously, we define the chronological past $I^-(p)$.

For a set S we define the chronological future $I^+(S) := \bigcup_{p \in S} I^+(p)$ and $J^+(p)$ is the causal future defined as the set which can be reached from p by timelike or lightlike curves.

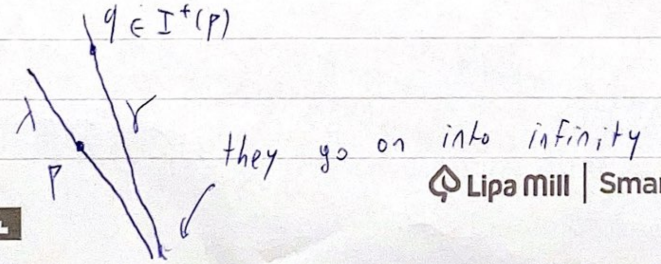
Corollary Let $q \in J^+(p) - I^+(p)$. Then every causal curve which connects p and q must be a lightlike geodesic.

Def Subset $S \subset M$ is achronal if $S \cap I^+(S) = \underbrace{\phi}_{\text{empty set}}$.

Theorem Let $S \subset M$. Then the boundary $\partial I^+(S)$ is an achronal, 3D submanifold.

Def $p \in M$ is a future endpoint of a future directed timelike or lightlike causal curve λ if for every neighborhood of p, O, there exists $t_0 \in \mathbb{R}$ such that $\forall t > t_0 \Rightarrow \lambda(t) \in O$.
λ is called future inextendible if it has no future endpoint. Analagous definitions follow for past endpoints and past inextendible curves.

For any past inextendible causal curve λ passing through p, we can choose $q \in I^+(p)$ and construct a timelike curve γ through q such that $\gamma \subset I^+(\lambda)$ and γ also past inextendible.



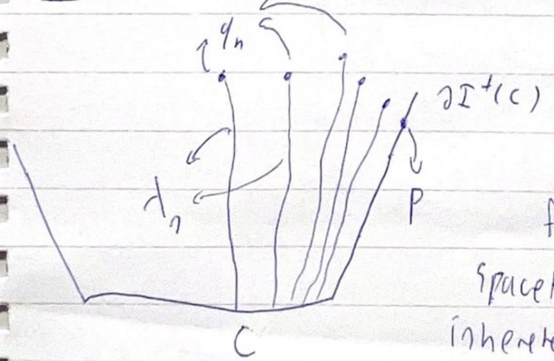
Def $p \in M$ is a <u>convergent point</u> of a set of curves $\{\lambda_n\}$ if there exists a neighborhood O of p and $N \in \mathbb{N}$ such that $\forall n > N$ we have $\lambda_n \cap O \neq \emptyset$.
A curve λ is a <u>convergent curve</u> of $\{\lambda_n\}$ if every point on λ is a convergent point.

Def $p \in M$ is a <u>limit point</u> if every neighborhood of p intersects infinitely many curves in $\{\lambda_n\}$.
A curve λ is a <u>limit curve</u> if there is a subsequence $\{\lambda_{n'}\}$ for which λ is a convergent curve.

So every convergent point is a limit point, but not conversely.
For a sequence $\{\lambda_n\}$ with a limit point $p \in M$, one can always construct its limit curve λ through p.

<u>Theorem</u> Let $C \subset M$ be a closed subset. Then every point $p \in \partial I^+(C)$ lies on a lightlike geodesic contained in $\partial I^+(C)$ which is either past inextendible or has a past endpoint on C.

<u>Proof:</u>



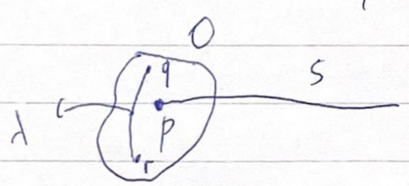
Choose a sequence $\{q_n\}$ in $I^+(C)$ which converges to p on $\partial I^+(C)$. Let $\{\lambda_n\}$ be a set of timelike curves from $\{q_n\}$ to C. Consider the spacetime $(M-C, g)$ with a topology inherited from (M, g) and with the induced metric.

because $M-C$ is open

In the new spacetime, all $\{\lambda_n\}$ are past inextendible, so there is also a past inextendible limit curve λ passing through p. If any point $r \in \lambda$ was inside $I^+(C)$, so would p which is not possible because it is ~~on~~ on the boundary. So $\lambda \subset J^+(C) - I^+(C)$, which means λ is a lightlike geodesic. Because it is past inextendible in $(M-C, g)$, that means in (M, g) it must either remain past inextendible or have a past endpoint on C.

