U S Naveen Balaji

Relativistically Speaking

An Introduction to General Relativity

To Leonard Susskind whose lectures and enthusiasm inspire me

Naveen Balaji U S

This book lays the mathematical foundations to General Relativity (GR) and presents the majestic grandeur and essence of the theory in a new and contemporary style. My fascination with the ideas of the theory date back to high school when I gained substantial knowledge from books and Youtube videos. As an undergraduate, I wanted to get rid of the pop-science noise and formally study the theory. As all beginners do, I too faced difficulty understanding the mathematics which would often lead to me giving up pursuing further into the physics. Although there are books which lay strong mathematical foundations, they often overwhelm a beginner. I eventually found out that it was one's imaginative thinking that enables them to understand deeply and appreciate theories such as GR and using this approach I mastered the basics so much so that I could conduct two summer schools teaching GR to highly-motivated undergraduates. It is these schools that encouraged me to write a book, a book in which I can communicate deep ideas in a contemporary style.

The summer schools usually consisted of about two complete weeks of the math required for formally studying and understanding the language of GR. A strong foundation in differential geometry was essential for understanding the concepts with ease although I did not give as much importance to differential topology as was required. I strongly believed that mastery of the index notation and the chain rule almost suffices because any serious student who aspires for a future in theoretical physics should be continually puzzled by the physics but not at all by the math. But beyond a point, math becomes a necessity rather than a requirement, it is necessary for someone interested in learning the theory with rigor for only when we truly have strong foundations will we be able to develop a holistic understanding of the subject. Almost an entire week was dedicated to the field equations of Einstein, we derived the equations and studied their properties from scratch. A similar approach was taken to the study of the Schwarzschild solution. The summer schools were long and an experience to remember, we spent five to six hours a day for five days a week and the entire course lasted for about six weeks. Now, that's approximately $150 - 180$ hours of content. The first chapter can be taught with two to three lectures per section and as for the rest of the book, each section would take up an entire lecture.

I have dedicated a quarter of the second chapter to laying foundations in Classical Mechanics (CM). This was done since most of the students who attended the summer schools knew CM but not to the precision required. Although it is a section that you can skip if you know the topic, I would advise you to go through it since it is constructed in a manner that is relevant to the future topics of the same chapter. The second chapter also consists of an entire section on Noether's theorem and its significance in GR, a section I personally enjoyed writing. I have consciously avoided discussions on topics such as the Kerr metric, the interior solution, the Oppenheimer-Synder collapse, and gravitational radiation since these are fairly advanced topics (from a physics point of view) for beginners. The book is written for beginners to obtain an advanced understanding of preliminary concepts with strong basics rather than obtain knowledge of advanced concepts with minimal basics. A quick note on the sign convention: I have used both $(-+++)$ and $(+---)$ metric signatures not to confuse the readers but to make them comfortable with the two approaches. An entire chapter has been dedicated to the study of embedding in N-dimensions, embedding diagrams and topics such as the Fujitani-Ikeda-Matsumoto embedding and the Schwarzschild-Tangherlini metric have been discussed in great mathematical detail. Chapter six is quite ambitious in both the concepts it covers and the type of approach taken, it introduces one to some of the advanced concepts of gravitational physics with the mathematical knowledge acquired in the previous five chapters.

On the whole, the aim of this book is to present with precision, but as intuitively as possible, the foundations and main consequences of GR and it is written for students of physics interested in exact mathematical formulations or for students of mathematics interested in intuitive understanding of physics—or indeed for anyone with a scientific mind irrespective of your educational field. The mathematical level of the six chapters of the book is that of undergraduates of mathematics or physics. The book assumes the reader to possess a fair knowledge of Set theory, Special Relativity and Electromagnetism and aims at communicating the concepts as intuitively as possible, constantly promoting the avant-garde and consciously avoiding the vicissitudes one faces in the conventionalist approach to physics and the labyrinthine choice of words in archaic texts.

I would like to thank Poojana Prasanna for her ideas, the truly original illustrations and for bearing my long conversations; my friends Sujan Kumar S and Vignesh T for their constructive criticism, for agreeing to edit the book and for pointing out potentially embarrassing mistakes; the students of the Entropi Gravitational Summer School of ′18- Tarun S R, Thejas Bhat, Anagha M, Pradyot Samartha and Mukund Balasubramanian-for accepting me as their teacher, for taking up the demanding course and providing honest reviews of my lectures which have shaped up this book; and Nishanth Shetty for his deep insights and honest reviews.

Naveen Balaji U S
January 2019

Contents

1

Essential mathematics

This chapter builds on the basic mathematical concepts required to understand Einstein's General Relativity. This chapter lays strong foundations of the study and physics on manifolds, tensors, and also contains a survey of the basic definitions of Riemannian and Lorentzian differential geometry that are necessary in General Relativity.

1.1 Manifolds: A Pedestrian Approach

Imagine that your car's tyres record of all the information regarding where you had been, have been and will be going, all the events that had happened on the road is recorded and stored in each thread of the tyre, and since you are unsure of the events of the future, making assumptions that the tyres never wear out (no matter what happens to it and lasts forever) and that they are super-elastic, we can safely assume that the tyres would possess infinite number of threads. One can access the information stored in each thread to check for the type of terrain the tyre has travelled upon, and the shapes of the localized deformations it had experienced. If there are infinite number of terrains, each unique (i.e. different from each other), upon which the tyres have travelled then there are an infinite number of deformation shapes (created on the tyres during travel). This is shown in figure 1.1

Hence, we can satisfactorily say that all the deformation shapes, although unique, are after all nothing but mere closed geometric figures. Now, if we randomly pick two tyres and name them A and B, we can find the set of deformations they have experienced and since the terrains are unique at each point and an infinite number of events have occurred, chances are that the sets of deformations of tyre A around an arbitrary point, p, are completely different from those of tyre B, around an arbitrary point, q. The deeper picture here is that although the set of deformations are different, we can perform few operations (strictly mathematical) in order to make the deformations look similar. Hence, let us propose the following method- For every deformation in tyre A

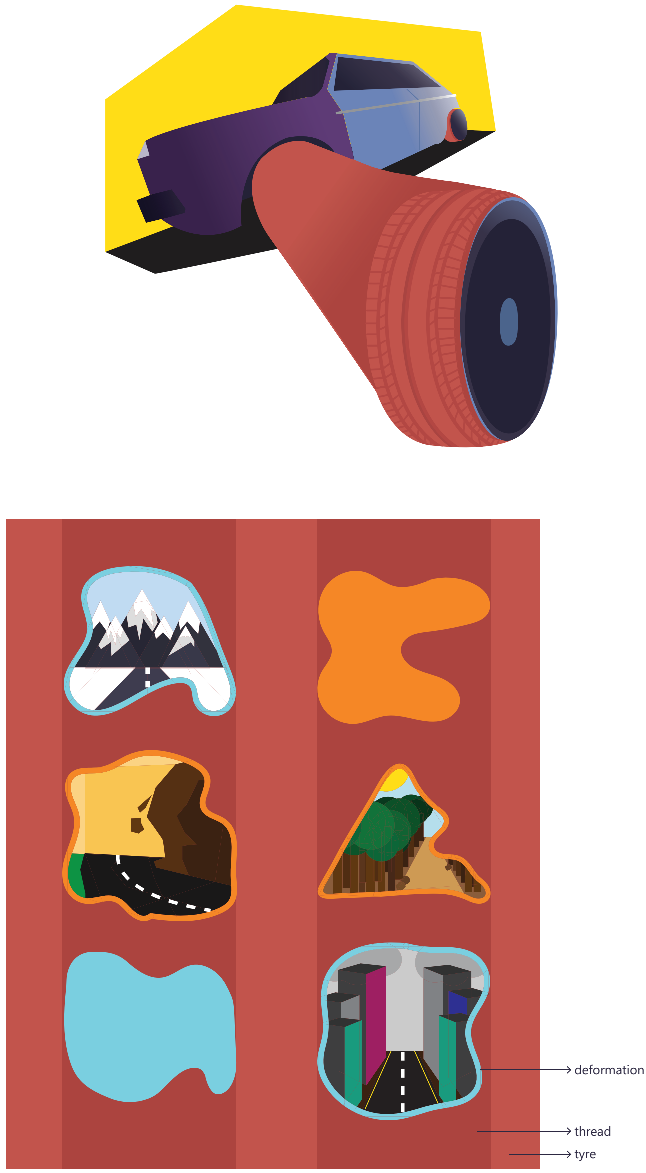


Fig. 1.1. A visual representation of the analogy presented in this chapter. The bottom figure shows the different deformation shapes created on the tyre surface that are stored in the threads along with the information of the type of terrain in which the deformation occurred.

there is another in B which has a very close resemblance, at the same time, for every deformation in A there are multiple other deformations in B since all of them after all fall under the category of closed figures. The converse of the above statements is also true (i.e. from tyre B to A). Let's recap before concluding, we picked two tyres at random and named them A and B. We observed that not only are the sets deformations around arbitrary points on the tyres not the same, but also that the deformations among the tyres had resemblances. At the end of the day, both the tyres had deformations which were nothing but closed shapes.

Alright then, let's conclude! A manifold is similar to our tyres, at localized points it represents Euclidean space just like how there were closed geometric deformations on localized points of the curved tyres, but ultimately it is nothing but a topological space because all deformations, although unique are nothing but closed shapes! More specifically a manifold is one in which at each point, has a neighbourhood which is *homeomorphic* to the Euclidean space. The deformations of the tyres could be made to look similar using the method I proposed, remember? When we say that for every deformation in tyre A there is another in B which has a very close resemblance, what it mathematically means is that every deformation in A has a *one-to-one* connection to another deformation in B. When we say that for every deformation in A there are multiple other, topologically similar deformations in B, it mathematically translates to stating that every deformation in A is connected *onto* multiple others in B. Since our deformations have both one-to-one and onto connections, mathematically we call them to be *bijective*. Do not forget that this bijective display of behaviour is limited only to the deformations in A, the converse is also true. In mathematics, there is a name assigned to the method we proposed- *homeomorphism* . If you still crave for a more *formal* definition, here it goes: Suppose $f : A \to B$ is a bijective (one-to-one and onto) function between topological spaces A and B. Since f is bijective, the inverse f^{-1} exists. If both f^{-1} and f are continuous functions, then f is called a homeomorphism. Since the topological spaces A and B are homeomorphic, we denote them by the following: $A \cong B$.

Since we are comfortable with the meaning of a manifold, let us go a bit deeper. We observed that the set of all deformations around an arbitrary point p on tyre A was not similar to that around point q on tyre B, in other words, the sets of deformations around localized points p and q are *non-interfering*. Also, in collection of the deformations of all tyres, there are many which might belong to tyre B. Mathematically, this is called an *open set*. In topology, the car is defined to be the entire set $\mathcal{C}$ and the tyres A, B, C,and D are its subsets, $\mathcal{C}$ is an open set if it is in a subset. Hence, we can distinguish the sets of deformations around the points p and q by two non-interfering open sets, this property is called *Hausdorff*. Moreover, each tyre comes with a family of deformations, each which are unique and universally belong to the entire set $\mathcal{C}$ and are home-

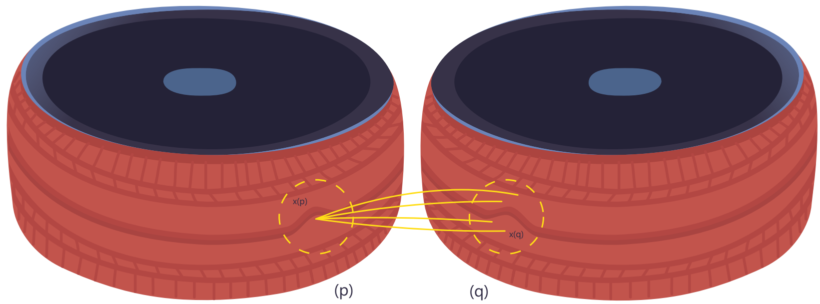


Fig. 1.2. Consider the immediate neighbourhoods of the deformation points p and q in tyres A and B. The deformation at point p in A has a *one-to-one* connection to another deformation point q in B and it has an *onto* connection to multiple others in the neighbourhood of q, i.e., in $x(q)$. The continuous and invertible function f that maps points in A to B, i.e., $f : A \to B$, is called a homeomorphism.

omorphic to other deformations in different tyres and hence to the whole set of deformations. Thus, these deformation families constitute the entire set $\mathcal{C}$, i.e., $\mathcal{C} = \cup_\alpha \nu_\alpha$, where ν_α is a subset (may be A, B, C or D depending upon α) and the homeomorphisms are depicted using maps, $\phi_\alpha : \nu_\alpha \to \mathbb{R}^n$, where $\mathbb{R}^n$ is the set of all Euclidean closed geometries in a Euclidean space of dimension n. If the above conditions are satisfied, i.e. if a topological space is Hausdorff, and comes with a family $\{(\nu_\alpha, \Phi_\alpha)\}$ with set ν_α being a subset of an open set $\mathcal{C}$ and homeomorphisms $\Phi_\alpha : \nu_\alpha \to \mathbb{R}^n$ such that $\mathcal{C} = \cup_\alpha \nu_\alpha$, we call such a topological space $\mathcal{C}$ to be a n-dimensional smooth manifold. The pairs $(\nu_\alpha, \Phi_\alpha)$ are called *charts*, the family $\{(\nu_\alpha, \Phi_\alpha)\}$ is called an *atlas*, and Φ_α is called a coordinate function.

In topology *smooth* corresponds to *differentiable*, and if our manifold has the metric signature, $(- \;+\; +\, +)$, we refer to it as a *Lorentzian manifold* $(\mathcal{C}, g)$, where $\mathcal{C}$ is our topological manifold and g is the metric tensor. A Lorentzian manifold is a type of *pseudo-Riemannian manifold* $(\mathcal{C}, g)$ which is a differentiable manifold $\mathcal{C}$ equipped with a non-degenerate, smooth, and symmetric metric tensor g. Spacetime is mathematically defined as a four-dimensional, smooth, connected Lorentzian manifold $(\mathcal{C}, g)$. Hence, if our car is a smooth, 4-D Lorentzian manifold then each localized deformation's frame of reference on this manifold is represented using coordinate charts. We adopt a more formal and mathematically rigorous, but intuitive approach to differential topology in the next section.

1.2 Manifolds: The One With The Rigorous Approach

Gravitational physics done in spacetime focuses on topological spaces $(\mathcal{M}, \Psi)$ that can be charted similar to how the surface of the Earth can be charted on an atlas. Thus, a topological space $(\mathcal{M}, \Psi)$ is known as a n-dimensional topological manifold if for all points d which belongs to the manifold M there not only exists an open set V in the topology Ψ, which contains the point d, but also exists an entire map z that takes every point in the set V to a subset in $\mathbb{R}^n$ in an invertible, one-to-one manner which in both directions is continuous. Representing this mathematically we write

$$\forall d \in \mathcal{M} : \exists V \in \psi : \exists z : V \to z(V) \subseteq \mathbb{R}^n, \tag{1.1}$$

where $z(V)$ is the image of the domain under the chart z. Note that this mapping of the point is done in such a way that:
1. z is invertible, i.e., there exists a map z^{-1} such that: $z^{-1} : z(V) \to V$,
2. z is continuous, and
3. z^{-1} is continuous.

1.2.1 Chart

Going back to the car analogy, we observe that the tyre of the car is actually a torus (the shape of a doughnut). We can claim that this tyre surface is the set $\mathcal{M}$ which is equipped with some topology Ψ in R^3, i.e., $M \subseteq \mathbb{R}^3$ (M is a subset of $\mathbb{R}^3$ equipped with a topology Ψ). Now, since the tyre is elastic, it deforms at points when it encounters rocks on the terrain. For a deformation point p, there exists an open subset V (which is its immediate neighbourhood and whose boundary is decided based on to what extent the neighbourhood is affected due to the rock encounter), and there exists a map z which maps the deformation point p and every point in its neighbourhood (i.e., in its open subset V), to some part of $\mathbb{R}^2$. This mapping done is bijective and generates an open region in the mapped part of $\mathbb{R}^2$, which is nothing but a set of real numbers. Mathematically, the previous statement is expressed as follows

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) | x, y \in R\}. \tag{1.2}$$

Say the deformation p occurred due to a rock R_1, then we can say that the position of the deformation made by the rock R_1 in $M \subseteq \mathbb{R}^3$ is at $z(p) \subseteq \mathbb{R}^2$, which contains two components: $z(p) = (z^1(p), z^2(p))$. What this means is that the position of the deformation of the rock R_1 on the three-dimensional surface of the tyre is mapped to a two-dimensional region and identified using two coordinates $z^1(p)$ and $z^2(p)$. The pair (V, z) is called a chart (see figure 1.3), where V is the boundary of the deformation effect of the rock in R^3, and $z(V)$ is its image existing in $\mathbb{R}^2$. The whole tyre is covered by charts that contain deformation points. Thus, for every point on the manifold, there exists a chart that contains the point, and the topological space $(\mathcal{M}, \Psi)$ is called a two-dimensional topological manifold.

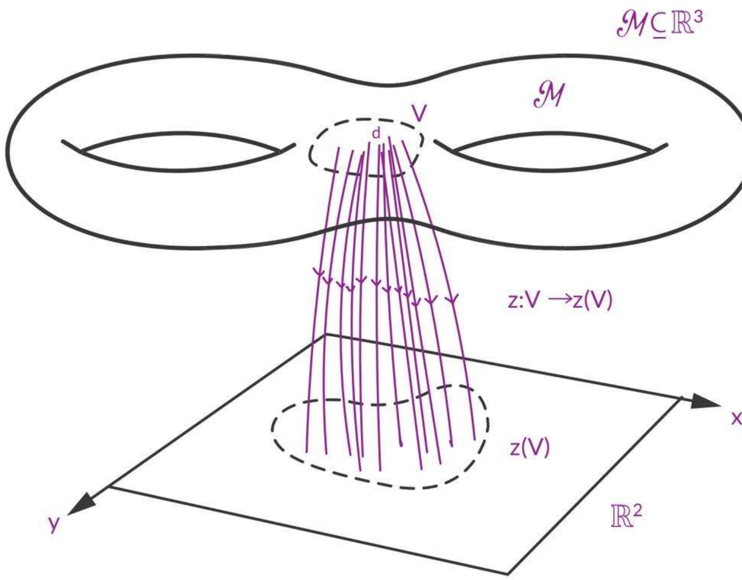


Fig. 1.3. Chart (V, z). All points in the open set V present in the manifold $\mathcal{M}$ equipped with a topology Ψ (which is similar to that of a double doughnut) are mapped to $z(V)$ by map z.

1.2.2 Atlas

(V, z) is a chart of $(\mathcal{M}, \Psi.)$ This manifold that contains a collection charts that can be classified under a set A which mathematically is

$$A = \{(V_\rho, z_\rho)|\rho \in I\}\,, \tag{1.3}$$

where ρ is a label which belongs to some arbitrary index set I, is called an atlas of $(\mathcal{M}, \Psi)$. In other words, the atlas comprises of a family of charts. The existence of the atlas A is subject to the condition that the union of all the charts domains must reproduce the original surface $\mathcal{M}$, i.e.,

$$\mathcal{M} = \cup_{\alpha\in I} V_\alpha . \tag{1.4}$$

What this means is that if we take all the images $z(V_\alpha)$ of all the open sets V_α existing in $\mathbb{R}^2$ and stitch them together, we must be able to reproduce the surface of the tyre again. The topological space $\mathcal{M}$ is said to be *paracompact* if for every atlas $A = \{(V_\rho, z_\rho)|\rho \in I\}$ there exists a locally finite atlas $B = \{(V_\sigma, z_\sigma)|\sigma \in I\}$ with each open set U_σ contained in some V_ρ.

1.2.3 Chart and Coordinate Maps

In general, for a n-dimensional topological manifold the function

$$z : V \to z(V) \subseteq \mathbb{R}^n, \tag{1.5}$$

is called a *chart map*. It is important to note that $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times ...$ is the set of n-tuples, thus, the image of a deformation point in $\mathbb{R}^n$ is represented using the coordinates $z(p) = (z^1(p), z^2(p), ..., z^n(p))$, where z^j is a map which takes a point in V and maps it to R (a real number), i.e., $z^j : V \to R$ for $j = 1, 2, ..., n$. A mathematical picture of this is as follows

$$z : V \to \mathbb{R}^n = \left\{ \begin{matrix} z^1 : V \to \mathbb{R} \\ \vdots \\ z^n : V \to \mathbb{R} \end{matrix} \right\} = z : V \to \mathbb{R}^n = \sum_{j=1}^{n} z^j, \tag{1.6}$$

where the individual charts z^j are called coordinate maps. What this means is that the deformation point $p(\in V)$ has its first coordinate at $z^1(p)$ present in the region $z(V)$ of the chart (V, z), its second coordinate at $z^2(p)$ present in the region $z(V)$ of the chart (V, z), and so on.

1.2.4 Chart Transition Maps

Consider two charts (V, z) and (U, w), with overlapping regions on the surface $\mathcal{M}$, equipped with a topology Ψ. Let the tyre encounter an arbitrary but small distribution of rocks (of the same shape and mass), the deformation points are all alike and exist within a region on the tyre. Let's call this as the *deformation region*. Consider the points to the deformation region's immediate right and immediate left. When the tyre encounters the rock distribution, the regions to the left and right are affected to some extent (whose boundary is set based on the magnitude of deformation). Let the open set of the left region containing the set of affected points be V and let the open set of the right region containing the set of affected points be U. Thus, we can conclude that the open sets contain a non-empty overlap (which is the deformation region itself), i.e., $V \cap U \neq \emptyset$. We know that V comes with a chart map z that takes any point in V and maps it to some region in $\mathbb{R}^n$, i.e., $z : V \to z(V) \subseteq \mathbb{R}^n$, and U contains with a chart map w that takes any point in U and maps it to some other region in $\mathbb{R}^n$, i.e., $w : U \to w(U) \subseteq \mathbb{R}^n$. Now a deformation point d present in the deformation region (which is the intersection of V and U) can be mapped to two regions of $\mathbb{R}^n$ using the chart maps z and w. Similarly, we can map all the points in the deformation region $(V \cap U)$ into $z(V)$ and $w(U)$ via chart maps z and w to obtain regions $z(V \cap U)$ and $w(V \cap U)$ in $\mathbb{R}^n$. What this implies is that a point d in the deformation region $(V \cap U)$ is mapped to two points in the regions $z(V \cap U)$ and $w(V \cap U)$ in $\mathbb{R}^n$. A natural question that arises is- how are these two points related? Let the mapped point of d in $z(V \cap U)$ be d', since we know that z is invertible we use this property to define d'', the mapped point of d in $w(V \cap U)$ as follows

$$d'' = w(z^{-1}(d')) = w \circ z^{-1}(d'). \tag{1.7}$$

Thus, $w \circ z^{-1}$ acts as a chart map which maps points of $z(V \cap U)$ existing in $\mathbb{R}^n$ to $w(V \cap U)$ existing in $\mathbb{R}^n$. Formally, this is known as a chart transition map. Chart transition maps contain the information of how to stitch together all the charts of an atlas. See figure 1.4 for a visualization of the mathematical concept.

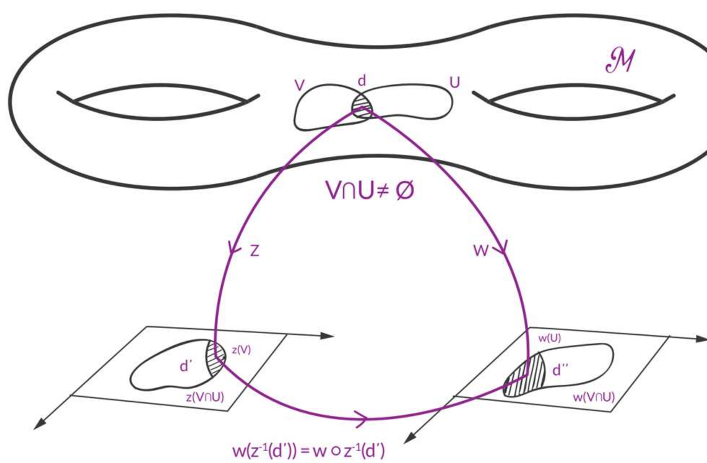


Fig. 1.4. Chart transition maps

Consider the tyre again of surface $\mathcal{M}$ equipped with a topology Ψ, now we specify that it is the Euclidean space of n-dimensions, $\mathbb{R}^n$. $\mathbb{R}^n$ is nothing but the set of all n-tuples $(x^1, x^2, ..., x^n)$, with $-\infty < x^j < \infty$. Let $\frac{\mathbb{R}^n}{2}$ be the *lower half* of $\mathbb{R}^n$, i.e., the lower half of the tyre for which $x^1 \leq 0$. Let p be the point existing on the lower half of the tyre and let V be the open set in which it is contained. The map z of the open set $V \subset \mathbb{R}^n$ (in $\frac{\mathbb{R}^n}{2}$) to the open set $V' \subset \mathbb{R}^m$ (in $\frac{\mathbb{R}^m}{2}$) is said to be of class C^k if the coordinates of the image point $z(p) = (z^1(p), z^2(p), ..., z^m(p)) = (x'^1, x'^2, ..., x'^m)$ in V' are k-times continuously differentiable functions (which refers to the existence of the k^{th} derivative which is continuous) of the coordinates $(x^1, x^2, ..., x^n)$ of the point p in the open set V on the tyre. A map is called C^∞ if it is C^k for all $k \geq 0$

and C^0 if it is a continuous.

A C^k n-dimensional manifold $\mathcal{M}$ is a set $\mathcal{M}$ together with a C^k atlas $A = \{(V_\rho, z_\rho)|\rho \in I\}$ where V_ρ are subsets of $\mathcal{M}$ and $z_\rho : V_\rho \to z_\rho(V_\rho) \subseteq \mathbb{R}^n$ are the one-to-one maps such that all the open sets V_ρ cover $\mathcal{M}$. If there exists a non-empty overlap between two open sets V_ρ (with map z_ρ) and U_σ (with map w_σ), i.e., $V_\rho \cap U_\sigma \neq \emptyset$, then the map $z_\rho \circ w_\sigma^{-1} : w_\sigma(V_\sigma \cap U_\sigma) \to z_\rho(V_\rho \cap U_\sigma)$ is a C^k map of an open subset of $\mathbb{R}^n$ to an open subset of $\mathbb{R}^n$.

1.2.5 Homeomorphism and Diffeomorphism

Let the images of V and U existing in $\mathbb{R}^n$ be v and u, i.e., let $z(V) = v$, and $w(U) = u$, and as stated previously, these are open sets since the chart maps are invertible. A mapping between the open sets of $\mathbb{R}^n$, $f : v \to u$, is called a homeomorphism if it is bijective and if f and its inverse f^{-1} are continuous. A differential homeomorphism is called a *diffeomorphism*. Analogous to how a homeomorphism is a bijection that is continuous and also possesses a continuous inverse, a diffeomorphism is a bijection which is differentiable with a differentiable inverse, i.e., if v and u are connected open subsets of $\mathbb{R}^n$ such that u is simply connected, a differentiable map $f : v \to u$ is a diffeomorphism if the differential $Df_p : \mathbb{R}^n \to \mathbb{R}^n$ is bijective at each point p in v. Another way to put this is to state that a mapping f between open sets of $\mathbb{R}^n : v \to u$ is called a diffeomorphism if it is bijective and if f and its inverse mapping f^{-1} are differentiable. Hence, every diffeomorphism is a homeomorphism, but not vice-versa. Generally, a bijective mapping is a C^k diffeomorphism if f and f^{-1} are of class C^k (see Table 1.1). Thus, the map z from $\mathcal{M}$ to $\mathcal{M}'$ is said to be a C^k diffeomorphism if it is a one-one C^k map and the inverse z^{-1} is a C^k map from $\mathcal{M}'$ to $\mathcal{M}$. We observed that the set of all deformations around an arbitrary point p on tyre A was not similar to that around point q on tyre B, in other words, the sets of deformations around localized points p and q are *non-interfering*. Also, in collection of the deformations of all tyres, there are many which might belong to tyre B. Consider a curve η present on the manifold $\mathcal{M}$. This curve η can be called k-times continuously differentiable if there exists a C^k atlas.

Once you are comfortable with the C^k classes, then, another definition to diffeomorphisms can be adopted. Two manifolds $\mathcal{M}$, $\mathcal{N}$ are said to be diffeomorphic if there exists a homeomorphism $f : \mathcal{M} \to \mathcal{N}$ such that f is a C^∞ function with a C^∞ inverse. f is called a diffeomorphism.

Mathematically, we know that this is called an open set. In topology, the car is defined to be the entire set $\mathcal{C}$ and the tyres A, B, C, and D are its subsets, $\mathcal{C}$ is an open set if it is in a subset. Hence, we can distinguish the sets of deformations around the points p and q by two non-interfering open sets, this property is called *Hausdorff*. Thus, in short, a topological space $\mathcal{M}$ is said to be a *Hausdorff space* if for two points a and b in $\mathcal{M}$, there exists disjoint open

Atlas	**Properties**
C^0	$C^0\ (\mathbb{R}^n \to \mathbb{R}^n)$ are continuous maps
C^1	$C^1\ (\mathbb{R}^n \to \mathbb{R}^n)$ are the maps that are once differentiable and continuous
C^k	$C^k\ (\mathbb{R}^n \to \mathbb{R}^n)$ are the maps that are k-times continuously differentiable
D^k	$D^k\ (\mathbb{R}^n \to \mathbb{R}^n)$ are the maps that are k-times differentiable but are not continuous
C^∞	$C^\infty\ (\mathbb{R}^n \to \mathbb{R}^n)$ are the maps that are many-times continuously differentiable
$C^{'\infty}$	$C^{'\infty}\ (\mathbb{R}^n \to \mathbb{R}^n)$ are the maps that are many-times continuously complex differentiable
C^ω	C^ω are the maps that can be Taylor expanded

Table 1.1. This table depicts the properties of class k atlases. $C^{'\infty}$ this is valid only for even dimensional manifolds under the condition that the chart maps satisfy the *Cauchy-Riemann* equations. C^ω stands for analytic; a function $f : \mathbb{R}^n \to R$ is analytic at $p \in \mathbb{R}^n$ if f can be expressed as a power series in the $(x^j - p^j)$ which converges in some neighbourhood of p.

sets V and U in M such that $a \in V$ and $b \in V$. This condition is sometimes called the *Hausdorff separation axiom* (see figure 1.5).

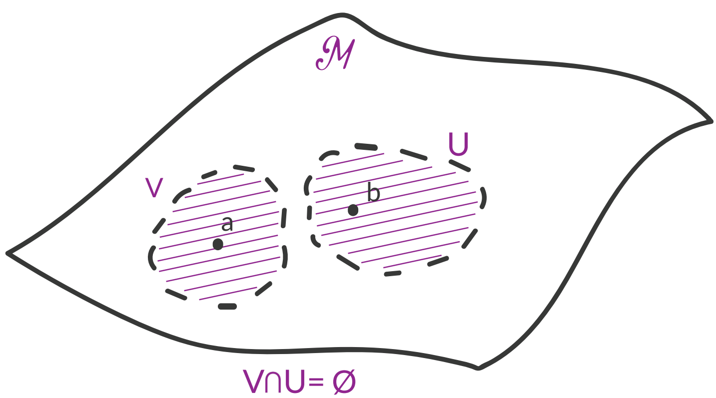


Fig. 1.5. Hausdorff separation axiom

Moreover, each tyre comes with a family of deformations, each which are unique and universally belong to the entire set $\mathcal{M}$ and are homeomorphic to other deformations in different tyres and hence to the whole set of deformations. Thus, these deformation families constitute the entire set $\mathcal{M}$, i.e., $\mathcal{M} = \cup_{\rho \in I} V_\rho$, where V_ρ is a subset (may be A, B, C or D) and the homeomorphisms are depicted using maps, $z_\rho : V_\rho \to zV_\rho \subseteq \mathbb{R}^n$, where $\mathbb{R}^n$ is the set of all Euclidean closed geometries in a Euclidean space of dimension n. If the above conditions are satisfied, i.e. if a topological space is Hausdorff, and comes with a family $\{(V_\rho, z_\rho) | \rho \in I\}$ with set V_ρ being a subset of an open set $\mathcal{M}$ and homeomorphisms $z_\rho : V_\rho \to z(V_\rho) \subseteq \mathbb{R}^n$ such that $\mathcal{M} = \cup_{\rho \in I} V_\rho$, we call such a topological space $\mathcal{M}$ to be a n-dimensional smooth manifold. All manifolds considered are assumed to paracompact, connected C^∞ Hausdorff manifolds without boundary.

1.2.6 Differential Manifold

Knowing the concept of a diffeomorphism, we can now reframe the concept of a topological manifold. An atlas bequeaths $\mathcal{M}$ with the structure of a topological manifold, of dimension n, if the mappings $w_\sigma \circ z_\rho^{-1}$ are homeomorphisms (i.e., continuous bijections) between open sets of $\mathbb{R}^n$, namely between $z_\rho(V_\rho \cap U_\sigma)$ and w_σ $(V_\rho \cap U_\sigma)$. If these mappings are diffeomorphisms, then the manifold can be called a differential manifold. Generally, the manifold is a *differential manifold* of class C^k if these mapping are C^k diffeomorphisms. Thus, the term smooth means that a class C^k with k large enough (in particular $k = \infty$). A differential manifold (or a smooth manifold) is often written as a C^∞ manifold. Thus, we can define a C^∞ manifold as to be a pair $(\mathcal{M}, A)$, where A is a maximal atlas for $\mathcal{M}$.

Consider two C^∞ manifolds, $(\mathcal{R}, V)$ and $(\mathcal{R}, U)$. These manifolds are called *isomorphic* if there exists an bijective (one-to-one and onto) function $f : \mathcal{R} \to \mathcal{R}$ such that $p \in V$ if and only if $p \circ f \in U$. Two C^∞ manifolds $(\mathcal{M}, A)$ and $(\mathcal{M}^{'}, B)$ are called *diffeomorphic* if there is a bijective function $f : \mathcal{M} \to \mathcal{M}^{'}$ such that $p \in B$ if and only if $p \circ f \in A$. Few books refer to the atlas for $\mathcal{M}$ as the *differential structure* for $\mathcal{M}$.

1.2.7 Tangent Space and Tangent Bundle

We take the tyre again and connect all the points across the threads which appear to possesses information regarding the deformations caused to the tyre. We connect the points (or events) with a continuous and smooth curve P. Upon examination, we observe that events on the curve occur at regular intervals (assumption made for simplicity) and using this fact we parameterize the curve in terms of χ. Hence the parameterized curve takes the form of $\frac{dP}{d\chi}$. Consider the very first event on the tyre Q and the very last event R. If the points were present within the same thread, then the curve $P(\chi)$ would actually be the straight line described by the equation- $P(\chi) = X + \chi(R - Q)$. The derivative of $P(\chi)$ can be written as follows

$$\frac{d}{d\chi}\left(X + \chi(R - Q)\right) = \mathbf{R} - \mathbf{Q} = \mathbf{v}_{QR}\mathbf{v}_{QR}\mathbf{v}_{QR}\mathbf{v}_{QR} = \left(\frac{dP}{d\chi}\right)_{\chi} = 0(1.8)$$

This is defined to be a *tangent vector*. More formally, a tangent vector v to the differential manifold $\mathcal{M}$ at a point $p \in \mathcal{M}$ is defined as $(V_\rho, z_\rho, v_{z_\rho})$, where (V_ρ, z_ρ) are charts which contain p and $v_{z_\rho} = v^j_{z_\rho}, j = 1, 2, ..., n$ are vectors in $\mathbb{R}^n$. What is more interesting is the resting place of this vector. This tangent vector does not lie on the manifold, i.e., it does not share the same home as that of the cure $P(\chi)$, rather it lies in a so called *tangent space* (see figure **??**) which touches or makes contact with the manifold only at $P(\chi = a)$, the point where $\frac{dP}{d\chi}$ was evaluated. Imagine that we take different colours of moulding clay and press them against the tyre, starting at specific points (events). Thus, all the tangent vectors of every event will be contained in specific bits of clay. Now, mould all these pieces of clay onto a single, larger, and continuous piece of clay. This is the tangent space which is a plane in which all the tangent vectors to all events are contained such that the plane is tangent to the tyre at every point.

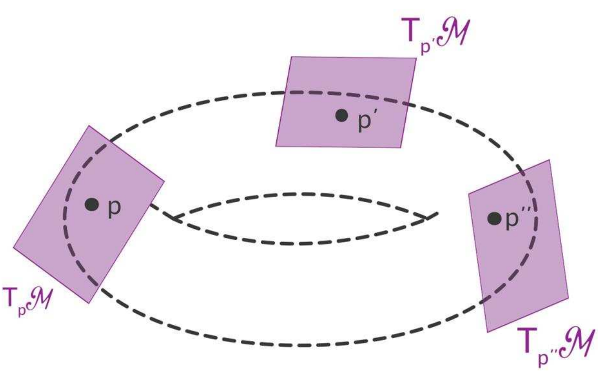


Fig. 1.6. The points p, p' and p'' lie in the tangent spaces $T_p\mathcal{M}$, $T_{p'}\mathcal{M}$ and $T_{p''}\mathcal{M}$ respectively, where $\mathcal{M}$ is a manifold equipped with a topology Ψ (that of a tyre).

All the tangent vectors at the point p constitute a *tangent space* (which is a vector space) to $\mathcal{M}^n$ at the point p. This tangent space is denoted by $T_p\mathcal{M}^n$ or simply a $T_p(\mathcal{M}^n)$. A tangent bundle is defined as the set of the pairs of the points and the tangent vector of that point, i.e., (p, v_p), where $p \in \mathcal{M}^n$

(point present in the n-dimensional manifold) and $v_p \in T_p\mathcal{M}^n$ (tangent vector contained in the tangent space of the n-dimensional manifold), denoted by $T\mathcal{M}^n$.

