Global Methods In General Relativity

Aum Rawal

August 23, 2024

MSc in Mathematical Physics

The University of Edinburgh

2024

Abstract

This thesis serves as a literature review of the renowned Penrose Singularity Theorem [9], with the primary goal of providing a largely self-contained proof of the theorem. The already existing literature, though brilliant, is sometimes too detailed to follow. Hence, this work tries to be as concise and comprehensible as possible, omitting results and definitions not directly needed for the Singularity theorem. We begin our discussion by studying the global causal structure of our Universe, which will be modeled by a 4-dimensional spacetime. This section addresses the topological aspects of the global structure of spacetime, and develop tools for constraining our spacetime to physically realistic conditions.

In the subsequent chapter, we examine the properties of congruences of curves, with a particular focus on geodesics. This investigation leads us to the Raychaudhuri equation, which governs the behavior of the expansion, θ of a congruence. A key result here is that the expansion tends to become singular within a finite affine parameter, suggesting the existence of a spacetime singularity.

Next, we delve into the concept of conjugate points and their critical role in the singularity theorem. We establish a relationship between the occurrence of conjugate points and the expansion of a geodesic congruence, demonstrating that conjugate points can only arise when θ diverges.

Finally, leveraging the results developed throughout this thesis, we present a proof of Penrose's Singularity Theorem

Declaration

I declare that this dissertation is entirely my own work.

As noted, this dissertation is a literature review, with the primary resource being General Relativity by R. Wald [10]. Many of the proofs presented here are inspired by this text and various other sources, which will be cited where relevant. In cases where original proofs are provided—either because they were skipped in the book or were too complicated to directly follow from another source—I try to provide a sketch of the proof. Any instance of original or adapted proofs will be clearly indicated in the text.

Since the calculations in this work are not particularly demanding, they were carried out manually, without the use of any software.

Personal Statement

My project began by working on the local calculations concerning geodesic congruences. I spent around a month to thoroughly understand and replicate Raychaudhuri equation, for both timelike and null case. It was the first time I was introduced to abstract index notations, and thus took some time to get used to them. My reference for this entire month was Wald (2010).[10] However, some proofs are omitted in Wald, and had to be done by myself. For instance, the section of null geodesic congruence (3.1.2), proposition (4.2.1), proposition (4.3.2) are entirely my work, although inspired from proofs of analogous theorems provided in Wald (2010). Theorem (4.3.1) was a tricky and lengthy proof to be involved in this text, so I resorted to give a sketch of the proof, aiming to provide some insight on the theorem. It is again my own work, although inspired from Hawking and Penrose (2010)[1].

In the following week, I started to work on topological aspects used to describe a global causal structure on spacetime. This involved learning about plenty new definitions and proofs. The proof I spent the most time on from this part was of theorem (2.1.2). I had to refer to multiple sources, namely, Wald (2010), Hawking and Ellis (1973)[2], and O' Neill (1983)[7]. The bit which involved proving continuous nature of an achronal boundary was the most challenging.

Towards the end of July, I had started working on the Penrose's Singularity theorem. My professor paved an interesting path for my project, that is, once I had the taste of the subject I jumped to prove the Singularity theorem, and then worked my way backwards to prove all the necessary results. This helped me have a clear goal in mind whenever I was trying to prove a proposition, and understand how it connects to the Singularity theorem. First two weeks of August were spent doing that.

Finally, last two weeks were spent in editing the work, making and labelling figures, correcting some proofs, and drinking plenty coffee.

Acknowledgements

Completing this dissertation has been a profound learning experience, and I am deeply grateful for the invaluable support of those who have contributed to its completion.

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. James Lucietti, for his invaluable guidance and support throughout this project. His detailed explanations of differential geometry and general relativity have been crucial in making this work possible. I deeply appreciate his cooperation and the time he dedicated to ensuring the success of this dissertation.

I would also like to thank my family and friends for their unwavering support throughout this process. In particular, I am deeply grateful to my mother for her constant encouragement and support.

Contents

List of Figures

Chapter 1

Introduction

The very first solution to Einstein's equations given by Karl Schwarzchild presented points (at $r = 0$ and $r = 2M$) where the metric diverged. Historically, these singularities were not paid much attention, as the Schwarzchild radius ($2M$) was very small relative to the size of a star, and hence cannot be reached by any experimental probe. For example, Schwarzchild radius for the Sun is calculated to be around 2.9 kms. It was later recognised by Kruskal that by using a different set of coordinates, the metric can become smooth at $r = 2M$ [6]. This begs the question that does the singularity at the center of Scwarzchild ($r = 0$) represent a physical singularity or is it just a mathematical artifact? The possibility of the existence of a physical singularity was strengthened when R. Oppenheimer and H. Snyder showed the collapse mechanism of a spherically symmetric star and thereby proving that for a sufficiently large object, physical singularity must occur at $r = 0$[8], as shown in figure 1.1. For a local comoving observer, the body passes within its Schwarzchild radius at $r = 2M$, and any kind of information about that object is lost to an outside observer, at least classically. Thus the existence of a singularity presents a serious problem for any complete discussion of the physics of the interior region.

There was still the matter of symmetry to be concerned with, and to respond to that Kerr proposed a rotating black hole solution, which also had a physical singularity. And since a high degree of symmetry is still present (and the solution is algebraically special), it might again be argued that this is not the best representation of the general situation.[9] On the other hand, the singularity theorem proposed by Penrose argues for the existence of singularity without assuming any symmetry. The aim of this report is to provide a (mostly) self-sufficient and cohesive proof of this theorem which should be easily understandable by anyone with a background in general relativity, geometry, and some topology. However, to begin our discourse on singularities, firstly we need to understand what a singularity is.

1.1 What is a Singularity ?

In *special relativity*, spacetime is the manifold $\mathbb{R}^4$ with a flat metric of lorentzian signature defined on it. And in case of Coulomb solution of Maxwell's equation for special relativity,

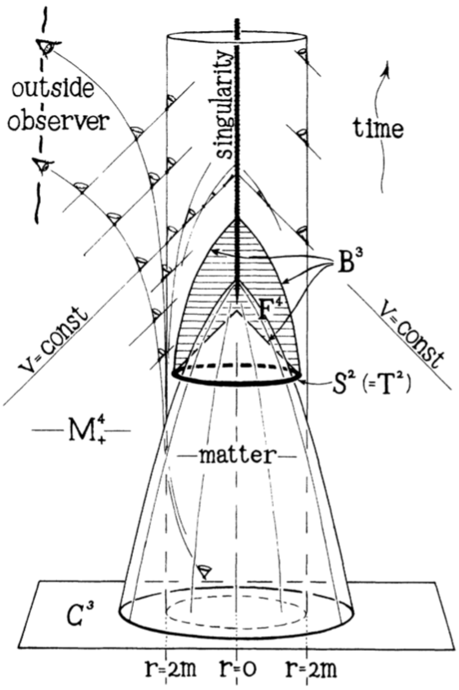

Figure 1.1: Image of a spherically symmetric collapsing matter[9]

we encounter a singularity at $r = 0$, which is defined in $\mathbb{R}^4$. Since this singularity is a point on the spacetime, we have no difficulty in saying that the event labeled by $r = 0$ is a singularity. However, in *general relativity*, we define our spacetime (M, g_{ab}) by a manifold M which admits a smooth metric *everywhere* on M. Thus, singularities we encounter in general relativity, for example, the "big bang" singularity of Robertson-Walker solution and Schwarzchild singularity at $r = 0$, are not considered to be part of the spacetime manifold; they are not at a "place" at some "time".

This definition fails to describe a singularity as a place in mathematical terms, nevertheless singularities do exist in Robertson-Walker and Schwarzchild spacetimes. To solve this issue, we will use the idea of "holes" left behind by the removal of singular points from spacetime. These "holes" will be evident if there exists *incomplete geodesics*. Now according to a fundamental hypothesis of general relativity : *The world lines of freely falling bodies in a gravitational field are simply the geodesics of the (curved) spacetime metric.* Thus, existence of a future (or past) incomplete geodesic would imply that a particle would cease to exist in a finite proper time (or it begin its existence a finite time ago in the past). This characteristic of incomplete geodesics establishes them as a satisfactory requirement for the existence of a physical singularity.

One possible objection to this is that we can artificially remove points from a spacetime manifold, which will cause geodesic incompleteness. One example of this is Schwarzchild spacetime in which we only consider point which have $r \geq 2M$. Such a spacetime is known as extendible. More precisely, A spacetime (M,g) is extendible if it is isometric to a proper subset of another spacetime (M',g'). Thus to avoid this possibility we will only consider spacetimes which are not extendible,i.e., *inextendible.*

In conclusion, the existence of a singularity in an inextendible spacetime can be proved by demonstrating that the spacetime contains incomplete geodesics. The Penrose (1965)[9] singularity theorem establishes that geodesic incompleteness is a necessary condition for a physically realistic spacetime. Hence, a singularity is a feature of the spacetime, irrespective of any symmetries imposed on its metric.

1.2 Mathematical context

Throughout the text we will assume that M is a 4-dimensional manifold with a lorentzian metric g_{ab},i.e., with signature $-+++$. *Spacetime is the manifold M on which a Lorentz metric, g_{ab}, is everywhere defined.* Also we will be using *abstract index notation* in the text. More precisely, a tensor of type (k,l) will be denoted by a letter followed by k covariant and l contravariant lower case latin indices, $T^{a_1,\ldots a_k}{}_{b_1,\ldots b_k}$. Reader should not confuse these with basis components. We will use greek letters, as usual, to denote any equation involving basis components. Moreover, for a tensor T_{ab} of type (0,2), we define

$$T_{(ab)} = \frac{1}{2}(T_{ab} + T_{ba}) \tag{1.1}$$

$$T_{[ab]} = \frac{1}{2}(T_{ab} - T_{ba}). \tag{1.2}$$

Curves, Connections and Geodesics

A *curve* (smooth), C, on a manifold M is simply a C^∞ map from $\mathbb{R}$ (or any interval of R) into M,

$$C : \mathbb{R} \to M.$$

Additionally, we will be using a torsion free, levi-civitta connection ∇_a,i.e,

$$\nabla_a g_{bc} = 0. \tag{1.3}$$

One should note that there are some theories of gravity which allows connections with torsion, but they won't be discussed in this work. Also, the reason for using a levi-civitta connection is to ensure that a null geodesic remains a null geodesic, physically implying that a massless particle always travel at the speed of light.
Given a connection, ∇_a, we can define a notion of parallel transport. A vector v^a is parellely transported along a curve with tangent t^a if $t^a\nabla_a v^b = 0$. *Geodesic* is a curve whose tangent vector,T^a, is parallely propagated along itself,i.e.,

$$T^a\nabla_a T^b = 0. \tag{1.4}$$

A more general equation will be $T^a\nabla_a T^b = \alpha T^b$, but due to reparameterization the right hand side can chosen to be 0. In the text, this condition 1.4 will be referred as the geodesic equation. Moreover, a covariant derivative of any vector v^a along a geodesic $\mu(s)$ will be written as

$$\frac{Dv^a}{ds} = T^b\nabla_b v^a.$$

Additionally, the lie derivative with respect to vector field v,i.e., $\mathcal{L}_v$ follows

$$\mathcal{L}_v \mu_a = v^b \nabla_b \mu_a + \mu_b \nabla_a v^b,$$

for an arbitrary one form ${\mu_a}$[1].

Remark In this text, a vector field will generally be referred to simply as a vector, provided it is defined everywhere.

Now, we will provide the definition of timelike and null curves as they will be frequently used throughout the text. A vector ξ^a is called timelike if $g_{ab}\xi^a\xi^b \leq 0$. A curve whose tangent is everywhere timelike is known as a *timelike curve*. Similarly, a null vector k^a satisfies $g_{ab}k^a k^b = 0$, and a curve whose tangent is everywhere null is known as a *null curve*. Finally, Riemann curvature tensor field, ${R_{abc}}^d$, of such a connection is defined as

$$\nabla_a \nabla_b \omega_c - \nabla_b \nabla_a \omega_c = {R_{abc}}^d \omega_d, \tag{1.5}$$

for all dual vector fields ω_c. A useful tensor that is constructed out of Reimann curvature tensor is the Ricci tensor, R_{ac},

$$R_{ac} = {R_{abc}}^b.$$

Ricci tensor is an essential element of the Einstein's equation, i.e.,

$$R_{ab} = 8\pi(T_{ab} - \frac{1}{2} g_{ab} T), \tag{1.6}$$

where T_{ab} is the *stress energy tensor* and T is the trace of the same.