$\square$

Def Let S be a closed, achronal set. Then the edge of S is a set $p \in S$ such that every neighborhood O of p contains points $q \in I^+(p)$ and $r \in I^-(p)$ and a timelike curve $\lambda \subset O$ which connects r and q without crossing S.

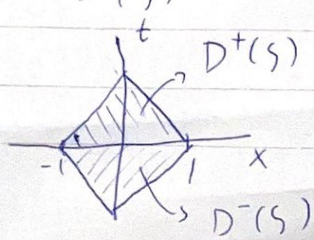


We can always "go around" S along its edge.

Def Future domain of dependence of a set S, $D^+(S)$, is a set of $p \in M$ such that every past inextendible causal curve through p must intersect S.
Analogous def. for past domain of dependence.

Full domain of dependence: $D(S) := D^+(S) \cup D^-(S)$

Ex Minkowski spacetime
$S = \{ t=0,\ x \in [-1, 1] \}$



closed, achronal set

Def Σ is a <u>Cauchy surface</u> if $D(\Sigma) = M$.

Intuitively, a domain of dependence is a set of events which are enteierly predictable from some initial data on S.

If a spacetime contains a Cauchy surface, it is called <u>globally hyperbolic</u>.

Def <u>Future Cauchy horizon</u> $H^+(S)$ is defined as

$$H^+(S) = \underbrace{\overline{D^+(S)}}_{\text{closure of } D^+(S)} - I^-(D^+(S))$$

Full Cauchy horizon:

$$H(S) = H^+(S) \cup H^-(S)$$

A Cauchy horizon represents an end of predictability of our theory since $H(S) = \partial D(S)$.

~~Theorem Every point $p \in H^+(S)$ lies on a past inextendible geodesic~~

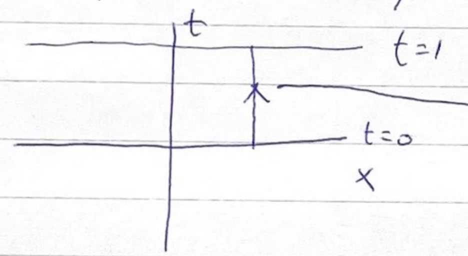
Theorem Let S be a closed, achronal set. Then every point $p \in H^+(S)$ lies on a geodesic contained in $H^+(S)$ and is either past inextendible or has an endpoint on the edge S.

Proof is very similar to the one for $\partial I^+(S)$ since they are both boundaries.

Causality conditions

We wish to avoid intersecting causal curves in order to prevent time travel paradoxes.

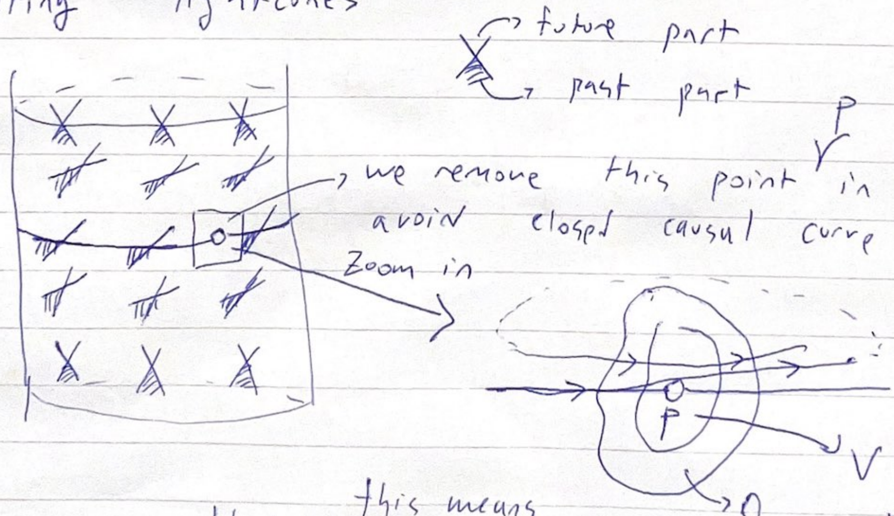
Example Take Minkowski spacetime and identify the hypersurfaces defined by $\{t=0\}$ and $\{t=1\}$.



