

General Relativity

Class 41

Bassel Saleh

E-mail: bassel.s.saleh@gmail.com

Contents

1	Motivation for Singularity Theorems	1
2	Geodesic focusing	1
3	Conjugate points	2
4	Causal Relations	4
4.1	Chronological future	4
4.2	Domain of dependence	4

In this lecture we introduce some motivation for looking at the singularity theorems of GR, and we introduce definitions and ideas that will be useful in stating and solving these theorems next time. Though this subject relies on some pretty formal mathematics, we will not be excessively rigorous in discussing it.

1 Motivation for Singularity Theorems

Recall Einstein's Field Equation (EFE),

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \quad (1.1)$$

We have spent much time studying how to interpret the physical meaning of this equation, as well as how to solve it in certain cases. The equation represents a set of nonlinear partial differential equations, and we can't say many meaningful things generally about the solutions. So far, we have either relied on symmetries or used heavy approximations to obtain workable solutions to these equations.

If we want to take a stab at understanding solutions to the EFE in a more general way, there are two options. We can try to formulate the equations in a way that makes them solvable on a computer, or we can prove some theorems about the structure of the solutions, even if we have no systematic way of producing them. It is this second approach that we will use by studying the singularity theorems. The first was proven by Penrose in 1965.

2 Geodesic focusing

We looked before at the notion of geodesic deviation, in which we considered a “bundle” (or *congruence*) of timelike geodesics with tangent vectors U^μ and a separation vector S^μ . We saw how to relate the evolution of S^μ along the geodesics to a tensor $B_{\mu\nu}$ that describes the behavior of the congruence. Specifically,

$$B_{\mu\nu} = \nabla_\nu U_\mu \quad (2.1)$$

$$U^\mu \nabla_\mu S^\nu = B^\nu{}_\mu S^\mu \quad (2.2)$$

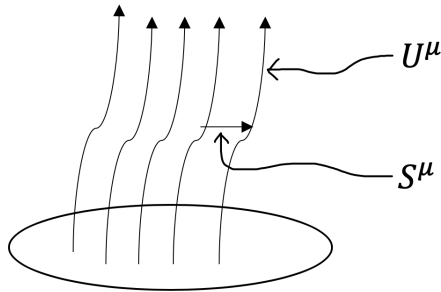


Figure 1. A congruence of geodesics, parameterized by tangent vector field U^μ and separation vector field S^μ .

This result came from the geodesic deviation equation, and it tells us that $B_{\mu\nu}$ acts as a transformation matrix that describes the evolution of S^μ . The congruence can expand, rotate, and shear, with this information being encoded in B . In particular,

- $B^\alpha{}_\alpha = \theta$ is the rate of change of the area of the bundle, representing *expansion*
- $\omega_{\mu\nu} = B_{[\mu\nu]}$ is the antisymmetric *twist tensor*
- $\sigma_{\mu\nu} = B_{(\mu\nu)} - \frac{1}{3}\theta\gamma_{\mu\nu}$ is the *shear tensor*, where $\gamma_{\mu\nu} = g_{\mu\nu} + U_\mu U_\nu$

We are going to focus on congruences that are hypersurface-orthogonal, which means setting $\omega_{\mu\nu} = 0$.

The evolution of θ is given by Raychaudhuri's equation¹,

¹ See Carroll Appendix F

$$\frac{d\theta}{d\tau} = U^\mu \nabla_\mu \theta = -\frac{1}{3}\theta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu}U^\mu U^\nu \quad (2.3)$$

Looking at the right hand side of Raychaudhuri's equation, the first term is negative definite (minus a squared quantity), the second term is negative definite ($\sigma_{\mu\nu}\sigma^{\mu\nu}$ is sort of a squared some of the eigenvalues of $\sigma_{\mu\nu}$), and the third term is zero by construction (we're choosing congruences such that $\omega_{\mu\nu} = 0$). This means the overall sign of $\frac{d\theta}{d\tau}$ is determined by the sign of the final term, $-R_{\mu\nu}U^\mu U^\nu$. This in turn depends on the stress-energy content of the spacetime. If, for example, we assume the Strong Energy Condition (SEC), then

$$R_{\mu\nu}t^\mu t^\nu \geq 0 \quad \forall \text{ timelike } t^\mu \quad (2.4)$$

$$\implies \frac{d\theta}{d\tau} < 0 \quad \text{i.e. gravity is attractive} \quad (2.5)$$