[1] Calculations to derive this can be found in Wald (2010) [10]

Chapter 2

Causal Structure

In special relativity, the causal structure of spacetime is described by a *light cone* associated with each event, p, in spacetime. This light cone is divided into two halves, labeled "future" and "past". Events that lie within the future of the light cone represent events which can be influenced by a massive field starting *at p*. And only the events lying in the past light cone of p can influence the event *at p* through a massive field. Events on the surface of the light cone can only be reached by a massless field, such as a light signal.

In general relativity, *locally* the causal structure has the same qualitative nature as in the flat spacetime of special relativity. However, due to nontrivial topology, *spacetime singularity*, or lack of an "orientation", global causal structure can be significantly different. In this chapter we aim to give an account of some basic definitions and results regarding causal structure of spacetime in general relativity. As we will see, the results from this chapter are crucial to prove the *singularity theorem*. Most of the discussion presented here are based on Wald (2010)[10], and Hawking and Ellis (1973)[2].

2.1 Basic definitions and Results

Let (M, g_{ab}) be a spacetime. We have that at each event $p \in M$, the tangent space V_p, is isomorphic to the Minkowski spacetime. In the following text we will refer to the light cone passing through the origin of V_p as the *light cone of p*. Thus, *light cone of p* is a subset of V_p not M. As in special relativity, we shall designate "future" and "past" to each halves of the light cone. However, it is not necessary that this designation is continuous as p varies over M. An example of such spacetime is shown in figure 2.1. This kind of spacetimes admit a property that we cannot consistently distinguish the notion of "going backward in time" as opposed to "going forward in time". Physically, however, it would seem reasonable to suppose there is a local thermodynamic arrow of time, given by the direction of increase of entropy, which defines a continuous designation of past and future. If this is the case, then we shall say that the spacetime is *time-orientable*. In the following, we will consider only time orientable spacetimes and thus will have a continuous designation of "future" and "past" halves of the light cone.

Definition 2.1.1. *A timelike or null vector lying on the future half of the light cone will*

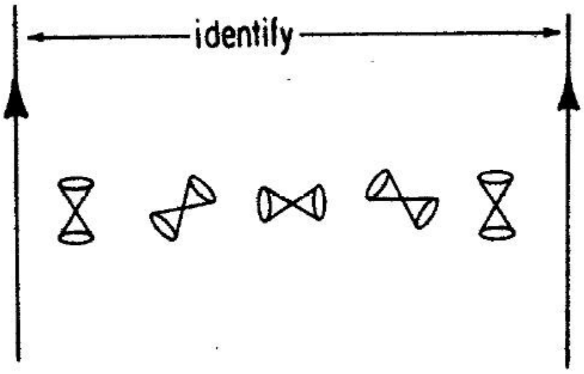

Figure 2.1: Example of a non-orientable spacetime[10]

be called future directed.

Analogous definitions follow for *past directed vector.*

Time orientable spacetimes satisfy an important property.

Lemma 2.1.1. *If (M, g_{ab}) is a time orientable spacetime, then there exists a smooth non vanishing timelike vector field t^a on M.*

Unfortunately, proof of this lemma had to be omitted but can be found in chapter eight of Wald (2010).[10]

Definition 2.1.2 (Future directed timelike / causal curve)**.** *A differentiable curve $\lambda(t)$ is said to be a future directed timelike curve if at each $p \in \lambda$ the tangent t^a is a future directed timelike vector.*
Similarly a curve is future directed causal if the tangent t^a at each point on the curve is either future directed timelike or null.

Analogous definitions apply to past directed timelike and causal curves.

Definition 2.1.3 (Chronological future)**.** *The set of events that can be reached by a future timelike curve starting from $p \in M$ is called chronological future of p, denoted $I^+(p)$,*

$$\begin{aligned} I^+(p) = \{q \in M \mid \exists\lambda(t) \text{ with } \lambda(0) = p \text{ and } \lambda(1) = q, \\ \text{where } \lambda \text{ is a future directed timelike curve}\} \end{aligned} \tag{2.1}$$

Physically, this is interpreted as the set of all events which can be influenced by what happens at p . One can note that a small deformation in the endpoint of a timelike curve still preserves the timelike nature of the curve, so for a $q \in I^+(p)$ we can find an open neighborhood $O \ni q$, and thus $I^+(p)$ *is an open set.* A formal proof can be given using theorem (2.1.1). For any subset $S \subset M$ we define

$$I^+(S) = \bigcup_{p \in S} I^+(p) \tag{2.2}$$

It follows that $I^+(S)$ is also an open set.
Analogous definition follows for *Chronological past*, and clearly it should also be an open set.

Definition 2.1.4 (Causal future). *Causal future of $p \in M$,denoted $J^+(p)$, is defined similarly as $I^+(p)$ except that we have "future directed causal curves" instead of timelike.*

In flat spacetime, $J^+(p)$ is a closed set but the same is not true for a general spacetime. For example, consider a spacetime in which a point from the future light cone of Minkowski spacetime is removed. In this case, $J^+(p)$ is not closed as can be seen form the figure. Moreover, in Minkowski spacetime, $I^+(p)$ consists of points that can be reached by future directed timelike geodesics from p and $\dot{I}^+(p)$ is precisely generated by future directed null geodesics starting from p. However, in general spacetime neither of these is true. But locally these properties are valid for convex complex neighborhoods.

Definition 2.1.5 (Convex normal neighborhood). *Convex normal neighborhood of p, $p \in M$, is an open set U with $p \in U$ such that for all q,$r \in U$ there exists a unique geodesic γ connecting q and r and staying entirely within U.*

Theorem 2.1.1.
1. *Let (M, g_{ab}) be an arbitrary spacetime, and let $p \in M$. Then there exists a convex normal neighborhood of p.*
2. *Furthermore, for any such U, $I^+(p)|_U$ consists of all points reached by future directed timelike geodesics starting from p and contained within U, where $I^+(p)|_U$ denotes the chronological future of p in spacetime (U,g_{ab}). In addition, $\dot{I}^+(p)|_U$ is generated by future directed null geodesics emanating from p.*

Proof.
1. This theorem is a remarkable result from differential geometry, and also non-trivial. It will be omitted from this work but interested readers can refer to page 180 of Hicks [5]. The theorem states that for any C^∞ manifold M and any C^∞ connection ∇ there exists a convex normal neighborhood for any point m of M.
2. This point is trivial but important as it is not essential for the unique geodesic connecting two points, p and q $\in$ U to lie in causal future of $I^+(p)$. But the proof is again non-trivial, and hence would be omitted. The proof is given in chapter 4 of Hawking and Ellis (1973)[2].

□

Corollary 2.1.1. *If $q \in J^+(p) - I^+(p)$, then any causal curve connecting p to q must be a null geodesic.*

Proof. Let λ be a causal curve connecting p and q. We can find a cover of convex normal neighborhood for this curve and as it is compact, extract a finite subcover $\{U_i\}$ (It is clear that we can assume that λ has no closed-loop parts, since redundant portions can be deleted). Suppose among them U_n contains q, thus the curve ends at q in U_n. Let p_n be the starting point of the curve in U_n. Suppose λ is not a null geodesic in U_n, then according to Theorem 2.1.1, $q \notin \dot{I}^+(p_n)|_{U_n}$ and $q \in I^+(p_n)|_{U_n}$ (as λ is a causal curve). This implies that there exists a unique timelike geodesic connecting p_n and q which stays entirely within U_n. That means one can change λ within each neighborhood while staying within it so that the tangent of the curve never becomes 0, making the new curve connecting p and $q$future oriented timelike. But as $q \notin I^+(p)$, p and q must be connected by a null geodesic.

One could also observe that if λ is assumed to be a null geodesic in U_n, then $q \notin I^+(p_n)$ and hence there won't be any timelike geodesic connecting the two points. Also, instead of using the neighborhood containing q, we could use any neighborhood and follow similar logic. □

Proof - As $q \notin I^+(p)$, any curve λ beginning at p and ending at q must be a null curve. Now λ is a compact set as it is a continuous image of a compact interval. Thus we can cover λ with convex normal neighborhoods and then extract a finite subcover. If λ fails to be a null geodesic in any such cover than using theorem (2.1.1), we can deform λ to be a null geodesic in that neighborhood. Upon repeating this for all neighborhoods in the subcover, we will obtain a null geodesic from p to q.

Definition 2.1.6 (Achronal)**.** *A subset $S \subset M$ is said to be achronal if there do not exist $p, q \in S$ such that $q \in I^+(p)$.*

Theorem 2.1.2. *Let $(M,\ g_{ab})$ be a time orientable spacetime, and let $S \subset M$. Then $\dot{I}^+(S)$ (if nonempty) is an achronal, three dimensional, C^0 - submanifold of M*

Proof. We will first prove the achronality of the boundary of the chronological future of S and then prove the existence of a manifold structure on the boundary and finally the continuous nature of the manifold.

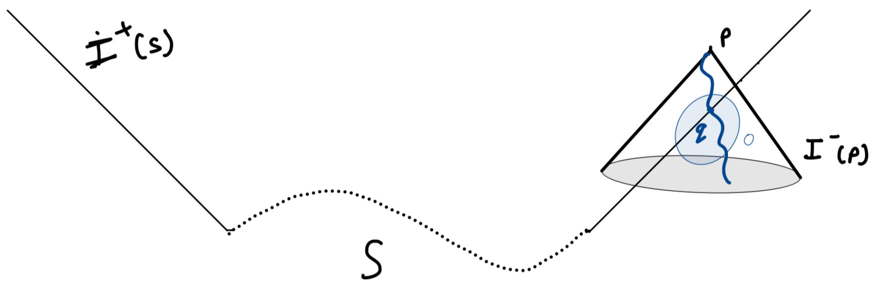

Figure 2.2: This figure shows the achronal boundary of S, and $q \in \dot{I}^+(S)$. We find a point $p \in I^+(q)$, therefore $q \in I^-(p)$, and O is such that $q \in O \subset \dot{I}^+(S)$

1. Consider a point $q \in \dot{I}^+(S)$. Now if $p \in I^+(q)$, then $q \in I^-(p)$. Moreover, since $I^-(p)$ is open, we can find an open neighborhood O of q, such that $O \subset I^-(p)$. Now since q is on the boundary of $I^+(S)$, any open subset of q will have elements which are in $I^+(S)$, therefore $I^+(S) \cap O \neq \emptyset$. Furthermore, by the continuous property of future timelike curves we have that $p \in I^+[O \cap I^+(S)] \subset I^+[I^+(S)] \subset I^+[S]$. This is true for all $p \in I^+(q)$,
$$\therefore \quad I^+(q) \subset I^+(S).$$

Now if $\dot{I}^+(S)$ is not achronal then we can have $q, r \in \dot{I}^+(S)$ such that $r \in I^+(q)$ which implies $r \in I^+(S)$. However, as $I^+(S)$ is open, we have $I^+(S) \cap \dot{I}^+(S) = 0$. This contradicts our assumptions, and thus $\dot{I}^+(S)$ is achronal.

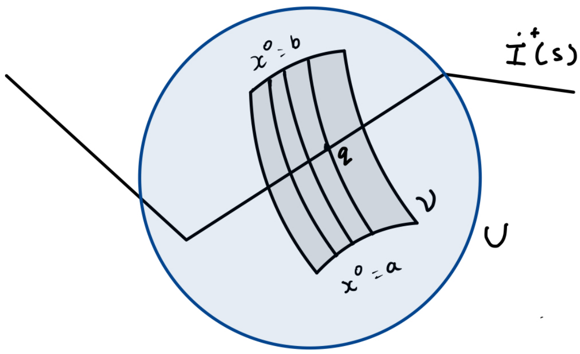

Figure 2.3: A diagram showing $q \in \dot{I}^+(S)$ such that a unique integral curve of $\partial/\partial x^0$ passes through q

2. To prove manifold structure of $\dot{I}^+(S)$, consider Riemann normal coordinates (x^0, x^1, x^2, x^3) in an open set U of $q \in \dot{I}^+(S)$, we will refer to this chart as ψ. Now using ψ we find $\mathcal{V}$ such that (1) $(\partial/\partial x^0)^a$ is everywhere future directed timelike, (2) $\psi(\mathcal{V}) \in \mathbb{R}^4$ is of the form, $(a, b) \times \mathcal{N} \subset \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$, and (3) the slice $x^0 = a$ of $\mathcal{V}$ is in $I^-(q) \cap U$ & the slice $x^0 = b$ of $\mathcal{V}$ is in $I^+(q) \cap U$. This means that all integral curves of $(\partial/\partial x^0)^a$ intersect $I^+(q) \subset I^+(S)$ and $I^-(q) \subset M - \overline{I^+(S)}$. Thus each of these integral curves must intersect $\dot{I}^+(S)$. Due to achronality of the boundary, timelike curves can only intersect $\dot{I}^+(S)$ once. Moreover, integral curves of $\partial/\partial x^0$ have constant (x^1, x^2, x^3), and thus can be characterised by these coordinates. Hence, we have established that each curve has a unique intersection point on $\mathcal{V} \cap \dot{I}^+(S)$ and each curve will have a unique set of coordinates in $\mathbb{R}^3$, which implies that in each such neighborhood we get a one-to-one association of points on $\dot{I}^+(S)$ with $\mathbb{R}^3$.