1.2.8 Immersions and Embeddings

An *immersion* is defined to be the function between differential manifolds whose derivative is everywhere injective (one-to-one), i.e., the function $h : \mathcal{M} \to \mathcal{M}^{'}$ is called an immersion between $\mathcal{M}$ and $\mathcal{M}^{'}$ (differential manifolds) if $D_p h : T_p\mathcal{M} \to T_{f(p)}\mathcal{N}$ is an injective function at every point p of M. In other words, an immersion simply means that the tangent spaces are mapped injectively, i.e., the map described above is injective.
Consider the map z from a C^k n-dimensional manifold $\mathcal{M}$ to a C^l o-dimensional manifold N. This map is called a C^r map ($r \leq k, r \leq l$) if, for the coordinates of the image point $z(p)$ in $\mathcal{N}$ are C^r functions of the coordinates of p in $\mathcal{M}$. A C^r map z ($r \geq 0$) is called an immersion if it and its inverse are C^r maps, i.e., if for each point $p \in M$ there exists an open set V such that the inverse map z^{-1} restricted to the image of the domain $z(V)$.

An immersion is called an *embedding* if it is a homeomorphism onto its image in the topology Ψ of a differentiable manifold $\mathcal{M}$. It is important to note that all embeddings are one-to-one immersions, however, the converse is not true.

1.2.9 Pseudo-Riemannian metrics

A metric g on a manifold M which is a symmetric covariant 2-tensor field is called a *pseudo-Riemannian metric* if the determinant $|g|$ with elements $g_{\alpha\beta}$, whose quadratic form it defines on contravariant vectors, $g(A, B)$, given in local charts by $g_{\alpha\beta}A^\alpha B^\beta$, does not vanish in any chart, i.e., it is non-degenerate. This definition is independent of the choice of charts because under a change of local coordinates $(x'^m) \to (x^m)$ it holds that

$$|g| = |g'|\frac{dx'}{dx}. \tag{1.9}$$

$(\mathcal{M}, g)$ is a diffeomorphism f which leaves g invariant, i.e., $f^*g = g$. Two pseudo-Riemannian manifolds $(\mathcal{M}, g)$ and $(\mathcal{M}', g')$ are called *locally isometric* if there exists a differential mapping f such that any point $p \in \mathcal{M}$ admits a neighbourhood $\mathcal{M}$, and $f(p)$ a neighbourhood $\mathcal{M}'$ with $(\mathcal{M}, g)$ and $(\mathcal{M}', g')$ isometric. It is important to note that pseudo-Riemannian manifolds can have different topologies although they possess the same dimension. Flat space is defined as a pseudo-Riemannian manifold is isometric with a pseudo-Euclidean space.

1.2.10 Lorentzian Manifold

In topology *smooth* corresponds to *differentiable*, and if our manifold has the metric signature, $(- \ + \ + +)$, we refer to it as a *Lorentzian manifold* $(\mathcal{M}, g)$, where $\mathcal{M}$ is our topological manifold and g is the metric tensor. A Lorentzian manifold is a type of pseudo-Riemannian manifold $(\mathcal{M}, g)$ which is a differentiable manifold $\mathcal{M}$ equipped with a non-degenerate, smooth, and symmetric metric tensor g.

Spacetime is mathematically defined as a four-dimensional, smooth, connected Lorentzian manifold $(\mathcal{M}, g)$. Here $\mathcal{M}$ is a connected four-dimensional Hausdroff C^∞ manifold and g is a Lorentz metric on $\mathcal{M}$. Hence, if our car is a smooth, 4-D Lorentzian manifold then each localized deformation's frame of reference on this manifold is represented using coordinate charts. Similarly, in the spacetime manifold, the coordinate charts are used to represent observers in reference frames. For a physicist the most preferred and useful definition is by identifying (locally) the manifold by $\mathbb{R}^n$.

1.2.11 Whitney and Nash Embedding Theorems

In short, the idea of the theorem is that any $C^{k\geq 1}$ atlas A of a topological manifold contains a C^∞ manifold, i.e., an atlas in which the chart transitional maps are at least once continuously differentiable we can *remove* more and more charts until we are left with a C^∞ atlas. What this implies is that we may always consider C^∞ manifolds (called *smooth manifolds*) from now on.
Let's formulate this mathematically. Let $\mathcal{M}$ be a smooth topological manifold of dimension n. The theorem roughly states that if $\mathcal{M}^n$ is a compact C^∞ manifold, then there is an embedding $z : \mathcal{M} \to \mathbb{R}^n$ for some N. The strongest version of the theorem is given below (without proof).

Theorem 1.1. *Any smooth manifold of dimension n can be immersed into $\mathbb{R}^{2n-1}$ and embedded into $\mathbb{R}^{2n}$*

Nash embedding theorem states that every Riemannian manifold can be isometrically embedded into some Euclidean space. Why this is interesting is because it mentions *isometric embedding*, i.e. preserving the length of curves in the manifold, whereas the Whitney theorem does not. According to this theorem 4-dimensional curved spacetime can be isometrically embedded in a flat spacetime of 39 dimensions or less[1].

[1] For the mathematically inclined see: Nash, J. (1954). C1 isometric imbeddings. Annals of mathematics, 383-396, and Nash, J. (1956). The imbedding problem for Riemannian manifolds. Annals of mathematics, 20-63.

1.2.12 Einsteinian Spacetime

A metric is called *Riemannian* if it's quadratic form, g, is positive definite. A pseudo-Riemannian metric g can be called a *Lorentzian metric* if the sign of the g is $(- \ + \ ... +)$. A spacetime of General Relativity is a pair $(\mathcal{M}, g)$, where $\mathcal{M}$ is a differentiable manifold and g is a Lorentzian metric on $\mathcal{M}$. Such a spacetime is called *Einsteinian* if there exists a physically meaningful stress-energy tensor (of rank two) T such that the following equations are satisfied on the differentiable manifold $\mathcal{M}$: $Einstein(g) = T$, and $\nabla T = 0$. [2].

1.3 Differential Forms and Tensors

1.3.1 1-Forms

A 1-form $\boldsymbol{\alpha}$ is a linear, real valued function of vectors. Consider a point p present in the manifold $\mathcal{M}$ equipped with a topology Ψ. If $\mathbf{R}$ is a vector at p, the number into which the 1-form maps $\mathbf{R}$ is expressed as $\langle \boldsymbol{\alpha}, \mathbf{R} \rangle$. This is nothing but the value of $\boldsymbol{\alpha}$ on $\mathbf{R}$ or simply, the contraction of $\boldsymbol{\alpha}$ on $\mathbf{R}$. For the sake of simplicity, let's call this the *carrot operator* and refer to it's operation as *carroting*[3]. In other words, we can define a 1-form $\boldsymbol{\alpha}$ at a point p in $(\mathcal{M}, \Psi)$ as a linear, real valued function on the tangent space T_p of the vectors at p. The condition of linearity allows us to arrive at two conclusions, the first one is to realize the mathematical property of linearity, and the other is to form a visual understanding via surfaces.

Linearity implies the following

$$\langle \boldsymbol{\alpha}, a\mathbf{R} + b\mathbf{S} \rangle = a \langle \boldsymbol{\alpha}, \mathbf{R} \rangle + b \langle \boldsymbol{\alpha}, \mathbf{S} \rangle \tag{1.10}$$

The tangent space T_p, for a given 1-form $\boldsymbol{\alpha}$, defined by $\langle \boldsymbol{\alpha}, \mathbf{R} \rangle = const$, is linear. We can imagine a 1-form as a set of planes in the tangent space. Imagine that each of these planes had a high-sensitive *vector-alarm* and as a vector pierces through a plane, the alarm would go off (which will be recorded). Also let's make the assumption that each alarm makes a unique sound (thus, enabling us to distinguish the panes). When $\langle \boldsymbol{\alpha}, \mathbf{R} \rangle = 0$, the vector (or more precisely, it's tip) touches the first plane and we hear the first alarm. Similarly when $\langle \boldsymbol{\alpha}, \mathbf{R} \rangle = n$, the vector pierces through the $(n+1)^{th}$ plane and in total we hear $(n+1)$ distinct alarm sounds (see figure 1.7).

---

[2] From Einstein's field equations we know that the first equation corresponds to $G^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu}$, and the second equation talks of the conservation of energy which is just a consequence of Einstein's field equations

[3] The credit for naming this operator goes to my students of the Gravitational Summer School $'18$

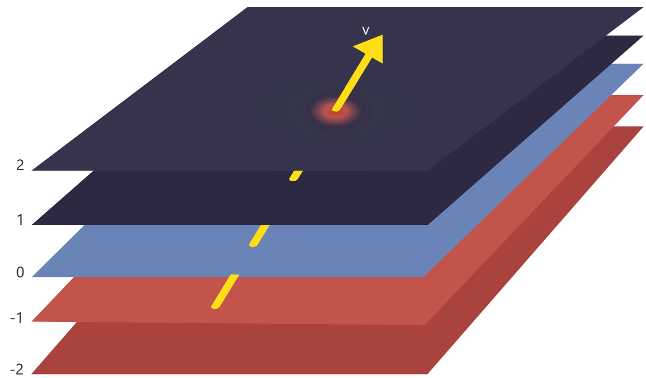


Fig. 1.7. Each of these planes have a high-sensitive *vector-alarm* and as a vector pierces through a plane, the alarm goes off. In this figure, the vector $\mathbf{v}$ pierces through the 1-form $\boldsymbol{\sigma}$ and $\langle \boldsymbol{\sigma}, \mathbf{v} \rangle = 4$, i.e., we hear 4 distinct alarms go off.

The simplest example of a 1-form is the gradient $\mathbf{dg}$ of a function g. Consider a vector u, and a curve $P(\eta)$ ($P(\eta) = \eta u + P_0$ which is parametrized in terms of η), and differentiate the function g along this curve.

$$\partial_{\mathbf{u}} g = \left(\frac{d}{d\eta} \right)_{\eta=0} g\left[P(\eta)\right] = \left(\frac{\mathbf{dg}}{d\eta} \right)_{P_0} \tag{1.11}$$

Observe that the operators $\partial_{\mathbf{u}} = \left(\frac{d}{d\eta} \right)_{\eta=0,P(\eta)}$, are related, i.e., the directional derivative and the gradient are related. Let the surfaces present in the tangent space T_p defined for the point p in $(\mathcal{M}, \Psi)$ be numbered with respect to g, (i.e., $g = 1$: surface one; $g = 2$: surface two; ...). Let the initial position of the vector (which starts from some arbitrary surface of g) be P_0. The first point of contact of the vector with a surface would be given as $\langle \mathbf{dg}, P - P_0 \rangle$, where $\mathbf{dg}$ is the stack of infinitesimal surfaces present between two g-surfaces. Thus, a generalized expression can be obtained: $g(P) = g(P_0) + \langle \mathbf{dg}, P - P_0 \rangle$. Since the relation between the directional derivative and the gradient is well established, let's apply $\partial_{\mathbf{u}}$ to $g(P)$ and evaluate the result at the take-off point P_0.

$$\partial_{\mathbf{u}} g = \left\langle \mathbf{dg}, \frac{dP}{d\eta} \right\rangle = \langle \mathbf{dg}, \mathbf{u} \rangle \tag{1.12}$$

In general $g(P)$ will have non-linear contributions of the order $\mathcal{O}(P - P_0)$. Similar to how vectors possesses a basis $\mathbf{e}_\beta$, 1-forms possess basis too. These basis 1-forms are denoted by $\boldsymbol{\omega}^\alpha$.

Lemma 1.2. *If the basis* 1*-forms set defined by* $\boldsymbol{\omega}^\alpha$ *and the basis vectors defined by* $\mathbf{e}_\beta$ *are duals of each other, then* $\langle \boldsymbol{\omega}^\alpha, \mathbf{e}_\beta \rangle = \delta^\alpha_\beta$

This lemma enables us to expand any arbitrary vector and 1-form in terms of their basis as follows: $e_\alpha, u = u^\alpha \mathbf{e}_{\boldsymbol{\alpha}}$ and $\rho = \rho_\beta \boldsymbol{\omega}^{\boldsymbol{\beta}}$. As an example we shall calculate the surfaces of $\boldsymbol{\rho}$ pierced by a basis vector $\mathbf{e}_\alpha$.

The piercing example

Example 1.3. $\langle \boldsymbol{\rho}, \mathbf{e}_{\boldsymbol{\alpha}} \rangle = \left\langle \boldsymbol{\rho}_{\boldsymbol{\beta}} \boldsymbol{\omega}^{\boldsymbol{\beta}}, \mathbf{e}_{\boldsymbol{\alpha}} \right\rangle = \rho_\beta \left\langle \boldsymbol{\omega}^{\boldsymbol{\beta}}, \mathbf{e}_{\boldsymbol{\alpha}} \right\rangle = \rho_\beta \delta^\beta_\alpha = \rho_\alpha$

Similarly, let's calculate the carroting $\langle \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \mathbf{u} \rangle$ for a vector $\mathbf{u} = \mathbf{e}_{\boldsymbol{\beta}} u^\beta$.

The carroting example

Example 1.4. $\langle \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \mathbf{u} \rangle = \left\langle \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \mathbf{e}_{\boldsymbol{\beta}} u^\beta \right\rangle = u^\beta \left\langle \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \mathbf{e}_{\boldsymbol{\beta}} \right\rangle = u^\alpha$

Well, what's the bigger picture here? Go on and multiply the piercing example with u^α, the carroting example with ρ_α and add both the equations to obtain the following result.

$$\begin{aligned}
&[\langle \boldsymbol{\rho}, \mathbf{e}_{\boldsymbol{\alpha}} \rangle u^\alpha + \langle \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \mathbf{u} \rangle \rho_\alpha] \\
&[\langle \boldsymbol{\rho}, \mathbf{e}_{\boldsymbol{\alpha}} u^\alpha \rangle + \langle \rho_\alpha \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \mathbf{u} \rangle] = \rho_\alpha u^\alpha + \rho_\alpha u^\alpha \\
&[\langle \boldsymbol{\rho}, \mathbf{u} \rangle + \langle \boldsymbol{\rho}, \mathbf{u} \rangle] = 2\rho_\alpha u^\alpha \\
&\langle \boldsymbol{\rho}, \mathbf{u} \rangle = \rho_\alpha u^\alpha
\end{aligned} \tag{1.13}$$

Thus, we have obtained a way of using components to calculate the coordinate-independent value of $\langle \boldsymbol{\rho}, \mathbf{u} \rangle$. Let's discuss a bit more on what the *dual* is (from the lemma). Since we can express the the 1-form $\boldsymbol{\alpha}$ at a point p in terms of it's basis $\boldsymbol{\alpha} = \boldsymbol{\omega}^{\mathbf{j}} \boldsymbol{\alpha}_{\mathbf{j}}$, the set of all 1-forms at p forms an n-dimensional vector space at p. This vector space is called the *dual space* of the tangent space T_p and is written as ${}^*T_p$. Let's revise the lemma a bit and re-state it as follows

Lemma 1.5. *For any* 1*-form* $\boldsymbol{\alpha} \in {}^*T_p$ *and any vector* $\mathbf{R} \in T_p$*, we can express the carroting* $\langle \boldsymbol{\alpha}, \mathbf{R} \rangle$ *in terms of the corresponding dual basis* $\boldsymbol{\omega}^{\mathbf{j}}$ *and* $\mathbf{e}_{\mathbf{j}}$ *by relations:*

$\langle \boldsymbol{\alpha}, \mathbf{R} \rangle = \left\langle \boldsymbol{\alpha}_{\mathbf{j}} \boldsymbol{\omega}^{\mathbf{j}}, \mathbf{R}^{\mathbf{j}} \mathbf{e}_{\mathbf{j}} \right\rangle = \alpha_j R^j$

Each function g on the manifold M defines a 1-form at a point p. This follows a rule which states that for each vector R, $\langle \mathbf{dg}, \mathbf{R} \rangle = \mathbf{Rg}$. Here, dg is called the differential of g.

1.3.2 Two Roads to Tensors: Road for Pedestrians

A tensor is like a slot machine (see figure 1.8), not just any ordinary one but rather a very modified machine. The generalized tensor slot machine has two slots (instead of one in the real machine) and two sub-slots. These slots are specific in what they accept. There are n first sub-slots and it accepts only 1-forms while there are m second sub-slots which accepts only vectors. Thus, we can mathematically represent a tensor slot machine S as follows

$$S(\underbrace{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}}_{n\ 1-forms}, \underbrace{\mathbf{u}, \mathbf{v}, ..., \mathbf{a}}_{m\ vectors}) \tag{1.14}$$

This S tensor is said to be of rank $\begin{pmatrix} \mathrm{n} \\ \mathrm{m} \end{pmatrix}$. It is important to note that most of the tensors do not remain the same if two slots of either 1-forms or vectors or both are interchanged, i.e., $S(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{v}, \mathbf{u}) \neq S(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{u}, \mathbf{v})$. Consider the following example which demonstrates how to work with tensors.

Example 1.6. Let F be a tensor of rank $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Define the tensor by inserting the basis vectors of the 1-forms and the vectors as follows:

$$\mathbf{F}^{\boldsymbol{\alpha\beta\gamma}}_{\boldsymbol{\delta\eta}} \equiv F\left(\boldsymbol{\omega}^{\boldsymbol{\alpha}}, \boldsymbol{\omega}^{\boldsymbol{\beta}}, \boldsymbol{\omega}^{\gamma}, \mathbf{e}_{\boldsymbol{\delta}}, \mathbf{e}_{\boldsymbol{\eta}}\right) \tag{1.15}$$

Now, the output can be calculated for the given input as follows

$$\begin{aligned} &F\left(\boldsymbol{\sigma}, \boldsymbol{\rho}, \boldsymbol{\nu}, \mathbf{u}, \mathbf{v}\right) = F\left(\sigma_\alpha \boldsymbol{\omega}^{\boldsymbol{\alpha}}, \rho_\beta \boldsymbol{\omega}^{\gamma}, \nu_\gamma \boldsymbol{\omega}^{\gamma}, u^\delta \mathbf{e}_{\boldsymbol{\delta}}, v^\eta \mathbf{e}_{\boldsymbol{\eta}}\right) \\ &\sigma_\alpha \rho_\beta \nu_\gamma u^\delta v^\eta F\left(\boldsymbol{\omega}^{\boldsymbol{\alpha}}, \boldsymbol{\omega}^{\boldsymbol{\beta}}, \boldsymbol{\omega}^{\gamma}, \mathbf{e}_{\boldsymbol{\delta}}, \mathbf{e}_{\boldsymbol{\eta}}\right) = \mathbf{F}^{\boldsymbol{\alpha\beta\gamma}}_{\boldsymbol{\delta\eta}} \sigma_\alpha \rho_\beta \nu_\gamma u^\delta v^\eta \end{aligned} \tag{1.16}$$

1.3.3 Two Roads to Tensors: Road for the Mathematically Inclined

We can form a *Cartesian product* from the tangent space T_p of vectors at point p and from the tangent space's dual ${}^*T_p$ of 1-forms at p as follows (The dual space ${}^*T_p$ is often called the *cotangent space.*)

$$\Pi^n_m = \underbrace{{}^*T_p}\, \times {}^*T_p \times ... \times {}^*T_{p_{\text{n factors}}} \times \underbrace{T_p \times T_p \times ... \times T_p}_{m\ factors} \tag{1.17}$$

this is the ordered set of 1-forms and vectors $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a})$. A tensor of rank $\begin{pmatrix} \mathrm{n} \\ \mathrm{m} \end{pmatrix}$ at a point p is a function on Π^n_m which is linear in each argument, i.e., if $\mathbf{T}$ is a tensor of rank $\begin{pmatrix} \mathrm{n} \\ \mathrm{m} \end{pmatrix}$ at p, the number into which $\mathbf{T}$ maps

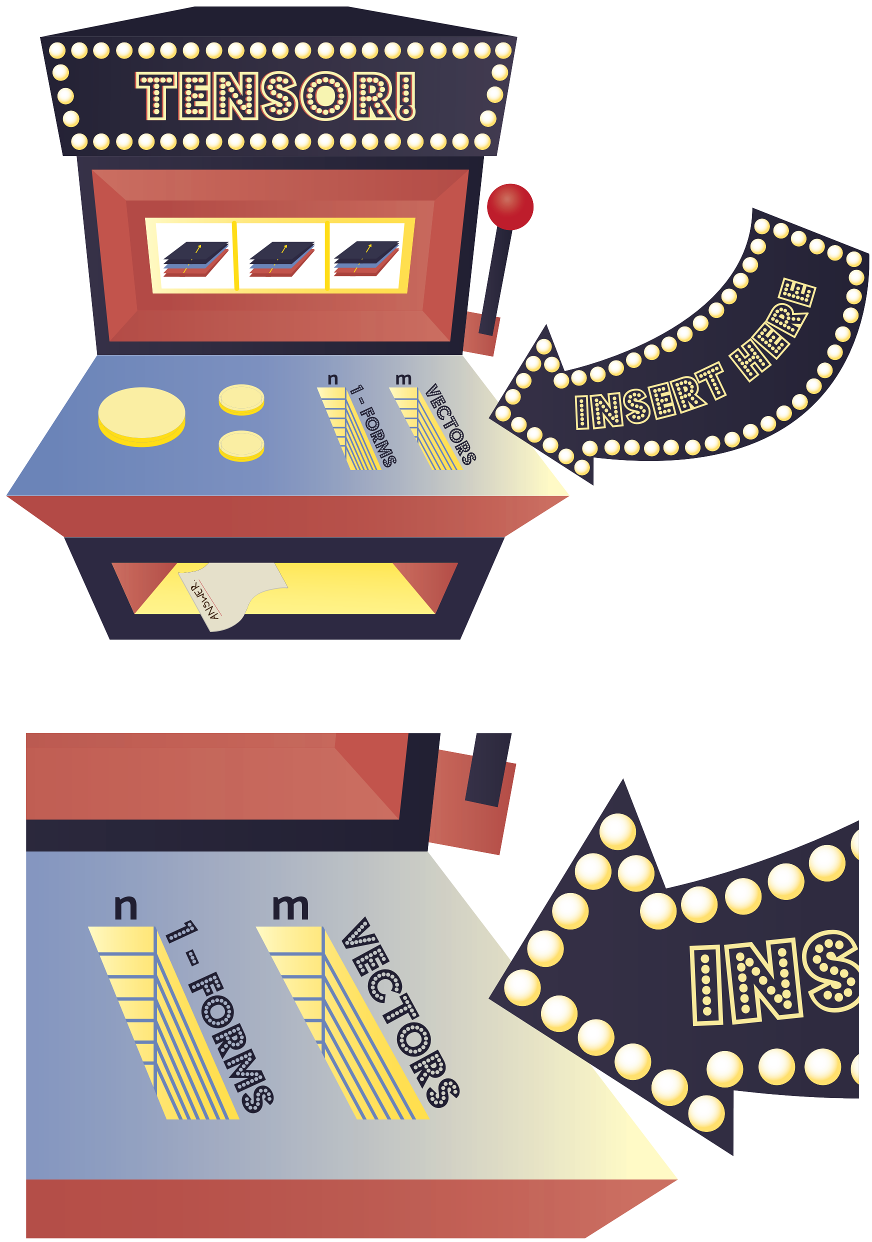


Fig. 1.8. The slot machine representation of tensors. The machine has n sub-slots accepting 1-forms and m sub-slots accepting vectors.

the element $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a})$ of Π_m^n as $T(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a})$. The property of linerity also applies to tensors (this is described below).

$$\begin{aligned} &T(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, a\mathbf{R} + b\mathbf{S}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) \\ &= aT(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{R}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) + bT(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{S}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) \end{aligned} \tag{1.18}$$

The space of all such tensors is called the *tensor product.*

$$\mathbf{T}^{\mathbf{n}}_{\mathbf{m}}(p) = \underbrace{{}^*T_p}\otimes{}^*T_p \otimes ... \otimes {}^*T_{p_{\text{n factors}}} \otimes \underbrace{T_p \otimes T_p \otimes ... \otimes T_p}_{m\ factors} \tag{1.19}$$

1.3.4 Tensor Operations: Addition

Let $\mathbf{S}$ and $\bar{\mathbf{S}}$ be tensors of rank $\begin{pmatrix} \mathrm{n} \\ \mathrm{m} \end{pmatrix}$, the addition of these tensors is defined by the following rule

$$\begin{aligned} &\left(S + \bar{S}\right)(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) \\ &= S(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) + \bar{S}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) \end{aligned} \tag{1.20}$$

1.3.5 Tensor Operations: Multiplication

The multiplication of the same tensor $\mathbf{S}$ considered in the previous example with a scalar ζ is shown below

$$\begin{aligned} &(\xi S)(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) \\ &= \xi \times S(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, ..., \boldsymbol{\zeta}, \mathbf{u}, \mathbf{v}, ..., \mathbf{a}) \end{aligned} \tag{1.21}$$

A *covariant* k-tensor at a point $p \in M$ is defined as a k-multilinear form on k direct products of the tangent space T_pM. Similarly, a *contravariant* k-tensor at a point $p \in M$ is defined as a k-multilinear form on k direct products of the cotangent space ${}^*T_pM$. The tensor product $\mathbf{S} \otimes \bar{\mathbf{S}}$ of a r-tensor $\mathbf{S}$ and a s-tensor $\bar{\mathbf{S}}$ is a $(r+s)$-tensor with components defined by products of components. Consider the product of a covariant 2-tensor $\mathbf{T}$ and a contravariant 3-tensor $\bar{T}$. The result is a mixed 4-tensor $\mathbf{T}\otimes\bar{\mathbf{T}}$ with the following components

$$\left(\mathbf{T} \otimes \bar{\mathbf{T}}\right)_{\alpha\beta}^{\ \ \gamma\delta} = T_{\alpha\beta}\bar{T}^{\gamma\delta} = W^{\gamma\delta}_{\alpha\beta} \tag{1.22}$$

When we refer $\mathbf{T}$ as a r-covariant tensor and $\bar{\mathbf{T}}$ as a s-contravariant tensor it implies that they are elements of the tensor product of r copies of T_p and s copies of $^*T_p$. The covariant and contravariant tensors can also be defined in terms of local coordinates as follows

$$W^{\gamma\delta}_{\alpha\beta} = \underbrace{\left(\frac{\partial y^{\zeta}}{\partial x^{\alpha}}\frac{\partial y^{\eta}}{\partial x^{\beta}}\right)}_{covariant\ part}\underbrace{\left(\frac{\partial x^{\gamma}}{\partial y^{\rho}}\frac{\partial x^{\delta}}{\partial y^{\xi}}\right)}_{contravariant\ part} W^{\rho\xi}_{\zeta\eta} \tag{1.23}$$

Let's generalize this and write down the local coordinate transformation of a n-tensor.

$$S^{\alpha_1\ \dots\ \alpha_m}_{\beta_1\ \dots\ \beta_n}(x) = \frac{\partial y^{\mu_1}}{\partial x^{\beta_1}} \dots \frac{\partial y^{\mu_n}}{\partial x^{\beta_n}} \frac{\partial x^{\alpha_1}}{\partial y^{\nu_1}} \dots \frac{\partial x^{\alpha_m}}{\partial y^{\nu_m}} S^{\nu_1\ \dots\ \nu_m}_{\mu_1\ \dots\ \mu_n}(y) \tag{1.24}$$

1.3.6 Tensor Operations: Contraction

Recollect the slot-machine definition of a tensor. Contraction is similar to shutting off of sub-slots (which contain both 1-forms and vectors) within the two main slots. Consider the mixed 4-tensor $\mathbf{R} = R(\boldsymbol{\sigma}, \mathbf{u}, \mathbf{v}, \mathbf{w})$ which is of rank $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$. We can shut off sub-slots 1 and 3 to reduce the tensor to a covariant 2-tensor, say $\mathbf{S}$. This operation is described below.

$$\begin{aligned} &S(u,w) = \textstyle\sum_{\alpha=0}^{3} R(\boldsymbol{\omega}^{\boldsymbol{\alpha}}, u, \mathbf{e}_{\boldsymbol{\alpha}}, w) \\ &S(u,v) = S_{\mu\nu}u^{\mu}w^{\nu} = R^{\alpha}_{\mu\alpha\nu}u^{\mu}w^{\nu} \\ &S_{\mu\nu} = R^{\alpha}_{\mu\alpha\nu} \end{aligned} \tag{1.25}$$

1.3.7 Tensor Operations:Symmetrization and Antisymmetrization

If the output of a tensor is unaffected by an interchange of 2 input vectors or 1-forms, then it is called a *symmetric tensor*, if not then it is called an *antisymmetric tensor.*

$$\begin{aligned} &\text{Symmetric} \\ &T(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = T(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) = T(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha}) = \dots \\ &\text{Antisymmetric} \\ &T(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = -T(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) = +T(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\alpha}) = -+\dots \end{aligned} \tag{1.26}$$

Let **F** be a 2-covariant tensor. The symmetrization and antisymmetrization of **F** can be represented as follows

$$\begin{aligned} &\text{Symmetrization:} \quad F_{(\alpha\beta)} = \tfrac{1}{2}\left(F_{\alpha\beta} + F_{\beta\alpha}\right) = S_{\alpha\beta} \\ &\text{Antisymmetrization:} \quad F_{[\alpha\beta]} = \tfrac{1}{2}\left(F_{\alpha\beta} - F_{\beta\alpha}\right) = A_{\alpha\beta} \end{aligned} \tag{1.27}$$

1.3.8 Tensor Operations: Wedge Product

Given any two vectors, we can construct their *bivector* by *wedging* them. This can also be done with 1-forms to obtain 2-*forms*. This concept can also be used to construct a *trivector* and 3-*forms*.

$$\begin{aligned} &\text{Bivector} \\ &\mathbf{a} \wedge \mathbf{b} \equiv \mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a} \\ &\text{2-form} \\ &\boldsymbol{\alpha} \wedge \boldsymbol{\beta} \equiv \boldsymbol{\alpha} \otimes \boldsymbol{\beta} - \boldsymbol{\beta} \otimes \boldsymbol{\alpha} \end{aligned} \tag{1.28}$$

This operator is interesting geometrically as it serves as a test of coplanarity[4]. Observe that the action of the wedge can be generalized- a wedge among p-forms produces a $(p+1)$-form. This raises a question-what maps these p-form fields to $(p+1)$-form fields, it this mapping linear?

1.3.9 Exterior differentiation

[5] The job of linearly mapping p-form fields to $(p+1)$-form fields is done by the exterior derivative d. If $z : \mathcal{M} \to \mathcal{N}$ is a C^r map and $\mathbf{\Lambda}$ is a C^k form field on $\mathcal{N}$, then $d(z^*\mathbf{\Lambda}) = z^* d(\mathbf{\Lambda})$. If $\varSigma$ is a function on N, the function $z^*\varSigma$ on $\mathcal{M}$ is defined by the mapping z as he function whose value at a point p on the manifold $\mathcal{M}$ is the value of $\varSigma$ at the image of the point $z(p)$, i.e., $z^*\varSigma(p) = \varSigma(z(p))$. What this implies is that z^* maps functions linearly from $\mathcal{N}$ to $\mathcal{M}$, similar to how z maps points from $\mathcal{M}$ to $\mathcal{N}$. Now, if $\zeta(t)$ is a curve existing in $\mathcal{M}$ and passing via the point p , then the image of this curve $z(\zeta(t))$ existing in N

[4] Consider 3 arbitrary vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$. If $\mathbf{a}$ and $\mathbf{b}$ are collinear, then $\mathbf{a} = \mathbf{b}\lambda$. This implies that $\mathbf{a} \wedge \mathbf{b} = \mathbf{b}\lambda \wedge \mathbf{b} = \mathbf{b}\lambda \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{b}\lambda = 0$. Now, if $\mathbf{c}$ is coplanar with $\mathbf{a}$ and $\mathbf{b}$, then $\mathbf{c}$ can be expressed as a linear combination (i.e., scalar multiplication followed by vector addition) of the other two vectors, i.e., $\mathbf{w} = \mathbf{b}\lambda + \mathbf{a}\epsilon$. What this implies is that $\mathbf{c} \wedge \mathbf{a} \wedge \mathbf{b} = 0$.

[5] This section is an optional read. Do skip this if this is your first read.

passes via the image point $z(p)$. Consider a tangent vector to the image curve, denoted by $z_* \left(\frac{\partial}{\partial t}\right)_\zeta |_p$. z_* acts as a linear map of the tangent space $T_p(\mathcal{M})$ into the tangent space $T_{z(p)}(\mathcal{N})$. For each C^r function g and vector $\mathbf{R}$ at $z(p)$, $R(z^* g(p)) = z^* R(f(z(p))$. We can make use of this mapping $z_* : \mathcal{M} \to \mathcal{N}$ to define a linear 1-form mapping as $z^* : T^*_{z(p)}(\mathcal{N}) \to T^*_p(\mathcal{M})$. This implies that an arbitrary p-form $\mathbf{\Lambda} \in T^*_{z(p)}$ is mapped onto the p-form $z^*\mathbf{\Lambda} \in T^*_p$, such that $\langle z^*\mathbf{\Lambda}, \mathbf{R}\rangle_p = \langle \mathbf{\Lambda}, z_*\mathbf{R}\rangle_{z(p)}$ is true for a vector $R \in T_p$. The fact that $d(z^*\mathbf{\Lambda}) = z^* d(\mathbf{\Lambda})$ holds is a consequence of of the previously mentioned result.

The *exterior derivative* acts on a function g which is merely a 0-form field to produce a 1-form field dg. Let's generalize this, let $\mathbf{\Lambda}$ be a p-form field defined by $\mathbf{\Lambda} = \Lambda_{\alpha\beta...\zeta} dx^\alpha \wedge dx^\beta \wedge ... \wedge dx^\zeta$. Now, take the exterior derivative of this p-form field to obtain a $(p+1)$-field as follows

$$d\mathbf{\Lambda} = d\Lambda_{\alpha\beta...\zeta} dx^\alpha \wedge dx^\beta \wedge ... \wedge dx^\zeta \tag{1.29}$$

This $(p+1)$-field is independent of the coordinates $\alpha, \beta, ..., \zeta = x^\alpha$ used in definition and to convince ourselves of this fact, consider another set of coordinates, say the barred counterparts given by $\bar{\alpha}, \bar{\beta}, ..., \bar{\zeta} = x^{\bar{\alpha}}$. In these coordinates we have $\mathbf{\Lambda} = \Lambda_{\bar{\alpha}\bar{\beta}...\bar{\zeta}} dx^{\bar{\alpha}} \wedge dx^{\bar{\beta}} \wedge ... \wedge dx^{\bar{\zeta}}$. The components of $\Lambda_{\bar{\alpha}\bar{\beta}...\bar{\zeta}}$ are given by

$$\Lambda_{\bar{\alpha}\bar{\beta}...\bar{\zeta}} = \frac{\partial x^\alpha}{\partial x^{\bar{\alpha}}} \frac{\partial x^\beta}{\partial x^{\bar{\beta}}} \cdots \frac{\partial x^\zeta}{\partial x^{\bar{\zeta}}} \Lambda_{\alpha\beta...\zeta} \tag{1.30}$$

and the new definition of $d\mathbf{\Lambda}$ in the coordinates of $x^{\bar{\alpha}}$ is the following [6]

$$\begin{aligned}
d\mathbf{\Lambda} &= d\Lambda_{\bar{\alpha}\bar{\beta}...\bar{\zeta}} dx^{\bar{\alpha}} \wedge dx^{\bar{\beta}} \wedge ... \wedge dx^{\bar{\zeta}} \\
&= d\left(\tfrac{\partial x^\alpha}{\partial x^{\bar{\alpha}}} \tfrac{\partial x^\beta}{\partial x^{\bar{\beta}}} \cdots \tfrac{\partial x^\zeta}{\partial x^{\bar{\zeta}}} \Lambda_{\alpha\beta...\zeta}\right) dx^{\bar{\alpha}} \wedge dx^{\bar{\beta}} \wedge ... \wedge dx^{\bar{\zeta}} \\
&= \tfrac{\partial x^\alpha}{\partial x^{\bar{\alpha}}} \tfrac{\partial x^\beta}{\partial x^{\bar{\beta}}} \cdots \tfrac{\partial x^\zeta}{\partial x^{\bar{\zeta}}} d\Lambda_{\alpha\beta...\zeta} dx^{\bar{\alpha}} \wedge dx^{\bar{\beta}} \wedge ... \wedge dx^{\bar{\zeta}} \\
&\quad + \tfrac{\partial^2 x^\alpha}{\partial x^{\bar{\alpha}} \partial x^{\bar{\chi}}} \tfrac{\partial x^\beta}{\partial x^{\bar{\beta}}} \cdots \tfrac{\partial x^\zeta}{\partial x^{\bar{\zeta}}} d\Lambda_{\alpha\beta...\zeta} dx^{\bar{\chi}} \wedge dx^{\bar{\alpha}} \wedge dx^{\bar{\beta}} \wedge ... \wedge dx^{\bar{\zeta}} \\
&= d\Lambda_{\alpha\beta...\zeta} dx^\alpha \wedge dx^\beta \wedge ... \wedge dx^\zeta
\end{aligned} \tag{1.31}$$

Consider the coordinate expression for dg, $dg = \frac{\partial g}{\partial x^\alpha} dx^\alpha$, observe that[6] $d(dg) = \frac{\partial^2 g}{\partial x^\alpha \partial x^\beta} dx^\alpha \wedge dx^\beta = 0.$. This implies that for any p-form field $\mathbf{\Lambda}$, $d(d\mathbf{\Lambda}) = 0$.

[6] The last line was arrived at as $\frac{\partial^2 x^\alpha}{\partial x^{\bar{\alpha}} \partial x^{\bar{\chi}}}$ is symmetric in $\bar{\alpha}$ and $\bar{\chi}$, but $dx^{\bar{\chi}} \wedge dx^{\bar{\alpha}}$ is skew symmetric.

1.3.10 General p-forms

The gradient of a scalar produces a 1-form, and the exterior derivative of this 1-form produces a 2-form, this chain continues. Thus the exterior derivative of a $(p-1)$-form produces a *p-form* defined as follows

$$\boldsymbol{\Xi} = \frac{1}{p!} \Xi_{i_1 \ i_2 \ \dots \ ip} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_p} \tag{1.32}$$

and the exterior derivative of $\boldsymbol{\Xi}$ is defined as follows

$$d\boldsymbol{\Xi} = \frac{1}{p!} \frac{\partial \Xi_{i_1 \ i_2 \ \dots \ ip}}{\partial x^{i_0}} dx^{i_0} \wedge dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_p} \tag{1.33}$$

It is important to note that $\Xi_{i_1 \ i_2 \ \dots \ ip}$ is antisymmetric under a 2 index interchange. The definitions provided here is an alternate one to the one in which the factor of $\frac{1}{p!}$ is not included[7] A *closed form* is defined as a form whose differential is zero while an *exact form* is defined as a form that is the differential of an exterior form (and it's an example of a closed form).