In particular we find that

$$\frac{d\theta}{d\tau} \leq -\frac{1}{3}\theta^2 \quad (2.6)$$

We can then saturate this bound and solve the differential equation to obtain a bound on θ itself. Doing so yields the solution²

² Again, this is true assuming the SEC holds and that $\omega_{\mu\nu} = 0$.

$$\theta^{-1} \geq \theta_0^{-1} + \frac{\tau}{3} \quad (2.7)$$

If at any point $\theta_0 < 0$, then θ will approach $-\infty$ in a finite amount of time, meaning the bundle has a negative infinite change in area. We conclude that if θ ever becomes negative, then the geodesics of the bundle *must* self-intersect in finite time $\tau \leq \frac{3}{|\theta_0|}$.³

³ See Wald 9.3.1

There is an analogous equation describing the expansion of null geodesics,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \hat{\sigma}_{\mu\nu}\hat{\sigma}^{\mu\nu} + \hat{\omega}_{\mu\nu}\hat{\omega}^{\mu\nu} - R_{\mu\nu}k^\mu k^\nu \quad (2.8)$$

where similarly $\hat{\omega}_{\mu\nu} = 0$ corresponds to hypersurface-forming null geodesics, and of course k^μ is just the tangent vector to these geodesics. This time, the Null Energy Condition (NEC) implies that $R_{\mu\nu}k^\mu k^\nu \geq 0 \quad \forall \text{ null } k^\mu$, which in turn implies that within $\lambda \leq \frac{2}{|\theta_0|}$ the expansion θ will approach $-\infty$ in finite time, provided that θ_0 was at some point negative.

The ultimate conclusion then of geodesic focusing is this: if nearby geodesics are initially moving toward each other (i.e. $\theta_0 < 0$), then gravity will be attractive and the geodesics will inevitably intersect. This is true for both timelike and null geodesics. A point at which geodesics come together like this is called a *caustic*.

3 Conjugate points

Let's introduce some definitions regarding **conjugate points**.

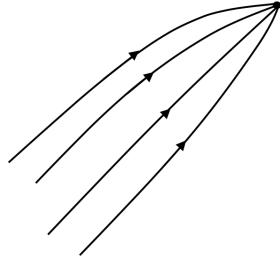


Figure 2. Convergence of a congruence of geodesics that are initially moving toward each other. They terminate at a caustic.

- Two points P and Q on a manifold are called *conjugate points* if two geodesics emanating from P come together at Q . In other words, the expansion θ of a bundle of geodesics originating at P is at first positive (meaning the geodesics expand away from each other) and then becomes negative (so the geodesics come back together at a caustic).
- A point P is conjugate to a hypersurface Σ if a congruence of geodesics normal to Σ comes together at P . These geodesics must originate at “nearby” points on Σ .
- The closeness of the starting points is crucial. If two geodesics initially far apart come together and intersect, the result does not contain a pair of conjugate points.

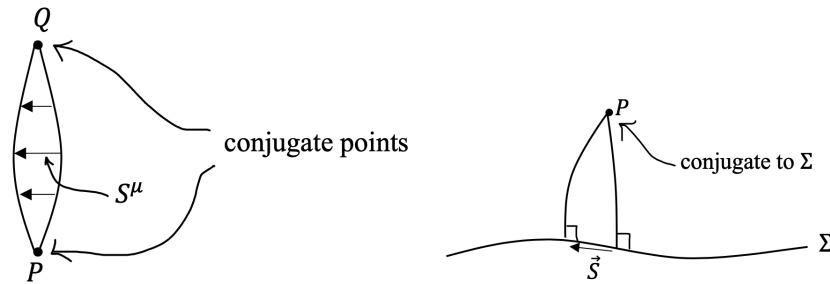


Figure 3. Picture of two conjugate points P and Q . Note that the separation vector S^μ grows and then shrinks. It is zero at each point but nonzero in between.