3. To prove the continuity of $\dot{I}^+(S)$, let $y \in \mathcal{N} \subset \mathbb{R}^3$, then the x^0 coordinate curve

$$s \mapsto \psi^{-1}(s, y)$$

should intersect $\dot{I}^+(S)$, and it will have a unique point of intersection. Let $h(y)$ be the x^0 coordinate of this point.

To prove continuity of the boundary, it is enough to prove continuity of h. Consider a converging sequence $\{y_n\}$ that converges in $\mathcal{N}$ to y. If $\{h(y_n)\}$ converges to $h(y)$, then h will be continuous. Since h gives value in a bounded interval of $\mathbb{R}$, for every sequence $\{h(y_n)\}$ there will be a convergent sub sequence $\{h(y_m)\}$ in $\mathbb{R}$(Bolzano - Weirstrass theorem). Let the convergent point of this subsequence be $r \neq h(y)$. This would imply that $\psi^{-1}(r, y) \notin \dot{I}^+(S)$, $\therefore \psi^{-1}(r, y) \in [(I^-(q) \cap \mathcal{V}) \cup (I^+(q) \cap \mathcal{V})]$, where $q = \psi^{-1}(h(y), y) \in \dot{I}^+(S)$. This is because the timelike curve mentioned before will connect q and $\psi^{-1}(r, y)$. However, for large enough n, $h(y_n)$ can get arbitrarily close to r, which implies $\psi^{-1}(h(y_n), y) \in [(I^-(q) \cap \mathcal{V}) \cup (I^+(q) \cap \mathcal{V})]$.

Since $\psi(h(y_n), r) \in \dot{I}^+(S)$, this causes a contradiction. Therefore, $h(y) = r$, which proves continuity of $\dot{I}^+(S)$. This proof was largely based on the one given in O'neill (1983)[7].

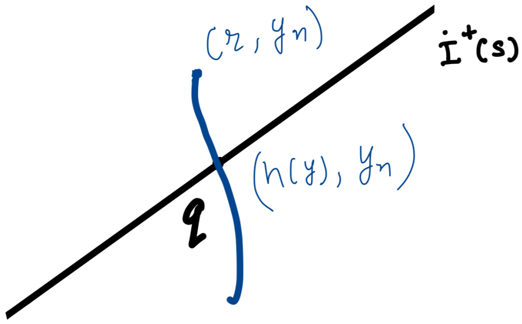

Figure 2.4: $q \in \dot{I}^+(S)$, r is the convergent point of $\{h(y_n)\}$, and by assumption $(r, y_n) \in I^+(q)$

$\square$

2.2 Global Hyperbolicity

Now, for a closed, achronal set, S, we define the *Future domain of dependence* of S, $D^+(S)$, by

$$D^+(S) = \{p \in M \mid \text{Every past inextendible causal curve through } p \text{ intersects } S\}$$

$D^+(S)$ is also known as the future Cauchy development, it is the region of spacetime that can be predicted from the data on S.

Similarly, we define *Past domain of dependence* of S by,

$$D^-(S) = \{p \in M \mid \text{Every future inextendible causal curve through } p \text{ intersects } S\}$$

$D(S) = D^+(S) \cup D^-(S)$ is called the (full) *domain of dependence* of S.

Definition 2.2.1. *A closed, achronal set Σ for which $D(\Sigma) = M$ is called a* ***Cauchy Surface****.*

Since Σ is achronal, we may think of Σ as representing an "instant of time" throughout the Universe. The existence of a Cauchy surface implies that there is no loss of information in the Universe, making it reasonable to assume that our Universe admits such a surface. This will be one of the hypothesis assumed while proving the Singularity theorem.

Definition 2.2.2. *A spacetime which admits a Cauchy surface is known as a* ***Globally hyperbolic spacetime****.*

Now we will state an important theorem which will be crucial in the proof of the singularity theorem.

Theorem 2.2.1. *Let* (M, g_{ab}) *be a globally hyperbolic spacetime. Then a global time function,* f*, can be chosen such that each surface of constant* f *is a Cauchy surface. Thus* M *can be foliated by Cauchy surfaces and the topology of* M *is* $\mathbb{R} \times \Sigma$*, where* Σ *denotes any Cauchy surface.*

Proof Unfortunately, proof of this theorem is not within the scope of this thesis and could be read from *General Relativity - R. Wald*[10].

Chapter 3

Geodesic congruences

To begin our discussion about singularity, we first need to understand the concept and properties of geodesic congruences. Roughly, a congruence of geodesic is a family of non-intersecting geodesics. Geodesic congruences are important as their properties help us understand the local behaviour of the spacetime, as we shall see.

First we pave our way through to understand timelike Raychaudhuri equation. This part is entirely from Wald (2010) [10]. In the following section, we deal with null case analogous of Raychaudhuri equation. This section is my work but inspired from analogous proof of the timelike case given in Wald (2010).

Definition 3.0.1. *Let (M, g_{ab}) be a spacetime and let $O \subset M$ be an open subset. A congruence in O is a family of curves such that through each $p \in O$ there passes precisely one curve in the family.*

Remark-In the above definition if we replace curves by geodesics then we get a congruence of geodesics.
It can be clearly seen that tangents to a congruence yield a vector field in O, and conversely every continuous vector field generates a congruence of curves.
Now, we will prove a crucial result needed to prove the singularity theorem.

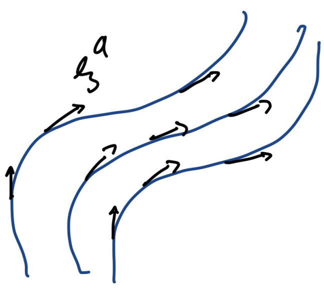

Figure 3.1: defining a vector field ξ^a by taking tangents at every point on a geodesic of a congruence

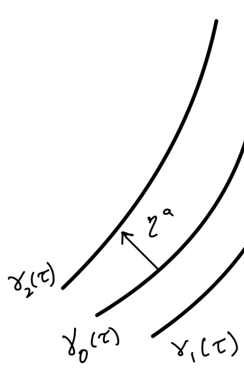

Figure 3.2: Deviation vector of a geodesic congruence

3.1 Raychaudhury's equation

3.1.1 Timelike case

Consider a smooth congruence of timelike geodesics, or geodesics whose tangent at every point is a timelike vector. As discussed above the congruence will give a vector field, ξ^a, of timelike vectors. Without loss of generality, we can assume that $\xi^a\xi_a = -1$. Using this we define a tensor field

$$B_{ab} = \nabla_b \xi_a.$$

And as $\nabla_b(\xi^a\xi_a) = 0$, we have $(\nabla_b\xi_a)\xi^a + \xi_a(\nabla_b\xi^a) = 0$,

$$\begin{aligned} \therefore (\nabla_b\xi_a)\xi^a &= 0 \\ \text{or } B_{ab}\xi^a &= 0. \end{aligned} \tag{3.1}$$

Also ξ^a must follow the geodesic equation, i.e.,

$$\begin{aligned} \xi^b\nabla_b\xi^a &= 0 \\ \text{or } B_{ab}\xi^b &= 0. \end{aligned} \tag{3.2}$$

To understand the physical significance of this tensor, we will need to refer to the concept of geodesic equation and deviation vectors (equation (A.3) in Appendix BA. Consider a smooth one parameter family of geodesics $\gamma_s(\tau)$ in the congruence, and let η^a be an orthogonal deviation vector on γ_0, see figure 3.2. We have

$$\begin{aligned} &\mathcal{L}_\xi \eta^a = 0, \\ \implies \ &\xi^b\nabla_b\eta^a = \eta^b\nabla_b\xi^a = B^a{}_b\eta^b. \end{aligned} \tag{3.3}$$

In an arbitrary spacetime it is not necessary for the deviation vector to satisfy parallel transport equation, therefore as we can see above, B^a_b measures the failure of η^a to be parallel transported. Physically, an observer on γ_0 will note nearby geodesics twist and stretch by $B^a{}_b$.

We assume that this congruence is hypersurface orthogonal, figure 3.3 , thus admits a

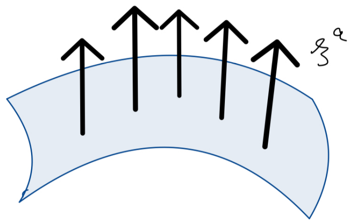

Figure 3.3: Hypersurface orthogonal to ξ^a

hypersurface (definition of a hypersurface can be found in appendix B.1). Now, we define a metric which acts as a projection operator onto this hypersurface, that is,

$$h_{ab} = g_{ab} + \xi_a \xi_b$$

which is also known as spatial metric. It is clear that

$$\begin{aligned} h_{ab}h^{ab} &= (g_{ab} + \xi^a \xi^b)(g_{ab} + \xi^a \xi^b) \\ &= 3. \end{aligned} \tag{3.4}$$

This is expected as h_{ab} is a metric on the hypersurface orthogonal to ξ^a, which is evidently 3-dimensional. Using the spatial metric we can define *expansion* θ, *shear* σ, and *twist* ω_{ab} as

$$\begin{aligned} \theta =& B^{ab} h_{ab}, \\ \sigma_{ab} =& B_{(ab)} - \frac{1}{3}\theta h_{ab}, \\ \omega_{ab} =& B_{[ab]}. \end{aligned}$$

We can verify that B_{ab} decomposes as follows

$$\begin{aligned} \frac{1}{3}\theta h_{ab} + \sigma_{ab} + \omega_{ab} &= \frac{1}{3}B^{cd}h_{cd}h_{ab} + B_{(ab)} - \frac{1}{3}B^{cd}h_{cd}h_{ab} + B_{[ab]} \\ &= \frac{1}{2}(B_{ab} + B_{ba}) + \frac{1}{2}(B_{ab} - B_{ba}) \\ &= B_{ab}. \end{aligned}$$

Furthermore, we would like to see how this tensor field changes along a geodesic. Using (3.1), and (1.5) we solve

$$\begin{aligned} \xi^c \nabla_c B_{ab} &= \xi^c \nabla_c \nabla_b \xi_a \\ &= \xi^c \nabla_b \nabla_c \xi_a + R^d_{cba} \xi^c \xi_d \\ &= \nabla_b (\xi^c \nabla_c \xi_a) - (\nabla_b \xi^c)(\nabla_c \xi_a) + R^d_{cba} \xi^c \xi_d \\ &= -B^c{}_b B_{ac} + R^d_{cba} \xi^c \xi_d, \end{aligned} \tag{3.5}$$

where $\xi^c \nabla_c \xi_a = \xi^c B_{ca} = 0$.