1.3.11 Parallel Transport and Covariant Differenriation

An *ordinary differential* of a vector $\mathbf{A}^\mu$ in a direction x^α is defined as follows

$$\frac{\partial \mathbf{A}^\mu}{\partial x^\alpha} dx^\alpha = \partial_\alpha \mathbf{A}^\mu dx^\alpha = \mathbf{A}^\mu_{,\alpha} \equiv \mathbf{A}^\mu(x + dx) - \mathbf{A}^\mu(x). \tag{1.34}$$

The ordinary differentials are defined by the difference between two vectors defined at two distinct points. In curved spacetime however, we need to account for the rotations undergone by the vector as it evolves with time. Thus, we introduce the quantity $\delta \boldsymbol{A^\mu}$ and subtract it from the ordinary differential to obtain the *covariant differential* $D_\alpha \mathbf{A}^\mu \equiv \mathbf{A}^\mu_{;\alpha}$. To observe this rotation due to curvature, we transport the vector $\mathbf{A}^\mu(x+dx)$ to the point x without changing it's direction (see figure 1.9). This is known as *parallel transport.*

$$\mathbf{A}^\mu_{;\alpha} \equiv \mathbf{A}^\mu(x + dx) - \left[\mathbf{A}^\mu(x) + \delta \mathbf{A}^\mu(x)\right]. \tag{1.35}$$

Let $\mathbf{S}$ be a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor. The covariant derivative $D_\mathbf{R}\mathbf{S}$ of $\mathbf{S}$ along a curve $P(\zeta)$, whose tangent vector $\mathbf{R} = \frac{dP}{d\zeta}$ is defined as follows

$$D_\mathbf{R}\mathbf{S}|_{P(0)} = lt_{\eta \to 0} \left[\frac{\mathbf{S}[P(\eta)]_{II^{el} \ transported \ toP(0)} - \mathbf{S[P(0)]}}{\eta} \right] \tag{1.36}$$

[7] i.e., in accordance to this definition, a 2-form is written as $dx^\alpha \wedge dx^\beta = \frac{1}{2}\left(dx^\alpha \otimes dx^\beta - dx^\beta \otimes dx^\alpha\right)$.

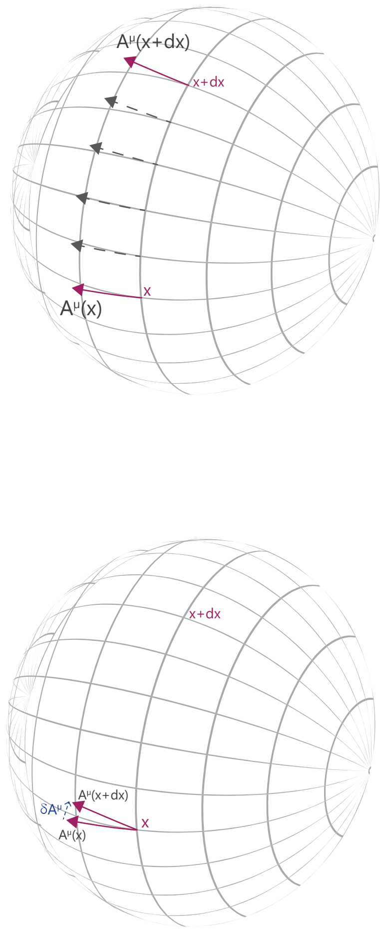


Fig. 1.9. The evolution of the vector along the curved path (L) and the parallel transport of $\mathbf{A}^{\mu}(x+dx)$ to the point x (R)

The covariant derivative, denoted by either ∇ or D is a *connection* at a point p on the manifold $\mathcal{M}$ which allots every vector field $\mathbf{R}$ at p a differential operator $D_{\mathbf{R}}$, such that the operator maps an arbitrary C^r vector field $\mathbf{S}$ into a vector field $D_{\mathbf{S}}$. Following are some of the algebraic properties of D.

1. $D_{\mathbf{R}}\mathbf{S}$ is a tensor in the argument $\mathbf{R}$. For arbitrary functions g, h and continuous, once-differentiable vectors fields, i.e., a C^1 vectors fields $\mathbf{R}, \mathbf{S}, \mathbf{Q}$,

$$D_{g\mathbf{R}+h\mathbf{S}}\mathbf{Q} = gD_{\mathbf{R}}\mathbf{Q} + hD_{\mathbf{S}}\mathbf{Q}. \tag{1.37}$$

2. $D_{\mathbf{R}}\mathbf{S}$ obeys the linearity condition.

$$D_{\mathbf{R}}\left(\mu\mathbf{S} + \nu\mathbf{Q}\right) = \mu D_{\mathbf{R}}\mathbf{S} + \nu D_{\mathbf{R}}\mathbf{Q}. \tag{1.38}$$

3. For any two C^1 vector fields of the same rank $\mathbf{R}, \mathbf{S}$

$$D_{\mathbf{R}}\mathbf{S} - D_{\mathbf{S}}\mathbf{R} = \left[\mathbf{R}, \mathbf{S}\right]. \tag{1.39}$$

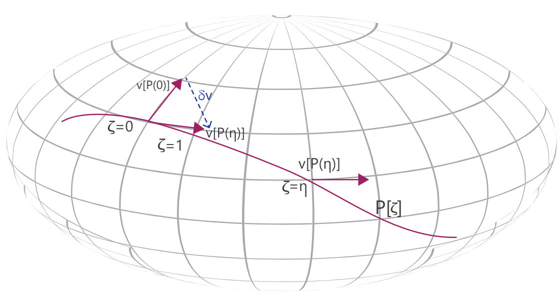


Fig. 1.10. Parallel transport of a vector on the parametrized curve $P(\zeta)$

In the third property, $[\mathbf{R}, \mathbf{S}]$ is called a *commutator*. Suppose $\mathbf{R}$ and $\mathbf{S}$ are tangent vectors fields, then it holds that $\mathbf{R} = \partial_{\mathbf{R}}$ and $\mathbf{S} = \partial_{\mathbf{S}}$ are true (from previous definitions). Thus, a commutator, which by itself is a tangent vector field is defined as follows

$$[\mathbf{R}, \mathbf{S}] \equiv [\partial_{\mathbf{R}}, \partial_{\mathbf{S}}] \equiv \partial_{\mathbf{R}}\partial_{\mathbf{S}} - \partial_{\mathbf{S}}\partial_{\mathbf{R}}. \tag{1.40}$$

Here, there is a need to define the commutation coefficients of a basis as the concept will come handy for a future topic. For the basis vectors $\mathbf{e}_{\boldsymbol{\alpha}}$ and $\mathbf{e}_{\boldsymbol{\beta}}$, the commutation coefficient $C^{\gamma}_{\alpha\beta}$ is obtained by commuting the basis vectors. It is a tensor-like coefficient which gives the difference between partial derivatives of two coordinates with respect to the other coordinate.

$$[\mathbf{e}_{\boldsymbol{\alpha}}, \mathbf{e}_{\boldsymbol{\beta}}] \equiv \partial_\alpha \mathbf{e}_{\boldsymbol{\beta}} - \partial_\beta \mathbf{e}_{\boldsymbol{\alpha}} \equiv C^{\gamma}_{\alpha\beta} e_\gamma. \tag{1.41}$$

If $C^{\gamma}_{\alpha\beta} = 0$ it is called a coordinate basis or *holonomic* if some $C^{\gamma}_{\alpha\beta} \neq 0$ then it is called a non-coordinate basis or *anholonomic*.

1.3.12 Connection Coefficients: An Introduction

From the example involving parallel transportation of the vector $\mathbf{A}^\mu$, for a small dx, $\delta\mathbf{A}^\mu(x)$ should be linear in dx and also in $\mathbf{A}^\mu(x)$ or in other words, it must be an output of some transformation of the vector $\mathbf{A}^\mu(x)$ at x. This is given below.

$$\delta\mathbf{A}^\mu(x) = B^\mu_\nu A^\nu(x). \tag{1.42}$$

Here, B^μ_ν is a matrix that transforms the vector during parallel transport. During parallel transport, the basis vectors and the basis 1-forms would twist, contract, expand, and turn according to the curvature, and this is quantified by the *connection coefficient*. The connection coefficient is defined as follows

$$\begin{aligned} \Gamma^{\gamma}_{\alpha\beta} &= \langle \boldsymbol{\omega}^\gamma, D_{\mathbf{e}_\alpha} \mathbf{e}_\beta \rangle = \langle \boldsymbol{\omega}^\gamma, D_\alpha \mathbf{e}_\beta \rangle \\ \Gamma^{\gamma}_{\alpha\beta} &= - \langle D_\alpha \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle . \end{aligned} \tag{1.43}$$

It is fairly easy to prove the latter equation (proof given below). The matrix mentioned previously is defined as $B^\mu_\nu = -\Gamma^\mu_{\nu\alpha} dx^\alpha$. Thus, we can conclude that the matrix accounts for all the contributions of the basis vectors and the basis 1-forms via the connection coefficients, over a small distance dx.

Proof. To Prove that $\Gamma^{\gamma}_{\alpha\beta} = - \langle D_\alpha \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle$

From lemma: $\langle \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle = \delta^\gamma_\beta$

$\mathrm{D}_\alpha \langle \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle = \partial_{\mathbf{e}_\alpha} \langle \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle = D_\alpha \left(\delta^\gamma_\beta \right) = 0$

Thus, $0 = \underbrace{(D_\alpha \boldsymbol{\omega}^\gamma) \otimes \mathbf{e}_\beta + \boldsymbol{\omega}^\gamma \otimes (D_\alpha \mathbf{e}_\beta)}_{D_\alpha(Contraction\ of\ \boldsymbol{\omega}^\gamma \otimes \mathbf{e}_\beta)}$

$$0 = \underbrace{\langle D_\alpha \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle + \langle \boldsymbol{\omega}^\gamma, D_\alpha \mathbf{e}_\beta \rangle}_{Contraction\ of\ [D_\alpha(\boldsymbol{\omega}^\gamma \otimes \mathbf{e}_\beta)]}$$

$$\Gamma^\gamma_{\alpha\beta} = -\langle D_\alpha \boldsymbol{\omega}^\gamma, \mathbf{e}_\beta \rangle .$$

When we take the covariant derivative of a tensor, we are to differentiate the tensor with respect to the arbitrary basis and also account for the twisting and turning of the 1-forms and the vectors present in the tensor's slots. Consider a tensor $\mathbf{S}$ of rank $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Upon covariant differentiation, we obtain the following terms.

$$S^\beta_{\alpha;\gamma} = S^\beta_{\alpha,\gamma} + \Gamma^\beta_{\nu\gamma} S^\nu_\alpha - \Gamma^\nu_{\alpha\gamma} S^\beta_\nu. \tag{1.44}$$

There are three thing to note here which will help in understanding how to take the covariant derivative for any arbitrarily ranked tensor. Firstly, it is to be noted that a + (positive) sign is used if the index being corrected is upstairs. In the example, an arbitrary summation index ν was used to correct β that resided upstairs, i.e., $+\Gamma^\beta_{\nu\gamma} S^\nu_\alpha$. The second point to be noted is the use of a $-$ (negative) sign. It is to be employed when the index being corrected is downstairs, -$\Gamma^\nu_{\alpha\gamma} S^\beta_\nu$. Lastly, we observe that the index being corrected shifts from the tensor $\mathbf{S}$ onto the connection Γ and is replaced on the tensor by a dummy summation index ν. Let's see a few examples to strengthen this concept. One way to check the correctness of your answer is to check for homogeneity- check if the indexes upstairs and downstairs are alike on either sides of the equation (this is shown in the examples).

1. $\underbrace{D_\gamma T_{\alpha\beta}}_{\{\frac{1}{\gamma\alpha\beta}\}} = \underbrace{\frac{\partial T_{\alpha\beta}}{\partial x^\gamma}}_{\{\frac{1}{\gamma\alpha\beta}\}} - \underbrace{\Gamma^\xi_{\gamma\alpha} T_{\beta\xi}}_{\{\frac{\xi}{\gamma\alpha\beta\xi}\}} - \underbrace{\Gamma^\xi_{\gamma\beta} T_{\alpha\xi}}_{\{\frac{\xi}{\gamma\alpha\beta\xi}\}}$
2. $\underbrace{D_\gamma T_\alpha}_{\{\frac{1}{\gamma\alpha}\}} = \underbrace{T_{\alpha,\gamma}}_{\{\frac{1}{\gamma\alpha}\}} - \underbrace{\Gamma^\beta_{\gamma\alpha} T_\beta}_{\{\frac{1}{\gamma\alpha}\}}$

1.3.13 Structure Coefficients

Let $\mathcal{M}$ be a manifold equipped with a topology. Consider the set of coordinates $\{x^\alpha\} = \left(x^1, x^2, ..., x^n\right)$ in the chart z. Now, the coordinates follow a lemma which states the following: $\frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta$. This lemma also implies that $d^2 x^\alpha = 0$. For moving frames, however, the differentials of the 1-forms $\boldsymbol{\Xi}$ do not vanish, i.e., the wedge product of two 1-forms do not yield a null result. This wedge product produces a 2-form given by[8]

[8] As mentioned before, in the alternate formalism the fraction $\frac{1}{2}$ is omitted.

$$d\Xi^m \equiv -\frac{1}{2} C^m_{ab} \Xi^a \wedge \Xi^b, \tag{1.45}$$

where C^m_{ab} is called the *structure coefficients* of the frame. The structure coefficient C^m_{ab} is antisymmetric in a and b.

1.3.14 Riemannian Connection

A *Riemannian connection* Ω is defined for a pseudo-Riemannian metric g. It is a linear connection obeys two conditions, the covariant derivative of the metric is zero, and the second condition requires that the second covariant derivatives of scalar functions to commute. The second condition implies that the connection has *vanishing torsion.*

Theorem 1.7. *The following conditions determine the Riemannian connection* $\Omega^\gamma_{\alpha\beta}$

$$\partial_\gamma g_{\alpha\beta} - \Omega^\lambda_{\gamma\beta} g_{\alpha\lambda} - \Omega^\lambda_{\gamma\alpha} g_{\lambda\beta} = 0. \tag{1.46}$$

Let h be a scalar function, the condition requires the following

$$D_\gamma \partial_\alpha h - D_\alpha \partial_\gamma h = 0. \tag{1.47}$$

The Riemannian connection is defined as follows

$$\begin{aligned} \Omega^\gamma_{\alpha\beta} &\equiv \Gamma^\gamma_{\alpha\beta} + g^{\gamma\lambda} \bar{\Omega}_{\alpha\beta,\lambda} \\ \bar{\Omega}_{\alpha\beta,\lambda} &\equiv \tfrac{1}{2} \left(g_{\lambda\xi} C^\xi_{\alpha\beta} - g_{\xi\beta} C^\xi_{\alpha\lambda} - g_{\alpha\xi} C^\xi_{\beta\lambda} \right). \end{aligned} \tag{1.48}$$

a coordinate basis for which the structure coefficients ($C^\xi_{\alpha\beta}$ and others) are zero, is called holonomic. A non-coordinate basis always has some non-zero structure coefficients, and is called anholonomic. In the holonomic case, the connection coefficients are called Christoffel symbols given by the following expression (will be proven later)

$$\Gamma^\gamma_{\alpha\beta} \equiv \frac{1}{2} g^{\gamma\lambda} \left(g_{\beta\lambda,\alpha} + g_{\alpha\lambda,\beta} - g_{\alpha\beta,\lambda} \right). \tag{1.49}$$

1.3.15 Revisiting the Metric Tensor

Using the slot machine definition of tensors, we can think of the metric tensor as a slot machine with two slots which accept only vectors as inputs: $g(\underbrace{\quad}_{vector1}, \underbrace{\quad}_{vector2})$. When the same vector is inserted into the slots, we get the square of the length of the vector as the output, $g(\mathbf{R}, \mathbf{R}) = \mathbf{R}^2$. When two different vectors are inserted, we obtain the scalar product of the vectors as the output. It is important to note that irrespective of the order of insertion of the vectors, the result remains unchanged. This is show below

$$g(\mathbf{R}, \mathbf{Q}) = g(\mathbf{Q}, \mathbf{R}) = \mathbf{RQ} = \mathbf{QR}. \tag{1.50}$$

The metric obeys the condition of linearity and in a specific coordinate system, its operation on the two input vectors is given by the following bilinear expression

$$g(\mathbf{R}, \mathbf{Q}) = g_{\mu\nu} R^{\mu} Q^{\nu}. \tag{1.51}$$

There exists a reason for the name metric tensor, at least for the case when the inner product is positive definite. Consider two points, z and $z + \Delta z$, infinitesimally close to each other. The square of the infinitesimal distance of the displacement vector with components Δz^m is represented as follows

$$(\Delta s)^2 \equiv g_{mn} \Delta z^m \Delta z^n. \tag{1.52}$$

Now, the metric tensor, just like any other tensor, would transform under a coordinate change as follows

$$g_{mn}(x) = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial \bar{x}^j}{\partial x^n} g_{ij}(\bar{x}), \tag{1.53}$$

from the above transformation, it is clear that the length Δs of the displacement vector is not dependent on the choice of coordinates, rather it is dependent only on the two points under consideration. This formula is nothing but the generalization of the Pythogorean theorem of Euclidean geometry, which states that

$$(\Delta s)^2 = \Delta x^2 + \Delta y^2 + \Delta z^2, \tag{1.54}$$

the emergence of the metric tensor is the starting point of Riemanninan geometry. Another interesting way to look at it, and one which we will be extensively using in future chapters, is to observe that the length of a curve $P[\zeta(t)]$ between two points $\zeta(t_1)$ and $\zeta(t_2)$ can be expressed as follows

$$s = \int_{t_1}^{t_2} \left[g_{mn} \frac{d\zeta^m}{dt} \frac{d\zeta^n}{dt} \right]^{\frac{1}{2}} dt. \tag{1.55}$$

Example 1.8. In this example, we demostrate a coordinate transformation for a metric from the toroidal to the Cartesian system. Toroidal coordinates are related to the usual Cartesian coordinates $\{x, y, z\}$ of Euclidean three-space $\mathbb{R}^3$ by

$$
\begin{aligned}
x &= a\tfrac{sinh\tau}{cosh\tau - cos\sigma}cos\phi, \\
y &= a\tfrac{sinh\tau}{cosh\tau - cos\sigma}sin\phi, \\
z &= a\tfrac{sin\sigma}{cosh\tau - cos\sigma}.
\end{aligned}
\tag{1.56}
$$

where a is a constant, $\sigma \in (-\pi, \pi]$, $\tau \geq 0$, and $\phi \in [0, 2\pi)$. We restrict our attention to the $y = 0$ plane and doing so we can see that $\phi = 0$. Thus, this makes our coordinate relations take the following form

$$
x = a\frac{sinh\tau}{cosh\tau - cos\sigma}, \quad a\frac{sin\sigma}{cosh\tau - cos\sigma}. \tag{1.57}
$$

Now, we first find the coordinate transformation matrix $M^{\mu}_{\mu'} = \partial x^{\mu}/\partial x^{\mu'}$ relating the Cartesian coordinates $\{x, z\}$ to toroidal coordinates $\{\tau, \sigma\}$ as follows

$$
M^{\mu}_{\mu'} = \begin{pmatrix} \frac{\partial x}{\partial \tau} & \frac{\partial x}{\partial \sigma} \\ \frac{\partial z}{\partial \tau} & \frac{\partial z}{\partial \sigma} \end{pmatrix} = -\frac{a}{(cos\sigma - cosh\tau)^2} \begin{pmatrix} (cos\sigma cosh\tau - 1) & sin\sigma sinh\tau \\ cos\sigma sinh\tau & cosh\tau sin\sigma \end{pmatrix}. \tag{1.58}
$$

Now, there are two ways we can find the metric line element ds^2. Before exploring the two methods, let us represent the line element in the Cartesian form

$$
ds^2 = dx^2 + dx^2, \tag{1.59}
$$

where,

$$
dx = \frac{\partial x}{\partial \tau}d\tau + \frac{\partial x}{\partial \sigma}d\sigma, \quad dz = \frac{\partial z}{\partial \tau}d\tau + \frac{\partial z}{\partial \sigma}d\sigma. \tag{1.60}
$$

The first method is to directly take the squares of dx and dz and substitute into the line element to obtain the line element in toroidal coordinates. This is straightforward computation. The other method is motivated by the fact that the metric tensor in Cartesian coordinates is diagonal and is equal to the identity matrix, i.e., $g_{\mu\nu} = \mathbb{I}$. Now, in order to obtain the metric in terms of the toroidal coordinates, we simply observe how the metric transforms under the coordinate changes as follows ($\bar{x} = \{\tau, \sigma\}$ and $x = \{x, z\}$)

$$g_{\mu'\nu'}(\bar{x}) = g_{\mu\nu}\frac{\partial x^\mu}{\partial \bar{x}^{\mu'}}\frac{\partial x^\nu}{\partial \bar{x}^{\nu'}} = \frac{\partial x^\mu}{\partial \bar{x}^{\mu'}} g_{\mu\nu}\frac{\partial x^\nu}{\partial \bar{x}^{\nu'}}$$

$$\Rightarrow \frac{a^2}{(cos\sigma - cosh\tau)^4}\begin{pmatrix}(cos\sigma cosh\tau - 1) & sin\sigma sinh\tau \\ cos\sigma sinh\tau & cosh\tau sin\sigma\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}(cos\sigma cosh\tau - 1) & sin\sigma sinh\tau \\ cos\sigma sinh\tau & cosh\tau sin\sigma\end{pmatrix}$$

$$= \mathcal{F}\begin{pmatrix}(cos\sigma cosh\tau - 1)^2 + cos\sigma sin\sigma sinh^2\tau & sin\sigma sinh\tau\left(cosh\tau\left(cos\sigma + sin\sigma\right) - 1\right) \\ cos\sigma sinh\tau\left(cosh\tau\left(cos\sigma + sin\sigma\right) - 1\right) & sin\sigma\left(cosh^2\tau sin\sigma + cos\sigma sinh^2\tau\right)\end{pmatrix}, \tag{1.61}$$

where $\mathcal{F} = a^2\left(cos\sigma - cosh\tau\right)^{-4}$. Now, to obtain the line element we simply compute

$$ds^2 = \begin{pmatrix}d\tau & d\sigma\end{pmatrix} g_{\mu'\nu'}\begin{pmatrix}d\tau\\ d\sigma\end{pmatrix}. \tag{1.62}$$

Notice that the line element in toroidal coordinates is not diagonal anymore.

Exercise 1

Find the coordinate transformation matrix and the corresponding line element in elliptic coordinates which are related to the Cartesian coordinates $\{x, y\}$ of Euclidean two-space $\mathbb{R}^2$ by

$$x = a\ cosh\mu\ cos\nu, \quad y = a\ sinh\mu\ sin\nu,$$

where $\mu \geq 0$ and $\nu \in [0, 2\pi]$ and the same for parabolic coordinates

$$x = \sigma\tau, \quad y = \frac{1}{2}\left(\tau^2 - \sigma^2\right).$$

Between these two, which has a diagonal line element?

1.3.16 Normal Coordinates

Consider the curve $P(\zeta)$ with well defined end points, say a and b, and let R^γ be the tangent vector. The tensor $\mathbf{S}$ is said to be parallelly transported along the curve $P(\zeta)$ if $\frac{D\mathbf{S}}{s\zeta} = 0$. The covariant derivative of the tangent vector can be expressed in the terms of the metric tensor as $D_\alpha \mathbf{R}^\gamma = g^{\mu\gamma} D_\alpha \mathbf{R}_\mu$. Now,

$$D_\alpha \mathbf{R}^\gamma = D_\alpha\left(g^{\mu\gamma}\mathbf{R}_\mu\right) = g^{\mu\gamma} D_\alpha \mathbf{R}_\mu + \mathbf{R}_\mu D_\alpha g^{\mu\gamma} \tag{1.63}$$

but, we know that $D_\alpha \mathbf{R}^\gamma = g^{\mu\gamma} D_\alpha \mathbf{R}_\mu$. Hence, we conclude that $\mathbf{R}_\mu D_\alpha g^{\mu\gamma} = 0$. This is not just something that we obtained from lousy reasoning. Observe the term carefully & you would realize that it is the very condition of that of parallel transport. What this means is that we want the inner product of two vector inputs, say **a** and **b**, $g(\mathbf{a}, \mathbf{b}) = \mathbf{ab} = g_{\mu\nu} a^\mu b^\nu$ to remain constant under parallel transport along a curve with tangent $\mathbf{R}^\gamma$. This gives rise to the following condition

$$R^\gamma D_\gamma \left(g_{\mu\nu} a^\mu b^\nu\right) = 0, \tag{1.64}$$

parallel transport requires $R^\gamma a^\mu b^\nu D_\gamma g_{\mu\nu}$ to be true for all R, a, b. The vanishing covariant metric derivative is not a consequence of using any connection, it's a condition that allows us to choose a specific connection $\Gamma^\rho_{\mu\nu}$. In principle, we could have connections for which $D_\gamma g_{\mu\nu} \neq 0$, but we specifically require a connection for which this condition is true because we want a parallel transport operation which preserves angles and lengths. In the local frame, which is the reference frame in the vicinity of an arbitrary point x_0 in which we can choose normal coordinates[9] such that at that point $g_{\mu\nu}(x_0) = \delta_{\mu\nu}$ and the derivative of the metric with respect to any component of the metric can be set to 0, i.e., $g_{\mu\nu,\alpha} = 0$ and also such that $\frac{\partial^2 g_{\mu\nu}}{\partial x^\mu \partial x^\nu} \neq 0$ (except when space is flat). The last condition implies that at the local point x_0, the connection vanishes (specifically, the Christoffel symbol vanishes), i.e., $\Gamma^\rho_{\mu\nu} = 0$.

Consider the locally flat coordinates (or normal coordinates) $\xi^i(x^\mu)$, it can be shown that $\frac{\partial^2 \xi^ß}{\partial x^\mu \partial x^\nu} = \Gamma^\rho_{\mu\nu} \frac{\partial \xi^ß}{\partial x^\rho}$[10]. It can be shown by the following calculation that the covariant derivative of the metric tensor vanishes.

$$\begin{aligned}
D_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - g_{\mu\sigma}\Gamma^\sigma_{\nu\rho} - g_{\sigma\nu}\Gamma^\sigma_{\mu\rho} \\
&= \partial_\rho \left(\frac{\partial \xi^i}{\partial x^\mu} \frac{\partial \xi^i}{\partial x^\nu} \right) - g_{\mu\sigma} \frac{\partial x^\sigma}{\partial \xi^i} \frac{\partial^2 \xi^i}{\partial x^\nu \partial x^\rho} - g_{\sigma\nu} \frac{\partial x^\sigma}{\partial \xi^i} \frac{\partial^2 \xi^i}{\partial x^\mu \partial x^\rho} \\
&= \frac{\partial^2 \xi^i}{\partial x^\rho \partial x^\mu} \frac{\partial \xi^i}{\partial x^\nu} + \frac{\partial \xi^i}{\partial x^\mu} \frac{\partial^2 \xi^i}{\partial x^\rho \partial x^\nu} - \frac{\partial \xi^j}{\partial x^\mu} \underbrace{\frac{\partial \xi^j}{\partial x^\sigma} \frac{\partial x^\sigma}{\partial \xi^i}}_{\delta^j_i} \frac{\partial^2 \xi^i}{\partial x^\nu \partial x^\rho} - \frac{\partial \xi^j}{\partial x^\sigma} \frac{\partial \xi^j}{\partial x^\nu} \frac{\partial x^\sigma}{\partial \xi^i} \frac{\partial^2 \xi^i}{\partial x^\mu \partial x^\rho} \\
&= 0.
\end{aligned} \tag{1.65}$$

1.3.17 Pfaffian Derivatives

Coframe on a manifold $\mathcal{M}$ is a system of 1-forms which form a basis of the *cotangent bundle* at every point (just to remind ourselves- The dual space $^*T_p$ is often called the *cotangent space*). The system of 1-forms Ξ^m used in defining the structure coefficients is a coframe. The *Pfaffian derivatives* ∂_m in the

[9] sometimes called *Gaussian normal coordinates*

[10] This equation will be proven in upcoming sections and it has a very deep physical meaning.

coframe Ξ^m of a function h are defined as follows

$$dh \equiv \partial_m h \; \Xi^m. \tag{1.66}$$

It is important to note that Pfaffian derivatives, unlike normal derivatives, do not commute. This can be seen from the analysis of the identy $dh \equiv 0$ as follows

$$\begin{aligned} d^2h &\equiv \tfrac{1}{2}\left[\partial_m\partial_n h - \partial_n\partial_m h - C^a_{mn}\partial_a h\right] \; \Xi^m \wedge \Xi^n \equiv 0 \\ \partial_m\partial_n h - \partial_n\partial_m h &= C^a_{mn}\partial_a h. \end{aligned} \tag{1.67}$$

The basis $\mathbf{e}_m$ which is the dual to Ξ^m satisfies the commutation conditions (in this formalism we include the fraction $\frac{1}{2}$, however, for all future purposes we will neglect this factor)

$$C^a_{mn} e_a = \left[\mathbf{e}_m, \mathbf{e}_n\right]. \tag{1.68}$$

1.3.18 Back to Connections

Consider a tensor $\mathbf{S}$ of rank $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. With the knowledge of the components of $S^\beta_{\alpha;\gamma}$ we can calculate the components of the covariant derivative $D_{\mathbf{R}}\mathbf{S}$ by a contraction into R^γ as follows (where $\mathbf{R} = \frac{dP}{d\zeta} = \frac{dx^\gamma}{d\zeta}$ is a tangent vector present on the curve $P(\zeta)$)

$$D_{\mathbf{R}}\mathbf{S} = \left(S^\beta_{\alpha;\gamma} R^\gamma\right) \mathbf{e}_\beta \otimes \boldsymbol{\omega}^\alpha. \tag{1.69}$$

The components of $D_{\mathbf{R}}\mathbf{S}$ are denoted by $\frac{DS^\alpha_\beta}{d\zeta}$. Thus, we obtain

$$\begin{aligned} \tfrac{DS^\beta_\alpha}{d\zeta} &\equiv S^\beta_{\alpha;\gamma} R^\gamma = S^\beta_{\alpha;\gamma} \tfrac{dx^\gamma}{d\zeta} \\ \tfrac{DS^\beta_\alpha}{d\zeta} &= \tfrac{dS^\beta_\alpha}{d\zeta} + \left(\Gamma^\beta_{\nu\gamma} S^\nu_\alpha - \Gamma^\nu_{\alpha\gamma} S^\beta_\nu\right) \tfrac{dx^\gamma}{d\zeta}. \end{aligned} \tag{1.70}$$

To find the connection coefficients for a given basis we first need to take metric coefficients in the given basis and then calculate their directional derivatives along the considered basis directions.

$$\begin{aligned} D_\gamma g_{\alpha\beta} &= g_{\alpha\beta,\gamma} - \Gamma^\nu_{\alpha\gamma} g_{\nu\beta} - \Gamma^\nu_{\beta\gamma} g_{\nu\alpha} = 0 \\ g_{\alpha\beta,\gamma} &- \Gamma_{\beta\alpha\gamma} - \Gamma_{\alpha\beta\gamma} = 0 \\ g_{\alpha\beta,\gamma} &= 2\Gamma_{(\alpha\beta)\gamma}. \end{aligned} \tag{1.71}$$

Let us now construct a metric for $\Gamma_{\nu\alpha\gamma}$.

$$\begin{aligned}\tfrac{1}{2}\left(g_{\nu\beta,\gamma}+g_{\nu\gamma,\beta}-g_{\beta\gamma,\nu}\right) &= \Gamma_{(\nu\beta)\gamma}+\Gamma_{(\nu\gamma)\beta}-\Gamma_{(\beta\gamma)\nu}\\ &= \tfrac{1}{2}\left(\Gamma_{\nu\beta\gamma}+\Gamma_{\beta\nu\gamma}+\Gamma_{\nu\gamma\beta}+\Gamma_{\gamma\nu\beta}-\Gamma_{\beta\gamma\nu}-\Gamma_{\gamma\beta\nu}\right) \qquad (1.72)\\ &= \Gamma_{\nu\beta\gamma}+\left(\Gamma_{\beta[\nu\gamma]}+\Gamma_{\gamma[\nu\beta]}-\Gamma_{\nu[\beta\gamma]}\right).\end{aligned}$$

Let $(\mathbf{R}=\mathbf{e}_\nu, \mathbf{Q}=\mathbf{e}_\lambda)$ be two basis vectors. We can now use them to construct structure coefficients by commuting the basis which was something we realized in the section on Pfaffian derivatives.

$$\begin{aligned}&[\mathbf{e}_\nu, \mathbf{e}_\lambda] = D_\nu \mathbf{e}_\lambda - D_\lambda \mathbf{e}_\nu = C^\rho_{\nu\lambda} e_\rho\\ &C^\rho_{\nu\lambda} e_\rho = \left(\Gamma^\rho_{\lambda\nu} - \Gamma^\rho_{\nu\lambda}\right) e_\rho = 2\Gamma^\rho_{[\lambda\nu]} e_\rho \qquad (1.73)\\ &\Gamma^\rho_{[\lambda\nu]} = -\tfrac{1}{2} C^\rho_{\nu\lambda} \quad \rightarrow \quad \Gamma_{\rho[\lambda\nu]} = -\tfrac{1}{2} C_{\nu\lambda\rho}.\end{aligned}$$

Combining the equations we obtain an expression for the connection coefficient.

$$\Gamma_{\nu\beta\gamma} = \frac{1}{2}[g_{\nu\beta,\gamma}+g_{\nu\gamma,\beta}-g_{\beta\gamma,\nu}+\underbrace{C_{\nu\beta\gamma}+C_{\nu\gamma\beta}-C_{\beta\gamma\nu}}_{=\,0\ \textit{for coordinate basis (holonomic)}}]. \qquad (1.74)$$

Thus, we obtain the Christoffel symbol which can be expressed as follows (after raising an index)

$$\Gamma^\alpha_{\beta\gamma} = g^{\alpha\nu}\Gamma_{\nu\beta\gamma}. \qquad (1.75)$$

1.3.19 Transformation Formula for Connections

Let us take $\mathbf{S}=\mathbf{e}_\alpha=\frac{\partial}{\partial x^\alpha}$ to be the basis vector field (whose components are constants), and let $\mathbf{R}=\mathbf{e}_\beta=\frac{\partial}{\partial x^\beta}$. We can now expand $D_\mathbf{R}\mathbf{S}$ in the basis and the coefficients of expansion, $\Gamma^\rho_{\beta\alpha}$ is given below. This relation between $D_\mathbf{R}\mathbf{S}$ and $\Gamma^\rho_{\beta\alpha}$ is established here.

$$D_{\frac{\partial}{\partial x^\beta}}\left(\frac{\partial}{\partial x^\alpha}\right) = \Gamma^\rho_{\beta\alpha}\frac{\partial}{\partial x^\rho}. \qquad (1.76)$$

The *transformation formula* can now be derived using the above definition.

$$
\begin{aligned}
D_{\frac{\partial}{\partial \bar{x}^\beta}}\left(\tfrac{\partial}{\partial \bar{x}^\alpha}\right) &= \bar{\Gamma}^\rho_{\beta\alpha}\tfrac{\partial}{\partial \bar{x}^\rho} \\
&= D_{\left(\frac{\partial x^\mu}{\partial \bar{x}^\beta}\right)\frac{\partial}{\partial x^\mu}}\left(\tfrac{\partial x^\gamma}{\partial \bar{x}^\alpha}\tfrac{\partial}{\partial x^\gamma}\right) \\
&= \tfrac{\partial x^\mu}{\partial \bar{x}^\beta} D_{\frac{\partial}{\partial x^\mu}}\left(\tfrac{\partial x^\gamma}{\partial \bar{x}^\alpha}\tfrac{\partial}{\partial x^\gamma}\right) \qquad (1.77)\\
&= \tfrac{\partial x^\mu}{\partial \bar{x}^\beta}\left[\tfrac{\partial x^\gamma}{\partial \bar{x}^\alpha} D_{\frac{\partial}{\partial x^\mu}}\left(\tfrac{\partial}{\partial x^\gamma}\right) + \left(\tfrac{\partial}{\partial x^\mu}\tfrac{\partial x^\gamma}{\partial barx^\alpha}\right)\tfrac{\partial}{\partial x^\gamma}\right] \\
&= \left[\tfrac{\partial x^\mu}{\partial \bar{x}^\beta}\tfrac{\partial x^\gamma}{\partial \bar{x}^\alpha}\Gamma^\sigma_{\mu\gamma} + \left(\tfrac{\partial^2 x^\sigma}{\partial \bar{x}^\beta \partial \bar{x}^\alpha}\right)\right]\tfrac{\partial}{\partial x^\sigma},
\end{aligned}
$$

where the dummy index γ is replaced with σ. To make the comparison between the last line and the first line of the derivation, we need to manipulate the partial factor $\frac{\partial}{\partial x^\sigma}$ as follows

$$\frac{\partial}{\partial x^\sigma} = \frac{\partial \bar{x}^\rho}{\partial x^\sigma}\frac{\partial}{\partial \bar{x}^\rho}. \qquad (1.78)$$

Thus, we obtain

$$
\begin{aligned}
\bar{\Gamma}^\rho_{\beta\alpha}\tfrac{\partial}{\partial \bar{x}^\rho} &= \left[\tfrac{\partial x^\mu}{\partial \bar{x}^\beta}\tfrac{\partial x^\gamma}{\partial \bar{x}^\alpha}\Gamma^\sigma_{\mu\gamma} + \left(\tfrac{\partial^2 x^\sigma}{\partial \bar{x}^\beta \partial \bar{x}^\alpha}\right)\right]\tfrac{\partial \bar{x}^\rho}{\partial x^\sigma}\tfrac{\partial}{\partial \bar{x}^\rho} \\
\bar{\Gamma}^\rho_{\beta\alpha}(\bar{x}) &= \tfrac{\partial \bar{x}^\rho}{\partial x^\sigma}\tfrac{\partial x^\mu}{\partial \bar{x}^\beta}\tfrac{\partial x^\gamma}{\partial \bar{x}^\alpha}\Gamma^\sigma_{\mu\gamma}(x) + \tfrac{\partial^2 x^\sigma}{\partial \bar{x}^\beta \partial \bar{x}^\alpha}\tfrac{\partial \bar{x}^\rho}{\partial x^\sigma}.
\end{aligned} \qquad (1.79)
$$

The three index notation seems to suggest that the Christoffel symbol is a tensor of rank three. This however is not true, the proof is in the extra term that appears in the transformation above. Due to this very same confusion Christoffel symbols, in older notations, were written as $\left\{ {\rho \atop \beta\alpha} \right\}$ instead of $\Gamma^\rho_{\beta\alpha}$.