every stationary observer lies on a closed timelike curve

We also want to prevent ~~...~~ causal curves coming arbitrarily close to intersecting themselves, since arbitrarily small perturbations would cause breaking of causality.

Example M = cylinder w/ metric g describing "tilting" lightcones



we remove this point in order to avoid closed causal curve

For every neighborhood O of p there exists $V \subset O$ such that V is intersected more than once

this means that causal curves passing near p can come very close to intersecting

Def Spacetime (M, g) respects the strong causality condition if for every $p \in M$ and every neighborhood O of p there exists $V \subset O$ such that no causal curve intersects V more than once.

This condition ensures there are no curves that almost intersect themselves, since that would be unstable.

While our cylinder example doesn't break causality via closed causal curves, it breaks the strong causality condition.

An important Lemma:

Lemma Let (M, g) be a strongly causal spacetime and $K \subset M$ compact subset. Then every causal curve $\lambda \subset K$ must have a future and past endpoint in K.

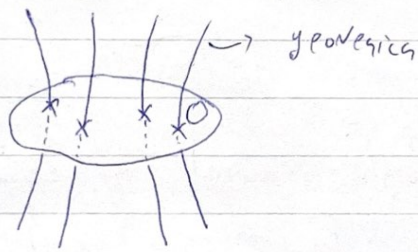
Proof: Without loss of generality let's assume the curve parameter ~~...~~ is $t \in \langle -\infty, \infty \rangle$. Let $\{t_i\}$ be a divergent sequence in $\mathbb{R}$, and construct a sequence in M $\{\lambda(t_i)\}$ where $\lambda : \mathbb{R} \to K$ is a causal curve. Because K is compact, $\{\lambda(t_i)\}$ has an accumulation point $p \in K$. Assume there exists a neighborhood O of p such that there is no $t_0 \in \mathbb{R}$ for which for all $t > t_0$ we have $\lambda(t) \in O$. The same must hold for $V \subset O$, $p \in V$. But because p is an accumulation point, V must contain infinitely many elements of $\{\lambda(t_i)\}$, which means that the curve λ must cross V more than once. $\square$

Another useful property:

<u>Lema</u> Let (M, g) be globally hyperbolic. Then (M, g) is strongly causal.

$\boxed{3}$ Geodesic congruences

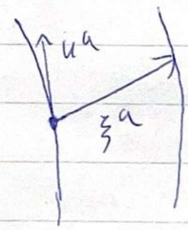
<u>Def</u> Let $O \subset M$ be open. A <u>congruence</u> of O is a family of geodesics such that through every $p \in O$ passes only one geodesic from the congruence.



Consider the mapping $(s, t) \longrightarrow \gamma_s(t)$ where $s, t \in \mathbb{R}$ get mapped to a point on the geodesic "marked" with s in the congruence

$u^a := \left(\frac{\partial}{\partial t}\right)^a \rightsquigarrow$ a tangent vector along a geodesic

$\xi^a := \left(\frac{\partial}{\partial s}\right)^a \rightsquigarrow$ deviation vector



We can always choose a Lie transported deviation vector, so $[u, \xi] = 0$.

We can separate the metric to a part parallel to u^a and a part transverse to u^a, say $h_{\mu\nu}$.