Next, by taking the trace of equation (3.5) with the spatial metric (3.1.1) we obtain

$$\xi^c (\nabla_c B_{ab}) h^{ab} = -B^c{}_b B_{ac} h^{ab} + R_{cba}{}^d \xi^c \xi_d h^{ab}$$
$$\implies \xi^c \nabla_c \theta - \xi^c B_{ab} \nabla_c \xi^a \xi^b = -B_{bc} B^{cb} - B^c{}_b B_{ac} \xi^a \xi^b + R_{bc}{}^{db} \xi^c \xi_d + R_{cba}{}^d \xi^c \xi_d \xi^a \xi^b. \quad (3.6)$$

Note that upon expanding $B_{bc}B^{cb}$ we get,

$$\begin{aligned} B_{ab} B^{ba} &= \frac{1}{9} \theta^2 h_{ab} h^{ab} + \sigma_{ab} \sigma^{ba} + \omega_{ab} \omega^{ba} \\ &+ \left(\frac{1}{3} \theta h_{ab} \sigma^{ba} + \frac{1}{3} \theta \sigma_{ab} h^{ba} \right) + \left(\frac{1}{3} \theta \omega_{ab} h^{ba} + \frac{1}{3} \theta h_{ab} \omega^{ba} \right) \\ &+ \sigma_{ab} \omega^{ba} + \omega_{ab} \sigma^{ba} \end{aligned}$$

To simplify this equation we can make the following steps -

1. Firstly, we have

$$\begin{aligned} \sigma_{ab} h^{ba} &= \left(B_{(ab)} - \frac{1}{3} \theta h_{ab} \right) h^{ba} \\ &= \frac{1}{2} \left(B_{ab} h^{ba} + B_{ba} h^{ba} \right) - \theta \\ &= \theta - \theta = 0. \end{aligned}$$

2. Also, $\omega_{ab} h^{ab} + \omega_{ba} h^{ab} = 0$, because of antisymmetry of ω_{ab} and similarly, $\sigma_{ab} \omega^{ba} + \omega_{ab} \sigma^{ba} = 0$.

And finally, upon using equation (3.4) we obtain

$$B_{ab} B^{ba} = \frac{1}{3} \theta^2 + \sigma_{ab} \sigma^{ba} + \omega_{ab} \omega^{ba}. \quad (3.7)$$

Looking back at equation (3.6), we note that the second term on the right-hand side approaches zero due to equation (3.1). Furthermore, expanding the second term on the left-hand side also yields zero. Additionally, since the Riemann tensor is antisymmetric in its first two indices, the last term in (3.6) also vanishes. Therefore from equation (3.6) and equation (3.7), we arrive at

$$\xi^c \nabla_c \theta = -\frac{1}{3} \theta^2 - \sigma_{ab} \sigma^{ba} + \omega_{ab} \omega^{ab} - R_{cd} \xi^c \xi^d \quad (3.8)$$

This is known as *Raychaudhuri's equation*, a fundamental equation that is crucial for understanding how the expansion of a geodesic congruence evolves along a geodesic.

This concludes our derivation of Raychaudhuri's equation for timelike geodesic congruences. In the next section we develop an analogous equation for a null geodesic congruence.

3.1.2 Null geodesic Congruence

Similar to timelike case, we consider a congruence of null geodesics. Thus we again have a null vector field k^a corresponding to this congruence. But in case of null geodesic congruence the task to derive an equation analogous to (3.8) is more complex. In timelike case we restricted our discussion to deviation vectors η^a orthogonal to ξ^a. This was because any arbitrary deviation vector could be decomposed into an orthogonal and a parallel (to ξ^a) component, and the parallel component would not change upon parallel transport (geodesic equation). But there is no natural way to do this decomposition in null case as the null vectors are orthogonal to themselves.
Also in timelike case we normalised the vectors by $\xi^a\xi_a = -1$. Basically that we parameterized the curve by the arc length. However this is clearly impossible with null curves, as they have zero arc length. Thus we can only choose an affine parameter (v) which satisfies the geodesic equation,

$$\frac{D}{dv}k^a = k^b\nabla_b k^a = 0.$$

Also, deviation vectors which differ only by a multiple of k^a again represent a displacement to the same nearby geodesic (Appendix). Thus the physically interesting quantity is an equivalence class of deviation vectors.
Let V_p be the vector space at $p \in M$. Consider a subspace $\tilde{V}_p$ orthogonal to k^a. It is clear that $\tilde{V}_p$ is a 3-dimensional space. Unlike timelike case, $\tilde{V}_p$ is not spanned by the deviation vectors orthogonal to k^a, as k^a is orthogonal to itself. Thus one is interested in a vector space of equivalence class $\hat{V}_p$ of $\tilde{V}_p$. For two vectors $x, y \in \tilde{V}$, we have

$$x^a \sim y^a \quad \text{if} \tag{3.9}$$
$$x^a - y^a = ck^a. \tag{3.10}$$

Thus we have,

$$\tilde{V}_p = \{v^a \in V_p | v^a k_a = 0\}$$
$$\hat{V}_p = \tilde{V}_p / \sim .$$

Proposition 3.1.1. *$\hat{V}_p$ is a 2 dimensional vector space.*

Proof. To prove this property of the hatted space we will consider a frame on V_p, namely, $e_a, e_b, e_1.e_2$, such that

$$g(e_a, e_b) = -1,$$
$$g(e_i, e_j) = \delta_{ij},$$
$$g(e_a, e_a) = 0,$$
$$g(e_b, e_b) = 0,$$
$$g(e_a, e_i) = 0,$$
$$g(e_b, e_i) = 0,$$
$$\text{And} \quad e_a = k^a.$$

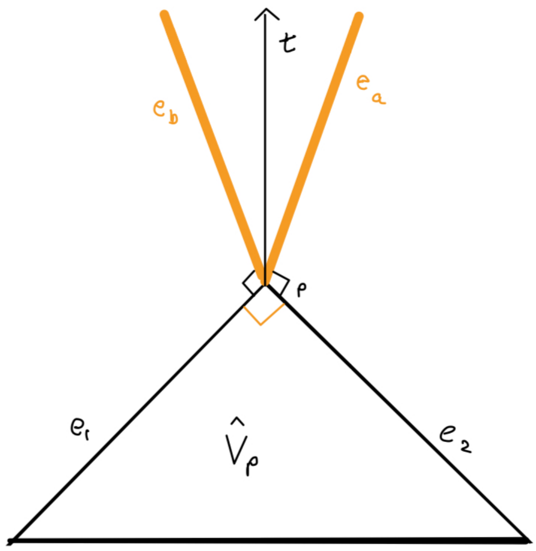

Figure 3.4: $\{e_a, e_b, e_1, e_2\}$ are the above mentioned frames on V_p, e_a, e_b represent the incoming and outgoing light rays, and $\hat{V}_p$ is the space spanned by $\{e_1, e_2\}$

Physically, e_a and e_b represent incoming and outgoing light rays at p. We can see from the properties of the basis that $\tilde{V}_p = span\ [e_1, e_2, e_a]$. Thus $\tilde{V}_p$ is a 3 dimensional space. Now upon taking a quotient of this space with the equivalence relation in (3.10), to get $\hat{V}_p$, we see that all the vectors spanned by e_a will become the zero vector of this space. And hence $\hat{V}_p$ will be spanned by $[e_1, e_2]$. □

Furthermore, we can define dual vectors in the hatted space. A dual vector $\mu_a \in V_p^*$ naturally gives rise to a dual vector $\tilde{\mu}_a \in \tilde{V}_p^*$ by restricting its action to vectors in $\tilde{V}_p$. However, $\tilde{\mu}_a$ can define $\hat{\mu}$ if and only if $\tilde{\mu}_a k^a = \mu_a k^a = 0$. This ensures that addition of a constant multiple of k^a to a vector $\hat{\eta}^a \in \hat{V}_p$ won't affect the dot product with the dual vector. More generally, a tensor $T^{a_1,\dots a_k}{}_{b_1,\dots b_l}$ can give a projection $\hat{T}^{a_1,\dots a_k}{}_{b_1,\dots b_l}$ over $\hat{V}_p$ if upon contracting any one of its indices with k^a or k_a gives 0.
This property is satisfied the tensors from previous section, namely, $B_{ab}, g_{ab}, \omega_{ab}, \sigma_{ab}$. Projection of g_{ab} over $\hat{V}_p$ shall be denoted by $\hat{h}_{ab}$.

Now we have sufficient construction to have a Raychaudhuri analogous equation for null geodesics. As before, we perform decomposition of $\hat{B}_{ab}$, i.e.,

$$\hat{B}_{ab} = \frac{1}{2}\theta\hat{h}_{ab} + \hat{\sigma}_{ab} + \hat{\omega}_{ab},$$

such that,

$$\theta = \hat{h}^{ab}\hat{B}_{ab}, \tag{3.11}$$

$$\hat{\sigma} = \hat{B}_{(ab)} - \frac{1}{2}\theta\hat{h}_{ab}, \tag{3.12}$$

$$\hat{\omega} = \hat{B}_{[ab]}. \tag{3.13}$$

Again $\theta, \hat{\sigma}, \hat{\omega}$ can be interpreted as expansion, shear, and twist of the geodesic congruence in consideration. As discussed $\hat{h}_{ab}$ is the metric of a two dimensional vector space, and moreover it is positive definite ($g(e_1, e_2) = \delta_{12}$). Thus the trace of this metric will equal to 2, as opposed to (3.4). This causes the change in the numerical factor of the term with expansion, i.e., $\frac{1}{2}\hat{h}_{ab}\theta$.

The same derivation as in the timelike case now leads to

$$k^c \nabla_c B_{ab} + B^c{}_b B_{ac} = R_{cba}{}^d k_d k^c.$$

All the tensors in this equation can be projected to the hat space, hence upon hatting this equation we obtain

$$k^c \nabla_c \hat{B}_{ab} + \hat{B}^c{}_b \hat{B}_{ac} = \widehat{R_{cba}{}^d k_d} k^c. \tag{3.14}$$

Upon contracting (3.14) with the hatted metric, we find that

$$k^c \nabla_c \theta = -\hat{B}^c_b \hat{B}^b_c - \hat{R}_{cd} k^c k^d.$$

Substituting the decomposition of $\hat{B}_{ab}$, we obtain

$$k^c \nabla_c \theta = \frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \hat{\sigma}_{ab}\hat{\sigma}^{ab} + \hat{\omega}_{ab}\hat{\omega}^{ab} - R_{cd} k^c k^d. \tag{3.15}$$

The nature of expansion for null geodesics congruence given by equation (3.15) is very similar to that of Raychaudhuri's equation (3.8).

3.2 Energy conditions

To further develop our arguments for the existence of singularity, we will need to take into account the matter distribution of the universe. This is done by assuming certain constraints on the stress energy tensor, which are known as the energy conditions.

Einstein's equation for a timelike vector ξ^a is

$$R_{ab}\xi^a\xi^b = 8\pi \left[T_{ab} - \frac{1}{2} T g_{ab} \right] \xi^a \xi^b = 8\pi \left[T_{ab}\xi^a\xi^b + \frac{1}{2} T \right] \tag{3.16}$$

where, $T_{ab}\xi^a\xi^b$ denotes the energy density of the Universe. It is generally believed to be non-negative, i.e.,

$$T_{ab}\xi^a\xi^b \geq 0. \tag{3.17}$$

This is known as the *weak energy condition.*

It is also assumed that stresses of matter, denoted by T will not become so large and negative to make right hand side of (3.16) negative, i.e.,

$$T_{ab}\xi^a\xi^b \geq -\frac{1}{2} T. \tag{3.18}$$

This is known as the *strong energy condition.*

Returning to the Raychaudhuri equation (3.8), if Einstein's equation holds, the strong energy condition is satisfied by T_{ab}, and the congruence is hypersurface orthogonal, which implies $\omega_{ab} = 0$ (appendix B.1), then the right hand side of Raychaudhuri equation will always be negative, as $\sigma_{ab}\sigma^{ab}$ is manifestly positive. Therefore, if these conditions are assumed then the expansion should follow

$$\frac{d\theta}{d\tau} + \frac{1}{3}\theta^2 \leq 0$$

$$\Rightarrow \frac{-1}{\theta^2}\frac{d\theta}{d\tau} \geq \frac{1}{3}.$$

Which leads to

$$\frac{d}{d\tau}(\theta^{-1}) \geq \frac{1}{3}$$

and hence,

$$\theta^{-1}(\tau) \geq \theta_0^{-1} + \frac{1}{3}\tau, \tag{3.19}$$

where θ_0 is the initial value of θ.

Note - Suppose that the congruence is initially converging, i.e., $\theta_0 < 0$. Then equation (3.19) implies that θ^{-1} must pass through zero, i.e., $\theta \mapsto -\infty$ within a (finite) proper time $\tau \leq 3/|\theta_0|$. Thus we proved the following lemma.