1.3.20 Torsion Tensor

The *torsion tensor* $T^\rho_{\alpha\beta}$ is a third-rank tensor, antisymmetric in the first two indices and with 24 independent components, i.e.,

$$T^\rho_{\alpha\beta} \equiv \Gamma^\rho_{[\alpha\beta]}. \qquad (1.80)$$

Consider the transformation of the Christoffel symbol again, considering the antisymmetric part of the transformation we can show that the torsion tensor does transform like a third-rank tensor. This bring us to a very important conclusion, that the torsion tensor cannot be eliminated locally due the reason that if a tensor vanishes at a particular point then it vanishes everywhere.

$$
\begin{aligned}
\Gamma^\rho_{\alpha\beta}(x) &= \tfrac{\partial x^\sigma}{\partial \bar{x}^\rho}\tfrac{\partial \bar{x}^\beta}{\partial x^\mu}\tfrac{\partial \bar{x}^\alpha}{\partial x^\gamma}\bar{\Gamma}^\sigma_{\mu\gamma} + \tfrac{\partial^2 x^\sigma}{\partial x^\beta \partial x^\alpha}\tfrac{\partial x^\sigma}{\partial \bar{x}^\rho} \\
\Gamma^\rho_{[\alpha\beta]} &= T^\rho_{\alpha\beta} = \bar{T}^\sigma_{\mu\gamma}\tfrac{\partial x^\rho}{\partial \bar{x}^\sigma}\tfrac{\partial \bar{x}^\mu}{\partial x^\alpha}\tfrac{\partial \bar{x}^\gamma}{\partial x^\beta}.
\end{aligned} \qquad (1.81)
$$

Let $\mathbf{R}$ and $\mathbf{S}$ be vectors. Without torsion $[\mathbf{R}, \mathbf{S}]$ and $D_{\mathbf{R}}\mathbf{S} - D_{\mathbf{S}}\mathbf{R}$ represent the same vector, i.e., $[\mathbf{R}, \mathbf{S}] = D_{\mathbf{R}}\mathbf{S} - D_{\mathbf{S}}\mathbf{R}$ & $D_{\mathbf{R}}\mathbf{S} - D_{\mathbf{S}}\mathbf{R} - [\mathbf{R}, \mathbf{S}] = 0$ represents a closed loop, and in the presence of torsion, there is no closure of loop (see figure 1.11). This is shown below

$$\mathbf{R}^{\alpha} D_{\alpha} \mathbf{S}^{\beta} - \mathbf{S}^{\alpha} D_{\alpha} \mathbf{R}^{\beta} - [\mathbf{R}, \mathbf{S}]^{\beta} = T^{\beta}_{\alpha\gamma} R^{\alpha} S^{\gamma}. \tag{1.82}$$

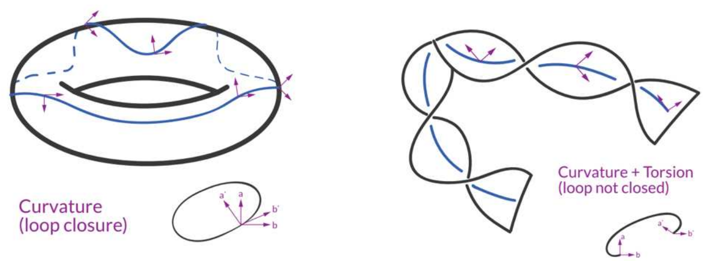


Fig. 1.11. Left:When there is curvature alone, we observe loop closure. Right: When torsion is introduced, there is no loop closure.

This describes the geometrical meaning of torsion, which is that torsion represents the failure of the loop to close. For all future calculations we shall assume a torsion-free connections, i.e., $\mathbf{T} = 0$.

1.4 Lie Algebra

1.4.1 Lie Bracket

Another useful operation is the *Lie bracket*, defined as

$$[X, Y](f) = X(Y(f)) - Y(X(f)). \tag{1.83}$$

We need to check that this Lie bracket does define a new vector field. One possible way is the use of local coordinates as follows

$$\begin{aligned}
[X,Y](f) &= X^{\mu}\partial_{\mu}\left(Y^{\nu}\partial_{\nu}f\right) - Y^{\mu}\partial_{\mu}\left(X^{\nu}\partial_{\nu}f\right) \\
&= X^{\mu}\left(\partial_{\mu}\left(Y^{\nu}\right)\partial_{\nu}f + Y^{\nu}\partial_{\mu}\partial_{\nu}f\right) - Y^{\mu}\left(\partial_{\mu}\left(X^{\nu}\right)\partial_{\nu}f + X^{\nu}\partial_{\mu}\partial_{\nu}f\right) \\
&= \left(X^{\mu}\partial_{\mu}Y^{\nu} - Y^{\mu}\partial_{\mu}X^{\nu}\right)\partial_{\nu}f + \underbrace{X^{\mu}Y^{\nu}\partial_{\mu}\partial_{\nu}f - Y^{\mu}X^{\nu}\partial_{\mu}\partial_{\nu}f}_{=X^{\mu}Y^{\nu}\underbrace{\left(\partial_{\mu}\partial_{\nu}f - \partial_{\nu}\partial_{\mu}f\right)}_{=0}} \\
&= \left(X^{\mu}\partial_{\mu}Y^{\nu} - Y^{\mu}\partial_{\mu}X^{\nu}\right)\partial_{\nu}f,
\end{aligned} \tag{1.84}$$

which is indeed a homogeneous first order differential operator. Here we have used the symmetry of the matrix of second derivatives of twice differentiable functions. Also, note that the last line of the equation gives an explicit coordinate expression for the commutator of two differentiable vector fields.

1.4.2 Lie Derivative

Given a vector field X, the lie derivative $\mathcal{L}_X$ is an operation on tensor fields. For a function f, we set

$$\mathcal{L}_X f = X(f), \tag{1.85}$$

and for a vector field Y, the Lie derivative coincides with the Lie bracket, i.e.,

$$\mathcal{L}_X Y = [X,Y]. \tag{1.86}$$

In terms of the carroting operation defined previously, if α is a 1-form and Y is a vector, then we define $\mathcal{L}_X$ to be that 1-form satisfying the following relation

$$\langle \mathcal{L}_X \alpha, Y\rangle = X\left[\langle \alpha, Y\rangle\right] - \langle \alpha, [X,Y]\rangle. \tag{1.87}$$

For a 1-form α, the Lie derivative is defined[11] as follows

$$\left(\mathcal{L}_X \alpha\right)(Y) = \mathcal{L}_X\left(\alpha(Y)\right) - \alpha\left(\mathcal{L}_X Y\right). \tag{1.88}$$

Let us check if the above equation transforms as a 1-form. Note that the RHS transforms in the desired way when Y is replaced with $Y_1 + Y_2$. Replacing Y with fY, where f is a function, we obtain the following

[11] note that this is just the Leibniz rule written the wrong-way round

$$
\begin{aligned}
(\mathcal{L}_X\alpha)(fY) &= \mathcal{L}_X(\alpha(fY)) - \alpha \underbrace{(\mathcal{L}_X fY)}_{=X(f)Y+f\mathcal{L}_XY} \\
&= X(f\alpha(Y)) - \alpha(X(f)Y + f\mathcal{L}_XY) \\
&= X(f)\alpha(Y) + fX(\alpha(Y)) - \alpha(X(f)Y) - \alpha(f\mathcal{L}_XY) \\
&= fX(\alpha(Y)) - f\alpha(\mathcal{L}_XY) \\
&= f((\mathcal{L}_X\alpha)(Y)).
\end{aligned}
\tag{1.89}
$$

Thus, $\mathcal{L}_X\alpha$ is a linear C^∞ map on vector fields (hence a covector field). In coordinate-components notation we have

$$
(\mathcal{L}_X\alpha)_\mu = X^\nu\partial_\nu\alpha_\mu + \alpha_\nu\partial_\mu X^\nu. \tag{1.90}
$$

For tensor products, the Lie derivative is defined by imposing linearity under addition together with the Leibniz rule as follows

$$
\mathcal{L}_X(\alpha\otimes\beta) = (\mathcal{L}_x\alpha)\otimes\beta + \alpha\otimes\mathcal{L}_X\beta. \tag{1.91}
$$

The Lie derivative along a vector field X is a differential operator that operates on tensor fields T converting them into tensors $\mathcal{L}_XT$. Since a tensor T is a is a sum of tensor products,

$$
T = T^{n_1\ldots n_p}_{m_1\ldots m_q}\partial_{n_1}\otimes\ldots\partial_{a_p}\otimes\ldots dx^{m_1}\otimes\ldots\otimes dx^{n_p}, \tag{1.92}
$$

requiring linearity with respect to addition of tensors gives thus a definition of Lie derivative for any tensor. Consider the following example where for a tensor T^n_m,

$$
\mathcal{L}_XT^n_m = X^a\partial_aT^n_m - T^a_m\partial_aX^n + T^n_a\partial_mX^a. \tag{1.93}
$$

Similarly, we have

$$
\begin{aligned}
\mathcal{L}_XR^{\mu\nu} &= X^\alpha\partial_\alpha R^{\mu\nu} - R^{\mu\alpha}\partial_\alpha X^\nu - R^{\nu\alpha}\partial_\alpha X^\mu, \\
\mathcal{L}_XW_{\mu\nu} &= X^\alpha\partial_\alpha W_{\mu\nu} + W_{\mu\alpha}\partial_\nu X^\alpha + W_{\nu\alpha}\partial_\mu X^\alpha.
\end{aligned}
\tag{1.94}
$$

These are all special cases for the more generalized formula of the Lie derivative, which is

$$\mathcal{L}_X T^{n_1\ldots n_p}_{m_1\ldots m_q} = X^a \partial_a T^{n_1\ldots n_p}_{m_1\ldots m_q} - T^{an_1\ldots n_p}_{m_1\ldots m_q} \partial_a X^{n_1} - \ldots - T^{n_1\ldots n_{p-1}a} \partial_a X^{n_p}$$
$$+ T^{n_1\ldots n_p}_{a\ldots m_q} \partial_{m_1} X^a + \ldots + T^{n_1\ldots n_p}_{m_1\ldots m_{q-1}a} \partial_{m_q} X^a \tag{1.95}$$

The following is a useful property of Lie derivatives

$$\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y] \,. \tag{1.96}$$

Applying this to a tensor **A**, we see

$$[\mathcal{L}_X, \mathcal{L}_Y] \, \mathbf{A} = \mathcal{L}_X \left(\mathcal{L}_Y \mathbf{A}\right) - \mathcal{L}_Y \left(\mathcal{L}_X \mathbf{A}\right) \tag{1.97}$$

1.4.3 The Geometric Approach to Lie Derivative

Consider a point p_0 on a manifold $\mathcal{M}$, every vector $Y \in T_{p_0}\mathcal{M}$ is tangent to some curve. To see this, let $\{y^i\}$ be an local coordinates near the point p_0, with $y^i(p) = y^i_0$, then Y can be written as $Y^i(p_0)\partial_i$. Now, if we set $\gamma^i(s) = y^i_0 + sY^i\,(p_0)$ (where $\gamma^i(s)$ is an arbitrary curve parameterized in terms of s between two points on the manifold), then $\dot{\gamma}^i(0) = Y^i\,(p_0)$ which establishes the claim. This observation shows that studies of vectors can be reduced to studies of curves. Let $\mathcal{M}$ and $\mathcal{N}$ be two manifolds and let $\phi : \mathcal{M} \to \mathcal{N}$ be a differentiable map between them. Given a vector $Y \in T_p\mathcal{M}$, the *push forward* $\phi_* Y$ of Y is a vector in $T_{\phi(p)}$ defined as follows (see figure 1.12)
let γ be any curve for which $Y = \dot{\gamma}(0)$, then

$$\phi_* Y = \left[\frac{d\,(\phi \circ \gamma)}{ds}\right]_{s=0} . \tag{1.98}$$

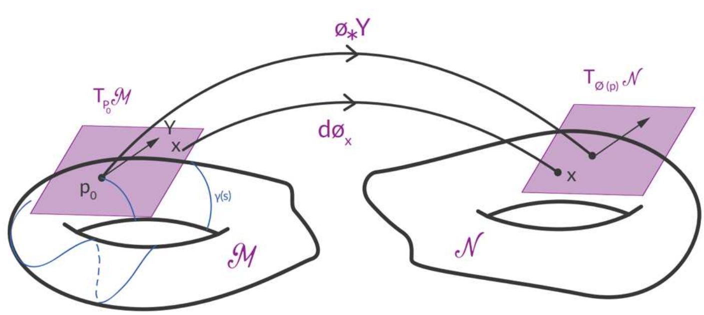


Fig. 1.12. If a map, ϕ, carries every point on manifold $\mathcal{M}$ to manifold $\mathcal{N}$ then the push forward of the map ϕ carries vectors in the tangent space at every point in $T\mathcal{M}$ to a tangent space at every point in $\mathcal{M}$.

In local coordinates x^j on N and y^i on $\mathcal{M}$, so that $\phi(x) = \left(\phi^j\left(y^i\right)\right)$, we find

$$(\phi_* Y)^j = \left[\frac{d\phi^j\left(\gamma^i(s)\right)}{ds}\right]_{s=0} = \left[\frac{\partial\phi^j\left(\gamma^i(s)\right)}{\partial y^i}\dot{\gamma}^i(s)\right]_{s=0} = \frac{\partial\phi^j\left(y^i\right)}{\partial y^i}Y^i. \quad (1.99)$$

Thus, this emphasizes that the definition is independent of the choice of the curve γ satisfying $Y = \dot{\gamma}(0)$. If this formula is applied to a vector field Y defined on $\mathcal{M}$ we get

$$(\phi_* Y)^j\left(\phi(y)\right) = \frac{\partial\phi^j}{\partial y^i}(y)Y^i(y). \quad (1.100)$$

The above equation demonstrates that if a point $x \in N$ has more than one pre-image, say $x = \phi(x_1) = \phi(x_2)$ with $x_1 \neq x_2$, then the equation will define more than one tangent vector at x in general. Thus, we can be certain that the push-forward of a vector field on $\mathcal{M}$ defines a vector field on $\mathcal{N}$ only when ϕ is a diffeomorphism. More generally, $\phi_* Y$ defines locally a vector field on $\phi(\mathcal{M})$ if and only if ϕ is a local diffeomorphism. In cases such as these, we invert ϕ and write the previous equation as follows

$$(\phi_* Y)^k(y) = \left(\frac{\partial\phi^j}{\partial y^i}\right)(y)\left(\phi^{-1}(x)\right). \quad (1.101)$$

When ϕ is understood as a coordinate change rather than a diffeomorphism between two manifolds, this is simply the standard transformation law of a vector field under coordinate transformations.
The push-forward operation can be extended to contravariant tensors by defining it on tensor products in the obvious way, and extending by linearity. Consider three vectors A, B, and C, then the push-forward operation is

$$\phi_*\left(A \otimes B \otimes C\right) = \phi_* A \otimes \phi_* B \otimes \phi_* C. \quad (1.102)$$

Consider a k-multilinear map ζ from $T_\phi(p_0)\mathcal{M}$ to $\mathbb{R}$. The *pull-back* $\phi^*\zeta$ of ζ is a multilinear map on $T_{p_0}\mathcal{M}$ defined as follows (see figure 1.13)

$$T_p M \ni (Y_1, ..., Y_k) \to \phi^*(\zeta)(Y_1, ..., Y_k) = \zeta\left(\phi^* Y_1, ..., \phi_* Y_k\right). \quad (1.103)$$

Let $\zeta = \zeta_\mu dx^\mu$ be a 1-form. If $Y = Y^\nu\partial_{nu}$ then

$$(\phi^*\zeta)(Y) = \zeta(\phi_* Y) = \zeta\left(\frac{\partial\phi^\beta}{\partial y^\alpha}Y^\alpha\partial_\beta\right) = \zeta_\beta\frac{\partial\phi^\beta}{\partial y^\alpha}Y^\alpha = \zeta_\beta\frac{\partial\phi^\beta}{\partial y^\alpha}dy^\alpha(Y). \quad (1.104)$$

Equivalently,

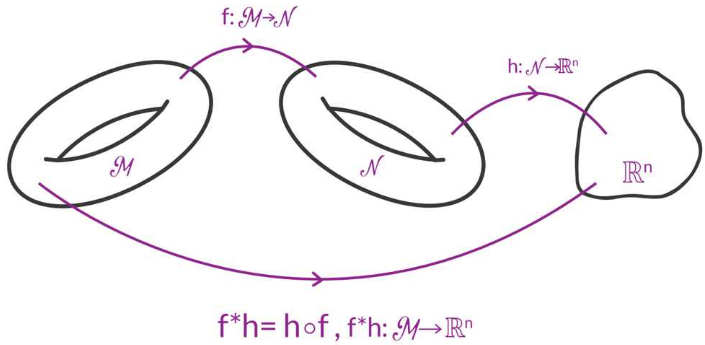


Fig. 1.13. Pull-back

$$\left(\phi^*\zeta\right)_\alpha = \zeta_\beta \frac{\partial \phi^\beta}{\partial y^\alpha}, \tag{1.105}$$

and if ζ is a 1-form field on $\mathcal{N}$, this reads the following

$$\left(\phi^*\zeta\right)_\alpha (y) = \zeta_\beta \left(\phi(y)\right) \frac{\partial \phi^\beta(y)}{\partial y^\alpha}. \tag{1.106}$$

It is to be noted here that $\phi^*\zeta$ is a field of 1-forms on M, irrespective of injective or surjective properties of ϕ[12]. For a function $\mathcal{R}$, the above equation reads the following

$$\left(\phi^* d\mathcal{F}\right)_\alpha (y) = \frac{\partial \mathcal{F}}{\partial x_\beta} \left(\phi(y)\right) \frac{\partial \phi^\beta(y)}{\partial y^\alpha} = \frac{\partial \left(\mathcal{F} \circ \phi\right)}{\partial y^\alpha}(y), \tag{1.107}$$

which is alternatively written as

$$\phi^* d\mathcal{F} = d\left(\mathcal{F} \circ \phi\right), \tag{1.108}$$

and using the notation

$$\phi^*\mathcal{F} = \mathcal{F} \circ \phi, \tag{1.109}$$

we can rewrite the alternate formulation of the equation for functions as

$$\phi^* d = d\phi^*. \tag{1.110}$$

In this context it is thus clearly of interest to consider diffeomorphisms ϕ, as

[12] Similarly, pull-backs of covariant tensor fields of higher rank are smooth tensor fields

then tensor products can now be transported in the following way; let $\hat{\phi}$ denote the associated map: Define $\hat{\phi}\mathcal{F} = \mathcal{F} \circ \phi$ for functions, $\hat{\phi} = \phi_*$ for covariant fields, and $\hat{\phi} = \left(\phi^{-1}\right)_*$ for contravariant tensor fields. We use the rule

$$\hat{\phi}\,(S \otimes R) = \hat{\phi}S \otimes \hat{\phi}R \tag{1.111}$$

for tensor products, and the definition is extended by linearity under multiplication by functions to any tensor fields. Thus, if Y is a vector field of 1-forms, we have

$$\hat{\phi}\,(Y \otimes \zeta) = \left(\phi^{-1}\right)_* Y \otimes \phi^* \zeta. \tag{1.112}$$

1.4.4 Isometries

Let $(\mathcal{M}, g)$ be a pseudo-Riemannian manifold. A map ξ is called an *isometry* if

$$\xi^* f = f, \tag{1.113}$$

where ξ^* is the pull-back defined in the previous section. The group $Iso(\mathcal{M}, g)$ of isometries of $(\mathcal{M}, g)$ carries a natural manifold structure; such groups are called *Lie groups*. It is to be noted that If $(\mathcal{M}, g)$ is Riemannian and compact, then $Iso(\mathcal{M}, g)$ is compact. Also, any element of the connected component of the identity of a Lie group $\mathcal{G}$ belongs to a one-parameter subgroup $\{\phi_q\}_{q \in R}$ of $\mathcal{G}$. This allows us to study actions of isometry groups by studying the generators of one-parameter subgroups, defined as

$$X(f)(x) = \left[\frac{d(f(\phi_q(x)))}{dq}\right]_{q=0}, \tag{1.114}$$

i.e.,

$$X = \left[\frac{d\phi_q}{dq}\right]_{q=0}. \tag{1.115}$$

These vector fields X is called *Killing vectors*. The knowledge of Killing vectors provides considerable amount of information on the isometry group. In General Relativity, it is of key importance that the dimension of the isometry group of $(\mathcal{M}, g)$ equals the dimension of the space of the Killing vectors.

1.4.5 Flows of Vector Fields

Let X be a vector field on M. For every $q_0 \in M$ consider the solution

$$\frac{dx^\alpha}{dt} = X^\alpha(x(t)), \quad x^\alpha(0) = x_0^\alpha. \tag{1.116}$$

There always exists a maximal interval I containing the origin on which the above equation has a solution. Both the interval and the solution are unique. This will always be the solution $I \ni t \mapsto x(t)$. The map

$$(t, x_0) \mapsto \phi_t[X]\,(x_0) = x(t), \tag{1.117}$$

where x^α is the solution of the equation [13], is called the *local flow* of X. We say that X generates $\phi_t[X]$. We will write ϕ_t for $\phi_t[X]$ when X is unambiguous in the context. X is called *complete* if $\phi_t[X](q)$ is defined for all $(t, q) \in R \times M$. The following properties are presented sans proof:
a. ϕ_0 is the identity map,
b. $\phi_t \circ \phi_r = \phi_{t+r}$,
c. The maps $x \mapsto \phi_t(x)$ are local diffeomorphisms; global if for all $x \in M$ the maps ϕ_t are defined for all $t \in R$,
d. ϕ_{-t} is generated by $-X$: $\phi_{-t}[X] = \phi_t[-X]$.
A family of diffeomorphisms satisfying property b above is called a 1-parameter group of diffeomorphisms.

1.4.6 Killing Vectors

Let ϕ_q be a 1-parameter group of isometries of $(\mathcal{M}, g)$, hence

$$\phi_q^* f = f \Longrightarrow \mathcal{L}_X f = 0. \tag{1.118}$$

Now, for the metric tensor, $g_{\alpha\beta}$,

$$\mathcal{L}_X g_{\alpha\beta} = X^\mu g_{\alpha\beta,\mu} + g_{\mu\beta} X^\mu_{,\alpha} + g_{\alpha\mu} X^\mu_{,\beta} \tag{1.119}$$

In a coordinate system where the partial derivatives of the metric vanish at a point p, the RHS equals $\nabla_\alpha X_\beta + \nabla_\alpha X_\beta$. But notice that the LHS is a tensor field, and two tensor fields equal in one coordinate system coincide in all coordinate systems. Thus, we have proved that generators of isometries satisfy the following equation

$$\nabla_\beta X_\alpha + \nabla_\alpha X_\beta = 0 \Longrightarrow \nabla_{(\alpha} X_{\beta)} = 0. \tag{1.120}$$

This can also be shown explicitly by carrying out a short calculation in which we substitute total derivatives in the RHS of the last two terms in 5.11 and obtain the following result

[13] the interval of existence of solutions depends upon x_0 in general

$$
\begin{aligned}
\mathcal{L}_X g_{\alpha\beta} &= X^\mu g_{\alpha\beta,\mu} + \left(\partial_\alpha\left(g_{\mu\beta}X^\mu\right) - X^\mu g_{\mu\beta,\alpha}\right) + \left(\partial_\beta\left(g_{\alpha\mu}X^\mu\right) - X^\mu g_{\alpha\mu,\beta}\right) \\
&= X^\mu g_{\alpha\beta,\mu} + X_{\beta,\alpha} - X^\mu g_{\mu\beta,\alpha} + X_{\alpha,\beta} - X^\mu g_{\alpha\mu,\beta} \\
&= 2\partial_{(\alpha}X_{\beta)} - X^\mu\left(g_{\alpha\mu,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu}\right) \\
&= 2\partial_{(\alpha}X_{\beta)} - 2\underbrace{\Gamma_{\alpha\beta\mu}}_{=g_{\mu\mu}\Gamma^\mu_{\alpha\beta}}X^\mu \\
&= 2\nabla_{(\alpha}X_{\beta)} = 0.
\end{aligned} \tag{1.121}
$$

From the calculation just carried out, the Lie derivative of the metric with respect to X vanishes. This means that the local flow of X preserves the metric. In other words, X generates local isometries of f. To make sure that X generates a 1-parameter group of isometries we need to make sure that X is complete; this requires separate considerations. We know that $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$, which implies that the commutator of two Killing vector fields is a Killing vector field, i.e.,

$$
\mathcal{L}_{[X,Y]}f = \mathcal{L}_X\underbrace{(\mathcal{L}_Y f)}_{=0} - \mathcal{L}_Y\underbrace{(\mathcal{L}_X f)}_{=0} = 0. \tag{1.122}
$$

We shall be revisiting these topics in explicit mathematical detail in the exercises and examples.

Example 1.9. In this example, we set out to show that if X, Y, and Z are vector fields, then the commutator of Lie derivatives of Z is given as

$$
[\mathcal{L}_X, \mathcal{L}_Y]\, Z = [[X, Y], Z]. \tag{1.123}
$$

This is a straightforward application of the Jacobi identity 1.126 (proof of which is an exercise). Expanding the commutator and using the definition of the Lie derivative we obtain

$$
\begin{aligned}
[\mathcal{L}_X, \mathcal{L}_Y]\, Z &= \mathcal{L}_X\left(\mathcal{L}_Y Z\right) - \mathcal{L}_Y\left(\mathcal{L}_X Z\right) \\
&= \mathcal{L}_X\left([Y, Z]\right) - \mathcal{L}_Y\left([X, Z]\right) \\
&= [X, [Y, Z]] - [Y, [X, Z]].
\end{aligned} \tag{1.124}
$$

Finally we bring the Jacobi identity to the above form and compare the equations to prove the identity a follows

$$
\begin{aligned}
&[X, [Y, Z]] - [Y, [X, Z]] + [Z, [X, Y]] = 0, \\
&\Rightarrow [X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z]
\end{aligned} \tag{1.125}
$$

Exercise 2

1. Show that for any three vector fields X, Y, and Z the following relation holds known as the Jacobi identity holds

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0. \tag{1.126}$$

2. Show that the Lie derivative of ω in a coordinate basis $\mathcal{L}_X\omega_\mu = X^\nu\partial_\nu\omega_\mu + \omega_\nu\partial_\mu X^\nu$ can be written in terms of the covariant derivative as follows: $\mathcal{L}_X\omega_\mu = X^\nu D_\nu\omega_\mu + \omega_\nu D_\mu X^\nu$.

3. Show that the commutator of Lie derivatives is the Lie derivative of the commutator as shown in 1.96.

1.5 The Three Types of Vectors

1.5.1 Non-Degeneracy of a Metric

A metric is said to be *non-degenerate* at a point p on a manifold $\mathcal{M}$ if there exists no non-zero vector $\mathbf{R} \in T_p(\mathcal{M})$ such that $g(\mathbf{R},\mathbf{Q}) = 0$ for all vectors $\mathbf{Q} \in T_p(\mathcal{M})$. We can now define a new metric tensor of rank $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ with components $g^{\mu\nu}$ with respect to a basis $\{x_\mu\}$ which is dual to the basis $\{x^\mu\}$, by the following expression

$$g^{\mu\nu}g_{\nu\xi} = \delta^\mu_\xi, \tag{1.127}$$

the matrix $g^{\mu\nu}$ is the inverse of the matrix $g_{\mu\nu}$, and these tensors can be used to provide an *isomorphism* between any contravariant tensor and a covariant one, i.e., to raise and lower indices. If $\mathbf{S}^{\mu\nu}$ are the components of a contravariant tensor, then we can lower its indices by making use of a metric tensors, and can also obtain mixed tensors as follows

$$\begin{aligned} S_{\mu\nu} &= g_{\mu\xi}g_{\nu\chi}S^{\xi\chi} \\ S^\mu_\nu &= g^{\mu\xi}S_{\xi\nu} \\ S^\nu_\mu &= g^{\nu\chi}S_{\chi\mu}. \end{aligned} \tag{1.128}$$

1.5.2 Timelike, Spacelike and Lightlike Vectors

Consider a Lorentzian metric g on a manifold $\mathcal{M}$ equipped with some topology. At a point p on the manifold, the non-zero vectors can be divided into three

Type	Condition
Timelike	$g(\mathbf{R}, \mathbf{R}) < 0$
Lightlike	$g(\mathbf{R}, \mathbf{R}) = 0$
Spacelike	$g(\mathbf{R}, \mathbf{R}) > 0$

Table 1.2. All the conditions are mentioned for a vector $\mathbf{R} \in T_p$

classes as given in table 1.2 (also see figure 1.14).

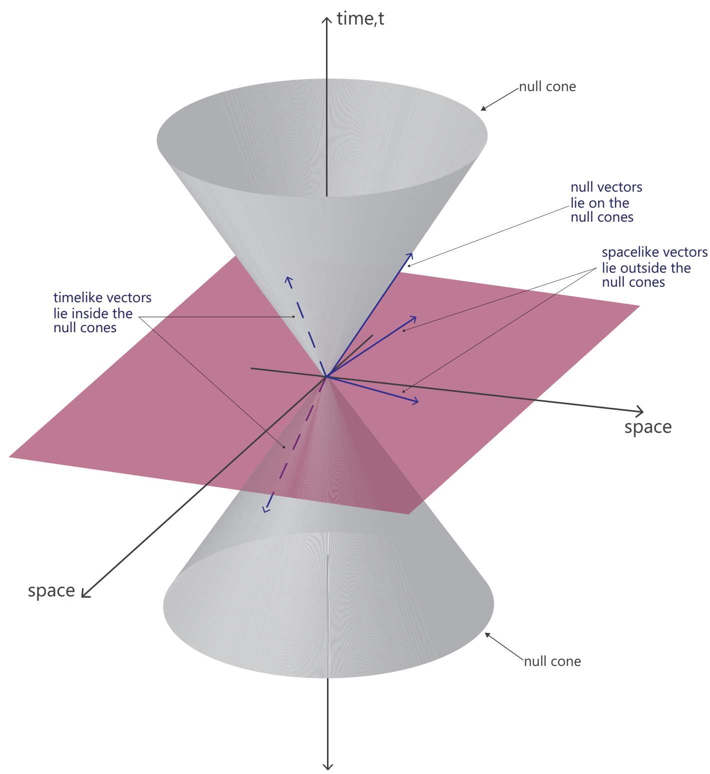


Fig. 1.14. This figure represents the null cones as defined by the Lorentz metric.

1.5.3 Null Cones

A vector $\mathbf{R} \in T_p$ is called *causal* if

$$g(\mathbf{R}, \mathbf{R}) \leq 0 \tag{1.129}$$

At each point p on a Lorentzian manifold $\mathcal{M}$, we can define a double cone C_p in the tangent space $T_p(\mathcal{M})$. This is called the *causal cone* and is expressed in terms of the following inequality

$$g(\mathbf{R}, \mathbf{R}) \leq 0,\ \ \mathbf{R} \in T_p(\mathcal{M}). \tag{1.130}$$

The boundary of the causal cone is called the *double cone*. The boundary is formed by the null or lightlike vectors in the tangent space of the Lorentzian manifold and this separates the timelike and the spacelike vectors.

1.5.4 Double Null Cones

If the vector $\mathbf{R}$ is either timelike or null, then it is called *causal* and if the vector $\mathbf{R} = \left(t, x^1, x^2, x^3\right)$ 3) is a null vector at a point q, then

$$t^2 = \left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2, \tag{1.131}$$

and hence $\mathbf{R}$ on cone with vertex at q, i.e., all null vectors at point q span a double cone, known as the *double null cone*.

1.6 Causality

For each point $p \in \mathcal{M}$, the linear space $(T_p\mathcal{M}, g)$ is isometric to the Minkowski spacetime $\mathbb{R}^{3+1}$ (three spacelike coordinates and one timelike coordinate) and hence there exists a basis (e_0, e_1, e_2, e_3) of the tangent space of $\mathcal{M}$, i.e., $T_p\mathcal{M}$ such that

$$g\left(e_\alpha, e_\beta\right) = M_{\alpha\beta}, \tag{1.132}$$

where $M_{\alpha\beta}$ is the Minkowski diagonal matrix $(-1, 1, 1, 1)$. Then, for any vector $\mathbf{R} \in T_p\mathcal{M}$ we have $X = \sum_\alpha X^\alpha e_\alpha$ and thus

$$g(\mathbf{R}, \mathbf{R}) = -\left(\mathbf{R}^0\right)^2 + \left(\mathbf{R}^1\right)^2 + \left(\mathbf{R}^2\right)^2 + \left(\mathbf{R}^3\right)^2. \tag{1.133}$$

If $\mathbf{R}$ is either timelike or null it is called causal. If $\mathbf{R}$ is null vector at p, then

$$\left(\mathbf{R}^0\right)^2 = \left(\mathbf{R}^1\right)^2 + \left(\mathbf{R}^2\right)^2 + \left(\mathbf{R}^3\right)^2, \tag{1.134}$$

and hence $\mathbf{R}$ lies on cone with vertex at p. It is to be noted that position of the vector depends upon the basis e_α. Thus, all null vectors at point p space a double cone, known as the *double null cone*. Let $\mathcal{N}_p$ denote the set of all null vectors in the tangent space of $\mathcal{M}$, then

$$\mathcal{N}_p = \{\mathbf{R} \in T_p\mathbb{R}^{3+1} : g(\mathbf{R}, \mathbf{R}) = 0\}, \tag{1.135}$$

let $\mathcal{I}_p$ denote the set of all timelike vectors in the tangent space of $\mathcal{M}$, then

$$\mathcal{I}_p = \{\mathbf{R} \in T_p\mathbb{R}^{3+1} : g(\mathbf{R}, \mathbf{R}) < 0\}, \tag{1.136}$$

and let $\mathcal{S}_p$ denote the set of all spacelike vectors in the tangent space of M, then

$$\mathcal{S}_p = \{\mathbf{R} \in T_p\mathbb{R}^{3+1} : g(\mathbf{R}, \mathbf{R}) > 0\}. \tag{1.137}$$

1.7 Geodesic equation

1.7.1 Introduction

A *geodesic* is defined as a spacetime curve that is the shortest distance between two points, straight and uniformly parametrized, or in other words, it is a curve whose distance between two points is stationary. Mathematically, a geodesic is a curve $P(\zeta)$ that parallel-transports its tangent vector, say $\mathbf{R} = \frac{dP}{d\zeta}$, along itself.

$$D_{\mathbf{R}}\mathbf{R} = 0. \tag{1.138}$$

In a local coordinate system $x^\eta[P(\zeta)]$, in which the tangent vector takes the form $R^\eta = \frac{dx^\eta}{d\zeta}$, the geodesic is expressed as follows

$$\frac{D\left(\frac{dx^\eta}{d\zeta}\right)}{d\zeta} = 0 = \frac{d\left(\frac{dx^\eta}{d\zeta}\right)}{d\zeta} + \left[\Gamma^\eta_{\alpha\beta}\frac{dx^\alpha}{d\zeta}\right]\frac{dx^\beta}{d\zeta}, \tag{1.139}$$

simplifying this gives us the *geodesic equation*,

$$\frac{d^2x^\eta}{d\zeta^2} + \Gamma^\eta_{\alpha\beta}\frac{dx^\alpha}{d\zeta}\frac{dx^\beta}{d\zeta} = 0. \tag{1.140}$$

1.7.2 Affine Parameter

If a geodesic is timelike,
1. It is a possible curve (or trajectory) for a freely falling observer, and
2. there exists a parameter ζ (called the *affine parameter*) which is a multiple of the observer's propertime, $\zeta = m\tau + c$

1.7.3 The Deeper Meaning: Part One

Let us try and reveal the deeper meaning that hides in plain sight. In normal coordinates, in a local frame (for a local observer), we know that the following conditions are satisfied: $g_{\mu\nu,\alpha} = 0$ and $\Gamma^{\mu}_{\nu\beta} = 0$. The the 4-velocity, which is the tangent vector of a timelike curve, is defined as $\mathbf{u} = \frac{dx^{\eta}}{d\tau}\mathbf{e}_{\eta}|_{\eta=0} = \frac{dx^{0}}{d\tau}\mathbf{e}_{0} = \mathbf{e}_{0}$. This is so because $\mathbf{u}$ and $\mathbf{e}_0$ both have unit length. Since the 4-velocity is constant, the 4-acceleration is zero, i.e.,

$$a = D_{\mathbf{u}}\mathbf{u} = D_0\mathbf{e}_0 = 0. \tag{1.141}$$

This equation is nothing but the previously defined geodesic equation. Comparing the equations we conclude the following

$$a = D_{\mathbf{u}}\mathbf{u} = D_0\mathbf{e}_0 = \Gamma^{\eta}_{00}\mathbf{e}_{\eta} = 0. \tag{1.142}$$

What this implies is that a freely falling observer experiences zero 4-acceleration, i.e., the observer moves along a geodesic with affine parameter equal to the observer's propertime. The geodesic equation for this observer in local coordinate is as follows

$$\frac{d^2x^{\eta}}{d\tau^2} + \Gamma^{\eta}_{\alpha\beta}\frac{dx^{\alpha}}{d\tau}\frac{dx^{\beta}}{d\tau} = 0. \tag{1.143}$$

Thus, we have proven that the observer's trajectory is a straight line![14]

1.7.4 The Deeper Meaning: Part Two

For the second reveal consider a particular spacelike coordinate, say $x^{\eta} = y$. Let the observer move slowly (i.e., at non-relativistic speeds), this assumption enables us to replace propertime with just time. Let α and β (indices present in the geodesic equation) be timelike components, this would mean that[15] $\frac{dx^{\alpha}}{d\tau} = \frac{dx^{\beta}}{d\tau} = \frac{dt}{d\tau} = 1$, and $\Gamma^{\eta}_{\alpha\beta} = \Gamma^{\eta}_{tt}$. Making all these changes in the geodesic

[14] more generally, freely falling particles move on straight lines

[15] although time and propertime are distinct conceptually, they are dimensionally the same quantity and thus cancel

equation, we obtain the following expression

$$\frac{d^2y}{dt^2} = -\Gamma^{\eta}_{tt}, \tag{1.144}$$

which looks freakishly like the expression for gravitational force. Well, hold on, let's dig a bit deeper.