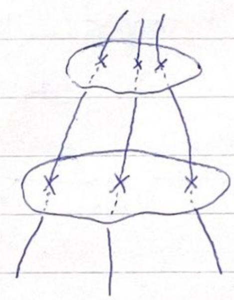
$$g_{\mu\nu} = \underbrace{h_{\mu\nu}}_{\text{transverse, effectively 3D}} + \underbrace{u_\mu u_\nu}_{\text{paralel}}$$

We can interpret $h := \det h_{\mu\nu}$ as describing the volume of the congruence in space

We define $\theta := \frac{1}{\sqrt{h}} \frac{d}{d\tau} \sqrt{h}$ called <u>expansion</u> where τ is the proper time of some referent geodesic.
We interpret θ as the fractional change in time of the 3-volume of the congruence. $\theta > 0$ means a growing congruence, and $\theta < 0$ a shrinking one

$\theta > 0$:

$\theta < 0$:



We can also define the <u>shear</u> $\sigma_{\mu\nu}$ describing the "squeezing" and <u>rotation</u> $\omega_{\mu\nu}$ of the congruence.
Evolution of θ is given by the <u>Raychaudhuri eq.</u>:

$$\boxed{\frac{d\theta}{d\tau} = -\frac{1}{3}\theta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu}u^\mu u^\nu}$$

(for more details, read Eric Poisson: A relativists toolkit)
Assume $\omega_{\mu\nu} = 0$ (this is true by Frobenius's theorem) and $R_{\mu\nu}u^\mu u^\nu \geq 0$ (strong energy condition). Then:

$$\frac{d\theta}{d\tau} \leq -\frac{1}{3}\theta^2 \Rightarrow \theta(\tau) \leq \frac{\theta(0)}{1 + \frac{1}{3}\theta(0)\tau}$$

We see for $\theta(0) < 0$ we have $\theta \to -\infty$ for some proper time $\tau \leq 3/|\theta(0)|$, which means geodesics focus to a point.

A point at which a congruence crosses ($\theta \to -\infty$) and the congruence is initially orthogonal to some ~~hyp~~ hypersurf. S is called <u>conjugate</u> to S.

A sufficient and necessary condition for a curve to maximize proper time from hypersurface S to point p is that the curve is a <u>geodesic</u> with <u>no conjugate points</u> between S and p. The congruence in this case can be understood roughly as perturbations of the referent curve. Concept of conjugate points can also be applied to a pair of points (see Jacobi field in Wald)

In a globally hyperbolic spacetime, there always exists a curve of maximal proper time.

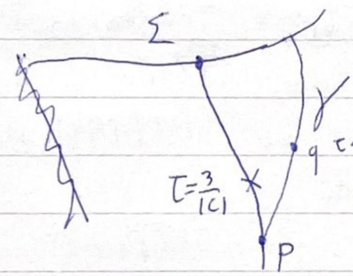
$\boxed{4}$ Singularity theorems

<u>Def</u> We say a curve λ is <u>incomplete</u> if it has no endpoints, but finite length (proper time). We say that a spacetime containing an incomplete curve is <u>singular</u>.

This definition of singularity avoids coordinate systems, and is sufficiently general.
It can be understood ~~as~~ in terms of observers who abruptly "fall through" a singularity with no endpoint.

Theorem Let (M, g) be globally hyperbolic spacetime satisfying the strong energy condition. Assume a Cauchy surface Σ such that for a congruence of Σ $\theta|_{\Sigma} \leq C = \text{const.} < 0$. Then no ~~th~~ past inextendible timelike curve from Σ has length greater than $3/|C|$.

Proof: Assume there exists a curve λ from Σ that is past directed with length greater than $3/|C|$. Let $p \in \lambda$ be a point lying past the length $3/|C|$.

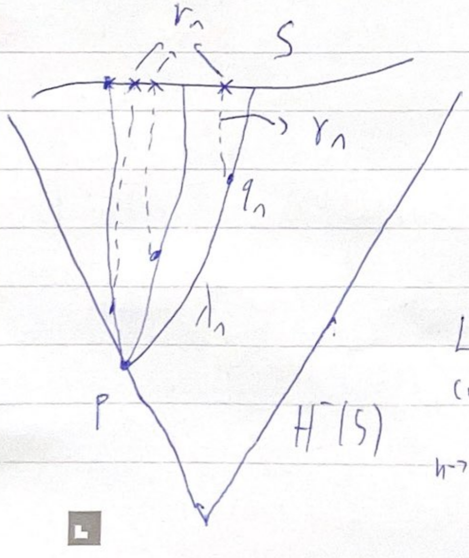


For a globally hyperbolic spacetime there always exists a curve of maximal proper time. Let γ be a curve of maximal proper time from Σ to p. Because it is maximal, p lies beyond $3/|C|$ proper time on γ as well. Because γ is maximal, it must be a geodesic with no conjugate points. But from Raychaudhuri's eq. we know there must be a conjugate point to Σ along γ at length less than $3/|C|$, and so γ cannot maximize proper time between Σ and p. We have reached a contradiction. $\square$

A great weakpoint of this theorem is the assumption of global hyperbolicity, but can we eliminate it?