Conjugate points create a problem with interpreting geodesics as curves of maximal (local) length on a manifold. In particular, consider two points P and R and a geodesic γ connecting them. Normally, we can interpret γ as a curve that maximizes the proper distance⁴ between P and R , which can be shown by taking the second variation of τ along γ and demonstrating that no local variations of γ can have a larger value of τ . However, if there is a point Q sitting on γ that is *conjugate* to P , it turns out we *can* construct a different curve from P to R that is longer than γ .

The resulting formal statement turns out to be the following: A timelike geodesic γ maximizes proper time τ between P and R if and only if γ contains *no* point conjugate to P between P and R .⁵ An analogous theorem exists for hypersurfaces; a geodesic between a point on Σ and a point P maximizes τ between them if and only if there are no conjugate points between them.

⁴ Proper time for timelike geodesics.

⁵ See Wald theorem 9.3.3

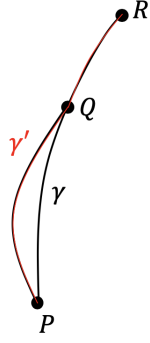


Figure 4. In this picture, γ is a geodesic between P and R , but the existence of Q , which is conjugate to P , lets us construct a new curve γ' which circumvents γ from P to Q , then continues along γ to R . It can be shown that γ' , which is of course not a geodesic, has a greater proper time than γ . We also know such a curve exists by virtue of P and Q being conjugate.

These theorems allow us to conclude an interesting consequence. If γ is a *null* geodesic between P and R , and there is a point Q along γ that is conjugate to P , then there must exist a timelike curve connecting P and R .

4 Causal Relations

4.1 Chronological future

Consider a set S of points on a spacetime manifold. We can define the *chronological future* of S , denoted $I^+(S)$, as the union of the future light cones of all points in S . So the chronological future of a single point P is the interior of P 's future light cone. This can be thought of as all events reachable from P by timelike observers (equivalently, all points that can be connected to P by a future-pointing timelike geodesic).

Intuitively, it seems the boundary of $I^+(P)$ must be made up of null curves. We find that, if S is a closed subset of the manifold, then the boundary of the chronological future of a S , denoted $\dot{I}^+(S)$, must be *ahronal*, meaning spacelike or null, and every point in $\dot{I}^+(S)$ lies on a null geodesic fully contained in $\dot{I}^+(S)$.⁶

⁶ See Wald theorems 8.1.3 and 8.1.6

We now want to explore the idea that if two null curves on $\dot{I}^+(S)$ collide, they must enter *into* $I^+(S)$. This can be seen pictorially for distinct null curves (those on the boundaries of light cones from two distinct points) in Figure 5. It turns out this is more general, as it applies also to nearby null curves that come together at a caustic. This is a consequence of our theorem from earlier, stating that if a null curve connects two points, with a conjugate point in between, then the two points are secretly timelike separated.

4.2 Domain of dependence

We define the *future domain of dependence* of a closed set S as the set of points for whom every past-directed causal curve hits S . We denote this $D^+(S)$. Every point for whom every future-directed causal curve hits S is the *past domain of dependence* of S , denoted $D^-(S)$. The full domain of dependence is simply the union,

$$D(S) := D^+(S) \cup D^-(S) \quad (4.1)$$

A spacetime manifold \mathcal{M} is *globally hyperbolic* if some ahronal surface Σ exists such that $D(\Sigma) = \mathcal{M}$. Such a surface is called a *Cauchy surface*. Minkowski and Schwarzschild