At large distances from the spherically symmetric gravitating object, spacetime is flat. Why? This is due to the fact that the influence of the gravitational field vanishes at large distances as it varies as $\approx r^{-2}$. Due to this, the ability of the tidal forces to curve spacetime at large distances from the gravitating object fades away thus resulting in a flat space. Note that when we talk about the gravitating object, we have assumed that there is no matter in the surroundings of our object, thus the vacuum field. We can represent this flat spacetime in terms of the line element as

$$ds^2 = dt^2 - \frac{1}{c^2}\left(dx^2 + dy^2 + dz^2\right). \tag{1.145}$$

Now, how would this metric change in the vicinity of our object? For this let us first write the basic form of the metric for a plane in polar coordinates. In flat space, the spatial distance between two points on a plane in polar coordinates is given by the following equation

$$ds^2 = r^2 d\theta^2 + dr^2. \tag{1.146}$$

now, let us modify this metric. We first start by making the replacements, $sin\theta \to sinh(\omega)$, and $cos\theta \to cosh(\omega)$. Here, ω is the angle with which the hyperbola increases with respect to the origin (a timelike coordinate). Thus, we have changed from polar coordinates to *hyperbolic coordinates*. In this frame, the acceleration along a particular hyperbola is the same, however, the acceleration along different hyperbolae are different. An analogous relation can be drawn to that of circular motion here, similar to the acceleration remaining the same along a particular hyperbola, the acceleration of a particle moving around a circle is uniform, however, the acceleration around another concentric circle of a different radius who definitely not be the same. We now make the transforms

$$X = rcosh(\omega),\ T = rsinh(\omega), \tag{1.147}$$

such that

$$X^2 - T^2 = r^2\left[cosh^2\omega - sinh^2\omega\right] = r^2. \tag{1.148}$$

Hence, producing the metric

$$ds^2 = r^2 d\omega^2 - dr^2. \tag{1.149}$$

This is the metric in which our gravitating object lies. Let us travel along a particular hyperbola and to determine the fate of the metric. Let the gravitating object, under consideration, be the super massive black hole located at the centre of our galaxy. Now, let us remove all the matter present outside this black hole and for the moment assume that the value of the energy density of the vacuum of space is zero[16]. By performing these actions, we have established the vacuum conditions. From a small distance from a hyperbola that is present next to where Earth was, just a moment ago, we compute the metric. The black hole is almost 26, 000 light years away from Earth. Placing the origin at the centre of the black hole, we re-define the position vector to be

$$r = R_{BH\to Hyp} + r', \tag{1.150}$$

where $R_{BH\to Hyp} = 26,000$ light years, is the distance between the black hole and the hyperbola trajectory which runs next to where Earth was a moment ago, and r' is the distance between the hyperbola and us. This distance is prone to vary since we are nothing but mere particles floating in space but would never exceed that of $R_{BH\to Hyp}$. Hence, $\frac{r'^2}{R^2_{BH\to Hyp}} \to 0$. Let us substitute this new relation into the metric and perform some manipulations.

$$\begin{aligned} ds^2 &= \left(R_{BH\to Hyp} + r'\right)^2 d\omega^2 - \left[d\left(R_{BH\to Hyp} + r'\right)\right]^2 \\ ds^2 &= \left(R^2_{BH\to Hyp} + r'^2 + 2R_{BH\to Hyp} r'\right) d\omega^2 - \left[dr'\right]^2 \\ ds^2 &\approx \left(1 + \frac{2r'}{R_{BH\to Hyp}}\right) R^2_{BH\to Hyp} d\omega^2 - dr'^2. \end{aligned} \tag{1.151}$$

We know that proper acceleration A, when the speed of light is set to unity is nothing but[17] $\frac{1}{R}$. Hence, here, $A = \frac{1}{R_{BH\to Hyp}} = \mathbf{g}$ (this $\mathbf{g}$ here is the acceleration due to gravity, not to be confused with the metric tensor). Define $R_{BH\to Hyp}\omega = t$, such that

$$ds^2 = (1 + 2r'\mathbf{g})\, dt^2 - dr'^2 = \left(1 + 2\frac{r'}{R}\right) dt^2 - dr'^2. \tag{1.152}$$

Now, we know that the Christoffel symbol is nothing but a combination of

[16] these conditions correspond to the *Ricci flatness condition* in which we assume the stress-energy tensor and the cosmological constant to have a null value.

[17] $A = \frac{c}{T} = \frac{c}{R\times c}\big|_{c\equiv 1} = \frac{1}{R}$

the partial derivatives of the metric tensor, with the assumptions made in the initial part of this section, it takes the following form

$$\Gamma^{\eta}_{tt} = \frac{1}{2} g^{\eta\gamma} \left(g_{t\gamma,t} + g_{\gamma t,t} - g_{tt,\gamma} \right). \tag{1.153}$$

From the metric (for the motion along the y-coordinate), $ds^2 = Adt^2 - dy^2$, $g^{y\gamma} = g^{yy} = 1$ (since η and γ assume a spacelike coordinate). Thus,

$$\Gamma^{y}_{tt} = \frac{1}{2} \left(g_{ty,t} + g_{yt,t} - g_{tt,y} \right) = -\frac{1}{2} \frac{\partial g_{tt}}{\partial y}. \tag{1.154}$$

Substituting this in the geodesic equation, we obtain an expression that in fact suggests the Christoffel symbol Γ^{y}_{tt} to be the force and g_{tt} to be the potential.

$$\frac{d^2 y}{dt^2} = -\Gamma^{y}_{tt} = \frac{1}{2} \frac{\partial g_{tt}}{\partial y}. \tag{1.155}$$

Replacing r' with y in the metric we obtained previously, we make an observation that confirms the statement made about the Christoffel symbol.

$$\begin{aligned} ds^2 &= \underbrace{(1 + 2y\mathbf{g})}_{=-g_{tt}} dt^2 - dy^2 \\ -g_{tt} &= (1 + 2y\mathbf{g}) \quad \rightarrow \quad \tfrac{\partial g_{tt}}{\partial y} = -2\mathbf{g} \\ \tfrac{d^2 y}{dt^2} &= -\Gamma^{y}_{tt} = -\mathbf{g}. \end{aligned} \tag{1.156}$$

Thus, the equation of motion of a geodesic in an accelerated coordinate frame (for non-relativistic speeds) is nothing but Newton's equation in an uniform gravitational field.

1.8 Curvature

A straight line has zero curvature and a circle of radius ρ has a curvature $\frac{1}{\rho}$. Well, why? A more important question is not why but how, how is the curvature defined for surfaces and based on what is it defined. The curvature of a curve is defined by how swiftly its unit normal vector $\mathbf{n}$ evolves as we move along the curve. For a small distance traversed on a circle, say ds, the infinitesimal change in the unit normal $|d\mathbf{n}|$ is equal to the angle $\frac{ds}{\rho}$. Similarly, the curvature of a straight line is zero because the normals are all parallel to each other and do not evolve with time. Hence, curvature is measured by the ratio of the infinitesimal change in the unit normal $|d\mathbf{n}|$ to the infinitesimal distance traversed by a point ds.

1.8.1 Gauss Curvature and Geodesic Deviation

Let $\mathcal{S}$ be a 2-dimensional surface with $\mathbf{n}(\xi)$ as unit normal. Let the local coordinates on this surface is denoted by $x(\xi) = (\xi^2, \xi^2)$ and the infinitesimal changes as a point traverses a distance be $d\xi^1, d\xi^2$. The infinitesimal distance traversed is then given by the expression given below which is tangential to the 2-dimensional surface

$$dx = \frac{\partial x}{\partial \xi^1} d\xi^1 + \frac{\partial x}{\partial \xi^2} d\xi^2 \equiv x_{,\alpha} d\xi^\alpha . \tag{1.157}$$

The unit normal is a vector that does not evolve as the point traverses because it is constant length $\mathbf{n} \cdot \mathbf{n} = 0$. This implies that the infinitesimal changes in it due to the parameters ξ are orthogonal to it. i.e., $d\mathbf{n} \cdot \mathbf{n} = 2\mathbf{n}d\mathbf{n} = 0$. Hence, we can conclude that $d\mathbf{n} = \mathbf{n}_{,\alpha} d\xi^\alpha$ is tangential to the surface. When the tangential vectors $d\mathbf{n}_1$ and $d\mathbf{n}_2$ are expanded in the basis vector we obtain the following expression

$$\mathbf{n}_{,\alpha} = L_\alpha^\beta x_{,\beta} . \tag{1.158}$$

The coefficients L_α^β define a mapping of the tangent vector dx into $d\mathbf{n}$, another tangent vector. The matrix L_α^β is called the *Weingarten matrix* and the mapping is called *Weingarten mapping*. This was the idea of curvature given by Carl Friedrich Gauss, which states that the curvature of a surface at a point is measured by the ratio of the area spanned by the infinitesimal components of the normal vector $d\mathbf{n}_1, d\mathbf{n}_2$, to the area spanned by the tangent vectors $dx_1 = x_{,1} d\xi^1$ and $dx_2 = x_{,2} d\xi^2$ on the surface.

Let $\mathbf{R}$ and $\mathbf{S}$ be two vectors, the area of the parallelogram spanned by these vectors is given as $|\mathbf{R} \times \mathbf{S}|$. The curvature is measured as follows

$$\begin{aligned} d\mathbf{n}_1 \times d\mathbf{n}_2 &= (\mathbf{n}_{,1} \times \mathbf{n}_{,2})\, d\xi^1 d\xi^2 \\ &= \left(L_1^1 x_1 + L_1^2 x_{,2}\right) \times \left(L_2^1 x_1 + L_2^2 x_{,2}\right) d\xi^1 d\xi^2 \\ &= \left(L_1^1 L_2^2 - L_1 L^2 L_2^1\right)(dx_1 \times dx_2) = |L|\,(dx_1 \times dx_2)\,, \end{aligned} \tag{1.159}$$

where $|L| = \left(L_1^1 L_2^2 - L_1 L^2 L_2^1\right)$ is called *Gauss curvature* (see figure 1.15). Hence, the curvature of a plane is zero since the normals are all in the same direction. Similarly, the curvature of a cylinder is zero too because the normal does not change in a direction parallel to the axis of the cylinder although it changes along the circular surface. The curvature of a sphere, however, is $\frac{1}{\rho^2}$ (where ρ is the radius of the sphere) because the soild angle $d\Omega$ that it subtends at the sphere's center is the same as that spanned by the normals.

Consider two geodesics that are initially parallel, separated by a distance χ_0.

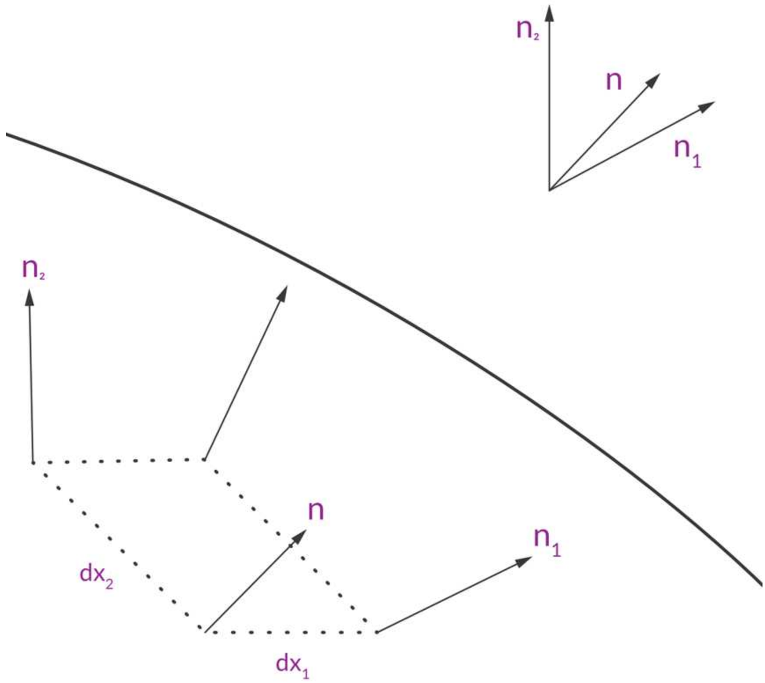


Fig. 1.15. Gaussian Curvature of a surface. The curvature of the surface shown at a point is the ratio of the area spanned by the increments $(d\mathbf{n})_1 = \mathbf{n}_1 - \mathbf{n}$ and $(d\mathbf{n})_2 = \mathbf{n}_2 - \mathbf{n}$ of the normal, to the area spanned by the tangential displacement vectors dx_1 and dx_2 on the surface.

They are no longer parallel when we traverse a distance s and their separation is measured by $\chi = \chi_0 cos\left(\frac{s}{\rho}\right)$, where ρ is the radius of the sphere. The separation follows a surprising equation.

$$\begin{aligned}
&\chi = \chi_0 cos\left(\tfrac{s}{\rho}\right)\\
&\tfrac{d\chi}{ds} = -\tfrac{\chi_0}{\rho} sin\left(\tfrac{s}{\rho}\right)\\
&\tfrac{d^2\chi}{ds^2} = -\tfrac{\chi_0}{\rho^2} cos\left(\tfrac{s}{\rho}\right) = -\tfrac{1}{\rho^2}\chi\\
&\tfrac{d^2\chi}{ds^2} + R\chi = 0,
\end{aligned} \tag{1.160}$$

where $R = \frac{1}{\rho^2}$ is the Gaussian curvature of the surface and the equation derived is called the equation of geodesic deviation which is nothing but the equation of simple harmonic motion!

1.8.2 Theorema Egregium

Carl Friedrich Gauss found out that the curvature of a surface can be measured exclusively in terms of quantities intrinsic to the surface, without any reference to how the surface is embedded in the surrounding (3-dimensional) space. The intrinsic quantities are the coefficients $g_{\mu\nu}$ of the line element ds which measures the distance between two close points present on the surface. Consider a pair of points separated by an infinitesimal distance $d\xi^{\alpha}$, $\left(\xi^1, \xi^2\right)$ and $\left(\xi^1 + d\xi^1, \xi^2 + d\xi^2\right)$. The separation between the points is expressed as follows

$$ds^2 = x_{,\mu} x_{,\nu} d\xi^{\mu} d\xi^{\nu} \equiv g_{\mu\nu} d\xi^{\mu} d\xi^{\nu}. \qquad (1.161)$$

This quadractic form of the metric is positive definite, symmetric and is non-singular (i.e., it's determinent is non-zero $|g| \neq 0$). The theorem states that it is the combination of the Weingarten matrices L^{μ}_{ν} that determines the Gauss curvature (also called total curvature), which is given in terms of the metric in an equation called the *Gauss equation*.

$$L^{\mu}_{\nu} L^{\kappa}_{\alpha} - L^{\mu}_{\alpha} L^{\kappa}_{\nu} = -g^{\mu\zeta} R^{\kappa}_{\zeta\nu\alpha}, \qquad (1.162)$$

where $R^{\kappa}_{\zeta\nu\alpha}$ is called the *Riemann curvature tensor*. The straightest possible curves on this surface, i.e., the geodesics, are expressed as function of $\xi^{\alpha}(s)$, where s is the distance measured along the curve, which satisfy the geodesic equation, $\frac{d^2\xi^{\kappa}}{ds^2} + \Gamma^{\kappa}_{\zeta\nu} \frac{d\xi^{\zeta}}{ds} \frac{d\xi^{\nu}}{ds} = 0$.

1.8.3 The Riemann Curvature Tensor

The *Riemann curvature tensor* is a higher-dimensional analogue of the Gaussian curvature. In 2-dimensions, the direction of the acceleration of one geodesic relative to another geodesic (called the *fiducial geodesic*) is fixed uniquely by the demand that their separation vector $\boldsymbol{\chi}$ be perpendicular to the fiducial geodesic. However, in higher-dimensions, the separation vector does not only remain perpendicular to the fiducial geodesic, but also rotates about it. In the slot machine analogy, the Riemann tensor is a machine that has three slots and in a coordinate system the components can be written as a *trilinear function* (it obeys linearity),

$$\mathbf{r} = R(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad \rightarrow \quad r^{\alpha} = R^{\alpha}_{\beta\lambda\xi} a^{\beta} b^{\lambda} c^{\xi}. \qquad (1.163)$$

The equation of geodesic deviation in higher dimensions replaces the Gaussian curvature with the Riemannian curvature tensor and the spatial distance with the propertime. Let the unit tangent vector or the 4-velocity be $u^\alpha = \frac{dx^\alpha}{d\tau}$, then the equation of of geodesic deviation is expressed as

$$\begin{aligned} &\tfrac{D^2\chi}{d\tau^2} + R(\mathbf{u}, \chi, \mathbf{u}) = 0 \\ &\tfrac{D^2\chi^\alpha}{d\tau^2} + R^\alpha_{\beta\lambda\xi}\tfrac{dx^\beta}{d\tau}\chi^\lambda\tfrac{dx^\xi}{d\tau} = 0. \end{aligned} \tag{1.164}$$

Thus, the Riemann tensor is an exterior 2-form taking values in the set of linear maps from the tangent plane to itself. The non-commutativity of covariant derivatives is a geometrical property of the metric. The commutation $(D_\alpha D_\beta - D_\beta D_\alpha)\, u^\mu$ of two covariant derivatives of a vector u is a mixed tensor with coefficients $R_{\alpha\beta}{}^{\mu}{}_{\nu}$ such that

$$(D_\alpha D_\beta - D_\beta D_\alpha)\, u^\mu = [D_\alpha, D_\beta]\, u^\mu = R_{\alpha\beta}{}^{\mu}{}_{\nu} u^\nu. \tag{1.165}$$

The components of the Riemann tensor in a coordinate basis is defined below and the proof follows.

$$\begin{aligned} R^\alpha_{\beta\lambda\xi} &= \langle \boldsymbol{\omega}^\alpha, [D_\lambda, D_\xi]\, \mathbf{e}_\beta \rangle \\ &= \Gamma^\alpha_{\beta\xi,\lambda} - \Gamma^\alpha_{\beta\lambda,\xi} + \Gamma^\alpha_{\mu\lambda}\Gamma^\mu_{\beta\xi} - \Gamma^\alpha_{\mu\xi}\Gamma^\mu_{\beta\lambda}. \end{aligned} \tag{1.166}$$

Proof. $\mathrm{R}^\mu_{\beta\lambda\xi} = \langle \boldsymbol{\omega}^\alpha, [D_\lambda, D_\xi]\, \mathbf{e}_\beta \rangle = \langle \boldsymbol{\omega}^\alpha, (D_\lambda D_\xi - D_\xi D_\lambda)\, \mathbf{e}_\beta \rangle$

$$= \langle \boldsymbol{\omega}^\alpha, (D_\lambda\,(D_\xi \mathbf{e}_\beta) - D_\xi\,(D_\lambda \mathbf{e}_\beta)) \rangle$$

$$= \left\langle \boldsymbol{\omega}^\alpha, \left(D_\lambda \left(\mathbf{e}_\mu \Gamma^\mu_{\beta\xi}\right) - D_\xi \left(\mathbf{e}_\mu \Gamma^\mu_{\beta\lambda}\right)\right)\right\rangle$$

$$= \left\langle \boldsymbol{\omega}^\alpha, \mathbf{e}_\mu \Gamma^\mu_{\beta\xi,\lambda} + \Gamma^\mu_{\beta\xi}\,(D_\lambda \mathbf{e}_\mu) - \mathbf{e}_\mu \Gamma^\mu_{\beta\lambda,\xi} - \Gamma^\mu_{\beta\lambda}\,(D_\xi \mathbf{e}_\mu)\right\rangle$$

$$= \left\langle \boldsymbol{\omega}^\alpha, \mathbf{e}_\mu \Gamma^\mu_{\beta\xi,\lambda} + \left(\mathbf{e}_\nu \Gamma^\nu_{\mu\lambda}\right)\Gamma^\mu_{\beta\xi} - \mathbf{e}_\mu \Gamma^\mu_{\beta\lambda,\xi} - \left(\mathbf{e}_\nu \Gamma^\nu_{\mu\xi}\right)\Gamma^\mu_{\beta\lambda}\right\rangle$$

$$= \left\langle \boldsymbol{\omega}^\alpha, \mathbf{e}_\mu \left(\Gamma^\mu_{\beta\xi,\lambda} - \Gamma^\mu_{\beta\lambda,\xi}\right) + \mathbf{e}_\nu \left(\Gamma^\nu_{\mu\lambda}\Gamma^\mu_{\beta\xi} - \Gamma^\nu_{\mu\xi}\Gamma^\mu_{\beta\lambda}\right)\right\rangle$$

$$= \left(\Gamma^{\mu}_{\beta\xi,\lambda} - \Gamma^{\mu}_{\beta\lambda,\xi}\right) \underbrace{\langle \boldsymbol{\omega}^{\alpha}, \mathbf{e}_{\mu} \rangle}_{=\delta^{\alpha}_{\mu}=1 \ if \ \alpha=\mu} + \left(\Gamma^{\nu}_{\mu\lambda}\Gamma^{\mu}_{\beta\xi} - \Gamma^{\nu}_{\mu\xi}\Gamma^{\mu}_{\beta\lambda}\right) \underbrace{\langle \boldsymbol{\omega}^{\alpha}, \boldsymbol{\nu} \rangle}_{=\delta^{\alpha}_{\nu}=1 \ if \ \alpha=\nu} .$$

Hence, $\mathrm{R}^{\mu}_{\beta\lambda\xi} = \Gamma^{\mu}_{\beta\xi,\lambda} - \Gamma^{\mu}_{\beta\lambda,\xi} + \Gamma^{\nu}_{\mu\lambda}\Gamma^{\mu}_{\beta\xi} - \Gamma^{\nu}_{\mu\xi}\Gamma^{\mu}_{\beta\lambda}.$

The Riemann curvature tensor is closely related to tidal forces, it represents the tidal force experienced by a particle moving along a geodesic.

1.8.4 Symmetries of the Riemann Tensor

The Riemann curvature tensor has, in 4-dimensions, $4\times4\times4\times4 = 256$ independent components. Observations reveals a variety of algebraic symmetries such as the *first skew symmetry*, the *second skew symmetry*, and the *block symmetry*.

$$R_{\alpha\beta\mu\nu} = R_{[\alpha\beta][\mu\nu]}, \quad R_{[\alpha\beta\mu\nu]} = 0, \quad R_{\alpha[\beta\mu\nu]} = 0. \tag{1.167}$$

All of the above symmetries reduce the Riemann tensor from 256 components to 20 independent components. The antisymmetry of $(\alpha\beta)$ and $(\mu\nu)$ in $R_{\alpha\beta\mu\nu}$ implies that there are $P = \frac{1}{2}n(n-1)$ different ways of choosing non-trivial pairs $(\alpha\beta)$ and P ways of choosing non-trivial pairs $(\mu\nu)$. The observation that the tensor is symmetric with respect to the interchange of $(\alpha\beta)$ and $(\mu\nu)$ implies that there are $\frac{1}{2}P(P+1)$ independent ways of choosing $\alpha\beta\mu\nu$ when the pair symmetries are considered. The last algebraic symmetry, called the *cyclic symmetry* can be written alternatively as

$$\begin{aligned} R_{\alpha[\beta\mu\nu]} &= \tfrac{1}{3!}\left(R_{\alpha\beta\mu\nu} - R_{\alpha\beta\nu\mu} + R_{\alpha\nu\beta\mu} - R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} - R_{\alpha\mu\beta\nu}\right) \\ R_{\alpha[\beta\mu\nu]} &= \tfrac{1}{3!}\left(R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta}\right) = \tfrac{1}{3!}\Delta_{\alpha\beta\mu\nu} = 0, \end{aligned} \tag{1.168}$$

where the pair symmetries guarantee that $\Delta_{\alpha\beta\mu\nu}$ is totally antisymmetric such that $\Delta_{\alpha\beta\mu\nu} = 0$ is trivial unless all the indices are distinct. The number of added constraints is then the number of combinations of 4 indices that can be taken from n indices is $\frac{n!}{(n-4)!4!}$. The number of independent components is then given by

$$\frac{1}{2}P(P+1) - \frac{n!}{(n-4)!4!} = \frac{n^2\left(n^2-1\right)}{12}. \tag{1.169}$$

In 4-dimensions, we then have $\frac{4^2\left(4^2-1\right)}{12} = 20$ independent components. $\Delta_{\alpha\beta\mu\nu}$ mentioned in the previous equations is called *Bianchi's first identity*. Besides the algebraic symmetries, there exists differential symmetries called *Bianchi identities* given by the following expression

$$R^{\alpha}_{\beta[\xi\mu;\nu]} = 0. \tag{1.170}$$

The contraction of the Riemann tensor is called the *Ricci tensor*[18], $R_{\mu\nu} = R^{\alpha}_{\beta\mu\nu} g^{\beta}_{\alpha}$ and the contraction of the Ricci tensor is the *curvature scalar*, $R = R_{\mu\nu} g^{\mu\nu}$.

$$R_{\mu\nu} = \partial_\rho \Gamma^{\rho}_{\nu\mu} - \partial_\nu \Gamma^{\rho}_{\rho\mu} + \Gamma^{\rho}_{\rho\lambda}\Gamma^{\lambda}_{\nu\mu} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\rho\mu}. \tag{1.171}$$

The Ricci tensor is symmetric, i.e., $R_{\mu\nu} = R_{\nu\mu}$, and out of the $\frac{1}{2}n^2\left(n^2-1\right)$ algebraically independent components of the Riemann tensor $\frac{1}{2}n\left(n+1\right)$ of them can be represented by the components of the Ricci tensor. For $n = 1$, $R_{\alpha\beta\mu\nu} = 0$; for $n = 2$, there exists only one independent component of the Riemann tensor, which is the curvature scalar; for $n = 3$, the Ricci tensor completely determines the curvature tensor.

Example 1.10. In this example, we revisit our claim 1.122 that the commutator of two Killing vectors is a Killing vector. Consider two Killing vectors, K and M which we expand in a basis $\{x^\alpha\}$ as $K = K^\alpha \partial_\alpha$ and $M = M^\alpha \partial_\alpha$. Both these vectors obey the Killing equation, i.e.,

$$D_{(\alpha} K_{\beta)} = 0, \quad D_{(\alpha} M_{\beta)} = 0. \tag{1.172}$$

It is easy to see that a linear combination (with constant coefficients c_1 and c_2) of these two Killing vectors, say

$$L = c_1 K + c_2 M, \ L = L^\alpha \partial_\alpha \tag{1.173}$$

is also a Killing vector since it obeys the Killing equation, i.e.,

$$D_{(\alpha} L_{\beta)} = c_1 D_{(\alpha} K_{\beta)} + c_2 D_{(\alpha} M_{\beta)} = 0. \tag{1.174}$$

Now consider the commutator of K and M with the former expanded in a basis $K = K^\alpha \partial_\alpha$ and the latter expanded in a basis $M = M^\beta \partial_\beta$. Let $L = L^\alpha \partial_\alpha$ be the commutator which is given as

$$\begin{aligned} L &= \left[K^\alpha \partial_\alpha, M^\beta \partial_\beta\right] = K^\alpha \partial_\alpha \left(M^\beta \partial_\beta\right) - M^\beta \partial_\beta \left(K^\alpha \partial_\alpha\right) \\ &= K^\alpha \left(\partial_\alpha M^\beta\right) \partial_\beta + K^\alpha M^\beta \partial_\alpha \partial_\beta - M^\beta \left(\partial_\beta K^\alpha\right) \partial_\alpha - M^\beta K^\alpha \partial_\alpha \partial_\beta \\ &= K^\alpha \left(\partial_\alpha M^\beta\right) \partial_\beta - M^\beta \left(\partial_\beta K^\alpha\right) \partial_\alpha. \end{aligned} \tag{1.175}$$

Now, using the definition of the covariant derivative, $D_\alpha M^\beta = \partial_\alpha M^\beta + \Gamma^{\beta}_{\alpha\lambda} M^\lambda$, we get

[18] can be defined as a covariant, contravariant, or a mixed tensor.

$$L^\alpha \partial_\alpha = K^\alpha D_\alpha M^\beta \partial_\beta - K^\alpha \Gamma^\beta_{\alpha\lambda} M^\lambda \partial_\beta - M^\beta D_\beta K^\alpha \partial_\alpha + M^\beta \Gamma^\alpha_{\beta\lambda} K^\lambda \partial_\alpha$$
$$= \left(K^\beta \left(D_\beta M^\alpha\right) - M^\beta \left(D_\alpha K^\alpha\right)\right) \partial_\alpha, \tag{1.176}$$

where the last line was obtained by swapping the indices and using the fact that $K^\alpha \Gamma^\beta_{\alpha\lambda} M^\lambda \partial_\beta = M^\beta \Gamma^\alpha_{\beta\lambda} K^\lambda \partial_\alpha$ due to repeated indices. Now that we have obtained an explicit expression for L we need to check if it obeys the Killing equation. Note that the terms below in boldface cancel

$$2D_{(\alpha}L_{\beta)} = D_\alpha \left(K^\mu \left(D_\mu M_\beta\right) - M^\mu \left(D_\mu K_\beta\right)\right) + D_\beta \left(K^\mu \left(D_\mu M_\alpha\right) - M^\mu \left(D_\mu K_\alpha\right)\right)$$
$$= \mathbf{D}_\alpha \mathbf{K}_\mu \mathbf{M}_\beta + K^\mu D_\alpha D_\mu M_\beta - \mathbf{D}_\alpha \mathbf{M}^\mu \mathbf{D}_\mu \mathbf{K}_\beta - M^\mu D_\alpha D_\mu K_\beta$$
$$+ \mathbf{D}_\beta \mathbf{K}^\mu \mathbf{D}_\mu \mathbf{M}_\alpha + K^\mu D_\beta D_\mu M_\alpha - \mathbf{D}_\beta \mathbf{M}^\mu \mathbf{D}_\mu \mathbf{K}_\alpha - M^\mu D_\beta D_\mu K_\alpha$$
$$= K^\mu D_\alpha D_\mu M_\beta - M^\mu D_\alpha D_\mu K_\beta + K^\mu D_\beta D_\mu M_\alpha - M^\mu D_\beta D_\mu K_\alpha. \tag{1.177}$$

Making use of the commutator relation $\left[D_\alpha, D_\mu\right] M_\beta = D_\alpha D_\mu M_\beta - D_\mu D_\alpha M_\beta$, we get terms with a common factor $D_{(\beta}M_{\alpha)}$ which is zero since M obeys the Killing equation. Now, using the definition of the Riemann tensor $\left[D_\alpha, D_\mu\right] M_\beta = R^\delta_{\beta\alpha\mu} M_\delta = g^{\lambda\delta} R_{\lambda\beta\alpha\mu} M_\delta$ to simplify the equation, we obtain

$$2D_{(\alpha}L_{\beta)} = K^{\mu\lambda} \left(R_{\lambda\beta\alpha\mu} + R_{\lambda\alpha\beta\mu}\right) - \underbrace{M^\mu K^\lambda \left(R_{\lambda\beta\alpha\mu} + R_{\lambda\alpha\beta\mu}\right)}_{\mu \longleftrightarrow \nu}$$
$$= M^\lambda K^\mu \left(R_{\lambda\beta\alpha\mu} + R_{\lambda\alpha\beta\mu} - R_{\mu\beta\alpha\lambda} - R_{\mu\alpha\beta\lambda}\right), \tag{1.178}$$

$$R_{abcd} = R_{cdab} = -R_{dcab} = R_{dcba}$$

$$\Rightarrow D_{(\alpha}L_{\beta)} = 0.$$

Since L obeys the Killing equation we can conclude that the commutator of two Killing vectors is a Killing vector.

Example 1.11. An important identity
We now will show that the following identity holds

$$D_\mu D_\sigma X^\rho = R^\rho_{\sigma\mu\nu} X^\nu. \tag{1.179}$$

Prior to calculation we already have a useful result we can make use of: $D_\sigma X_\rho = -D_\rho X_\sigma$ from which we have

$$D_\mu D_\sigma X_\rho = -D_\mu D_\rho X_\sigma. \tag{1.180}$$

We rewrite this in terms of the commutator $[D_\rho, D_\mu]$ to get

$$\begin{aligned} D_\mu D_\sigma X_\rho &= -D_\sigma D_\mu X_\sigma + [D_\rho, D_\mu]\, X_\sigma \\ &= D_\rho D_\sigma X_\mu + [D_\rho, D_\mu]\, X_\sigma \\ &= ([D_\rho, D_\sigma]\, X_\mu + D_\sigma D_\rho X_\mu) + [D_\rho, D_\mu]\, X_\sigma \\ &= -D_\sigma D_\mu X_\rho + [D_\rho, D_\sigma]\, X_\mu + [D_\rho, D_\mu]\, X_\sigma, \\ &= -\left([D_\sigma, D_\mu]\, X_\rho + D_\mu D_\sigma X_\rho\right) + [D_\rho, D_\sigma]\, X_\mu + [D_\rho, D_\mu]\, X_\sigma \\ &= -D_\mu D_\sigma X_\rho + [D_\sigma, D_\mu]\, X_\rho + [D_\rho, D_\sigma]\, X_\mu + [D_\rho, D_\mu]\, X_\sigma. \end{aligned} \tag{1.181}$$

From this we obtain the following expression

$$D_\mu D_\sigma X_\rho = \frac{1}{2}\left([D_\sigma, D_\mu]\, X_\rho + [D_\rho, D_\sigma]\, X_\mu + [D_\rho, D_\mu]\, X_\sigma\right), \tag{1.182}$$

and making use of the definition of the Riemann tensor $[D_a, D_b]\, X^c = R^c_{dab} X^d$, the expression simplifies to the following

$$\frac{2}{X^\nu} D_\mu D_\sigma X_\rho = \underbrace{R_{\rho\nu\mu\sigma}}_{=-R_{\nu\rho\mu\sigma}} + \underbrace{R_{\mu\nu\rho\sigma}}_{=R_{\nu\mu\sigma\rho}} + \underbrace{R_{\sigma\nu\rho\mu}}_{-R_{\nu\sigma\rho\mu}}. \tag{1.183}$$

We can now make use of the cyclic symmetry of the Riemann tensor to see that $R_{\nu\rho\mu\sigma} + R_{\nu\sigma\rho\mu} = -R_{\nu\mu\sigma\rho}$ to finally get

$$\begin{aligned} D_\mu D_\sigma X_\rho &= \underbrace{R_{\nu\rho\mu\sigma}}_{=R_{\mu\sigma\nu\rho}=R_{\sigma\mu\rho\nu}} X^\nu, \\ \Rightarrow D_\mu D_\sigma X^\rho &= R^\rho_{\sigma\mu\nu} X^\nu. \end{aligned} \tag{1.184}$$

Exercise 3

1. Using the various symmetries of the Riemann tensor show that
a. $R_{\alpha\beta\mu\nu} = \frac{2}{3}\left(R_{\alpha(\beta\mu)\nu} - R_{\alpha(\beta\nu)\mu}\right)$.
b. $R_{\alpha\beta\mu\nu} + 2\left(R_{\alpha[\beta\mu]\nu} - R_{\alpha[\beta\nu]\mu}\right) = 0$.

2. Show that the Killing vector X satisfies the following identities
a. $D_\alpha D_\beta X^\alpha = R_{\alpha\beta} X^\alpha$.
b. $[D_\alpha, D_\beta]\, D^{[\alpha} X^{\beta]} = 0$.
c. $X^\alpha D_\alpha R = 0$.

d. $D_\gamma D_\beta K_\alpha - D_\beta D_\gamma K_\alpha = R^\lambda_{\alpha\beta\gamma} K_\lambda$.
e. Use the identity of the previous question and the fact that $R^\lambda_{[\alpha\beta\gamma]} = 0$ to show

$$D_\beta \left(D_\gamma K_\alpha - D_\alpha K_\gamma\right) + D_\gamma \left(D_\alpha K_\beta - D_\beta K_\alpha\right) + D_\alpha \left(D_\beta K_\gamma - D_\gamma K_\beta\right) = 0, \tag{1.185}$$

and hence use the Killing equation show that $D_\beta D_\alpha K_\gamma = R^\rho_{\beta\alpha\gamma} K_\rho$.

3. Consider the two-dimensional metric $ds^2 = \left(1 + \frac{u^2+v^2}{4l^2}\right)^{-1} \left(du^2 + dv^2\right)$, where l is a constant. Let X be a Killing vector which has components $X^u = -v$ and $X^v = u$. Show that X satisfies the Killing equation $D_{(\mu} X_{\nu)} = 0$, $\mu, \nu = 0, 1$.

4. Show that the following identity holds

$$D^\kappa D_\kappa R_{\mu\nu\alpha\beta} = 2R^\lambda_{\mu\kappa\beta} R^\kappa_{\alpha\lambda\nu} - 2R^\lambda_{\nu\kappa\beta} R^\kappa_{\alpha\lambda\mu} - R^\lambda_{\kappa\beta\alpha} R^\kappa_{\lambda\mu\nu}.$$

This is the Penrose equation for a vacuum spacetime. Hint: Use the Bianchi identity.

5. Let $\kappa = \mathcal{L}_X g$ using which we define

$$\Omega^\alpha_{\mu\nu} = \frac{1}{2} g^{\alpha\beta} \left(D_\mu \kappa_{\nu\beta} + D_\nu \kappa_{\mu\beta} - D_\beta \kappa_{\mu\nu}\right).$$

Notice that this is similar to how the Christoffel symbol is defined.
a. Show that $\left[D_{[\gamma}, \mathcal{L}_X\right] R_{\mu\nu]\alpha\beta} = \Omega^\delta_{\alpha[\gamma} R_{\mu\nu]\delta\beta} + \Omega^\delta_{\beta[\gamma} R_{\mu\nu]\alpha\delta}$.

b. Show that simplification of $\kappa^\alpha_{\mu\nu}$ yields the Killing identity, i.e., obtain $\Omega^\alpha_{\mu\nu} = \frac{g^{\alpha\beta}}{2} \left(D_\mu D_\nu X_\beta + R_{\gamma\mu\nu\beta} X^\gamma\right)$. The Killing vectors satisfy the Killing identity here since $\kappa = \mathcal{L}_X g = 0$ and in-turn Ω which is a linear combination of the Lie derivatives of g vanishes.