Lemma 3.2.1. *Let ξ^a be the tangent field of a hypersurface orthogonal timelike geodesic congruence. Suppose $R_{ab}\xi^a\xi^b \geq 0$, as will be the case if Einstein's equation holds and the strong condition is satisfied by the matter. If the expansion θ takes the negative value θ_0 at any point on a geodesic in the congruence, then θ goes to $-\infty$ along the geodesic within proper time $\tau \leq 3/|\theta_0|$.*

Now turning our attention to null geodesics, we notice from (3.16) that using Einstein's equation we will obtain

$$R_{ab}k^ak^b = 8\pi T_{ab}k^ak^b.$$

Thus all we need to ensure that the last term of equation (3.15) is non positive is that for all null k^a,

$$T_{ab}k^ak^b \geq 0. \tag{3.20}$$

If the strong energy condition, equation (3.18), holds then we have $T_{ab}\xi^a\xi^b - \frac{1}{2}T\xi^a\xi_a \geq 0$. And, from continuity of ξ^a, (3.18) will hold for all null vector fields k^a. Thus, the strong energy condition will satisfy (3.20). Similarly, if the weak energy condition (3.17) holds, then equation (3.20) will also be satisfied. Furthermore, similar to timelike case, $\hat{\sigma}_{ab}\hat{\sigma}^{ab}$ will be negative, and $\hat{\omega}_{ab}$will vanish for a hypersurface-orthogonal geodesic congruence. Therefore under either *weak or strong* energy condition, and assuming hypersurface orthogonality, we find that

$$\frac{d\theta}{d\lambda} + \frac{1}{2}\theta^2 \geq 0,$$

and hence, similar to timelike case,

$$\theta^{-1}(\lambda) \geq \theta_0^{-1} + \frac{1}{2}\lambda,$$

where θ_0 is again the initial value of θ. Thus, we arrive at the following lemma.

Lemma 3.2.2. *Let k^a be the tangent field of a hypersurface orthogonal null geodesic congruence. Suppose $R_{ab}k^a k^b \geq 0$, as will be the case if Einstein equation is followed and suppose the matter in spacetime satisfies either weak or strong energy condition. If the expansion θ takes negative value along the geodesic in the congruence, then θ goes to $-\infty$ along that geodesic within affine length $\lambda \leq 2/|\theta_0|$.*

In this section we started with the definition of geodesic congruence , then derived Raychaudhuri's equation which describes the evolution of expansion of a congruence. And then by using strong energy condition, we were able to prove that initially converging geodesics, either timelike or null, should converge to a singular point within finite amount of proper time. This provides a hint at a possible singularity but to be more rigorous we will need the help of conjugate points.

Chapter 4

Conjugate points

In this chapter, we obtain a criteria for when a timelike geodesic fails to be a local maximum of proper time between two points. Similarly, we obtain a criteria for when a null geodesic fails to remain on the boundary of the future of a point or a two-dimensional surface.

This chapter is again mostly taken from Wald(2010).[10]. However proof of theorem 4.3.1 was too complicated to be included here, so instead a sketch of the proof is provided. This sketch is my own work, based on my intuitive understanding from reading Hawking and Penrose (2010)[1].

Consider a timelike geodesic γ with v^a as the tangent vector, then η^a is known as a jacobi field if it is a solution to the geodesic deviation equation, i.e.

$$v^a \nabla_a (v^b \nabla_b \eta^c) = -R_{abd}{}^c \eta^b v^a v^d \tag{4.1}$$

Definition 4.0.1 (Conjugate points). *$p, q \in \gamma$ are known as conjugate points if $\exists$ a jacobi field η^a which is not identically 0 but vanishes at both p and q.*

Unlike deviation vector, jacobi field is defined without a congruence of geodesics. Thus, there need not be two geodesics passing through a point if the jacobi field vanishes at that point. Now consider a congruence of timelike geodesics γ with tangent ξ^a and let $p \in \gamma$. Then every jacobi field which vanishes at p must be a deviation vector of this congruence.

4.1 Conjugate points of a congruence

In this section we will explore how the expansion of a congruence behaves at conjugate points. Intuitively, we can expect expansion to behave pathologically at these points but in this section we aim to give a proof of the same.

Proposition 4.1.1. - *Let γ be a geodesic of a timelike geodesic congruence and let $p \in \gamma$, a point $q \in \gamma$ lying to the future of p will be conjugate to p if and only if the expansion, θ of the congruence approaches $-\infty$ at q.*

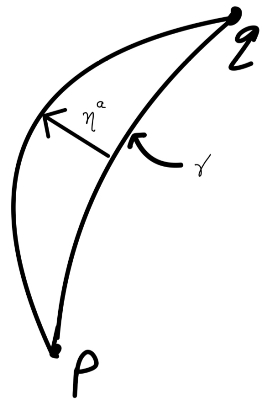

Figure 4.1: p and q are conjugate points of a congruence, and η^a is the jacobi field. In this case it can also be understood as a deviation vector.

Proof. To begin with, we will introduce an orthonormal basis of spatial vectors e_1^a, e_2^a, e_3^a orthogonal to ξ^a, and parallelly propagated along γ. The components η^μ of the deviation vector must satisfy the following differential equation

$$\frac{d^2\eta^\mu}{d\tau^2} = -R_{\alpha\beta\gamma}{}^{\mu}\xi^\alpha\eta^\beta\xi^\nu. \tag{4.2}$$

This is a second order, linear differential equations for η^μ, hence the value of η^μ at time τ must depend linearly on the initial condition $\eta^\mu(0)$ and $d\eta^\mu/d\tau(0)$ at p. Although, by definition, $\eta^\mu(0) = 0$, therefore we must have

$$\eta^\mu(\tau) = A^\mu{}_\nu(\tau)\frac{d\eta^\nu}{d\tau}(0), \tag{4.3}$$

where $A^\mu{}_\nu$ satisfies

$$\frac{d^2A^\mu{}_\nu}{d\tau^2} = -R_{4\beta 4}{}^{\mu}A^\beta{}_\nu. \tag{4.4}$$

Now for q to be a conjugate point we need $\eta^\mu = 0$ at q, and non- triviality requires that $\frac{d\eta^\mu}{d\tau}(0) \neq 0$. Suppose $\gamma(a) = q$, then upon applying equation (4.3) at a we get,

$$\eta^\mu(a) = A^\mu{}_\nu(a)\frac{d\eta^\nu}{d\tau}(0). \tag{4.5}$$

Therefore,

$$\eta^\mu(a) = 0 \Leftrightarrow detA^\mu{}_\nu(a) = 0.$$

It is evident from the definition of conjugate points that between p and q we must have non-zero $detA^\mu{}_\nu$, and so the inverse of $A^\mu{}_\nu$ exists between these two points.

We turn our attention, now, to find a relation of the matrix $A^\mu{}_\nu$ with the tensor field $B_{ab} = \nabla_b\xi_a$. Consider a component of the dual frame e^μ. Upon expanding, η^μ can be written as $\eta^\mu = (e^\mu)_b \eta^b$. We note, from equation (3.3), that

$$\begin{aligned}\frac{d\eta^\mu}{d\tau} &= \xi^a \nabla_a \eta^\mu = \xi^a \nabla_a [(e^\mu)_b \eta^b] \\ &= (e^\mu)_b \eta^a \nabla_a \xi^b \\ &= (e^\mu)_b B^b{}_a \eta^a \\ &= B^\mu{}_a \eta^a,\end{aligned}$$

where in the second line we used the fact that the basis is parallelly propagated along γ, and that deviation vector is orthogonal to ξ^a. Moreover, from equation (4.5), we have

$$\frac{d\eta^\mu}{d\tau} = \frac{dA^\mu{}_\nu}{d\tau} \frac{d\eta^\nu}{d\tau}(0).$$

Thus upon comparing the above two equation we obtain

$$\frac{dA^\mu{}_\nu}{d\tau} \frac{d\eta^\nu}{d\tau}(0) = B^\mu{}_\alpha A^\alpha{}_\nu \frac{d\eta^\nu}{d\tau}(0).$$

This holds true for all initial conditions. Therefore, in matrix notation we can write it as

$$\begin{aligned}\frac{dA}{d\tau} &= BA, \\ \Rightarrow B &= \frac{dA}{d\tau} A^{-1}.\end{aligned}$$

This is true because A is non singular everywhere except at p and q. Now upon taking the trace of this equation we find

$$\theta = trB = tr\left[\frac{dA}{d\tau} A^{-1}\right].$$

Using the property of the trace, namely,

$$tr\left[\frac{dA}{d\tau} A^{-1}\right] = \frac{1}{detA} \frac{d}{d\tau}(detA),$$

it implies that

$$\theta = \frac{d}{d\tau}(ln|detA|). \tag{4.6}$$

Now from equation (4.4) we can see that $d(det\ A)/d\tau$ will not go to infinity anywhere along the geodesic γ. Therefore $\theta \to -\infty$ at q $\iff$ $det\ A \to 0$ at q. Observe, from equation (4.5), that $detA$ vanishes if and only if the jacobi field vanishes.

□

We are now able to prove the following proposition regarding existence of conjugate points.

Proposition 4.1.2. *Let (M, g_{ab}) be a spacetime satisfying $R_{ab}\xi^a\xi^b \geq 0$ for all timelike ξ^a. Let γ be a timelike geodesic and let $p \in \gamma$. Suppose the convergence of the congruence of timelike geodesics emanating into the future point from p attains the negative value θ_0 at $r \in \gamma$. Then within proper time $\tau \leq 3/|\theta_0|$ from r along γ there exists a point q conjugate to p, assuming that γ extends that far.*

Proof. From lemma (3.2.1), we clearly have that if the expansion is negative at some point then within finite amount of proper time the expansion will go to ∞. And by using proposition (4.1.1) we get the desired result. $\square$

Now we focus on the congruent points of a null geodesic congruence. A Jacobi field for null geodesic congruence again must follow the jacobi equation (4.1). Let k^a be the tangent along a null geodesic μ and η^a be the Jacobi field, then the jacobi equation is written as

$$k^a\nabla_a(k^b\nabla_b\eta^c) = -R^c_{abd}\eta^b k^a k^d. \tag{4.7}$$

Using the geodesic equation, it follows that we can move the tangent vector k^c in and out of the derivative, therefore $k_c k^a\nabla_a(k^b\nabla_b\eta^c) = k^a\nabla_a(k^b\nabla_b\eta^c k_c)$. But from the definition of Reimann curvature, contracting right side of equation (4.7) with k_c will give us 0. Thus we have,

$$k^a\nabla_a[k^b(\nabla_b\eta^c)] = 0. \tag{4.8}$$

This shows that $k^a\eta_a$ is a linear function of the affine parameter $\implies$

$k^a\eta_a$ cannot vanish at two points unless $k^a\eta_a$ vanish at all points.

Proposition 4.1.3. *If η^a is a Jacobi field then so is $\eta^a + (a + b\lambda)k^a$.*

Proof. Let η^a be a Jacobi field and thus it must follow equation (4.7). Upon substituting η^c with $\eta^c + (a + b\lambda)k^c$ in the left side of equation (4.7), we can see that the extra term added to the deviation vector follows

$$k^a\nabla_a(k^b\nabla_b(a + b\lambda)k^c) = ak^a\nabla_a(k^b\nabla_b k^c) + bk^a\nabla_a(k^b\nabla_b(\lambda k^c)).$$

The last term on the right hand side can be solved using leibniz rule and $\nabla_k\lambda = 1$, therefore it is evident that the above line of equation will be zero, and hence the left side of (4.7) will remain unchanged.

Now looking at the right side of (4.7), as before $R^c_{abd}k^b k^a k^d = 0$. Therefore upon substituting the jacobi field with the one proposed above, the jacobi equation will remain the same and thus the new field should also be a jacobi field. $\square$

This gives us the following corollary.

Corollary 4.1.1. *p and q are conjugate points $\iff$ $\exists\eta^a$ such that $\eta^a_p = dk^a$ and $\eta^a_q = ck^a$ for some real c and d.*

Proof. - By the definition of conjugate points, there exists a jacobi field which vanishes at p and q. So it follows from the above proposition that at those points there should exist a jacobi field which is a multiple of k^a.

If there is a Jacobi field dk^a at p and ck^a at q, then $dk^a + (a+b\lambda)k^a$ and $dk^a + (a+b\lambda)k^a$ are also Jacobi fields at the respective points. Thus upon using the freedom of the affine parameter, we can get Jacobi fields which vanish at both p and q. And thus imply that p and q are conjugate. $\square$

So now we have two properties that jacobi fields of a null geodesic congruence should satisfy -

1. $k^a \eta_a = 0$ for all points on a null geodesics μ.
2. if η^a is a jacobi field at a point then so is $\eta^a + ck^a$.

Therefore η^a will have a projection in $\hat{V}_p$ introduced in Section (1.1.2). Let us denote this projection by $\hat{\eta}^a$. In the projection space at p, $(\hat{V}_p)$, $ck^a \sim 0$. Similar holds true at q. Therefore using the corollary 4.1.1, we proved the following proposition.