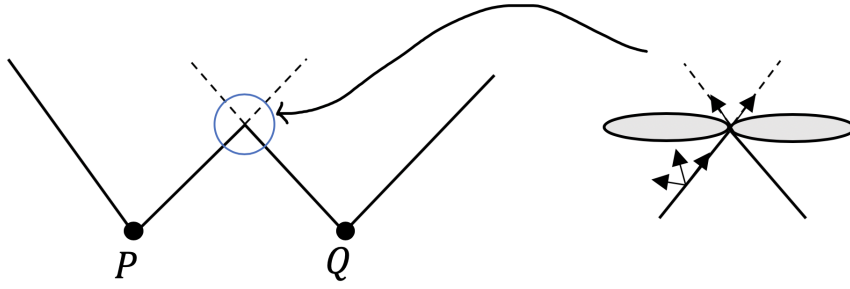


Figure 5. At any point on one of the light cones leading up to the intersection, there is only one direction for null vectors to point. All other directions will be into the light cone. At the intersection then, the only place for the curves to go is into the interior of one of the light cones, and hence into $I^+(\{P, Q\})$.

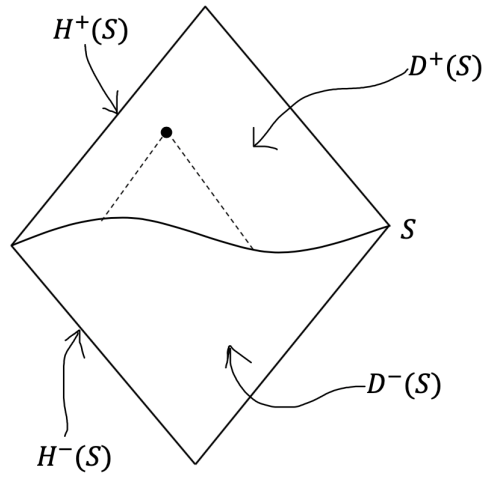


Figure 6. $D^+(S)$ contains every point whose entire past intersects with S , and $D^-(S)$ contains every point whose entire future will intersect with S . The boundaries of these domains, denoted $H^+(S)$ and $H^-(S)$ respectively, are called *Cauchy horizons*.

spacetimes are both examples of globally hyperbolic manifolds. In fact, these types of spacetimes are of particular interest because they allow for “physics” to be defined on them in a sensible way, namely by allowing one to place initial data on a Cauchy surface Σ for $g_{\mu\nu}$ and all relevant fields, then evolving them into the future to obtain the full evolution of these quantities in \mathcal{M} .

Globally hyperbolic spacetimes have the following property:

- If you have a vector field t^μ emanating from Σ , the integral curves of t^μ pierce surfaces of constant time exactly *once*. This allows the spacetime to be foliated into surfaces of constant time, where data can be evolved from one surface to the next.⁷

⁷ See Wald 8.3.12

We can conclude with a few properties of globally hyperbolic spacetime that are the result of this idea.

- A globally hyperbolic manifold \mathcal{M} has the topology of $\Sigma \times \mathbb{R}$
- The curves provide a map for points on these surfaces to each other. In other words, every point on a surface of constant time can be uniquely mapped to a point on any

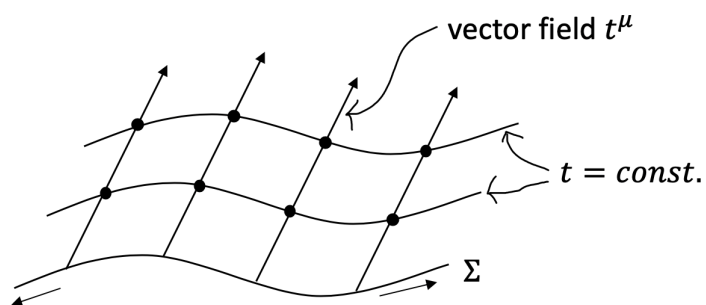


Figure 7. Surfaces of constant t are pierced by the integral curves of t^μ exactly once.

of the other surfaces of constant time.

- Each constant time surface is a Cauchy surface.