6. Consider a metric $g_{\alpha\beta}$. If this metric, when a Lie derivative is taken over a Killing vector K, satisfies the Killing equation does it imply that the Lie derivative of the Affine connection vanishes, i.e., $\mathcal{L}_K \Gamma^\mu_{\alpha\beta} = 0$? Would it also imply that the Lie derivative of the Curvature vanishes, i.e., $\mathcal{L}_K R_{\mu\nu\alpha\beta} = 0$ (this is known as curvature collineation)? Would this hold if there is torsion?

1.8.5 Weyl Tensor

For values of $n > 3$ in $\frac{1}{2} n^2 \left(n^2 - 1\right)$, the components of the Riemann curvature tensor apart from it's own components are represented by the *Weyl tensor* $W_{\alpha\beta\mu\nu}$.

$$W_{\alpha\beta\mu\nu} = R_{\alpha\beta\mu\nu} - \frac{2}{n-2} \left(g_{\alpha[\nu} R_{\mu]\beta} + g_{\beta[\mu} R_{\nu]\alpha}\right) + \frac{2}{(n-1)(n-2)} R g_{\alpha[\mu} g_{\nu]\beta} \tag{1.186}$$

The Weyl tensor also possesses all three algebraic symmetries and in addition it can be thought of as that part of the curvature tensor such that all contractions vanish, i.e., $W^{\alpha}_{\beta\alpha\nu} = 0$. The Weyl tensor is a measure of the curvature of spacetime or, more generally, of a pseudo-Riemannian manifold. It can be shown that the Weyl tensor of a three-dimensional pseudo-Riemannian manifold $(\mathcal{M}, g)$ is identically zero. Like the Riemann curvature tensor, it expresses the tidal force that a body feels when moving along a geodesic. It differs from the Riemann curvature tensor in that it does not convey information on how the volume of the body changes, but rather only how the shape of the body is distorted by the tidal force.

The Weyl tensor is equal to the Riemann tensor in a Ricci-flat space ($R_{\mu\nu} = 0$). It is considered that the Weyl tensor embodies in some sense the non-Newtonian properties of the gravitational field, in particular its radiation properties. This point of view is supported by the fact that the equations for massless fields, at least in four spacetime dimensions, are *conformally invariant* (this is a concept that will be explained soon). Since W is obviously zero for a flat metric, it is also zero if the metric is conformal to a flat metric. It can be proved that if $n > 3$, then the identical vanishing of the Weyl tensor implies that the metric is locally *conformally flat.* Consider two metrics, g and $\bar{g}$. These metrics are said to be conformal if and only if

$$\bar{g} = \omega^2 g, \tag{1.187}$$

where ω is a non-zero differentiable function. If this condition is satisfied, then for any vectors $\mathbf{R}, \mathbf{Q}, \mathbf{S}, \mathbf{V}$ at a point p on the manifold $\mathcal{M}$,

$$\frac{g(\mathbf{R}, \mathbf{Q})}{g(\mathbf{S}, \mathbf{V})} = \frac{\bar{g}(\mathbf{R}, \mathbf{Q})}{\bar{g}(\mathbf{S}, \mathbf{V})}, \tag{1.188}$$

so angles and lightlike world-lines are preserved under conformal transformations. The null cone structure in the tangent space $T_p(\mathcal{M})$ is preserved by conformal transformations since for a vector $\mathbf{R} \in \mathcal{M}$,

$$g(\mathbf{R}, \mathbf{R}) > 0, = 0, < 0 \;\Rightarrow\; \bar{g}(\mathbf{R}, \mathbf{R}) > 0, = 0, < 0. \tag{1.189}$$

The metric components are related as follows

$$\bar{g}_{\mu\nu} = \omega^2 g_{\mu\nu}. \tag{1.190}$$

This concept of conformal factors and how helps it to select a relevant two-dimensional part of a spacetime and to make it's *stereographic projection* on a compact space is studied in *Penrose-Carter diagrams.* The idea behind these diagrams is that under conformal maps (when the conformal factor is dropped) lightlike or null world lines and angles between them do not change.

Revisiting the cases for different values of n, we conclude the following: In the case of $n = 1$, it is implied there are zero truly independent components in the metric. The implication of the case when $n = 2$, is that in any 2-dimensional Riemann manifold, it is a standard result that locally we can always choose coordinates to make the metric conformally flat. The implication of the case when $n = 3$, is that in any 3-dimensional Riemann manifold, it is a standard result that locally we can always choose coordinates to make the metric diagonal, i.e., $g_{mn} = Diag\,(g_{11}, g_{22}, g_{33})$ (all the non-diagonal elements are zero), i.e., Riemann 3-manifolds have metric that are always locally *diagonalizable.*

Example 1.12. A full-blown metric calculation
Before proceeding further, we shall put the concepts introduced thus far to work and calculate all the algebraic quantities for a given metric. Consider the Poincaré half-plane model which describes hyperbolic geometry in 2-dimensions. It has the following line element

$$ds^2 = \frac{1}{y^2}\left(dx^2 + dy^2\right), \tag{1.191}$$

which can be represented in the matrix form as follows

$$g_{\mu\nu} = \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix}, \; g^{\mu\nu} = \begin{pmatrix} y^2 & 0 \\ 0 & y^2 \end{pmatrix} \tag{1.192}$$

The Christoffel symbols for this metric can be calculated as follows

$$\begin{aligned}
\Gamma^x_{xx} &= g^{xx}\Gamma_{xxx} = g^{xx}\tfrac{1}{2}\left(g_{xx,x} + g_{xx,x} - g_{xx,x}\right) = \tfrac{1}{2}y^2\partial_x y^{-2} = 0, \\
\Gamma^x_{xy} &= \Gamma^x_{yx} = g^{xx}\Gamma_{xyx} = g^{xx}\tfrac{1}{2}\left(g_{xx,y} + g_{xy,x} - g_{xy,x}\right) = \tfrac{1}{2}y^2\left(-\tfrac{2}{y^3}\right) = -\tfrac{1}{y}, \\
\Gamma^y_{yy} &= g^{yy}\Gamma_{yyy} = g^{yy}\tfrac{1}{2}\left(g_{yy,y} + g_{yy,y} - g_{yy,y}\right) = \tfrac{1}{2}y^2\left(-\tfrac{2}{y^3}\right) = -\tfrac{1}{y}, \\
\Gamma^y_{xy} &= \Gamma^y_{yx} = g^{yy}\Gamma_{yxx} = g^{yy}\tfrac{1}{2}\left(g_{xy,x} + g_{xx,y} - g_{yx,x}\right) = \tfrac{1}{2}y^2\left(-\tfrac{2}{y^3}\right) = -\tfrac{1}{y}, \\
\Gamma^x_{yy} &= g^{xx}\Gamma_{yyx} = g^{xx}\tfrac{1}{2}\left(g_{xy,y} + g_{xy,y} - g_{yy,x}\right) = 0, \\
\Gamma^y_{xx} &= g^{yy}\Gamma_{xxy} = g^{yy}\tfrac{1}{2}\left(g_{yx,x} + g_{yx,x} - g_{xx,y}\right) = -\tfrac{1}{2}y^2\left(-\tfrac{2}{y^3}\right) = \tfrac{1}{y}.
\end{aligned} \tag{1.193}$$

We can represent the components of the Christoffel symbols in a matrix form as follows with the matrix having a x and y labels for the rows and columns and the individual columns having a x and y label

$$\begin{pmatrix} \begin{pmatrix} \Gamma^x_{xx} \\ \Gamma^x_{xy} \end{pmatrix} & \begin{pmatrix} \Gamma^x_{yx} \\ \Gamma^x_{yy} \end{pmatrix} \\ \\ \begin{pmatrix} \Gamma^y_{xx} \\ \Gamma^y_{xy} \end{pmatrix} & \begin{pmatrix} \Gamma^y_{yx} \\ \Gamma^y_{yy} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 0 \\ -y^{-2} \end{pmatrix} & \begin{pmatrix} -y^{-2} \\ 0 \end{pmatrix} \\ \\ \begin{pmatrix} y^{-2} \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ -y^{-2} \end{pmatrix} \end{pmatrix} \tag{1.194}$$

Using the information of the Christoffel symbols we can now find out the expressions for the Riemann tensor. For the given metric, we will have 4 Riemann tensors for a given upper index, i.e., say x is the upper index then we have R^x_{xxx}, R^x_{xxy}, R^x_{xyx}, and R^x_{xyy}. Since we have a 2-dimensional matrix, we have a Riemann tensor whose components span a 2×2 matrix whose individual components are 2×2 matrices. The upper index is read as the label of the global row, the first lower index is the label of the global column, the second lower index is the label of local row, and the third lower index is the label of the local column. The smart way to proceed here is to calculate one component in each local 2×2 matrix are try and fix the other components using the knowledge of the symmetries. Let's first calculate R^x_{xxy}

$$R^x_{xxy} = \Gamma^x_{xx,y} - \Gamma^x_{xy,x} + \underbrace{\Gamma^x_{\beta y}\Gamma^\beta_{xx}}_{=\Gamma^x_{xy}\Gamma^x_{xx}+\Gamma^x_{yy}\Gamma^y_{xx}} + \underbrace{\Gamma^x_{\beta x}\Gamma^\beta_{xy}}_{=\Gamma^x_{xx}\Gamma^x_{xy}+\Gamma^x_{yx}\Gamma^y_{xy}} = 0, \tag{1.195}$$

since this component vanishes, we can write the Riemann tensor in a covariant form as $R^x_{xxy} = g^{xx} R_{xxxy}$ and use the first & second skew symmetries and block symmetry to find

$$\begin{aligned} R_{xxxy} &= -R_{xxyx} = 0 \Rightarrow R^x_{xyx} = 0, \\ R_{xxxy} &+ R_{xxyx} + R_{xyxx} = 0 \Rightarrow R^x_{yxx} = 0, \\ R_{xxxy} &= \underbrace{R_{xyxx}}_{=-R_{yxxx}} \Rightarrow R^y_{xxx} = 0. \end{aligned} \tag{1.196}$$

Proceeding the same way for other components we find that the only non-zero components are R^x_{yxy}, R^x_{yyx}, R^y_{xxy}, and R^y_{xyx}, which take the following form

$$R^x_{yxy} = R^y_{xyx} = -\frac{1}{y^2}, \; R^x_{yyx} = R^y_{xxy} = \frac{1}{y^2}. \tag{1.197}$$

Using this information, the Ricci tensor components can be calculated quite easily as follows

$$\begin{aligned} R_{xx} &= g^{yy} R_{xyxy} = g^{yy} g_{xx} R^x_{yxy} = -\tfrac{1}{y^2}, \; R_{xy} = 0, \\ R_{yy} &= g^{xx} R_{xyxy} = g^{xx} g_{xx} R^x_{yxy} = -\tfrac{1}{y^2}, \; R_{yx} = 0. \end{aligned} \tag{1.198}$$

Finally , with all the above expressions, we can calculate the components of the Weyl tensor. Again the smart way to proceed is to realize that the Weyl tensor possesses the same symmetries as that of the Riemann tensor. After some algebra (which you are encouraged to do) it can be seen that all the components of the Weyl tensor vanish.

Exercise 4

1. Explicitly show that for the following metric, all the components of the Weyl tensor vanish identically

$$ds^2 = \frac{l^2}{z^2}\left(-dt^2 + dx^2 + dz^2\right). \tag{1.199}$$

This called the Poincaré patch of AdS_3 spacetime, l here is called the AdS radius, and $t \in (-\infty, \infty)$, $z > 0$ or $z < 0$ (divides hyperboloid into two charts). Of course the realization that the Weyl tensor has 0 components in $d = 3$ may aid in making future calculations much simpler. This is so since the Weyl tensor, for values $n > 3$, has $\frac{n^2(n^2-1)}{12} - \frac{n(n+1)}{2}$ components in n-dimensions[19]. Return to this question after studying Einstein's field equations. Einstein's equations in vacuum (with the cosmological constant, Λ) take the form: $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 0$, taking the trace of this equation gives

$$\begin{aligned}
&g^{\mu\nu}R_{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu}g_{\mu\nu} + \Lambda g^{\mu\nu}g_{\mu\nu} = 0;\ g^{\mu\nu}g_{\mu\nu} = \delta^\mu_\mu = \textstyle\sum_{i=0}^n \delta^i_i = n,\\
&R - \tfrac{n}{2}R + n\Lambda = 0, \\
&\Rightarrow R = \tfrac{2\Lambda n}{n-2},
\end{aligned} \tag{1.200}$$

where n is the dimension of spacetime. For $n = 3$, the Riemann and Ricci have 6 independent components and hence, they can be expressed in terms of the other as follows

$$\begin{aligned}
R_{\alpha\beta\mu\nu} &= 2\left(g_{\alpha[\nu}R_{\mu]\beta} + g_{\beta[\mu}R_{\nu]\alpha}\right) - Rg_{\alpha[\mu}g_{\nu]\beta} \\
&= \Lambda\left(g_{\alpha\mu}g_{\nu\beta} - g_{\alpha\nu}g_{\mu\beta}\right).
\end{aligned} \tag{1.201}$$

This form actually corresponds to a maximally symmetric solution. Calculate the Riemann tensor components making use of the Christoffel symbols and using this developed expression. Compare the two to obtain the value of the cosmological constant Λ. Does this value seem weird? What could this value tell us about this particular spacetime?

2. Consider the metric of a 4-sphere which has the following line element

[19] this is just a linear combination of the expressions which give the number of independent components of the Riemann and Ricci tensors respectively

$$ds^2 = r^2 \left(d\tau^2 + sin^2\tau \; d\Omega_3^2\right), \tag{1.202}$$

where $d\Omega_3^2 = d\psi^2 + sin\psi \; d\Omega_2^2$ is the metric of a 3-sphere and $d\Omega_2^2$ is the metric of a 2-sphere. This metric is diagonal and hence, we may apply the developed formalism.

3. Show that in $n = 2$ the Riemann tensor takes the following form

$$R_{\alpha\beta\mu\nu} = \frac{R}{2}\left(g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\nu}g_{\beta\mu}\right). \tag{1.203}$$

1.8.6 The Kulkarni-Nomizu Product

The Kulkarni-Nomizu product of any two covariant 2-tensors A and B is defined as follows

$$\begin{aligned}(A \owedge B)_{\mu\nu\alpha\beta} &= A_{\alpha\mu}B_{\nu\beta} + A_{\nu\beta}B_{\alpha\mu} - A_{\nu\alpha}B_{\beta\mu} - A_{\beta\mu}B_{\nu\alpha} \\ &= 2A_{\mu[\alpha}B_{\beta]\nu} + 2A_{\nu[\beta}B_{\alpha]\mu},\end{aligned} \tag{1.204}$$

which is a covariant 4 tensor that we can refer to as $\mathcal{P}_{\mu\nu\alpha\beta}$. An easier way to represent this and understand this product is to write it as the sum of the determinant of two separate matrices

$$(A \owedge B)_{\mu\nu\alpha\beta} = \begin{vmatrix} A_{\mu\alpha} & A_{\mu\beta} \\ B_{\nu\alpha} & B_{\nu\beta} \end{vmatrix} + \begin{vmatrix} B_{\mu\alpha} & B_{\mu\beta} \\ A_{\nu\alpha} & A_{\nu\beta} \end{vmatrix} \tag{1.205}$$

The advantage of the Kulkarni-Nomizu product is that when the tensors A and B are symmetric then the symmetries of the product is exactly the symmetries of the Riemann tensor, i.e., that are as follows
1. Antisymmetric in the first two indices, $(A \owedge B)_{\mu\nu\alpha\beta} = -(A \owedge B)_{\nu\mu\alpha\beta}$
2. Antisymmetric in the last two indices, $(A \owedge B)_{\mu\nu\alpha\beta} = -(A \owedge B)_{\mu\nu\beta\alpha}$
3. Symmetric in paired indices, $(A \owedge B)_{\mu\nu\alpha\beta} = (A \owedge B)_{\alpha\beta\mu\nu}$
4. Satisfies the Bianchi identity, $(A \owedge B)_{\mu\nu\alpha\beta} + (A \owedge B)_{\nu\alpha\mu\beta} + (A \owedge B)_{\alpha\mu\nu\beta} = 0$.

This seems like it would certainly help simplify the expression of the Weyl tensor which also possesses the weird antisymmetric parts such as the one shown above. Consider the Kulkarni-Nomizu product of two metric tensors $g_{\mu\nu}$ and $g_{\alpha\beta}$

$$\begin{aligned}(g \owedge g)_{\mu\nu\alpha\beta} &= g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha} + g_{\nu\beta}g_{\mu\alpha} - g_{\nu\alpha}g_{\mu\beta} \\ &= 2\left(g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha}\right) \\ &= 2g_{\mu[\alpha}g_{\beta]\nu},\end{aligned} \tag{1.206}$$

where we have used the fact that the metric tensor is symmetric. Using this result let us first simplify the antisymmetric piece attached to the Ricci scalar in the expression for the Weyl tensor

$$Rg_{\mu[\alpha}g_{\beta]\nu} = \frac{R}{2}\left(g \owedge g\right)_{\mu\nu\alpha\beta} = \frac{R}{2}\bar{g}_{\mu\nu\alpha\beta}, \tag{1.207}$$

where $\bar{g}_{\mu\nu\alpha\beta}$ is a short-hand notation for the product. Before considering the next piece, let's create a short-hand notation. Let $\bar{A}_{\mu\nu\alpha\beta}$ be the Kulkarni-Nomizu product of the covariant 2-tensor A and the metric tensor, i.e.,

$$\bar{A}_{\mu\nu\alpha\beta} = \left(A \owedge g\right)_{\mu\nu\alpha\beta}. \tag{1.208}$$

Now, substituting the tensor A in the above definition with the Ricci tensor, we obtain the simplification of the second antisymmetric piece as follows (using Ric instead of R to represent the Ricci tensor in this product in order to avoid confusion with the Ricci scalar)

$$\begin{aligned} g_{\mu[\alpha}R_{\beta]\nu} - g_{\nu[\alpha}R_{\beta]\mu} &= R_{\mu\alpha}g_{\nu\beta} + R_{\nu\beta}g_{\mu\alpha} - R_{\mu\beta}g_{\nu\alpha} - R_{\nu\alpha}g_{\mu\beta} \\ &= \left(Ric \owedge g\right)_{\mu\nu\alpha\beta} = \bar{R}_{\mu\nu\alpha\beta}. \end{aligned} \tag{1.209}$$

Note here that $\bar{R}_{\mu\nu\alpha\beta}$ is the representation of the Kulkarni-Nomizu product and is not to be confused with the Riemann tensor. We can now use these expressions to simplify the Weyl tensor as follows

$$\begin{aligned} W_{\mu\nu\alpha\beta} &= R_{\mu\nu\alpha\beta} - \tfrac{1}{(n-2)}\left(R_{\mu\alpha}g_{\nu\beta} + R_{\nu\beta}g_{\mu\alpha} - R_{\mu\beta}g_{\nu\alpha} - R_{\nu\alpha}g_{\mu\beta}\right) \\ &\quad + \tfrac{1}{(n-1)(n-2)}R\left(g_{\mu\alpha}g_{\beta\nu} - g_{\mu\beta}g_{\alpha\nu}\right) \\ &= R_{\mu\nu\alpha\beta} - \tfrac{1}{(n-2)}\left(Ric \owedge g\right)_{\mu\nu\alpha\beta} + \tfrac{R}{2(n-1)(n-2)}\left(g \owedge g\right)_{\mu\nu\alpha\beta} \\ &= R_{\mu\nu\alpha\beta} - \tfrac{1}{(n-2)}\bar{R}_{\mu\nu\alpha\beta} + \tfrac{R}{2(n-1)(n-2)}\bar{g}_{\mu\nu\alpha\beta}. \end{aligned} \tag{1.210}$$

We can further simplify by introducing the Schouten tensor which is defined as follows

$$S_{\mu\nu} = \frac{1}{(n-2)}\left(R_{\mu\nu} - \frac{R}{2(n-1)}g_{\mu\nu}\right), \tag{1.211}$$

which simplifies the Weyl tensor representation to the following expression

$$W_{\mu\nu\alpha\beta} = R_{\mu\nu\alpha\beta} - \left(S \owedge g\right)_{\mu\nu\alpha\beta}. \tag{1.212}$$

Exercise 5

1. Consider the global AdS_5 metric

$$ds^2 = -\left(1+\frac{r^2}{l^2}\right)dt^2 + \left(1+\frac{r^2}{l^2}\right)^{-1} dr^2 + r^2\left(d\theta^2 + sin^2\theta\ d\phi^2 + cos^2\theta\ d\psi^2\right). \tag{1.213}$$

a. Show that the Ricci scalar is $R = -20/l^2$.
b. Verify that the components of the Weyl tensor vanish identically.
c. Show that by replacing explicitly in the equation $W_{\mu\nu\alpha\beta} = 0$, that $\bar{g}_{\mu\nu} = e^{2\kappa(x)} g_{\mu\nu}$ (note that $\omega^2(x) = e^{2\kappa(x)}$ here) is a valid solution, where $\kappa(x)$ is a function of all the coordinates, i.e., $\kappa(x) = \kappa(t, r, \theta, \phi, \psi)$.
d. Find the components of the Schouten tensor.

2. Simplify the following products
a. $\mathcal{L}_K A_{\mu\nu} \owedge g_{\alpha\beta}$
b. $\mathcal{L}_K A_{\mu\nu} \owedge \mathcal{L}_K B_{\alpha\beta}$.
c. $[\mathcal{L}_X, \mathcal{L}_Y] A_{\mu\nu} \owedge \mathcal{L}_{[X,Y]} B_{\mu\nu}$.
d. $\mathcal{L}_X S_{\mu\nu} \owedge g_{\alpha\beta}$, where $S_{\mu\nu}$ is the Schouten tensor.

3. Consider the conformally flat metric $g_{\mu\nu} = \omega^2(x)\eta_{|mu\nu}$ on the domain $\mathbb{R}^{1+n}$ and set $\omega(x) = \frac{1}{1+\frac{K}{4}\sigma}$, where $\sigma = \eta_{\mu\nu}x^\mu x^\nu = -\left(x^0\right)^2 + \sum_{i=1}^n \left(x^i\right)^2$ and K is a constant.
a. Consider the case $K = 1$. Find the domain in $\mathbb{R}^{1+n}$ where the metric is defined. This metric is called the de Sitter metric.
b. Consider the case $K = -1$. Find the domain in $\mathbb{R}^{1+n}$ where the metric is defined. This metric is called the anti de Sitter metric.

4. The traceless Ricci curvature is defined as

$$\mathcal{R}^T_{\mu\nu} = R_{\mu\nu} - \frac{1}{n} R g_{\mu\nu}. \tag{1.214}$$

a. Show that the trace of this is indeed null so that it lives up to its name.
b. Show that the Schouten tensor can be written in terms of $\mathcal{R}^T_{\mu\nu}$ as follows

$$S_{\mu\nu} = \frac{1}{(n-2)}\mathcal{R}^T_{\mu\nu} + \frac{1}{2n(n-2)} R g_{\mu\nu}. \tag{1.215}$$

c. Consider the Poincaré patch of AdS_3 spacetime as given in 1.199. Find all the components of the Schouten tensor for the same.
d. Since AdS_3 spacetime is of dimension $n = 3$, all the components of the Weyl

tensor vanishes identically. Calculate the Cotton tensor. Is this spacetime locally conformally flat?
e. Show that the following holds

$$R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} = W_{\mu\nu\alpha\beta}W^{\mu\nu\alpha\beta} + \frac{4}{(n-2)}R_{\mu\nu}R^{\mu\nu} - \frac{2}{(n-1)(n-2)}R^2, \quad (1.216)$$

and show that this quantity, called the Kretschmann invariant, for the Poincaré patch of the AdS_3 metric is equal to $12/l^4$.

1.8.7 The Cotton Tensor

We have now established that in three-dimensions the Weyl tensor vanishes identically, and the Riemann and thew Ricci tensors have only 6 independent components. However, it is to be noted that not all 3-dimensional spacetimes are conformally flat. Hence, we require a replacement for the Weyl tensor, that would confirm conformal flatness of a spacetime and here is where the Cotton tensor comes into the picture. The Cotton tensor is a covariant 3-tensor that is given by

$$C_{\mu\nu\alpha} = D_\alpha \left(R_{\mu\nu} - \frac{1}{2(n-1)}Rg_{\mu\nu} \right) - D_\nu \left(R_{\mu\alpha} - \frac{1}{2(n-1)}Rg_{\mu\alpha} \right), \quad (1.217)$$

which is trace-free, i.e.,

$$g^{\mu\nu}C_{\mu\nu\alpha} = g^{\mu\alpha}C_{\mu\nu\alpha} = g^{\nu\alpha}C_{\mu\nu\alpha} = 0. \quad (1.218)$$

The skew symmetry in the second and third indices and Bianchi's first identity for the Cotton are as follows

$$C_{\mu\nu\alpha} = -C_{\nu\mu\alpha}, \quad C_{\mu\nu\alpha} + C_{\nu\alpha\mu} + C_{\alpha\mu\nu} = 0. \quad (1.219)$$

We can now make an important observation. Consider the covariant derivative of $C_{\mu\nu\alpha}$ and $C_{\nu\alpha\mu}$

$$D_\alpha C_{\mu\nu\alpha} = D_\alpha C_{\nu\alpha\mu}, \quad (1.220)$$

from which, and the covariant derivative of the first Bianchi identity, we can immediately conclude that

$$D_\alpha C_{\alpha\mu\nu} = 0. \quad (1.221)$$

We now want to establish a relation between the Weyl tensor and the Cotton tensor. To do this we connect the two via the Schouten tensor. First we express the Cotton tensor in terms of the Schouten tensor as follows

$$\begin{aligned} C_{\mu\nu\alpha} &= (n-2)\left(D_\alpha P_{\mu\nu} - D_\nu S_{\mu\alpha}\right) \\ &= 2\,(n-2)\,D_{[\alpha}S_{\nu]\mu}, \end{aligned} \tag{1.222}$$

where we have made use of the fact that the Schouten tensor is symmetric in its indices (since it's a linear combination of the Ricci tonsor and the metric tensor which are symmetric). Now, taking the covariant derivative of the Weyl tensor $W_{\beta\mu\nu\alpha}$, it is easy to obtain the following expression (do this! hint: take the trace of the Bianchi identity)

$$D_\beta W^\beta_{\mu\nu\alpha} = 2\,(n-3)\,D_{[\alpha}S_{\nu]\mu}. \tag{1.223}$$

Thus, we finally obtain the relation

$$D_\beta W^\beta_{\mu\nu\alpha} = \frac{n-3}{n-2}C_{\mu\nu\alpha}. \tag{1.224}$$

This implies that for $n \geq 4$, the assumption that the manifold is locally conformally flat, i.e., $W = 0$, the Cotton tensor is identically zero also in this case, but this is only a necessary condition. Thus, a manifold can be referred to as locally conformally flat if and only if for $n \geq 4$, the Weyl tensor vanishes and for $n = 3$, the Cotton tensor vanishes. For $n < 3$, the Cotton tensor is identically zero. What we have stated here is called the Weyl-Schouten theorem. Formally put, this states that

Theorem 1.13. *Weyl-Schouten*
A Manifold $\mathcal{M}$ of dimension n is conformally flat if and only if

- $n = 2$,
- $n = 3$ *and the Cotton tensor vanishes, i.e.,* $C_{\mu\nu\alpha} = 0$, *or*
- $n \geq 4$ *and the Weyl tensor vanishes, i.e.,* $W_{\mu\nu\alpha\beta} = 0$

Example 1.14. Conformal flatness in the global AdS spacetime and a little more Consider a spacetime $\mathbb{R}^{2,2}$, i.e., with coordinates $\{x^0, x^1, x^2, x^3\}$. This is a metric with signature $(--++)$ whose line element takes the form

$$ds^2 = -\left(dx^0\right)^2 + \left(dx^1\right)^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2. \tag{1.225}$$

AdS_3 is the submanifold of this space which is given by the constraint

$$-\left(x^0\right)^2 - \left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 = -l^2, \tag{1.226}$$

where l is called the AdS radius. The following parameterizations satisfy the constraint equation

$$x^0 = l\ cosh\psi\ cos\left(\tfrac{t}{l}\right), \quad x^1 = l\ cosh\psi\ sin\left(\tfrac{t}{l}\right),$$
$$x^2 = l\ sinh\psi\ cos\phi, \quad x^3 = l\ sinh\psi\ sin\phi, \tag{1.227}$$

where $t \in [0, 2\pi l)$, $\phi \in [0, 2\pi)$, and $\psi \in [0, \infty)$. It is easy to show that the induced metric reads

$$ds^2 = l^2\left(-cosh^2\psi\ dt^2 + d\psi^2 + sinh^2\psi\ d\phi^2\right). \tag{1.228}$$

This is the AdS_3 metric in global coordinates (with $x^0 = t$, $x^1 = \psi$, $x^2 = \phi$). We now need to check if this is conformally flat. It is easy to check that all the components of the Weyl tensor vanishes identically (do this!). But we need to tread carefully here since we are in $n = 3$ where the vanishing of the Weyl tensor is deceptive. In accordance with the Weyl Schouten theorem, we can call this spacetime conformally flat if and only if the Cotton tensor vanishes. So let's check for this. We first compute the Schouten tensor, which in $n = 3$, reads

$$S_{\mu\nu} = R_{\mu\nu} - \frac{1}{4}Rg_{\mu\nu} = -\frac{1}{2}\begin{pmatrix} -cosh^2\psi & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & sinh^2\psi \end{pmatrix}. \tag{1.229}$$

Now, using the Cotton-Schouten relation derived in 1.222, we find the components of the Cotton tensor to be

$$\begin{aligned}
&C_{ttt} = C_{\psi\psi\psi} = C_{\phi\phi\phi} = 0, \\
&C_{t\psi\phi} = 0 \Rightarrow C_{\psi t\phi} = 0,\ C_{\phi t\psi} = 0, \\
&C_{\psi t\psi} = C_{t\psi\psi} = C_{\psi\psi t} = 0,\ C_{\psi\phi\psi} = C_{\phi\psi\psi} = C_{\psi\psi\phi} = 0, \\
&C_{t\psi t} = -D_\psi S_{tt} = -\partial_t S_{tt} + 2\Gamma^t_{\psi t}S_{tt} = 0,\ \Gamma^t_{\psi t} = tanh\psi, \\
&C_{\phi\psi\phi} = -D_\psi S_{\psi\psi} = -\partial_\psi S_{\phi\phi} + 2\Gamma^\phi_{\psi\phi}S_{\phi\phi} = 0,\ \Gamma^\phi_{\psi\phi} = coth\psi.
\end{aligned} \tag{1.230}$$

We find that all the components of the Cotton tensor do indeed vanish and hence, AdS_3 is conformally flat. This metric can also be expressed in Poincaré coordinates (t, x, z) with the following parameterization

$$x^0 = \frac{lt}{z},\ x^1 = \frac{l^2}{z},\ x^2 + x^3 = \frac{lx}{z},\ x^2 - x^3 = \frac{-t^2 + x^2 + z^2}{z}. \tag{1.231}$$

The resulting metric is called the Poincaré patch of the AdS_3 (see 1.199) and this was something that was solved as an example problem previously. AdS_3 is a maximally symmetric spacetime. The global AdS_3 metric can also be parameterized by representing the it as follows

$$g = \frac{1}{l^2}\begin{pmatrix} x^0 - x^2 & -x^1 + x^3 \\ x^1 + x^3 & x^0 + x^2. \end{pmatrix} \tag{1.232}$$

We can here note two properties of this type of representation. Firstly, all the entries of the matrix are real numbers, i.e., $x^0, x^1, x^2, x^3 \in \mathbb{R}$ and secondly, the determinant of the matrix is unity, i.e.,

$$\begin{aligned} det|g| &= \tfrac{1}{l^2}\left(\left(x^0\right)^2 - \left(x^2\right)^2 - \{\left(x^3\right)^2 - \left(x^1\right)^2\}\right) \\ &= cosh^2\psi\left(sin^2\left(\tfrac{t}{l}\right) + cos^2\left(\tfrac{t}{l}\right)\right) - sinh^2\psi\left(sin^2\phi + cos^2\phi\right) = 1. \end{aligned} \tag{1.233}$$

These observations were important to make since we can now compare this to a symmetry group. The symmetry group which possesses the very same properties is $SL\left(2, \mathbb{R}\right)$ which is read as special linear group of degree 2 over the field of real numbers. It is defined as the group of 2×2 matrices with entries from the field of real numbers and a unit determinant, under matrix multiplication. It is represented as follows

$$SL\left(2, \mathbb{R}\right) \equiv \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,\ b,\ c,\ d \in \mathbb{R},\ ad - bc = 1\right\} \tag{1.234}$$

To observe the type of curvature singularity in global AdS_3, we compute the Kretschmann scalar. Since the Weyl tensor vanishes, the expression for the invariant gets simplified, from 1.216 with $n = 3$ we have

$$K = 4R_{\mu\nu}R^{\mu\nu} - R^2, \tag{1.235}$$

where the Ricci tensor components and the Ricci scalar read

$$R_{\mu\nu} = \begin{pmatrix} 2cosh^2\psi & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2sinh^2\psi \end{pmatrix}, \quad R^{\mu\nu} = g^{\mu\mu}g^{\nu\nu}R_{\mu\nu} = \tfrac{1}{l^4}\begin{pmatrix} 2sech^2\psi & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2cosech^2\psi \end{pmatrix},$$

$$R = -\tfrac{6}{l^2}. \tag{1.236}$$

Substituting this into 1.235, we obtain

$$K = \frac{12}{l^2}, \tag{1.237}$$

which is a constant and this implies that there exists no curvature singularity.

Exercise 6

1. Simplify the following (here X is a Killing vector)
a. $\left[D_\alpha, \mathcal{L}_X\right] S_{\mu\nu}$, where $S_{\mu\nu}$ is the Schouten tensor.

b. $[D_e, \mathcal{L}_X]\, R_{abcd}$, where R_{abcd} is the Riemann tensor.
c. $[D_\alpha, \mathcal{L}_X]\, C_{\mu\nu\beta}$, where $C_{\mu\nu\beta}$ is the Cotton tensor.
d. $[D_e, \mathcal{L}_X]\, W_{abcd}$, where W_{abcd} is the Weyl tensor.

2. Consider two conformally equivalent metrics $g_{\mu\nu}$ and $\hat{g}_{\mu\nu} = \omega^2(x) g_{\mu\nu} = e^{2\kappa(x)} g_{\mu\nu}$. Now, consider scaling all lengths by a constant factor $l > 0$ so that $\omega(x) = l$ and $\kappa(x) = ln(l)$. Show that the following hold
a. $\hat{\Gamma}^\alpha_{\mu\nu} = \Gamma^\alpha_{\mu\nu}$.
b. $\hat{R}_{\mu\nu\alpha\beta} = l^2 R_{\mu\nu\alpha\beta}$, $\hat{R}_{\mu\nu} = R_{\mu\nu}$, and $\hat{R} = l^{-2} R$.
c. $\hat{S}_{\mu\nu} = S_{\mu\nu}$ and $\hat{C}_{\mu\nu\alpha} = C_{\mu\nu\alpha}$.

3. Consider a spacetime with coordinates $\{u, v, x, y\}$ given by the metric

$$ds^2 = f(u)^2 dx^2 + g(u)^2 dy^2 - 2dudv. \tag{1.238}$$

This describes a plane-fronted gravitational wave moving in the u direction where $u = t - z$ and $v = t + z$. Perform your calculations in the $\{u, v, x, y\}$ coordinate system.
a. Show that this metric has zero scalar curvature.
b. Show that this metric has a null Kretschmann scalar.
c. Show that the only non-zero component of the Schouten tensor is S_{uu}.
d. Use this information to show that the all the Cotton tensor components vanish identically.
e. Is this a conformally flat spacetime?

4. Consider a field defined by $\varphi(x)$ which satisfies the massless Klein-Gordon equation in flat spacetime which reads

$$\Box \varphi(x) = 0,$$

where $\Box \equiv \eta^{\mu\nu} \partial_\mu \partial_\nu$ is the d'Alembert operator. Now consider a metric $g_{\mu\nu}$ which is related to the flat space metric by a conformal transformation $g_{\mu\nu} = e^{2\kappa(x)} \eta_{\mu\nu}$.
a. Show that the expression for Ricci scalar in n-dimensional spacetime for a metric which is conformally-related to the Minkowski metric is

$$R = 2(1-n) e^{-2\kappa} \Box \kappa - (2-n)(1-n) e^{-2\kappa} \eta^{\alpha\beta} \partial_\alpha \kappa \partial_\beta \kappa.$$

b. Show that the transformed field $\bar{\varphi}(x)$ is related to the original field by $\bar{\varphi} = e^{\lambda\kappa(x)} \varphi(x)$ and satisfies the Klein-Gordon equation in n-dimensional space which reads

$$g^{\mu\nu} D_\mu D_\nu \bar{\varphi}(x) - \omega R \bar{\varphi}(x) = 0, \tag{1.239}$$

where $\lambda = \frac{2-n}{2}$ and $\omega = \frac{1}{4}\frac{n-2}{n-1}$.

c. Write down the Klein-Gordon equation in $n = 4$ and find it's solution.