Proposition 4.1.4. *Along a null geodesic μ, the points $p, q \in \mu$ will be conjugate if and only if a vector $\hat{\eta}^a$ in $\hat{V}$ satisfies the geodesic deviation equation and vanishes at p and q.*

Now consider a null geodesic congruence containing the two-dimensional family of null geodesics emerging from p. Then all the $\hat{\eta}^a$ (from the above proposition) will be deviation vectors of such a congruence. We can now prove the following proposition about existence of congruent points for null geodesics.

Proposition 4.1.5. *Let (M, g_{ab}) be a spacetime satisfying $R_{ab}k^a k^b \geq 0$ for all null k^a. Let μ be a null geodesic and let $p \in \mu$. Suppose the convergence, θ, of the null geodesics from p attains the negative value θ_0 at $r \in \mu$. Then within affine length $\lambda \leq 2/|\theta_0|$ from r, there exists a point q conjugate to p along μ, assuming that μ extends that far.*

Proof. The proof will be very much analogous to the one in timelike case,i.e, proposition (4.1.1). But instead of the 3 dimensional vector space, orthogonal to a timelike vector ξ^a, we will use 2 dimensional vector space $\hat{V}_p$ and corresponding vectors $\hat{\eta}^a$. Consider the basis, from section 3.1.2, $\{e_1, e_2\}$ of $\hat{V}_p$ which are parallelly propagated along γ. As before, the components of deviation vector satisfy,

$$\frac{d^2\hat{\eta}^i}{d\lambda^2} = -R_{\alpha\beta\gamma}{}^i k^\alpha \hat{\eta}^\beta k^\gamma, \tag{4.9}$$

where $i = 1, 2$, and λ is an affine parameter. One can observe that this is again a second order, linear differential equation, and hence $\hat{\eta}(\tau)$ must be a linear function of $\hat{\eta}(0)$ and $d\hat{\eta}^\mu/d\tau(0)$. Since, by construction $\hat{\eta}(0) = 0$, we must have

$$\hat{\eta}^i(\lambda) = \hat{A}^i{}_j(\lambda)\frac{d\hat{\eta}^j}{d\lambda}(0), \tag{4.10}$$

where $\hat{A}^i{}_j$ is a 2x2 matrix. Clearly, $d\hat{A}^i{}_j(0)/d\lambda = \delta^i_j$. Substituting the above equation in equation (4.9), we see that $A^\mu{}_\nu$ satisfies

$$\frac{d^2\hat{A}^i{}_j}{d\lambda^2} = -R_{\alpha l\sigma}{}^i k^\alpha k^\sigma \hat{A}^l{}_j. \tag{4.11}$$

Now to have a conjugate point q we need a deviation vector which vanishes at q but is non trivial at other points along the geodesic, thus $d\hat{\eta}/d\lambda(0) \neq 0$. Hence it is clear from (4.10) that to have $\hat{\eta}(\lambda_1) = 0$, for some λ_1, we should have $det\hat{A}^i{}_j(\lambda_1) = 0$. Moreover,

$det\ \hat{A} = 0$ is the necessary and sufficient condition for the existence a conjugate point to p.

To find a relation between the matrix $\hat{A}$ and the tensor $B_{ab}(= \nabla_a k_b)$ we note that

$$\frac{d\hat{\eta}^i}{d\lambda} = k^a \nabla_a \hat{\eta}^i. \tag{4.12}$$

Now using the above mentioned basis, a component of the deviation vector $\eta^a \in \hat{V}_p$ can be written as

$$\hat{\eta}^i = (e^i)_a \hat{\eta}^a,$$

where $i = 1, 2$.

Also the tensor B_{ab} can be projected to the dual hat space. First we can restrict B_{ab} at a point (lets say p) to the space of vectors orthogonal to k^a, i.e., $\tilde{V}_p$. Moreover, it is clear from the geodesic equation that $B_{ab}k^a = 0$, therefore $B_{ab}(\eta^a + ck^a) = B_{ab}\eta^a$. Thus the tensor is defined for the projection space and can be denoted as $\hat{B}_{ab} \in \hat{V}^*$ at all points along the geodesic.

Going back to equation (4.12), we note that

$$\begin{aligned}\frac{d\hat{\eta}^i}{d\lambda} =& k^a \nabla_a [(e))_b \hat{\eta}^b] \\ =& (e^i)_b k^a \nabla_a \hat{\eta}^b \\ =& (e^i)_b \hat{\eta}^a \nabla_a k^b \\ =& (e^i)_b \hat{\eta}^a \hat{B}_a{}^b \\ =& \hat{B}_a{}^i \hat{\eta}^a = \hat{B}^i{}_a \hat{\eta}^a,\end{aligned} \tag{4.13}$$

where in the second line we used parallel transport property of the frame. Also, in the last line we assumed that $\hat{\omega} = 0$,see equation (3.13),(which is the case for hypersurface orthogonal geodesic congruence), which implies symmetry of $\hat{B}_{ab}$. And in the third equality we used the orthogonality equation of a deviation vector, namely $\nabla_\eta k^a = \nabla_k \eta^a$. In this case it works for hatted vector as can be checked in the calculations below.

Orthogonal deviation vectors follow

$$\nabla_k \eta^a = \nabla_\eta k^a.$$

Now let $\eta'^a = \eta^a + ck^a$. Substituting this on the left hand side of above equation gives

$$\nabla_k \eta'^a = \nabla_k \eta^a.$$

And substituting it on the right hand side gives,

$$\nabla_{\eta'} k^a = (\eta^b + ck^b)\nabla_b(k^a) = \nabla_\eta k^a.$$

Thus we have,

$$\nabla_{\eta'} k^a = \nabla_k \eta'^a.$$

η' is a representation of an element of the equivalence class $\hat{\eta}$, therefore we have

$$\nabla_{\hat{\eta}} k^a = \nabla_k \hat{\eta}^a.$$

This result proves our claim, and hence we have equation (4.13).

Now $\hat{B}^i{}_a$ from equation (4.13) is a one form, thus its inner product with $\hat{\eta}^a$ can be written in matrix form,$\hat{B}^i{}_j\hat{\eta}^j$, where (i,j = 1,2). Thus by using (4.13) and (4.10) we obtain

$$\frac{d\hat{A}^i_j(\lambda)}{d\lambda}\frac{d\hat{\eta}^j}{d\lambda}(0) = \hat{B}^i_k\hat{A}^k{}_j(\lambda)\frac{d\hat{\eta}^j}{d\lambda}(0). \tag{4.14}$$

This is true for all initial conditions, hence ,in matrix notation, we have

$$d\hat{A}/d\lambda = \hat{B}\hat{A}. \tag{4.15}$$

where both $\hat{B}$ *and* $\hat{A}$ are 2×2 matrices.
Upon following the same steps as before we get to the following equation for null geodesics,

$$\theta = \frac{d}{d\lambda}(ln\ |det\ \hat{A}|). \tag{4.16}$$

Thus we have that $\theta \to -\infty \iff det\hat{A} \to 0$ and as established before, $det\hat{A} = 0 \iff \hat{\eta} = 0$. Hence, this calculation along with lemma (3.2.2) proves the desired proposition. □

4.2 Hypersurface conjugacy

In above sections we defined conjugacy of two points. Now in a similar manner we can define conjugacy between a point and a smooth spacelike hypersurface. Consider a null geodesic congruence orthogonal to a hypersurface Σ. Thus this congruence is manifestly hypersurface orthogonal.

Definition 4.2.1. *A point p on a timelike geodesic μ of the geodesic congruence orthogonal to Σ is said to be conjugate to Σ along μ if there exists an orthogonal deviation vector η^a of the congruence which is nonzero on Σ but vanishes at p.*

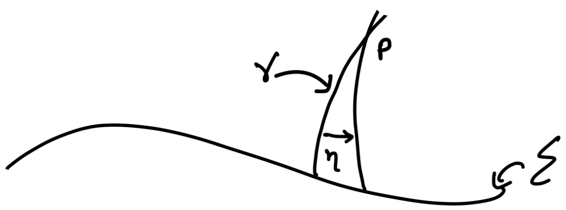

Figure 4.2: This figure shows a point, p, conjugate to the hypersurface Σ

For null geodesics, a notion of conjugacy can be defined for a point and a two-dimensional spacelike surface S. But firstly, we define the notion of congruence of geodesics orthogonal to S. For all points $q \in S$, there will exist exactly two future directed null geodesics k_1^a, k_2^a spanned by basis e_b *and* e_a . If we consider S to be orientable, then we can find a continuous choice of k_1^a and k_2^a over S, and thus define two families of null geodesics, which can be referred as "incoming" and "outgoing" families. We will now refer to each of these families as a congruences.

Definition 4.2.2. *Let μ be a null geodesic in a congruence. A point $p \in \mu$ is said to be conjugate to S along μ, if along μ there exists a deviation vector $\hat{\eta}^a$ of the congruence which is non zero on S but vanishes at p.*

Now we can prove a proposition analogous to proposition (4.1.5) but for a two dimensional surface S and a point.

Proposition 4.2.1. *Let (M, g_{ab}) be a spacetime satisfying $R_{ab}k^a k^b \geq 0$ for all null k^a. Let S be a smooth two-dimensional spacelike submanifold such that the expansion, θ, of the "outgoing" null geodesics has the negative value θ_0 at $q \in S$. Then within affine parameter $\lambda \leq 2/|\theta_0|$, there exists a point p conjugate to S along the null geodesic μ passing through q.*

Proof. - To prove this we have to amend equation (4.10) and then follow the steps from the proof of proposition (4.1.5). As the deviation vector does not vanish on the surface, we will have one more initial condition depending on $\hat{\eta}(0)$. Hence, now

$$\hat{\eta}^i(\lambda) = \hat{A}^i{}_j(\lambda)\frac{d\hat{\eta}^j}{d\lambda}(0) + \hat{H}^i{}_j\hat{\eta}^j(0). \tag{4.17}$$

Substituting this in the geodesic deviation equation, we get -

$$\frac{d^2\hat{A}^i{}_j}{d\lambda^2}\frac{d\hat{\eta}^j}{d\lambda}(0)+\frac{d^2\hat{H}^i_j}{d\lambda^2}\hat{\eta}^j(0) = -R_{\alpha l\sigma}{}^i k^\alpha k^\sigma \hat{A}^l j\frac{d\hat{\eta}^j}{d\lambda}(0) - R_{\alpha l\sigma}{}^i k^\alpha k^\sigma \hat{H}^i_j\hat{\eta}^j(0) \tag{4.18}$$

One can note that equation (4.13) still follows and (4.18) is true for all $\hat{\eta}(0)$, and $d\hat{\eta}/d\lambda(0)$. Therefore, the matrix $\hat{A}$ still follows (4.15).
Hence, now we can follow the same step as in proposition (4.1.5) and get the desired result. □

4.3 Extremum length curves

The idea of conjugate points is an essential one for the extremal length properties of time-like geodesics. In this chapter, we first prove that geodesics will be the curve of extremal length in case of timelike curves. But as we shall see, conjugate points characterize the stage in spacetimes at which a timelike geodesic fail to be a local maximum of proper time between two points and moreover, a null geodesic fails to remain on the boundary of the future of a point.

Let $p, q \in M$ and consider a smooth one-parameter family of smooth timelike curves $\lambda_\alpha(t)$ from p to q. Therefore for each $\alpha \in \mathbb{R}$, λ_α would be a timelike curve. Moreover, the curve parameter t is chosen such that for all α we have $\lambda_\alpha(a) = p$ and $\lambda_\alpha(b) = q$. Let the tangent vectors $(\partial/\partial t)^a$, by T^a and the deviation vectors, $(\partial/\partial\alpha)$, by X^a. Then X^a vanishes at both p and q and, we also have $\mathcal{L}_T X^a = T^b\nabla_b X^a - X^b\nabla_b T^a = 0$ everywhere. Length of a curve can be defined as

$$\tau(\alpha) = \int_a^b f(\alpha, t)dt, \tag{4.19}$$

where $f = (-T^aT_a)^{1/2}$. Now we prove that the necessary and sufficient condition for the curve γ to extremize τ for all possible smooth families λ_α with $\lambda_0 = \gamma$ is that γ be a geodesic. We have

$$\begin{aligned}
\frac{d\tau}{d\alpha} &= \int_a^b \frac{\partial f}{\partial\alpha}dt \\
&= \int_a^b X^a\nabla_a(-T^bT_b)^{1/2}dt \\
&= -\int_a^b \frac{1}{f}T_bX^a\nabla_aT^bdt \\
&= -\int_a^b \frac{1}{f}T_bT^a\nabla_aX^bdt \\
&= -\int_a^b T^a\nabla_a\left[\frac{1}{f}T_bX^b\right]dt + \int_a^b X^bT^a\nabla_a(T_b/f)dt \\
&= \int_a^b X^bT^a\nabla_a(T_b/f)dt,
\end{aligned}$$

where in the second we used definitions of X^a and f, and for the last equation we note that the first term of the previous equation is 0 as $X^a = 0$ at the end points. Therefore $d\tau/d\alpha = 0$ for all X^a if and only if $T^a\nabla_a(T_b/f) = 0$ for $\alpha = 0$, which is precisely the geodesic equation for an arbitrary parameter. Thus for a curve to extremize the length (proper time), it must follow the geodesic equation and therefore must be a geodesic.
Although, it is not necessary that a given geodesic will be a curve of extremal length. This where the idea of conjugate points play a role.