Theorem Let (M, g) be strongly causal and respect the strong energy condition. Assume there exists a compact, achronal, smooth, spacelike, edgeless hypersurface S such that for some congruence normal to S we have $\theta|_S \leq C = \text{const.} < 0$. Then at least one timelike geodesic that is past inextendible cannot have length greater than $3/|C|$.

Proof: Our strategy will be to show the compactness of the past Cauchy horizon $H^-(S)$ and then use properties of Cauchy horizons to come to a contradiction.
Assume all past inextendible timelike geodesics from S have length greater than $3/|C|$.
Consider the globally hyperbolic spacetime $(\text{int}[D(S)], g)$. Because it is globally hyp., we can apply the previous theorem and conclude that all timelike geodesics must exit $D^-(S)$ before they reach length $3/|C|$. That means they must intersect $H^-(S)$, so it is not empty.



Let $p \in H^-(S)$. Since all timelike curves have proper time less than $3/|C|$, all causal curves ~~geodesics~~ ~~also~~ have an upper bound of τ_0 on their length.
Let $\{\lambda_n\}$ be a sequence of timelike curves from S to p such that $\lim_{n \to \infty} \tau[\lambda_n] = \tau_0$.

Choose a sequence of points $\{q_n\}$ such that $q_n \in \lambda_n$ and which converges to p. Since $q_n \in I^+(p) \subset \text{int}[D(S)]$, there exists a maximal curve γ_n from S to q_n which is a geodesic. Since $\{q_n\}$ converges to p, we have $\lim_{n \to \infty} \tau[\gamma_n] = \tau_0$.

Let r_n be the intersection of γ_n and S. Because S is compact, $\{r_n\}$ has an accumulation point $r \in S$. Let γ be the maximal geodesic from r to p. Because of continuous dependence of a geodesic's proper time on its initial conditions, we must have
$\tau[\gamma] = \tau_0$.

We now repeat the argument for a sequence $\{p_n\}$, $p_n \in H^-(S)$. For each p_n there is a maximal geodesic $\tilde{\gamma}_n$ from S to p_n. Consider set of intersections $\tilde{r}_n$ of $\tilde{\gamma}_n$ and S. It again has an accumulation point $\tilde{r}$, and we again construct $\tilde{\gamma}$ maximal geodesic from $\tilde{r}$ to some point p, which will be an accumulation point of $\{p_n\}$ contained in $H^-(S)$. It follows that $H^-(S)$ is compact, because $\{p_n\}$ was an arbitrary sequence.

We know that $H^-(S)$ is generated by geodesics contained in $H^-(S)$ which are either future inextendible or have endpoints on the edge of S. Since S is with no edge, they must be future inextendible. But this is not possible because $H^-(S)$ is compact, which means all curves in it have endpoints. We reached a contradiction. $\square$

An example of such a set S for Schwarzschild would be any sphere with radius less than 2M.

There are many other singularity theorems for specific situations, like gravitational collapse. All singularity theorems have 3 things in common:

1) conditions on geometry - often stated via energy conditions like the strong energy, conditions used for focusing of geodesics, it is actually a geometric condition on the Ricci tensor

2) causality condition - i.e. global hyperbolicity or strong ~~[illegible]~~ causality condition to eliminate ~~[illegible]~~ unphysical causal structures

3) initial conditions - a ~~[illegible]~~ bounded negative expansion ensures conjugate points

While great at identifying singular spacetimes, ~~they~~ singularity theorems tell us nothing about the singularity itself. It could be a benign regular singularity removable by embedding into a larger non-singular spacetime, or they could contain highly pathological quasiregular singularities. (For more on this: Ellis and Schmidt "Singular Space-Times")