1.9 Hypersurfaces

In a 4-dimensional spacetime manifold, a *hypersurface* is a 3-dimensional submanifold that can be either spacelike, timelike, of null. A particular hypersurface Σ is selected by giving parametric equations of the form

$$x^\mu = x^\mu \left(y^\alpha\right), \tag{1.240}$$

where y^α $(\alpha = 1, 2, 3)$ are coordinates intrinsic to the hypersurface. Consider a 2-sphere in a 3-dimensional flat space, it can be described either by $\zeta(x, y, z) = x^2 + y^2 + z^2 - r^2 = 0$, where r is the radius of the sphere, or by $x = r\ sin\theta\ cos\phi$, $y = r\ sin\theta\ sin\phi$, and $z = r\ sin\phi$, where θ and ϕ are intrinsic coordinates. In general, a hypersurface is a manifold of dimension $n - 1$ embedded in n-dimensional space. Hence, a hypersurface is therefore the set of solutions to a single equation

$$\Lambda \left(x_1, ..., x_n\right) = 0. \tag{1.241}$$

1.9.1 Normal on Hypersurfaces

The normal of a hypersurface Σ whose equation is of the form $\Sigma \equiv r - C = 0$, where C is a constant is

$$n_\mu = \Sigma_{,\mu} = (0, 1, 0, 0). \tag{1.242}$$

At any point on a generic hypersurface we can introduce a locally inertial frame, and rotate it in such a way that the components of the normal vector are

$$n^a = \left(n^0, n^1, 0, 0\right) \ \&\ n_a n^a = \left(n^1\right)^2 - \left(n^0\right)^2. \tag{1.243}$$

Consider a vector v^a tangent to the surface at the same point. v^a must be orthogonal to n^b

$$n_a v^a = -n^0 v^0 + n^1 v^1 = 0, \tag{1.244}$$

which implies

$$\frac{v^0}{v^1} = \frac{n^1}{n^0}. \tag{1.245}$$

Thus,

$$v^a = \zeta \left(n^1, n^0, \alpha, \beta\right), \tag{1.246}$$

where α, β, ζ are constant and arbitrary. Now, the norm of the tangent vector is

$$v_a v^a = \zeta^2 \left(-\left(n^1\right)^2 + \left(n^0\right) + \left(\alpha^2 + \beta^2\right)\right) = \zeta^2 \left(-n_a n^a + \left(\alpha^2 + \beta^2\right)\right). \tag{1.247}$$

We have the following cases:
1. If $n_\mu n^\mu < 0$, the hypersurface is called spacelike, and the tangent vector is necessarily a spacelike vector.
2. If $n_\mu n^\mu = 0$, the hypersurface is called null, and the tangent vector can be spacelike or null.
3. If $n_\mu n^\mu > 0$, the hypersurface is called timelike, and the tangent vector can be spacelike or null.
Consider a point q on a surface $\Sigma = 0$. If the surface is spacelike, the tangent vectors of the surface lie all outside the lightcone in q. Therefore, a particle passing through the point q, whose velocity vector lies inside the cone, can cross the surface only in one direction. If the surface is null, the situation is nearly the same: the tangent vectors to the surface lie inside to the lightcone in q, or are tangent to it, thus a particle can cross the surface in one direction only. It the surface is timelike, some tangent vectors of the surface are inside the cone, some others are outside, i.e. the surface cuts the cone, and then a particle passing through q can cross the surface in both directions.

Example 1.15. Let us consider the AdS_3 metric in global coordinates 1.228. We now consider a hypersurface Σ and want to find the metric of the hypersurface i.e., the induced metric on the hypersurface. Let a vector $\partial_\mu f$ be the normal to a hypersurface. If the hypersurface is non-null, we can now introduce a unit normal n_{mu} such that, as previously mentioned, $n^\mu n_\mu = \epsilon$ takes different values based on the whether the hypersurface is spacelike or timelike. We can now write down an expression of n_μ as follows

$$n_\mu = \frac{\epsilon \partial_\mu f}{|g^{\alpha\beta} \partial_\alpha f \partial_\beta f|^{\frac{1}{2}}}, \tag{1.248}$$

Note that when Σ is null, we have $g^{\alpha\beta} \partial_\alpha f \partial_\beta f = 0$. Thus, under the special case of a null hypersurface we define the normal vector $\bar{n}_\mu$ as follows

$$\bar{n}_\mu = -\partial_\alpha f. \tag{1.249}$$

Now, as an example consider the AdS_3 metric in global coordinates 1.228. Let us consider a spacelike hypersurface (where $\psi = \Psi =$ constant) where the two-dimensional induced metric $\gamma_{\alpha\beta}$ reads

$$\gamma_{\mu\nu} = \begin{pmatrix} -l^2 cosh^2\Psi & 0 \\ 0 & l^2 sinh^2\Psi, \end{pmatrix} \tag{1.250}$$

and the normal vector to this hypersurface with $f = \psi$ is

$$\begin{aligned} n_t &= \frac{\partial_t\psi}{|g^{tt}\partial_t\psi\partial_t\psi|^{\frac{1}{2}}} = 0, \\ n_\psi &= \frac{\partial_\psi\psi}{|g^{\psi\psi}\partial_\psi\psi\partial_\psi\psi|^{\frac{1}{2}}} = \frac{1}{\left|\sqrt{g^{\psi\psi}}\right|} = l, \\ n_\phi &= \frac{\partial_\phi\psi}{|g^{\phi\phi}\partial_\phi\psi\partial_\phi\psi|^{\frac{1}{2}}} = 0, \\ &\Rightarrow n_\mu = l\delta^\psi_\mu = (0, l, 0) \end{aligned} \tag{1.251}$$

Similarly, we find that if Σ is a timelike hypersurface the normal then takes the form

$$n_\mu = -lcosh^2\psi\ \delta^t_\mu = \left(-lcosh^2\psi, 0, 0\right). \tag{1.252}$$

Exercise 7

1. Consider the Bertotti-Kasner metric which has the following line element with $\Lambda > 0$

$$ds^2 = -dt^2 + e^{2t\sqrt{\Lambda}}dr^2 + \frac{1}{\Lambda}\left(d\theta^2 + sin^2\theta\ d\phi^2\right). \tag{1.253}$$

a. Show that this spacetime is not conformally flat.
b. Show that the Ricci and the Kretschmann scalar are 4Λ and $8\Lambda^2$ respectively.
c. Consider a spacelike hypersurface. Find the normal to the hypersurface and the induced metric.
d. Is the induced metric conformally flat?
e. Show that the Ricci and the Kretschmann scalars of the induced metric are 2Λ and $4\Lambda^2$ respectively.
f. Consider a timelike hypersurface and repeat the exercise.

2. Find the induced metric on a timelike hypersurface to the BTZ metric 3.57 and find the normal vector to the hypersurface.

2

Einstein's Field Equations

2.1 Newton v Einstein: The missing Sun

Before we divulge into the details and study the field equations, let us take a step back to put things into perspective. We must now ask ourselves an important question, why a different theory? I mean we are all cool with Newton's stuff, and as if his theories weren't enough we are moving towards a much more complicated one. We are in a very tricky situation now, upon studying the field equations one may either give up complaining that the math is just too much or one may ignore the math for a moment and focus on understanding the elegance of the equations. Of course, my aim is to try and stimulate the latter. I strongly believe that in order to understand a theory born out of the power of sheer imagination it is our responsibility to try and appreciate it using our own.

Let's perform a *Gedankenerfahrung*, which might possibly explain the need for a new theory. Imagine that for the moment both Newton and Einstein are alive (of course they are in every physicist's heart!) and that they are participants in a debate hosted by Wolfgang Pauli (of course this weird choice comes with a reason).

Since its obvious that there would arise tensions (strictly egoistical) in a room of physicists, the argument the two would have is almost inevitable. Let's not take sides as of yet and try and review what each of them have to say. In Newton's version, he states that there is a potential (let's call it ϕ) everywhere in space and it varies from place to place. This variation of the potential, or better, the differential variation of the potential in space gives rise to a field, i.e., $E = \nabla\phi\,(r)$ (I'll let you know what ∇ stands for later). The field instructs the particles how to move and decides their acceleration, i.e., $F = ma = -m\nabla\phi\,(r)$. Hence, we obtain the following relation

$$a = -\nabla\phi\,(r)\,. \tag{2.1}$$

Now, what equation would instruct the field and tell it how to behave? We are to find the differential change in the field in order to predict its characteristics and the differential change in space (volume differentiation) is known as divergence and is represented as follows

$$\nabla E = \left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}\right) E. \tag{2.2}$$

We are to bear in mind that the field E unlike the potential ϕ is a vector quantity, and in order solve for the same we make use of *Gauss theorem* which states the following

$$\int \nabla E dx\ dy\ dz = \int d\nu\ E_n, \tag{2.3}$$

were ν is the surface area and E_n is the normal vector to the surface area. In one-dimension we can write

$$\begin{aligned} &\textstyle\int d\nu\ E_n = \left(-\frac{Gm}{r^2}\right)\int dv = -\frac{Gm}{r^2}4\pi r^2 = -4\pi Gm \\ &\textstyle\int \nabla E\ d^3x = \int d\nu\ E_n = -4\pi Gm, \end{aligned} \tag{2.4}$$

thus,

$$\nabla E = -4\pi G\left(\frac{\Delta m}{\Delta V}\right) = -4\pi G\rho = \nabla^2\phi. \tag{2.5}$$

Mass instructs the field how to curve via the density, ρ, this equation is known as the *Poisson equation.* Thus, Newton would argue that since his theory makes sense, if the Sun were to suddenly vanish, then the Earth would immediately be flung in a direction tangential to its orbit almost analogous to how a stone attached to a string would fly in a tangential path when released from rotations about an axis. After this rather elegant conclusion, Pauli would be quite convinced with the reasoning and one would require a counter of epic proportions in order to even sound sane. In Einstein's version, he starts off by stating that Newton is simply wrong and that he is ignoring a very important concept. In Einstein's version, he states that it is well known that light takes time to travel from the Sun and reach Earth, and since the Sun is approximately 149.27 million kilometres away from us, it would take close to 8.3 minutes for a ray of light to reach us from the Sun (Thus the Sun you observe while reading this book is what it was 8.3 minutes ago, furthermore, all that we see around us is the past!). He goes on to argue that when light itself (the fastest thing known to us) takes time to completely its journey, how can gravity be any faster? Thus, if the Sun were to suddenly go missing, we would get to know the same only after 8.3 minutes, and some $8.3 + x$ minutes later, the Earth

would fly off tangentially. Thus, Pauli would announce Einstein the winner and let Newton know that he was "not even wrong"!
In General Relativity, the equation: $a = -\nabla\phi(r)$, is replaced by a statement. The statement tells us that once we gain knowledge of the geometry (i.e. g_{00} here), the rule is that particles move on spacetime geodesics. It is quite interesting how this ceases to be true.

2.2 Stress-energy tensor: the messenger of mass

Space tells mass how to move and mass tells space how to curve. Prior to observing the curvature, we are to probe for a quantity that will enable us to understand how much mass-energy is present in a unit volume. This quantity is the *stress-energy tensor*. Spacetime possesses multiple contributions of 4-momentum from all sorts of particles from different fields. The contributions also pour in from the electromagnetic fields, neutrino fields, etc. Thus, we can view spacetime as an ocean of 4-momentum and the flow of water in the ocean is described by the stress-energy tensor **T**. Since **T** is a tensor, it has a slot machine definition. The stress-energy tensor program is a linear, and symmetric slot machine which accepts two vector inputs, i.e., $T(\underbrace{\quad}_{IP_1},\underbrace{\quad}_{IP_2})$. The output, for a given input, of **T** are as follows
1. Input a 4-velocity v of Mr. Absolute Zero and leave the other space sans any input. This produces the following output

$$T(\ \ ,v) = T(v,\ \ ) = -\left\{\frac{dp}{dV}\right\}, \tag{2.6}$$

where, the RHS denotes the density of 4-momentum, i.e., the 4-momentum per unit volume as measured in Mr. Absolute Zero's local Lorentz frame. In the component form, we have the following expression for him with 4-velocity u^μ,

$$T^\mu_\nu u^\nu = T^\mu_\nu v^\nu = -\left(\frac{dp^\mu}{dV}\right). \tag{2.7}$$

2. Now, enter the 4-velocity of Mr. Absolute Zero as the second input and enter any arbitrary unit vector n as the second input. The program displays the following output

$$T(u,n) = T(n,u) = -n\frac{dp}{dV}, \tag{2.8}$$

where, the RHS denotes the component of the 4-momentum density, as measured in Mr. Absolute Zero's Lorentz frame. In the component form, we have the following expression

$$T_{\mu\nu}v^{\mu}n^{\nu} = T_{\mu\nu}n^{\mu}v^{\nu} = -n_{\alpha}\frac{dp_{\alpha}}{dV}, \tag{2.9}$$

3. Enter the 4-velocity of Mr. Absolute Zero for either of the inputs.

$$T(v,v) = \{\text{density of mass} - \text{energy measured in his Lorentz frame}\}. \tag{2.10}$$

4. Now, select two spacelike basis vectors for Mr. Absolute Zero, in his Lorentz frame, e_i and e_j. Input the basis vectors to the tensor program T. The output is the i,j component of the stress as measured by Mr. Absolute Zero, it can be expressed as follows

$$\begin{aligned}\mathrm{T}_{ij} &= T(e_i,e_j) = T_{ji} = T(e_j,e_i)\\ &= \left\{\begin{array}{l}\mathrm{i-component\ of\ force\ acting\ from\ side\ x^j} - \delta \mathrm{\ to\ side\ x^j}+\\ \delta,\ \mathrm{across\ a\ unit\ surface\ area\ with\ perpendicular\ direction\ e_j}\end{array}\right\}\\ &= \left\{\begin{array}{l}\mathrm{j-component\ of\ force\ acting\ from\ side\ x^i} - \delta \mathrm{\ to\ side\ x^i}+\\ \delta,\ \mathrm{across\ a\ unit\ surface\ area\ with\ perpendicular\ direction\ e_i}\end{array}\right\}.\end{aligned} \tag{2.11}$$

Now, since you know how to construct the stress-energy tensor for Mr. Absolute Zero, lets probe further into its physical significance. A stress-energy tensor $T^{\alpha\beta}$ is the flux of the α^{th} component of 4-momentum across a surface of constant x^{β}, thus, $T^{\mu 0}$.
a. T^{00}: The flux of the 0^{th} component of 4-momentum across the time surface,i.e., it indicates the density of energy.

b. $T^{k0} = T^{0k}$: Energy flux across the surface at constant x^k, i.e., indicates the flow of energy along x^k.

c. $T^{kd} = T^{dk}$: Flux of k-momentum across d-surface, i.e., indicates stress.

2.3 Conservation: What Does it Really Mean?

We know from Maxwell's electrodynamic equation that the derivative of the *Faraday tensor* is proportional to the 4-current[1], this can be expressed as

$$d\,{}^{*}F = 4\pi\,{}^{*}J, \tag{2.12}$$

[1] i.e., the number of Maxwell tubes that end in an elementary volume is equal to the amount of electric charge present in that volume

where $^*J$ and $^*F$ is the 4-current dual and the *Faraday dual* (i.e. the *Maxwell tensor*) respectively, and as previously discussed the Maxwell tensor is a 2-form while its exterior derivative, the 4-current dual is a 3-form. Thus, we could say that in a region filled with charge, Maxwell's tubes take the origin and the density is described by the 3-form $^*J$. In general, the 4-current tensor has four components, with the first one indicating charge density and the other three indicating current density. With this picture in mind how does one define charge conservation? If you are to conserve a particular charge, does it mean that you draw a boundary over the distribution and prohibit it from moving? Or does it mean that you transfer that specific charge density into an imaginary box of finite volume and move it to infinity?

If we follow the first definition, then we are to still deal with the charge density present in the room whenever we do physics but pretend as though it doesn't exist (leading to an awkward situation). The second case seems legit, right? Nah, not really because if we move the charge density to infinity it would mean that we are moving the charge box over a time interval, leading to the creation of a current and as it passes via different areas it takes the form of current density (leading to a messy situation). So, is the fate of conservation bound to be awkward or messy? The answer is neither, we are to change our perspective a bit here. Let's start viewing the charge density from the perspective of the field its present in. Let's define the field and "connect it" to the source (the charge here) in such a way that the conservation of the source shall be an automatic consequence of some condition imposed on the field. Assume a hypercube in the 4-D spacetime to be the volume element in which an event occurs. There is a mathematical theorem that states that the two-dimensional boundary of the three-dimensional boundary of the four-dimensional cube is zero.

Rather than divulging into the mathematics, let's try and understand this intuitively—we have previously observed the unique working of the exterior derivative, let me remind you, a 1-form α which is a gradient $\alpha = df$, must satisfy $d\alpha = 0$ (since if α is a 1-form then f is a 0-form because its exterior derivative is equal to α; now, $d\alpha = ddf$, which is zero). It is to be noted that not all 1-forms follow this relation. If a 1-form does α satisfies $d\alpha = 0$, then it follows that *locally* it has the form $\alpha = df$ for some f.This is an instance of the *Poincaré lemma*, which says that if a f-form γ satisfies $d\gamma = 0$, then *locally* γ has the form $\gamma = d\epsilon$, for some $(f-1)$-form ϵ. Now, consider a f-form β in a coordinate patch, with coordinates $x^1, x^2, x^3, \ldots,\ x^n$, there exists an asymmetrical set of components $\beta_{p\ldots u} (= \beta_{[p\ldots u]})$ to represent β. This representation can be expressed as follows

$$\beta = \sum \beta_{p\ldots u} dx^p \wedge \cdots \wedge dx^u. \tag{2.13}$$

The exterior derivative of a f-form is a $(f+1)$-form that is written as $d\beta$, and which has components

$$(d\beta)_{qp\ldots u} = \frac{\partial}{\partial x^{[q}} \beta_{p\ldots u]}, \tag{2.14}$$

the notation is all messed up due to the antisymmetrization which extends over all $(f+1)$ indices, including the one on the derivative symbol. Thus, we can formulate the *fundamental theorem of exterior calculus*. For a f-form ζ as follows

$$\int_A d\zeta = \int_{\partial A} \zeta, \tag{2.15}$$

where A is some *compact* $(f+1)$-dimensional *oriented* region whose oriented f-dimensional boundary (which is also compact) is ∂A.
What is going on here? Well let's see if we can understand the physical meaning of the integral and try and do something with it. First off, the meaning of *compact* here is a region with a specific property that any infinite sequence of points lying in A must *accumulate* at some arbitrary point that exists within A. An *accumulation point* z has a specific property associated to it that every open set in A which contains z ,must also contain members of an infinite sequence, such that the points of the sequence get closer and closer to the point z, without any limit. Examples of compact surfaces include the surfaces of a 2-sphere and that of a torus. However, the Euclidean plane is a non-compact surface. We now move on to the other term, *oriented*, which refers to the allotment of a consistent sign convention at every point of A. For a 0-manifold, this orientation allots a positive or a negative sign to each point. For a 1-manifold, this orientation associates a direction to the curve via a symbol (arrow). For a 2-manifold, this orientation is the circulation of the tangent vector at each point. A great example for a non-orientable surface is the *Mobius strip*. Thus, the boundary ∂A of a compact oriented $(f+1)$-dimensional region A consists of those points of A that do not lie within itself. If A is well-behaved, then ∂A is a compact oriented f-dimensional region (which might be possibly empty). Its boundary $\partial\partial A$ is empty (and thus $\partial\partial = \partial^2 = 0$, which makes sense because we know that $dd = d^2 = 0$). Examples of this "phenomenon" include the boundary of the closed unit disc in the complex plane is the unit circle; the boundary of the 2-sphere is empty, etc.

Similarly, taking the example of a cube in 3-dimensions−the boundary of a cube is its faces (2-dimensional), and the boundary of the each of the faces are composed of four edges (1-dimensional), and all edges are used up in uniting one face to another (i.e. no edges are left out). We can conclude that the 1-dimensional boundary of the 2-dimensional boundary of the 3-dimensional cube is identically zero, i.e., the boundary of a boundary vanishes. We can extend this concept to a hypercube and state that the 2-dimensional boundary of a 3-dimensional boundary of a 4-dimensional cube is identically zero.

From Maxwell's equations (presented here without proof), it is known that $d\,{}^*F = 4\pi\,{}^*J$, describes the features of the field ${}^*F$. Here, the equations are expressed in a coordinate-free geometric form where $\mathbf{F}$ is called the *Faraday tensor* is a mathematical object that describes the electromagnetic field in spacetime, and ${}^*F$ is called the dual of the tensor[2]. The Faraday tensor is associated with the antisymmetric matrix of six electromagnetic fields as follows (note the difference between $\mathbf{F}$ and ${}^*\mathbf{F}$)

$$\begin{aligned} F^{\mu\nu} &= \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \\ {}^*F^{\mu\nu} &= \begin{pmatrix} 0 & -E_z & +E_y & +E_x \\ +E_z & 0 & -E_x & -B_y \\ -E_y & -B_z & 0 & -B_z \\ +B_x & +B_y & +B_z & 0 \end{pmatrix} \end{aligned} \tag{2.16}$$

Observing the above matrices we can arrive at a possible equation for the relation between $\mathbf{F}$ and ${}^*\mathbf{F}$ as

$${}^*F_{\alpha\beta} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{\mu\nu}, \tag{2.17}$$

where $\epsilon_{\mu\nu\alpha\beta}$ is called the *Levi-Civita* symbol in 4-dimensions. It is defined as follows

$$\epsilon_{\mu\nu\alpha\beta} = \left\{ \begin{array}{l} +1, \text{ if } (\mu,\nu,\alpha,\beta) \text{ is an even permutation of } (1,2,3,4), \\ -1, \text{ if } (\mu,\nu,\alpha,\beta) \text{ is an odd permutation of } (1,2,3,4), \\ 0, \text{ otherwise} \end{array} \right\} \tag{2.18}$$

This form of Maxwell's equation is useful as it contains, within itself, the electrostatic and the electromagnetic equations, i.e., $(\nabla\mathbf{E} = 4\pi\rho)\,,\left(\frac{\partial\mathbf{E}}{\partial t} - \nabla\times\mathbf{B} = -4\pi\mathbf{J}\right)$ respectively. Observe that the 4-current tensor contains the elements of the RHS of either equations, i.e., $J^\alpha = \left(J^0, J^1, J^2, J^3\right) = (\rho, \{J\})$ and as explained previously, since the Faraday tensor describes the electromagnetic field, it must hence contain the electric and magnetic fields as its components. Thus, a single geometric law is used to express the two Maxwell's equations as follows

$$F^{\mu\nu}_{\ ,\nu} = 4\pi J^\mu, \tag{2.19}$$

an equivalent formalism of the coordinate-independent law would be to express the equation as $\nabla\mathbf{F} = 4\pi\mathbf{J}$. The conservation of the source, i.e., $d^*J = 0$, is

[2] this is not to be confused with the dual lemma presented in introductory chapters. Here, the meaning of dual is associated to the *Hodge star operator*.

a direct consequence of the identity $dd^*F = 0$. Thus, conservation is a direct consequence of the vanishing of a boundary of a boundary. Conservation literally demands no creation or destruction of the source inside the 4-dimensional cube. It is also to be noted that the integral of an event leading to a creation, i.e. of d^*J, over this 4-dimensional region is to be zero. Thus, conservation means zero creation of charge in a 4-dimensional region. Mathematically speaking, the application of the exterior derivative to either side of the second Maxwell equation, i.e., $d\ {}^*F = 4\pi\ {}^*J$, and using the fact that $d^2 = 0$, we can deduce that the 4-current $\mathbf{J}$ satisfies the vanishing boundary condition $d^*J = 0, \quad \text{or } \nabla_\alpha J^\alpha = 0$ since

$$dd^*F = 4\pi d^*J = 0, \tag{2.20}$$

is true. This vanishing divergence of the 4-current yields a conservation law for electric charge. From the fundamental theorem of exterior calculus, we can write the conservation law as follows

$$\int_A d^*J = \int_{\partial A} {}^*J = 0, \tag{2.21}$$

the same law expressed in the "boundary of a boundary vanishes" language is the following

$$4\pi \int_{\partial A} {}^*J = \int_{\partial A} d^*F = \int_{\partial\partial A} {}^*F = 0. \tag{2.22}$$

So, is it possible to use a similar reasoning to prove that there exist no magnetic monopoles? Indeed, we can. Magnetic charge is linked with the electromagnetic field via the equation $4\pi J_{magnetic} = dF$. Thus, if any magnetic charge is not present then it would imply that the integral of $J_{magnetic}$ over any 3-volume A is zero (the fundamental theorem of exterior calculus); or

$$\int_A dF = \int_{\partial A} F = \{\text{magnetic flux passing } via\ \partial A\} = 0, \tag{2.23}$$

and as mentioned previously this can be expressed in terms of the "boundary of a boundary vanishes" language by introducing a 4-potential, V (called vector potential). Thus, we express the Faraday tensor as $F = dV$ (V is a 1-form and its exterior derivative produces the 2-form, F), and have (this automatically leads to conservation, something that we wanted when we started off this topic)

$$\int_{\partial A} F = \int_{\partial A} dV \int_{\partial\partial A} V \equiv 0. \tag{2.24}$$

Thus, this concept of the "boundary of a boundary vanishes" is utilised to extend to the concept of conservation of either of Maxwell's equations, i.e.,

$dF = 0$, and $d^*J = 0$. Now, it is only natural to think if such a conservation is valid in gravitational physics, and if it leads us to laws. Yup! It's true, conservation is a key concept in gravitational physics too but here we make use of a so called "double dual" $^*R^*$ of the Riemann tensor which has the following relation with the Einstein tensor G and the stress-energy tensor T

$$G = Tr^*R^* = 8\pi T, \tag{2.25}$$

where 'Tr' is the *trace* of the matrix (of the Riemann double dual tensor here). The reason for the 8π would become clear once we derive the field equations. The conservation of the source here is expressed as $d^*T = 0$, and it's a consequence of $d^*G = 0$, which is called *the contracted Bianchi identity.* However, unlike what the meaning of the vanishing boundary meant for charge, the meaning of the same in gravitational physics is expressed via net moment of rotation of a hypercube[3]. Conservation of the stress-energy tensor for a hypercube can be expressed by making use of the fundamental theorem of exterior as follows

$$\int_A d^*T = \int_{\partial A} {}^*T = 0. \tag{2.26}$$

2.4 Conservation Leads to Continuity?

Imagine there exists a charge distribution in the room (see figure 2.1) you are presently in (if you're outside then assume an imaginary box around the size of your room), now suppose you decide to send the charge off to some distant place, then as it leaves your room, the charge flows as a current through the walls of the room (since a moving charge generates a current). Not only does the charge generate a current, but also generates a current density as the current flows out, say the door of your room, i.e., across a particular area. This leads to the idea of *continuity.*

Let the amount of charge present in your room (i.e., charge per unit volume of your room or charge density) be $\rho \left(= \frac{q}{A_{room}}\right)$. As the charge passes via the walls, the amount of charge leaving the room per unit time is $-\dot{\rho}(t)$ (negative sign indicates the fact that it's leaving your reference frame, i.e., your room). Now, you already know that this leaving charge is travelling through the walls in the form of current (current density actually since charge density

[3] See Misner, C. W., Thorne, K. S., Wheeler, J. A., & Kaiser, D. I. (2017). Gravitation. Princeton University Press; Chapter 15 for more information.

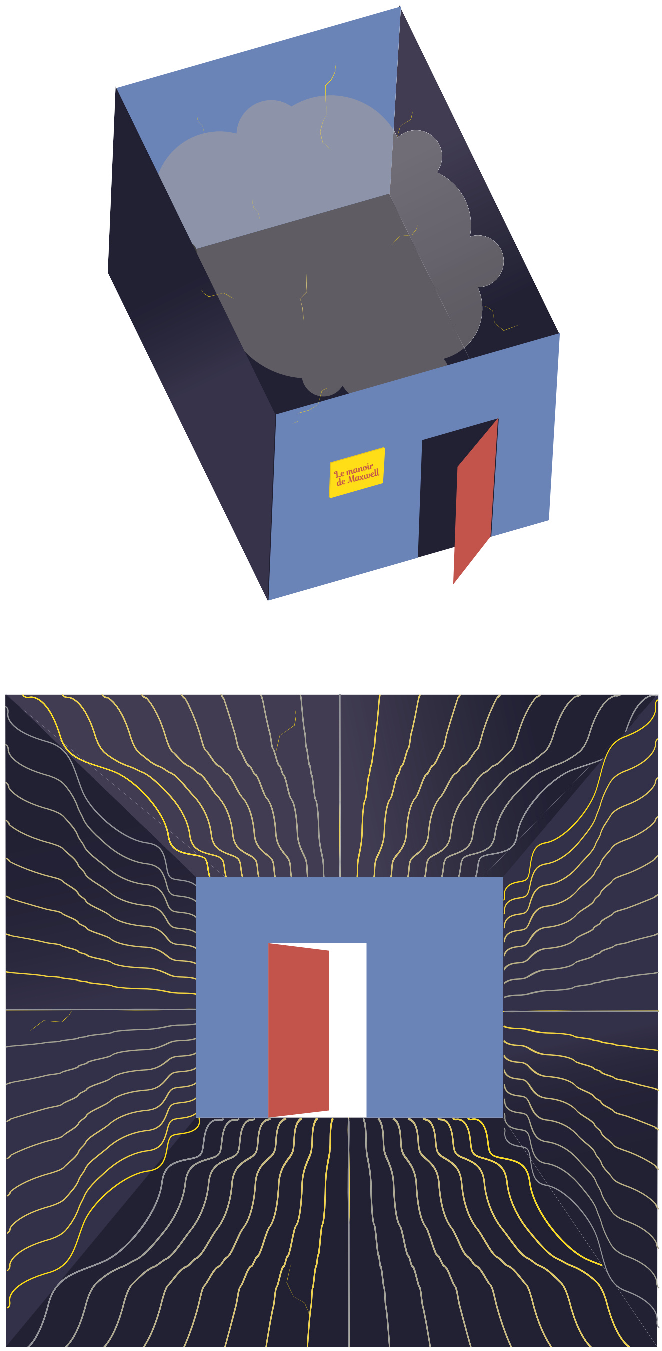


Fig. 2.1. A visual representation of the analogy presented in this chapter. The figure (top) shows the presence of charge density in the room and the bottom figure shows the current as charge moves towards the door.

is travelling not charge), thus giving rise to a changing current in the x, y, & z coordinates, i.e., giving rise to a diverging current density which can be expressed as ∇J. This relation can be expressed as follows

$$\begin{aligned} -\dot{\rho}(t) &= \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right) J = \nabla J, \\ \dot{\rho}(t) &+ \sum_{\alpha=1}^{3} \frac{\partial J^{\alpha}}{\partial x^{\alpha}} = 0. \end{aligned} \tag{2.27}$$

We already know that $\mathbf{J}$ is the 4-current, and its components are $\left(J^0, J^1, J^2, J^3\right) = (\rho, J_x, J_y, J_z)$, hence we can express the above equation as follows

$$\frac{\partial J^{\alpha}}{\partial x^{\alpha}} = 0, \tag{2.28}$$

this is the continuity equation. This form of the continuity is valid in your room which is situated in flat space but in curved spacetime, continuity takes a different form. The difference is (quite obvious)- replace the partial derivative with a covariant derivative to take the form

$$\frac{DJ^{\alpha}}{\partial x^{\alpha}} = J^{\alpha}_{;\alpha} = 0. \tag{2.29}$$

What is to be noted here is that electric charge by itself is an invariant, i.e., it does not change no matter how it moves. The same cannot be said for charge and current density because they are components of a 4-current vector. A similar continuity equation can be proposed for gravitational physics in terms of the stress-energy tensor as follows

$$\frac{DT^{\alpha\beta}}{\partial x^{\beta}} = T^{\alpha\beta}_{;\beta} = 0. \tag{2.30}$$

What this equation implies is that the amount of energy E passing through the room you're in, per unit time is the energy current, $T^{0\alpha}$. For example, if $\alpha = 2$, then the flow of energy along $x^2 = y$ direction is equal to T^{02}, while T^{00} denotes the energy density. In the coordinate-independent sense, this yields a conservation law for the stress-energy tensor as

$$\nabla \mathbf{T} = 0. \tag{2.31}$$

2.5 And Finally, The Field Equations...

Let's place a source with mass (that curves spacetime) in the room you reside in, we know from the previous section that this geometric object's stress-energy tensor $\mathbf{T}$ must have zero divergence $\nabla \mathbf{T} = 0$, because of the conservation of energy-momentum. This source which has mass will communicate to the space around it and give it instructions on how to curve, i.e., it is responsible for the *generation* of gravity. Now, we know that when mass tells space how to curve, space in response will tell matter how to move. Thus, there exists a completely geometrical *object* which is proportional to the stress-energy tensor. This *object* must possess similar characteristics as **T**, i.e., it must be symmetric and also have its own personalized conservation law (i.e., it must be divergence-free). This *object* happens to be the *Einstein tensor* **G** (because as mention in conservation, the conservation law $d^*T = 0$ is a consequence of the contracted Bianchi identity $d^*G = 0$). From the above reasoning, we can express the relation as follows

$$\mathbf{G} = \zeta \mathbf{T}, \tag{2.32}$$

where ζ is the proportionality factor, which will be revealed later. Remember that we wanted to make a connection between the field and the source in such a way that the conservation of the source shall be an automatic consequence of some condition imposed on the field. The zero divergence of the Einstein tensor is the condition here and this automatically leads to the conservation of the stress-energy tensor. Now, if **G** is providing subtitles for the conversation between the geometry and mass, then it's language (i.e., the field's language) must be in terms of the metric tensor and curvature tensor. To see this relation and the conversation between mass and geometry in curved spacetime we are to first understand how the conversation plays out in flat spacetime, i.e., in Newton's version. We go back to the debate that Einstein and Newton had, and now Pauli questions them,

> How does matter tell space to curve? And how does space tell matter to move?