Theorem 4.3.1. *Let γ be a smooth timelike curve connecting two points $p, q \in M$. Then the necessary and sufficient condition that γ locally maximize the proper time between p and q over smooth one parameter variations is that γ be a geodesic with no points conjugate to p between p and q.*

Sketch of proof. A detailed proof of this theorem is beyond the scope of this report, but I will try to provide an intuitive understanding of the theorem.
In case of reimannian geometry, curves of extremum length are shortest curves, which are geodesics as proved above. But the inverse is not always satisfied. To illustrate that consider two points p and q on the surface of the Earth (positive definite metric). Without loss of generality, consider p to be at the north pole. Thus the curve of minimal length connecting these points will be a longitude going through q, which is a geodesic. But there will be another longitude connecting p and q, which passes through the south

pole. Now the south pole can be considered as the conjugate point in our space. Both geodesics are stationary points of length under first order variation. But the second order variation of the geodesic going through a conjugate point can give a shorter curve. This can also be seen from the figure below.

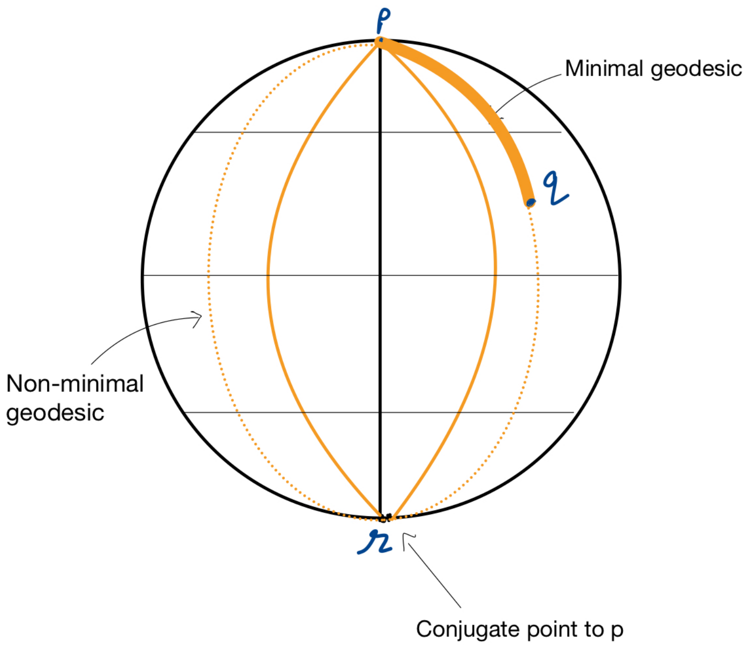

Figure 4.3: Non-minimal geodesic

But in case of lorentzian geometry, shortest length does not make sense as all curves could be deformed into a null curve, and hence have 0 length. Thus in case of lorentzian geometry, curves of extremal length refers to curves of maximum length. And analogous to the discussion above, geodesics with a conjugate point will fail to be a geodesic of maximal length, as the above theorem suggests. Using this theorem we can prove an essential result for the singularity theorem.

Theorem 4.3.2. *Let S be a smooth two-dimensional spacelike submanifold and let μ be a smooth causal curve from S to p. Then the necessary and sufficient condition that μ cannot be smoothly deformed to a timelike curve connecting S and p is that μ be a null geodesic orthogonal to S with no point conjugate to S between S and p.*

Proof To prove this theorem we need to prove following points -

- μ is a null geodesic
- μ is orthogonal to S
- μ does not have any points conjugate to S before reaching p

Now to have that μ cannot be deformed to a timelike curve implies that μ is curve of extremal length, and hence a geodesic. Moreover, μ should not be timelike to begin with,

hence it should be a null geodesic.
As S is a spacelike surface, all the null vectors on S are orthogonal to it. Thus μ will be orthogonal to S.
Finally, if μ had a conjuagate point between S and p, then by making analogous arguments as in the proof of theorem (4.3.1) we can find a geodesic (upon second order deviation of μ) which has a larger (non-zero) proper time, and hence be timelike. Thus μ must not have a conjugate point.

Chapter 5

Singularity theorem

In this chapter, we have built upon the results obtained thus far, and we are now in a position to prove Penrose's Singularity Theorem. This proof will be largely based on Wald (2010) [10], and Penrose(1965)[9].

5.1 Trapped surface

Consider a two dimensional spacelike submanifold S such that $\forall p \in S, \exists$ two future directed null vectors U_1^a *and* U_2^a orthogonal to S. The integral curves corresponding to these vectors can be considered as the incoming and outgoing geodesics respectively at S.

Definition 5.1.1. *A compact, 2-dimensional, smooth spacelike submanifold, T, such that expansion θ, of both sets of future directed null geodesics (i.e., incoming and outgoing) orthogonal to T is everywhere negative is called a 'trapped surface'.*

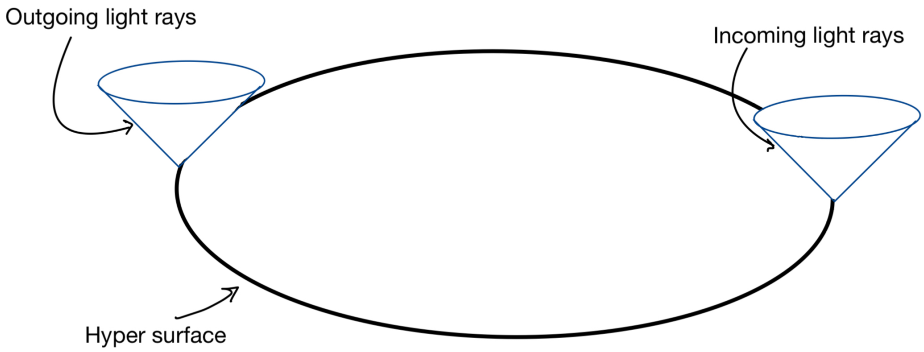

Figure 5.1: Outgoing and incoming null geodesics in Minkowski spacetime

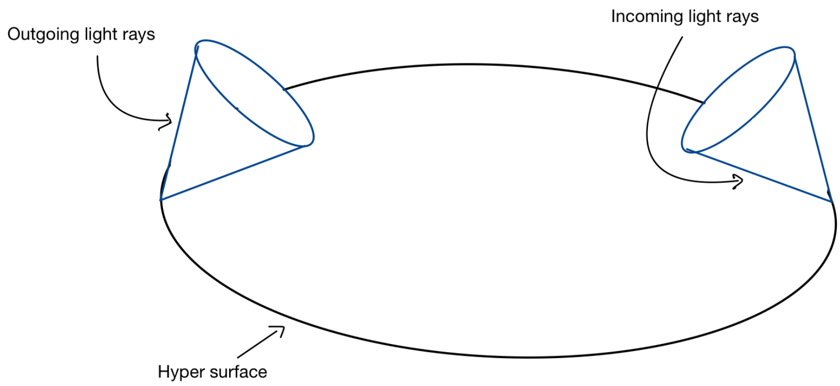

Figure 5.2: Outgoing and incoming light rays in a trapped surface

We are now able to prove the final theorem needed for singularity theorem.

Theorem 5.1.1. *Let (M, g_{ab}) be a globally hyperbolic spacetime and let K be a compact, orientable, two-dimensional spacelike submanifold of M. Then every $p \in \dot{I}^+(K)$ lies on a future directed null geodesic starting from K which is orthogonal to K and has no point conjugate to K between K and p.*

Proof. Using theorem (2.1.1), if $p \in \dot{I}^+(K)$ then p must lie on a null geodesic emanating from K. Moreover, $p \notin I^+(K)$, therefore there cannot be a timelike curve from K to p. Hence by using theorem (4.3.2) we can conclude that the null geodesic must be orthogonal to K and should not have a conjugate point before p. □

Now to further the discussion, consider a trapped surface T and consider a function $f^+ : T \times [0, \frac{2}{|\theta_0|}] \to \lambda(t)$, where $t \in [0, \frac{2}{|\theta_0|}]$ and λ is a future directed outgoing null geodesic orthogonal to T. Therefore $\lambda(q, t)$ will be a point "t" affine parameter away from S such that $\lambda(0) = q$. Similarly consider $f^- : T \times [0, \frac{2}{|\theta_0|}] \to \lambda(t)$, here λ will be a future directed incoming null geodesic.

Proposition 5.1.1. *Consider*

$$A = f^+\left\{T \times \left[0, \frac{2}{|\theta_0|}\right]\right\} \cup f^-\left\{T \times \left[0, \frac{2}{|\theta_0|}\right]\right\}$$

Then

$$\dot{I}^+(T) \subset A$$

Proof. Let $p \in \dot{I}^+(T)$ but $p \notin A$ and according to theorem (5.1.1), all points on the boundary of causal future of a compact surface lies on a future directed null geodesic starting from the surface and T is compact. Thus p can be connected to T by a null geodesic but as $p \notin A$, $p = \lambda(t) \quad | \quad t > \frac{2}{\theta_0}$. According to theorem (4.2.1), there will be a conjugate point between p and T. This contradicts (5.1.1), hence $p \in A \implies \dot{I}^+(T) \subset A$ □

5.2 Penrose singularity theorem

Theorem 5.2.1. *Let (M, g_{ab}) be a connected globally hyperbolic spacetime with a non-compact Cauchy surface Σ. Suppose $R_{ab}k^a k^b \geq 0$ for all null k^a, as will be the case if the spacetime is a solution to Einstein's equation. Suppose further that M contain a trapped surface T, Let $\theta_0 < 0$ denote the maximum value of θ for both sets of orthogonal geodesics on T. Then at least one inextendible future directed orthogonal null geodesics from T has affine length no greater than $\frac{2}{|\theta_0|}$.*

Proof. Suppose all future directed null geodesics from T have affine length $\geq \frac{2}{\theta_0}$. Then by proposition (5.1.1), we have, $\dot{I}^+(T) \subset A$. Furthermore, $T \times [0, 2/|\theta_0|$ is a compact set and f^+ *and* f^- are continuous functions, hence, A must be compact. $\dot{I}^+(T)$, being the boundary set has to be closed and it is a subset of a compact set A. Therefore $\dot{I}^+(T)$must be compact as well.

Now, we show that compactness of $\dot{I}^+(T)$ contradicts the existence of a nonompact Cauchy surface Σ. Using lemma (2.1.1), we choose a smooth timelike vector field t^a on M. According to the definition of Σ 2.2.1, each integral curve of t^a will intersect Σ precisely once, while each integral curve of t^a will intersect $\dot{I}^+(T)$ at most once, due to achronality (proposition (2.1.2)). Hence we may define a map $\Psi : \dot{I}^+(T) \to \Sigma$ by following the integral curve of t^a from $\dot{I}^+(T)$ to Σ. Let $S \subset \Sigma$ denote the image $\Psi[\dot{I}^+(T)]$, and let S be given the topology induced by Σ. Thus $\Psi : \dot{I}^+(T) \to S$ is an homeomorphism. As $\dot{I}^+(T)$ is compact, S is also compact. And as a subset of Σ, S must be closed.
Meanwhile, from proposition (2.1.2), we have that $\dot{I}^+(T)$ is a C^0 - manifold . Hence, each point of $\dot{I}^+(T)$ has a neighborhood homeomorphic to an open ball in $\mathbb{R}^3$. Since Ψ is a homeomorphism, this holds for S as well. As each point of S is homeomorphic to an open ball, S must be open as a subset of Σ. But from theorem (2.2.1), Σ is connected, therefore the only open and closed subset of Σ can be either $\emptyset$ or Σ itself. And

$\dot{I}^+(T) \neq \emptyset \implies S = \Sigma$. This proves that Σ is compact, which contradicts our assumption and hence all null geodesics starting from a trapped surface cannot extend to $\frac{2}{\theta_0}$. This proves the existence of a singularity. $\square$

Chapter 6

Conclusion

The Singularity Theorem established in the previous chapter does not assume any specific symmetry of spacetime, thereby providing a general proof for the existence of singularities. This demonstrates that singularities are properties of the structure of spacetime itself rather than being artefacts of a particular metric. However, the theorem is based on certain assumptions regarding the properties that spacetime must satisfy. These assumptions are justified by the argument that any physically realistic spacetime should adhere to them. In this chapter, we will quote another singularity theorem, in the context of cosmology, and summarize the foundational assumptions and explore additional singularity theorems that operate under weaker conditions. One should note, however, that these theorems give no information about the nature of singularities of whose existence they prove.

6.1 Cosmological singularity

Here we will state another singularity theorem that relies on the same basic assumptions as the one in section 5.2.1. This theorem can be interpreted as showing that if the universe is globally hyperbolic and at one instant of time is expanding everywhere at a rate bounded away from zero, then the universe must have a singularity a finite time ago, in its beginning state, which is the famous *Big Bang* singularity.

Theorem 6.1.1. *Let (M, g_{ab}) be a globally hyperbolic spacetime with $R_{ab}\xi^a\xi^b \geq 0$ for all timelike ξ^a. Suppose there exists a smooth spacelike Cauchy surface, Σ for which the trace of the extrinsic curvature (or the expansion of the past directed orthogonal geodesic congruence) satisfies $K \geq C < 0$ everywhere, where C is a constant. Then no past directed timelike curve from Σ can have a length greater than $3/|C|$. In particular, all past directed timelike geodesics are incomplete.*

Unfortunately, due to time restrictions, the proof of this theorem was be included in this dissertation. But interested readers can find a proof in Wald (2010)[10].

6.2 Further discussion

Apart from physically obvious assumptions such as connectedness of spacetime, which is always possible by taking a connected subset of an otherwise disconnected space, and the energy condition, we have the following -

1. Orientability of spacetime
2. Global hyperbolicity of (M, g_{ab})
3. Non compact Cauchy surface
4. Existence of a trapped surface
5. Negative expansion of a congruence

Non-compactness of the Cauchy surface is the only additional hypothesis which hasn't been discussed in the text before. As mentioned, Σ is understood as an instance of time, and a non-compact Σ implies that the Universe is not closed, which is a reasonable assumption. Our aim here is to understand if a spacetime would admit singularities if any one of these above mentioned assumptions were to be removed.

The first theorem we will mention is a theorem due to Hawking[3]. This theorem removes the assumption of a globally hyperbolic spacetime, however the price paid for this is the additional hypothesis that that Cauchy surface,Σ, be compact and we get a significantly weakened conclusion that only at lest one past directed timelike geodesic (rather than all past directed timelike curves) must be incomplete.

Theorem 6.2.1. *Let* (M, g_{ab}) *be a spacetime satisfying Einstein's equations and holds strong energy condition for matter. Suppose there exists a compact, edgeless, smooth spacelike hypersurface* S *such that for the past directed normal geodesic congruence form* S *we have* $K < 0$ *everywhere on* S*. Let* C *denote the maximum value of* K*, so* $K \geq C < 0$ *everywhere on* S*. then t least one inextendible past directed timelike geodesic from* S *has length no greater than* $3/|C|$*.*

Once more the proof of this theorem will be referred to relevant sources, Wald (2010), Hawking (1967).[3]

Similarly. theorem (5.2.1) contains the hypothesis that (M, g_{ab}) is globally hyperbolic. However, this hypothesis again can be eliminated with some additional assumptions. We shall not attempt to give a proof of this theorem here, referring readers to Hawking and Penrose (1970) [4], and Hawking and Ellis (1973) [2] for a proof.

Theorem 6.2.2. *Suppose a spacetime* (M, g_{ab}) *satisfies following four conditions.*

1. $R_{ab}v^a v^b \geq 0$ *for all timelike and null* v^a*, as will be the case if Einstein's equation is satisfied with the strong energy condition holding for matter.*
2. *The timelike and null generic conditions are satisfied,i.e* $R_{abcd}v^a v^d \neq 0$ *for at least one point for each timelike or null geodesic, where* v^a *is the tangent vector.*

3. No closed timelike curves exist.

4. At least one of the following properties holds : (a) (M, g_{ab}) possesses a compact achronal set without edge; i.e. it is a closed universe, (b) (M, g_{ab}) possesses a trapped surface, or (c) there exists a point $p \in M$ such that the expansion of the future (or past) directed null geodesics emanating from p becomes negative along each geodesic in this congruence.

Then (M, g_{ab}) must contain at least one incomplete timelike or null geodesic.

6.3 Concluding remarks

This concludes our discussion of singularity theorems. As discussed in the introduction, our path to prove existence of singularities was to prove existence of incomplete geodesics. One should remember that this argument works only because we considered inextendible spacetimes, which is always possible. We first developed the idea of a global causal structure that a physical spacetime should abide to, followed by how a geodesic congruence should behave, under reasonable assumptions. Thereafter, we introduced a crucial idea of conjugate points and how their existence affects maximal length properties of a geodesic. Finally, using the results from these sections we proved that if the spacetime follows some physically accepted assumptions (including existence of Cauchy surface) then the spacetime admits at least one incomplete geodesic, hence admits a singularity. Moreover, we also glanced at other singularity theorems by Hawking and Penrose which admits weaker assumptions but also provides weaker results.

Bibliography

[1] Stephen Hawking and Roger Penrose. *The nature of space and time*, volume 40. Princeton University Press, 2010.

[2] Stephen W Hawking and George FR Ellis. *The large scale structure of space-time*. Cambridge university press, 2023.

[3] Stephen William Hawking. The occurrence of singularities in cosmology. . causality and singularities. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 300(1461):187–201, 1967.

[4] Stephen William Hawking and Roger Penrose. The singularities of gravitational collapse and cosmology. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 314(1519):529–548, 1970.

[5] Noel J Hicks. *Notes on differential geometry*, volume 1. van Nostrand Princeton, 1965.

[6] M. D. Kruskal. Maximal extension of schwarzschild metric. *Phys. Rev.*, 119:1743–1745, Sep 1960.

[7] Barrett O'neill. *Semi-Riemannian geometry with applications to relativity*. Academic press, 1983.

[8] J Robert Oppenheimer and Hartland Snyder. On continued gravitational contraction. *Physical Review*, 56(5):455, 1939.

[9] Roger Penrose. Gravitational collapse and space-time singularities. *Physical Review Letters*, 14(3):57, 1965.

[10] Robert M Wald. *General relativity*. University of Chicago press, 2010.

Appendices

Appendix A

Geodesic deviation

Let M be a manifold with a connection ∇. Consider a 1 - parameter family of geodesics such that $\gamma : I \times I' \to M \quad I, I' \in R$
i) for a fixed s, $\gamma_s(t)$ is a geodesic with affine parameter t, ii) $(s,t) \mapsto \gamma_s(t)$ is smooth and one-to-one with a smooth inverse. This one parameter family of geodesics form a 2 dimensional surface $\Sigma \subset M$.

Now consider a tangent vector field T^a on the geodesics, and thus satisfies

$$T^a \nabla_a T^b = 0.$$

The vector field $X^a = (\partial/\partial s)^a$ represents the displacement to an infinitesimally nearby geodesic, and is called the *deviation vector*. It should be noted that any change in the affine parameter $t \to t' = t - c(s)$ will cause

$$\begin{aligned} \frac{\partial}{\partial s} &= \frac{\partial t'}{\partial s}\frac{\partial}{\partial t'} + \frac{\partial s'}{\partial s} + \frac{\partial}{\partial s'} \\ X^a &= -\frac{\partial c}{\partial s} T^a + X'^a \\ \implies X'^a &= X^a + \frac{\partial c}{\partial s} T^a. \end{aligned} \tag{A.1}$$

Hence X'^a is a deviation vector pointing to the same geodesic as X^a.

Since X^a and T^a are coordinate vector fields, they commute:

$$\begin{aligned} \mathcal{L}_X T &= 0 \\ T^b \nabla_b X^a &= X^b \nabla_b T^a. \end{aligned} \tag{A.2}$$

We shall now derive an equation that relates $a^a = \nabla_T \nabla_T X^a$ to the Reimann tensor. We have

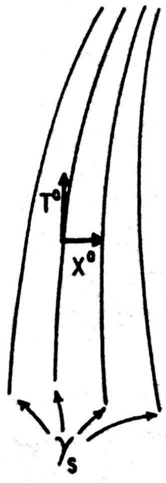

Figure A.1: A one-parameter family of geodesics γ_s, with tangent T^a and deviation vector X^a. [10]

$$\begin{aligned}
a^a &= \nabla_T \nabla_T X^a \\
&= \nabla_T \nabla_X T^a \\
&= \nabla_X \nabla_T T - R(T, S, T)^a \\
&= -R(T, S, T)^a,
\end{aligned}$$

where $R(T, S, T)^a = R_{cbd}{}^a X^b T^c T^d$. Thus we have,

$$T^c \nabla_c (T^b \nabla_b X^a) = -R_{cbd}{}^a X^b T^c T^d. \tag{A.3}$$

This is known as the *geodesic deviation equation.*

Appendix B

Frobenius's theorem

In this appendix, I wish to give Frobenius's theorem and how it relates to the discussion of geodesic congruences

B.1 Integral submanifolds

Consider a manifold M of n dimensions. At each point $x \in M$ have a subspace $W_x \subset T_pM$. We denote the collection of subspaces W_x for all $x \in M$ by W.

Definition B.1.1. *An integral submanifold of W is a submanifold $N \subset M$ such that at every point $p \in N$, the tangent space of N at p, T_pN, coincided with W_x.*

In this appendix we will deal with if it is possible to have an integral submanifold. An important case of it is when we a metric on M and want to know if a vector field ξ^a is orthogonal o a family of hypersurfaces,i.e., whether the (n-1) dimensional subspace, W, orthogonal to ξ^a are integrable.

Definition B.1.2 (Hypersurface)**.** *Hypersurface is an embedded submanifold of dimension (n-1).*

Theorem B.1.1. *Frobenius's theorem A necessary and sufficient condition for a smooth specification, W, of m-dimensional subspaces of the tangent space at each $x \in M$ to possess integral submanifolds is that W be involute, i.e., for all $Y^a, Z^a \in W$ we have $[Y, Z]^a \in W$.*

We won't get into the proof of this theorem. Moreover, it can also be expresses in dual formalism -
Given $W_x \in T_xM$ as above, we can consider one-forms $\omega \in T_xM^*$ which satisfy

$$\omega_a X^a = 0 \tag{B.1}$$

for all $X^a \in W_x$. It is clear that such ωs span an (n- m) dimensional subspace, $V_x^* \subset T_pM^*$. Collection of V_x^* is V^*.

Theorem B.1.2. *Let V^* be a smooth specification of an (n - m) dimensional subspace of one-forms. Then the associated m - dimensional subspace W of the tangent space admits integral submanifolds if and only if for all $\omega \in V^*$ we have $d\omega = \sum_\alpha \mu^\alpha \wedge \nu^a$, where each $\mu^* \in V^*$ and ν^a is any arbitrary one-form.*

This formalism gives a useful criterion for when a vector field ξ^a is hypersurface orthogonal. This is equivalent to -

$$\xi^a \text{ is hypersurface orthogonal } \iff \xi_{[a}\nabla_b\xi_{c]} = 0$$

$$\begin{aligned} \xi_{[a}\nabla_b\xi_{c]} &= \xi_a\omega_{bc} + \xi_b\omega_{ca} + \xi_c\omega_{ab} = 0 \\ \implies \xi^a\xi_a\omega_bc &+ \xi_b\omega_{ca}\xi^a + \xi_c\omega_{ab}\xi^a = 0 \\ &\implies \omega_{bc} = 0 \end{aligned}$$

where we used the definition of ω_{ab} from section (1.1.2) and equations (3.1) and (3.2). This calculation shows that for a geodesic congruence to be hypersurface orthogonal, ω_{ab} must vanish.