Newton argues that since potential ϕ exists everywhere in space and since it varies from place to place, the field tells particles how to move by putting a limit on acceleration as follows

$$\begin{aligned} &\mathbf{F} = m\mathbf{a} = -m\nabla\phi\,(x) \\ &\mathbf{a} = -\nabla\phi\,(x) \quad \text{[Field informs particles how to move]}, \end{aligned} \tag{2.33}$$

he further proceeds to state that it is the *Poisson equation* that informs the gravitational field how to curve and concludes his thought.

$$\nabla^2\phi\,(x) = 4\pi G\rho\,(x) \quad \text{[Mass tells space how to curve via density]} \tag{2.34}$$

When its Einstein's turn he says that he would like to build on Newton's argument and he establishes that for a spherically symmetric source, the solution to the Poisson equation outside the gravitating source is as follows

$$\phi = -\frac{MG}{r}. \tag{2.35}$$

He then proceeds to draw parallels between this finding and his friend, Karl Schwarzschild's formula that explains the geometry outside a gravitating source (i.e. the Schwarzschild metric) as follows

$$\begin{aligned} ds^2 &= \left(1 - \tfrac{2MG}{r}\right) dt^2 - \left(1 - \tfrac{2MG}{r}\right)^{-1} dr^2 - r^2 d\Omega^2 \\ g_{tt} &= g_{00} = \left(1 - \tfrac{2MG}{r}\right) = (1 + 2\phi) . \end{aligned} \tag{2.36}$$

Now that he has established a relation between the metric tensor and the Newtonian potential, he goes on to observe that there exists a Poisson-like equation in terms of the time component of the metric tensor as follows

$$\nabla^2 g_{00} = 2\nabla^2\phi = 2\,(4\pi G\rho) = 8\pi G\rho. \tag{2.37}$$

Einstein finally concludes that matter, via density, tells space how to curve by affecting the geometry (in terms of the metric). Well, Einstein is absolutely correct, but let us try and understand this Poisson-like equation that Einstein had built in his head. Einstein's generalization of the above equation first involves accounting for an *object* that is built out of derivatives of the metric tensor, which implies that it has to account for the geometry and not the matter; this turns out to be the *Einstein tensor* **G**. The other generalization involves accounting for all the components of the source, i.e., accounting for all the messengers (like density) of the source that communicate to space; this turns out to be the stress-energy tensor **T**. Thus, in component form, we have

$$G^{\mu\nu} = 8\pi G T^{\mu\nu}. \tag{2.38}$$

Our aim now is to arrive at a plausible representation for the Einstein tensor. Since **G** is proportional to the stress-energy tensor **T**, it must also possess properties of the same. Thus, a suitable candidate for the Einstein tensor must be a two-tensor, must be symmetric, and must possess derivatives of the metric tensor. The reason for the last condition is that the metric tensor by itself has the physical meaning of potential, and its derivative has the physical meaning of a field, similarly, the dynamical motion of the field can be described by the derivative of the field which is nothing but the second order derivative of the metric tensor. In short, we need a candidate which must represent the geometry of the field and be able to communicate to matter, informing it how to move. Now, we know that the Christoffel symbol possesses first order derivatives and has the form of $\Gamma \sim \frac{1}{2} g^{-1}\partial g$, but since it's not a tensor, it is not a suitable candidate. What about the Riemann tensor? Well, it takes up the form: $R \sim \partial\Gamma + \Gamma\Gamma - \ldots$, and since the Christoffel symbols contain the first order derivatives of the metric tensor, it automatically implies that the Riemann tensor possesses the second order derivatives of the metric tensor by taking up such a form: $R \sim \partial^2 g + (\partial g)^2 + \dots$, and thus making it a suitable candidate. But wait! The Riemann tensor is not a two-tensor, it has four indices. Not to worry because it can always be contracted to obtain the Ricci tensor (which is a two-tensor) which takes up the following form

$$\begin{aligned} R_{\mu\nu} \equiv R^{\alpha}_{\mu\alpha\nu} &= \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\beta\alpha}\Gamma^{\beta}_{\mu\nu} - \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}{}_{\mu\alpha} \\ &= g^{\delta\alpha} R_{\delta\mu\alpha\nu} \\ &= \tfrac{1}{2} g^{\delta\alpha} \left(g_{\delta\nu,\mu\alpha} + g_{\mu\alpha,\delta\nu} - g_{\delta\alpha,\mu\nu} - g_{\mu\nu,\delta\alpha} \right) \\ &+ g^{\delta\alpha} g_{\lambda\xi} \left(\Gamma^{\lambda}_{\mu\alpha}\Gamma^{\xi}_{\delta\nu} - \Gamma^{\lambda}_{\mu\nu}\Gamma^{\xi}_{\delta\alpha} \right). \end{aligned} \tag{2.39}$$

From the above equation we clearly see that the Ricci tensor is composed of second order derivatives of the metric tensor hence making this the perfect candidate. It is important to note that the Ricci tensor by itself can be expressed in a further simplified form as $R^{\mu\nu} = g^{\mu\nu} R$, where R is called the *scalar curvature*. Enter the dilemma due to this simplification made−what candidate is suitable and which to choose, it is here that we resort to energy conservation. We are well aware of the conservation imposed of the stress-energy tensor (as a consequence of the contracted Bianchi identity, remember?), and hence, in the component form we write that the covariant derivative of the stress-energy tensor is zero, i.e., $D_\mu T^{\mu\nu} = 0$ (and just like $d^*T = 0$ followed from $d^*G = 0$, this equation follows from $D_\mu G^{\mu\nu} = 0$). Following from this, we have

$$D_\mu \left(g^{\mu\nu} R \right) = g^{\mu\nu} \left(D_\mu R \right) + R \left(D_\mu g^{\mu\nu} \right). \tag{2.40}$$

We know from the local flatness condition (i.e., in the local Minkowski reference frame where we applied Gaussian normal coordinates) that $D_\mu g^{\mu\nu} = 0$. Thus, the second term on the RHS vanishes leaving us with $D_\mu \left(g^{\mu\nu} R \right) = g^{\mu\nu} \left(D_\mu R \right)$. Oh, wait! I did mention to you previously that of all the two-rank tensors in the universe that we can form by contracting the Riemann tensor, it is only the Einstein tensor **G** that retains part of the Bianchi identities (and I also went on to mention the equation $G^{\mu\nu}_{;\nu} = 0$). Taking this as our hint, let's begin with the Bianchi identity and see if we can land up with a comfortable expression for **G**. First, let's start by writing down the Bianchi identity for the Riemann tensor

$$D_\mu R_{\alpha\beta\chi\xi} + D_\xi R_{\alpha\beta\mu\chi} + D_\chi R_{\alpha\beta\xi\mu} = 0. \tag{2.41}$$

Upon multiplying by $g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi}$ on either side of the equation (since the metric derivatives vanish, these act as constants and can be taken inside the derivative), we obtain the following

$$D_\mu g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\alpha\beta\chi\xi} + D_\xi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\alpha\beta\mu\chi} + D_\chi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\alpha\beta\xi\mu} = 0,$$

$$D_\mu g^{\nu\mu} R + D_\xi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\alpha\beta\mu\chi} + D_\chi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\alpha\beta\xi\mu} = 0, \tag{2.42}$$

making use of the block symmetry $\left(R_{[abcd]} = 0\right)$ of the Riemann curvature tensor we express $R_{\alpha\beta\mu\chi} = R_{\mu\chi\alpha\beta}$, and $R_{\alpha\beta\xi\mu} = R_{\xi\mu\alpha\beta}$; hence we obtain

$$D_\mu g^{\nu\mu} R + D_\xi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\mu\chi\alpha\beta} + D_\chi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\xi\mu\alpha\beta} = 0, \tag{2.43}$$

making use of the first and second skew symmetries (i.e., $R_{abcd} = -R_{bacd}$, and $R_{abcd} = -R_{abdc}$), we express $R_{\mu\chi\alpha\beta} = -R_{\chi\mu\alpha\beta}$, and $R_{\xi\mu\alpha\beta} = -R_{\xi\mu\beta\alpha}$; hence we obtain

$$D_\mu g^{\nu\mu} R - D_\xi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\chi\mu\alpha\beta} - D_\chi g^{\nu\mu} g^{\alpha\chi} g^{\beta\xi} R_{\xi\mu\beta\alpha} = 0, \tag{2.44}$$

using the Ricci tensor definition $\left(R^{ab} = g^{ac} g^{bd} R_{cd}\right)$, we write the following

$$D_\mu g^{\nu\mu} R - D_\xi g^{\nu\mu} g^{\beta\xi} R_{\mu\beta} - D_\chi g^{\nu\mu} g^{\alpha\chi} R_{\mu\alpha} = 0,$$

$$D_\mu g^{\nu\mu} R - D_\xi R^{\nu\xi} - D_\chi R^{\nu\chi} = 0, \tag{2.45}$$

replacing dummy variables ξ, and χ with μ we have the following equation

$$\left[D_\mu g^{\nu\mu} R - D_\mu R^{\nu\mu} - D_\mu R^{\nu\mu}\right] = 0. \tag{2.46}$$

We can now factorize the derivative to obtain

$$D_\mu g^{\nu\mu} R - 2 D_\mu R^{\nu\mu} = 0,$$

$$D_\mu \left(R^{\nu\mu} - \tfrac{1}{2} g^{\nu\mu} R\right) = 0. \tag{2.47}$$

On comparison of this result to the identity that we obtained via energy conservation we can conclude the following

$$D_\mu R = \tfrac{1}{2} g^{\mu\nu} \partial_\mu R \overset{leads\ to}{\Longrightarrow} D_\mu R^{\mu\nu} = \tfrac{1}{2} D_\mu \left(g^{\mu\nu} R\right)$$

$$\overset{which\ is}{\Longrightarrow} D_\mu \left(R^{\nu\mu} - \tfrac{1}{2} g^{\nu\mu} R\right) = 0 \tag{2.48}$$

Observing the equation, we find out that this object possesses the second order derivatives of the metric tensor is a two-tensor (which is symmetric), and most importantly, retains a part of the Bianchi identity (this identity that is being conserved is called *Bianchi's first identity*). We also observe that the divergence of the tensor is null. Thus, this is the perfect candidate for our Einstein tensor $G^{\mu\nu} = \left(R^{\nu\mu} - \frac{1}{2} g^{\nu\mu} R\right)$. We can now write the Einstein's field equations in their complete form as follows

$$R^{\nu\mu} - \frac{1}{2} g^{\nu\mu} R \equiv G^{\mu\nu} = 8\pi T^{\mu\nu}. \tag{2.49}$$

Remember the proportionality constant ζ, which related the stress-energy tensor $\mathbf{T}$ to the Einstein tensor $\mathbf{G}$? Now, by comparison, we can conclude that $\zeta = 8\pi$. What this field equation told Einstein was that the source of the gravitational field is not limited to energy density but also depends upon the flow of energy, the flow of momentum, and the momentum density. The components of the stress-energy tensor have the dimensions of energy density, i.e., $T^{\mu\nu} \sim ml^2t^{-2} \sim mc^2l^{-3}$; the Christoffel symbol has the dimensions of inverse length, i.e., $\Gamma^{\alpha}{}_{\mu\nu} \sim l^{-1}$; and the Ricci tensor has the dimensions of inverse square length, i.e., $R^{\mu\nu} \sim l^{-2}$. Thus, in order to maintain dimensional homogeneity, we require the constant κ to have the following dimensions: $\kappa \sim m^{-1}lc^{-2}$. This implies that the κ, in terms of the fundamental constants, has the following form: $\kappa \sim G/c^4$. Previously, we did not include this constant since we were working in natural units (i.e., $G \equiv c \equiv 1$). We can now rewrite the equation in terms of this constant (called *seminal Newton's constant*) as follows

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \equiv G^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu} \tag{2.50}$$

The equation is also written with the additional term of Λg, where Λ is called the *cosmological constant*. Einstein introduced the cosmological constant when looking for a stationary model for the cosmos and did not include it in he equation (which he later called *the greatest blunder of his life*). The equations with the cosmological constant included read

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu}. \tag{2.51}$$

2.6 An action-based approach to the field equations

Is there a more efficient way to derive Einstein's field equations, a method which is strictly mathematical but would effortlessly lead us to the equation? Well, yes and this method uses a principle that forms the heart of physics—*the principle of least action*. The *action* of a system is a mathematical construct that takes the trajectory of the system as its argument and gives out a real number for the result. It is an attribute of the dynamics of a physical system from which the equations of motion of the system can be derived.

2.6.1 A Brief History of Classical Mechanics

Why is it that when we throw a ball, the ball always seems to trace the path of a parabola? When we throw the ball, the ball's trajectory is influenced by the force with which we throw, the angle at which we throw, the air resistance it encounters, etc. You might counter my point by stating that when a ball is being passed between two persons who are separated by a short distance, its trajectory is straight and not parabolic. Yes, it seems straight to the naked eye, but there is a slight course correction and the trajectory becomes parabolic as the ball, which possesses mass, is pulled down due to gravity. What I'm asking you to think about is the reason why the trajectory of the ball is parabolic without even considering the factors affecting it, i.e., why is the trajectory of a projectile a parabola? If you are a high schooler who proudly professes knowledge in math, you will easily churn out a proof. But then again, what I am asking you is that why is the ball taking the parabolic path and not, say a wavy, zig-zag path? The answer to this is provided by the principle of least action. As the name says it all we minimize this action which is analogous to the minimization of a function. Suppose Mr. Absolute Zero (at point P_{AZ}) is throwing the ball to his cousin, Mr. Zero Entropy (at point P_{ZE}), then the ball would start from point P_{AZ} at a time t_0 and reach point P_{ZE} at a time t_1 (see figure 2.2).

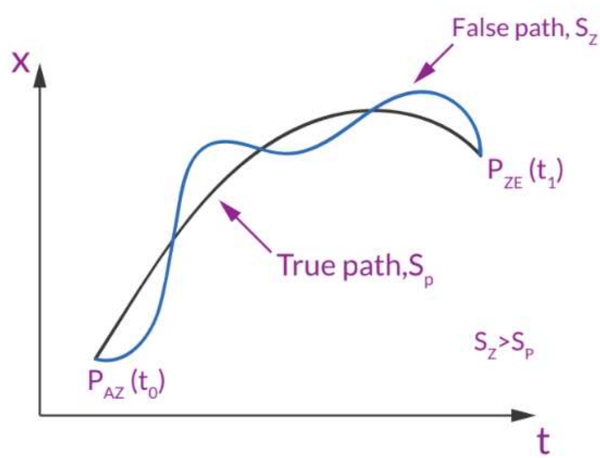


Fig. 2.2. The true and the false path taken by the ball that Mr.Absolute Zero throws from P_{AZ} towards his cousin Mr.Zero Entropy located at P_{ZE}.

Let's consider two separate cases, one when the trajectory of the ball is a parabola, and the second when the trajectory of the ball is a crazy zig-zag curve, but in both cases the ball travels from P_{AZ} to P_{ZE} in the time interval $t_1 - t_0$. We know for a fact that when the ball travels its kinetic and potential energies would vary (variation in kinetic energy due to variation in the ball's velocity, and variation in potential energy due to variation in the height between the ball and the ground during the course of the ball between the two points). If we calculate the kinetic energy at every point in the ball's path, take away the potential energy, and integrate this quantity over the time during the whole path, we would get a real number as the answer. What this implies is that the difference between the average kinetic energy $\left(T = \frac{1}{2}m\dot{x}^2\right)$ and the average potential energy $(V = mgx)$ (i.e. $T - V$) is as little as possible for the path of the ball from P_{AZ} to P_{ZE}. Representing this mathematically, we can define action $\mathcal{S}$ as follows

$$\mathcal{S} \equiv \int_{t_0}^{t_1} (T - V)dt = \int_{t_0}^{t_1} \left(\frac{1}{2}m\dot{x}^2 - mgx\right)dt. \tag{2.52}$$

Following this definition, for the trajectory of the ball to be a parabola it's action must be the least in comparison to the actions of the billions of possible paths between the points P_{AZ} and P_{ZE}. In order to prove this, we need to adopt an appropriate mathematical formulation. You might immediately shout, "calculus!", but it is important to note that we are not probing for a point on a path, we are probing for the path itself! This sort of a problem is dealt with by making use of *calculus of variations*.

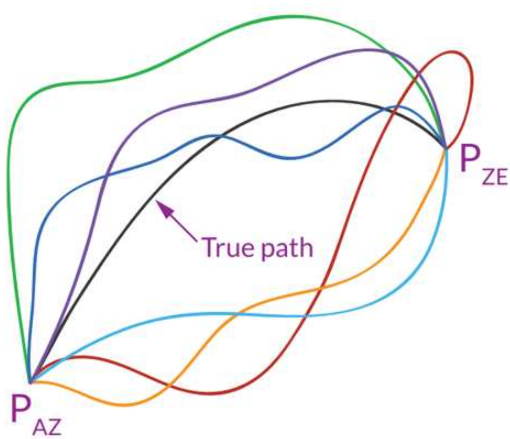


Fig. 2.3. the possible paths between Mr.Absolute Zero at P_{AZ} and his cousin Mr.Zero Entropy at P_{ZE}.

Consider the two paths we initially discussed, i.e., the parabolic one $(\mathcal{S}_P)$ and the crazy zig-zag one $(\mathcal{S}_z)$, following the least action principle, we know that the action of the true path is to be the least, i.e., $\mathcal{S}_P < S_z$. Let's raise the stakes, suppose we are not able to view the event of the ball being passed between Mr. Absolute Zero and Mr. Zero Entropy, and we are to find the true path of the ball and the equation of motion of the ball along this path. One method that immediately comes to mind is to calculate the action for the billions of possible paths that exist between points P_{AZ} and P_{ZE} and hence find the least (see figure 2.3), but this obviously is not possible as we humans have everything but time to waste. Let's try and approach this problem mathematically−when we have a quantity that has a minimum, one of it's properties is that when we move away from the minimum in first order, the deviation of the function from its minimum value is only second order, i.e., take a function $g(x)$ which at its minima has the following property $g^{'}(x) = 0$. Thus, the infinitesimal change (or the deviation) of the function as we move away from the minima is $g^{''}(x)$. Let the true path that the ball takes be $x(t)$, let us take another fiducial path $\xi(t)$ (see figure 2.4) which starts and ends at the same point the true path does but differs from $x(t)$ by an amount $\zeta(t)$ in between (i.e., $\zeta(t_0) = \zeta(t_1) = 0$). Now, let us calculate the action for the true path ($\mathcal{S}_T$) as follows

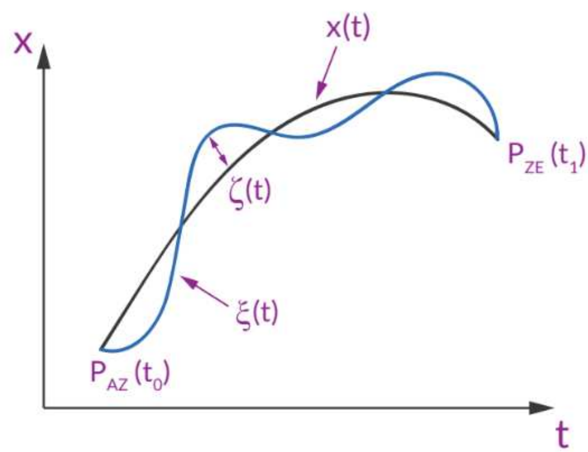


Fig. 2.4. $\xi(t) = x(t) + \zeta(t)$ is the fiducial path that differs from the true path $x(t)$ by an amount $\zeta(t)$.

$$\mathcal{S}_T = \int_{t_0}^{t_1} \left(\frac{m}{2}\dot{x}^2 - V(x)\right) dt, \tag{2.53}$$

and replace $x(t)$ with $\xi(t) = x(t) + \zeta(t)$ to get

$$\begin{aligned}\mathcal{S}_{T'} &= \int_{t_0}^{t_1}\left(\tfrac{m}{2}\left[\tfrac{dx}{dt}+\tfrac{d\zeta}{dt}\right]^2 - V(x+\zeta)\right)dt \\ &= \int_{t_0}^{t_1}\left(\tfrac{m}{2}\left[\left(\tfrac{dx}{dt}\right)^2+\left(\tfrac{d\zeta}{dt}\right)^2+2\tfrac{dx}{dt}\tfrac{d\zeta}{dt}\right] - V(x+\zeta)\right)dt,\end{aligned} \tag{2.54}$$

as discussed previously, since the deviation from the minima is only of second order, we can approximate them out. Thus, the kinetic energy takes the following form

$$T = \frac{m}{2}\left[\left(\frac{dx}{dt}\right)^2 + 2\frac{dx}{dt}\frac{d\zeta}{dt}\right] + (\text{second and higher order terms})\,. \tag{2.55}$$

Upon expanding the potential energy (Taylor expansion), it takes the following form

$$\begin{aligned}V(x+\zeta) &= V(x) + \zeta V^{'}(x) + \tfrac{\zeta^2}{2!}V^{''}(x) + \ldots \\ &= V(x) + \zeta V^{'}(x) + (\text{second and higher order terms})\,.\end{aligned} \tag{2.56}$$

Thus, the action integral takes the form

$$\begin{aligned}\mathcal{S}_{T'} &= \int_{t_0}^{t_1}\left(\tfrac{m}{2}\left(\tfrac{dx}{dt}\right)^2 + m\tfrac{dx}{dt}\tfrac{d\zeta}{dt} - V(x) - \zeta V^{'}(x) + \mathcal{O}\left(V^{''}(x)\right)\right)dt, \\ \mathcal{S}_{T'} &= \int_{t_0}^{t_1}\left(\tfrac{m}{2}\left(\tfrac{dx}{dt}\right)^2 - V(x) + m\tfrac{dx}{dt}\tfrac{d\zeta}{dt} - \zeta V^{'}(x) + \mathcal{O}\left(V^{''}(x)\right)\right)dt, \\ \mathcal{S}_{T'} &= \mathcal{S}_T + \int_{t_0}^{t_1}\left(m\tfrac{dx}{dt}\tfrac{d\zeta}{dt} - \zeta V^{'}(x)\right)dt\ , \\ \delta\mathcal{S} &= \mathcal{S}_{T'} - \mathcal{S}_T = \int_{t_0}^{t_1}\left(m\tfrac{dx}{dt}\tfrac{d\zeta}{dt} - \zeta V^{'}(x)\right)dt\ .\end{aligned} \tag{2.57}$$

Thus, we have found out the variation in the action due to a ζ variation in path. To calculate this variation, we need to integrate the RHS as follows

$$\delta\mathcal{S} = m\frac{dx}{dt}\,\zeta(t)\,|_{t_0}^{t_1} - \int_{t_0}^{t_1}\frac{d}{dt}\left(m\frac{dx}{dt}\right)\zeta(t)\,dt - \int_{t_0}^{t_1}\left(\zeta(t)\,V^{'}(x)\right)dt, \tag{2.58}$$

we know that $\zeta(t_0) = \zeta(t_1) = 0$, hence, inputting this condition we get the following result

$$\delta\mathcal{S} = \int_{t_0}^{t_1}\left(-m\frac{d^2x}{dt^2} - V^{'}(x)\right)\zeta(t)\,dt, \tag{2.59}$$

in order to obtain the true path, the deviation between the fiducial path and the true path must be zero (so that the paths align one on top of each other). What this implies is that under the limit $\zeta(t) \to 0$, the fiducial path tends to the true path and the deviation in their action tends to zero ($\delta S \to 0$). Using these conditions, we finally obtain the following

$$\delta\mathcal{S} = 0 = \int_{t_0}^{t_1} \left(-m\frac{d^2x}{dt^2} - V^{'(x)} \right) \zeta\,(t)\, dt \overset{implies}{\Longrightarrow} -m\frac{d^2x}{dt^2} - V^{'}\,(x) = 0. \quad (2.60)$$

What we obtain by applying the least action is the following equation of motion

$$m\frac{d^2x}{dt^2} = -V^{'}\,(x)\,, \quad (2.61)$$

which is nothing but Newton's law, $(F = ma)$. Hence, the action integral will be minimum for the path that satisfies the differential equation: $-m\frac{d^2x}{dt^2} - V^{'}\,(x) = 0$. The trajectory of a projectile is a parabola because the parabola happens to be the trajectory along which the action of the system is least! Thus, stated otherwise, the least action principle says that the actual evolution of a system is such that its action $\mathcal{S}$ attains an extremum value. Did you ever wonder, in the above experiment, why we calculated the kinetic energy less the potential energy, i.e., $T - V$? Why was it not $T + V$, or T^V, or anything else? To start off this quantity, $T - V$, is called the *Lagrangian.* If we to calculated the action with the sum of the energies, the equation would yield the wrong answer. Since the Lagrangian is composed of the kinetic and the potential energies, it depends upon the position, velocity, and time, i.e. $\mathcal{L} \to \mathcal{L}\,(x, \dot{x}, t)$. Thus, we can redefine the action integral as follows

$$\mathcal{S} = \int_{t_0}^{t_1} \mathcal{L}(x, \dot{x}, t)dt. \quad (2.62)$$

It is to be noted that in Newtonian mechanics we use the rectangular coordinate system and consider all the constraint forces. Lagrange's approach completely avoids the consideration of these constraints by adopting to "generalized coordinates" like the radial distance r, and polar angle ϕ, etc., which are consistent with the constraint relations. The number of generalized coordinates employed are the same as the number of degrees of freedom of the system under consideration. Thus, the main advantage in Lagrangian mechanics is that we don't have to consider the forces of constraints, and just by having information of the kinetic and potential energies we can choose some generalized coordinates and blindly calculate the equations of motion totally analytically. In Lagrangian mechanics we never concern ourselves with the constraints and the geometrical nature of the system. Now back to the question−Why is the Lagrangian represented as $T - V$? The answer, believe it or not, has to do with the ticking of watches. Consider a flat circular plate, rotating about an axis passing through it's centre. It rotates at an angular velocity of ω. Let's now place Mr. Absolute Zero (with frame of reference ζ_{AZ}) at the centre of the plate $(r = 0)$, and his cousin, Mr. Zero Entropy (with frame of reference ζ_{ZE}), at the periphery $(r = R)$. ζ_{ZE} is boosted with respect to ζ_{AZ} by some velocity $v\,(t)$ $(= \omega r)$. Let's now compare their watches and see how much time

$\Delta\tau$ will elapse in Mr. Zero Entropy's moving watch. We know that the line interval $ds^2 = -c^2dt^2 + dx^2$ has the same values in all Lorentz frames, we can evaluate it in ζ_{AZ} and ζ_{ZE} and equate the results. Since Mr. Zero Entropy's frame, ζ_{ZE}, is co-moving, the line interval takes the form: $ds^2 - c^2d\tau^2$. Upon equating both the line intervals we obtain the following relation

$$d\tau = \int \left(1 - \frac{v^2(t)}{c^2}\right)^{1/2} dt. \tag{2.63}$$

This is propertime and is an invariant quantity. Letting the time lapse in Mr. Absolute Zero's watch be $\Delta t(0)$, and that of Mr. Zero Entropy's watch be $\Delta t(r)$, we obtain upon relation to the previous equation,

$$\Delta t(r) = \left(1 - \frac{v^2(t)}{c^2}\right)^{\frac{1}{2}} \Delta t(0) = \left(1 - \frac{\omega^2 r^2}{c^2}\right)^{\frac{1}{2}} \Delta t(0). \tag{2.64}$$

Mr. Absolute Zero, at $r = 0$, feels a centrifugal acceleration equal to $\omega^2 r$ but would not be able to distinguish this from the gravitational acceleration that arises from a gravitational potential on the flat circular plate. We know that the potential satisfies the following relation

$$\frac{\partial\phi}{\partial r} = -\omega^2 r \overset{upon\ integration}{\Longrightarrow} \phi = -\frac{1}{2}\omega^2 r^2. \tag{2.65}$$

Since the laws of physics must be covariant in all frames (i.e., the principle of equivalence), Mr. Zero Entropy must also face a similar centrifugal acceleration. Using this reasoning, we obtain the following relation

$$\Delta t(\phi) = \left(1 - \frac{\omega^2 r^2}{c^2}\right)^{\frac{1}{2}} \Delta t(0) = \left(1 + \frac{2\phi}{c^2}\right)^{\frac{1}{2}} \Delta t(0), \tag{2.66}$$

this tells us that the flow of time depends on the gravitational potential at which the watches are located (i.e., the time that ticks for the brothers located at different positions on the flat circular plate). We can use the above result to modify the line interval, in the presence of a gravitational field as follows

$$ds^2 = -c^2d\tau^2 = -\left(1 + \frac{2\phi}{c^2}\right) c^2dt^2 + dx^2, \tag{2.67}$$

stationary, un-accelerated watches with $dx = 0$ will have a time lapse that depends upon the potential energy they are located at similar to what was expressed previously. The action of a relativistic particle can be expressed as follows

$$\mathcal{S} = -mc^2 \int d\tau. \tag{2.68}$$

Now, in the presence of a weak gravitational field the action for a particle must have the same form as formulated earlier. This can be expressed as follows

$$\mathcal{S} = -mc^2 \int \sqrt{\frac{1}{c^2}\left(1 + \frac{2\phi}{c^2}\right) c^2 dt^2 - \frac{1}{c^2} dx^2}\ , \tag{2.69}$$

multiplying and dividing by dt^2, we obtain,

$$\begin{aligned} \mathcal{S} &= -mc^2 \int \sqrt{\tfrac{1}{c^2}\left(1 + \tfrac{2\phi}{c^2}\right) c^2 - \tfrac{1}{c^2}\dot{x}^2 dt} \\ &= -mc^2 \int \sqrt{1 + \left(\tfrac{2\phi - v^2}{c^2}\right)} dt \\ \mathcal{S} &\cong -mc^2 \int \left(1 - \tfrac{v^2}{2c^2} + \tfrac{\phi}{c^2}\right) dt. \end{aligned} \tag{2.70}$$

This action can be further simplified to obtain

$$\mathcal{S} \cong \int \left(\frac{mv^2}{2} - m\phi - mc^2\right) dt = \int \left(-mc^2 + (T - V)\right) dt = \int \mathcal{L} dt. \tag{2.71}$$

Notice that except for a constant $\left(-mc^2\right)$, we obtained the exact same result of the Lagrangian. Thus, we can conclude that the Lagrangian is defined as $T - V$ and not anything else because gravity affects the rate of flow of watches!

Consider an action with a Lagrangian $\mathcal{L}$. Let the change in path be: $\alpha(t)$. This is the false path and is related to the true path $x(t)$ by the following relation: $\alpha_\epsilon(t) = x(t) + \eta(t)$. Let this false path's velocity be $\beta(t)$ such that $\beta_\epsilon\ (t) = \dot{x}(t) + \dot{\eta}(t)$. As always, the conditions imposed on the variation $\eta(t)$ are: $\eta(t_0) = 0$, & $\eta(t_1) = 0$, i.e., $\eta(t)$ is a differentiable function, and ϵ is small. The action can be written as

$$\mathcal{S} = \int_{t_0}^{t_1} \mathcal{L}\left(x(t) + \epsilon\eta(t), \dot{x}(t) + \epsilon\dot{\eta}(t), t\right) dt, \tag{2.72}$$

which is extremal with respect to ϵ such that

$$\begin{aligned} \left(\tfrac{d}{d\epsilon} S\right)_{\epsilon=0} &= \tfrac{d}{d\epsilon} \int_{t_0}^{t_1} \mathcal{L}\left(x(t) + \epsilon\eta(t), \dot{x}(t) + \epsilon\dot{\eta}(t), t\right) dt \\ &= \tfrac{d}{d\epsilon} \int_{t_0}^{t_1} \mathcal{L}\left(\alpha_\epsilon(t), \beta_\epsilon(t), t\right) dt = \int_{t_0}^{t_1} \tfrac{d\mathcal{L}_\epsilon}{d\epsilon} dt \end{aligned}$$

$$\int_{t_0}^{t_1} \left(\tfrac{dx}{d\epsilon}\tfrac{\partial \mathcal{L}_\epsilon}{\partial x} + \tfrac{d\alpha_\epsilon}{d\epsilon}\tfrac{\partial \mathcal{L}_\epsilon}{\partial \alpha_\epsilon} + \tfrac{d\beta_\epsilon}{d\epsilon}\tfrac{\partial \mathcal{L}_\epsilon}{\partial \beta_\epsilon}\right) dt \tag{2.73}$$

$$= \int_{t_0}^{t_1} \left(\tfrac{d\alpha_\epsilon}{d\epsilon}\tfrac{\partial \mathcal{L}_\epsilon}{\partial \alpha_\epsilon} + \tfrac{d\beta_\epsilon}{d\epsilon}\tfrac{\partial \mathcal{L}_\epsilon}{\partial \beta_\epsilon}\right) dt$$

$$= \int_{t_0}^{t_1} \left(\tfrac{\partial \mathcal{L}}{\partial x}\eta + \tfrac{\partial \mathcal{L}}{\partial \dot{x}}\dot{\eta}\right) dt.$$

We can now integrate by parts to obtain

$$\begin{aligned} &\int_{t_0}^{t_1} \left(\tfrac{\partial \mathcal{L}}{\partial x}\eta + \tfrac{\partial \mathcal{L}}{\partial \dot{x}}\dot{\eta}\right)dt \\ &= \int_{t_0}^{t_1} \left(\tfrac{\partial \mathcal{L}}{\partial x}\eta\right)dt + \tfrac{\partial \mathcal{L}}{\partial \dot{x}}\eta\,(t)\,|_{t_0}^{t_1} - \int_{t_0}^{t_1} \left(\tfrac{d}{dt}\tfrac{\partial \mathcal{L}}{\partial \dot{x}}\right)\eta\,(t)dt \\ &= \int_{t_0}^{t_1} \left(\tfrac{\partial \mathcal{L}}{\partial x} - \tfrac{d}{dt}\tfrac{\partial \mathcal{L}}{\partial \dot{x}}\right)\eta\,(t)\,dt = 0. \end{aligned} \tag{2.74}$$

This yields the *Euler-Lagrange equation*:

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 0 \tag{2.75}$$

This is a second order, partial differential equation. The evolution of a physical system is described by the solutions to the Euler-Lagrange equation for the action of the system. It is important to observe the time dependence of the Lagrangian. The two sources of time dependence are the generalized coordinate and velocity represented by $q\,(t)$ and $\dot{q}\,(t)$. Thus, we can express the time dependence of the Lagrangian as follows (for j particles)

$$\frac{d\mathcal{L}}{dt} = \sum_j \left(\frac{\partial \mathcal{L}}{\partial q_j}\dot{q}_j + \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\ddot{q}_j\right), \tag{2.76}$$

however, if the Lagrangian has an explicit time dependence, then there would be an addition term $\frac{\partial \mathcal{L}}{\partial t}$. If we observe the RHS of above equation, the first term is: $\frac{\partial \mathcal{L}}{\partial q_j}\dot{q}_j = \dot{p}_j\dot{q}_j$, and the second term is: $\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\ddot{q}_j = p_j\ddot{q}_j$. Hence, the time rate of change of the Lagrangian with explicit time dependence can be expressed as follows

$$\begin{aligned} \tfrac{d\mathcal{L}}{dt} &= \tfrac{d}{dt}\textstyle\sum_j \left(p_j\ \dot{q}_j\right) + \tfrac{\partial \mathcal{L}}{\partial t} \\ \tfrac{\partial \mathcal{L}}{\partial t} &= -\tfrac{d}{dt}\left(\textstyle\sum_j \left(p_j\ \dot{q}_j\right) - \mathcal{L}\right), \end{aligned} \tag{2.77}$$

now, define the quantity $\sum_j \left(p_j\ \dot{q}_j\right) - \mathcal{L} = \mathcal{H}$, which is called the *Hamiltonian*, which varies with time if and only if the Lagrangian has an explicit time dependence. This is the central piece in the new formulation of mechanics called the *Hamiltonian formulation of classical mechanics*. What does this $\mathcal{H}$ represent? Let's try and find out. Consider the standard Lagrangian, $\mathcal{L} = \frac{m\dot{q}^2}{2} - V\,(q) = m\dot{q}\frac{\dot{q}}{2} - V\,(q) = p\frac{\dot{q}}{2} - V\,(q)$. Let's integrate this quantity into our definition of the Hamiltonian. Upon doing so we observe the following

$$\mathcal{H} = \left(\sum_j \left(p_j\ \dot{q}_j\right) - \mathcal{L}\right) = p\dot{q} - \left(p\frac{\dot{q}}{2} - V\,(q)\right) = p\dot{q} + V\,(q) = T + V\,(q) = \text{Energy}. \tag{2.78}$$

What is more important to notice here is that if there is no explicit dependence on the Lagrangian, then the energy $\mathcal{H}$ is conserved, i.e., $\frac{\partial \mathcal{L}}{\partial t} = \frac{d\mathcal{H}}{dt} = 0$. In the Hamiltonian formulation, we do away with all the $\dot{q}$'s and replace them with the momenta, i.e., the p's. Thus, we replace $\dot{q}$ with $\frac{p}{m}$. We have the Hamiltonian as follows

$$\mathcal{H} = T + V(q) = \frac{m\dot{q}}{2} + V(q) = \frac{p^2}{2m} + V(q) = \mathcal{H}(p, q). \tag{2.79}$$

When we take partial derivatives of the Hamiltonian, we observe the following

$$\frac{\partial \mathcal{H}}{\partial q} = \frac{dV(q)}{dq} = -F(q) = -\dot{p}, \quad \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{m} = \dot{q}. \tag{2.80}$$

In the first equation the fact that the rate of change of momentum is the force is exploited. These are called the Hamilton's equations and are first order partial differential equations. What is important to notice is that we have simplified a n-second order partial differential system (the Euler-Lagrange equations) to a $2n$-first order partial differential system (Hamilton's equations). In the action (or the Lagrangian) formulation of classical mechanics the trajectory followed by the system was described via the generalized coordinates $q(t)$, and the equations are second order differential equations. The time dependent coordinate space used in the Lagrangian formalism is known as the *configuration space* . In the Hamiltonian formulation, however, the focus is on the space of coordinates q_j and its conjugate momenta p_j. This space is referred to as *phase space* . The evolution of the system is described as a trajectory in phase space, and this trajectory is obtained via the two first order equations. Let's examine the time derivative of the Hamiltonian. When the Hamiltonian has an explicit time dependence we express It as: $\mathcal{H} = \mathcal{H}(p, q, t)$. The value of $\mathcal{H}$ varies with time because of its explicit time dependence, and also because q and p are themselves functions of time. The total time derivative of the Hamiltonian is expressed as follows

$$\frac{d\mathcal{H}}{dt} = \sum_j \frac{\partial \mathcal{H}}{\partial q_j}\frac{dq_j}{dt} + \sum_j \frac{\partial \mathcal{H}}{\partial p_j}\frac{dp_j}{dt} + \frac{\partial \mathcal{H}}{\partial t} = \sum_j \frac{\partial \mathcal{H}}{\partial q_j}\dot{q}_j + \sum_j \frac{\partial \mathcal{H}}{\partial p_j}\dot{p}_j + \frac{\partial \mathcal{H}}{\partial t}. \tag{2.81}$$

We can now express $\dot{q}$ and $\dot{p}$ in terms of the derivatives of $\mathcal{H}$ by making use of Hamilton's equations to obtain the following

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t} + \sum_j \left(\frac{\partial \mathcal{H}}{\partial q_j}\frac{\partial \mathcal{H}}{\partial p_j} - \frac{\partial \mathcal{H}}{\partial p_j}\frac{\partial \mathcal{H}}{\partial q_j} \right) = \frac{\partial \mathcal{H}}{\partial t}. \tag{2.82}$$

For a conserved system, neither the kinetic nor the potential energy is to contain any explicit time dependence. Thus, $\frac{\partial \mathcal{H}}{\partial t} = 0$, and we obtain

$$\frac{d\mathcal{H}}{dt} = 0, \tag{2.83}$$

this implies that there exists a law of conservation of energy ($\mathcal{H} = T + V = E = const$).

2.6.2 Lagrangian Formulation

There exists a more elegant method to obtain the equations of motion of the system in which we probe for a special generating function $\mathcal{S}$ such that the new Hamiltonian is $\mathcal{H}^{'} = 0$.
Since you are comfortable with the action principle let's try and create an action integral that would yield Einstein's field equation when solved for. Consider a field present in spacetime, say $\psi(r,t)$, here $r \in (x,y,t)$. We can write an action for this field as

$$\mathcal{S} = \int \mathcal{L}(\psi, \psi_{,\alpha}) dr dt, \tag{2.84}$$

where $\mathcal{L}(\psi, \psi_{,\alpha})$ is called the *Lagrange density* and it is a function of the field ψ and the derivatives of the same with respect to an arbitrary basis α. The field that we considered can be of any type (scalar, vector, tensor, spinor), for simplicity we shall restrict our observations to only a scalar field. In the Lagrangian formulation of a field theory, we have an arbitrary region ζ of a spacetime manifold which is bounded by a closed hypersurface $\partial\zeta$ and the Lagrangian density $\mathcal{L}(\psi, \psi_{,\alpha})$ which is a scalar function of the field and its derivatives. Thus, the action functional has the form

$$\mathcal{S}[\psi] = \int \mathcal{L}(\psi, \psi_{,\alpha}) \sqrt{-g} d^4x. \tag{2.85}$$

Dynamical equations for the field ψ are obtained by the introduction of a variation $\delta\psi(x^\alpha)$. It is to be mentioned that this introduction is to be arbitrary within ζ but vanishes everywhere on $\partial\zeta$, i.e.,

$$\delta\psi|_{\partial\zeta} = 0. \tag{2.86}$$

We need to build an action integral such that it is invariant everywhere and the simplest invariant available is of the following form

$$I = \int \sqrt{|g|} d^4x, \tag{2.87}$$

where $d^4x = dx^0 dx^1 dx^2 dx^3 = dt dx dy dz$, and $|g| = |\det(g_{\mu\nu})|$. $\sqrt{-g} d^4x$ is the volume element of a 4-dimensional parallelepiped with edges $dt, dx, dy, \&\ dz$. If we multiply this with any scalar field $\psi(x,y,z,t)$ the integral would still remain an invariant. Hence, we make use of the Ricci scalar, R to form the action as follows

$$\mathcal{S} = \alpha \int \sqrt{-g} d^4x + \beta \int R\sqrt{-g} d^4x, \tag{2.88}$$

where α and β are some constants. This equation describes only gravity and does not account for matter. Hence, we add some action which contains matter, fields or particles, and describes the interactions between them and the metric (and hence with gravity). With the inclusion of this matter action $\mathcal{S}_M$, and fixing constants α and β we obtain

$$\mathcal{S}_{EH} = -\frac{1}{16\pi\kappa} \int \sqrt{-g} d^4x \,(R + \mathit{\Lambda}) + \mathcal{S}_M. \tag{2.89}$$

This is the *Einstein-Hilbert action*. Here, $\kappa = G/c^4$ is the *seminal Newton's constant* and Λ is called the *cosmological constant*.

2.6.3 Variation of the Einstein-Hilbert Action

We apply the principle of least action and vary the action with respect to the metric to obtain equations of motion as follows

$$\begin{aligned} \delta_g \mathcal{S}_{EH} &= [\mathcal{S}(g + \delta g) - \mathcal{S}(g)] \\ &= -\tfrac{1}{16\pi\kappa} \delta_g \int \sqrt{-g} d^4x \,(R + \Lambda) + \delta_g \mathcal{S}_M = 0, \end{aligned} \tag{2.90}$$

$$-\frac{1}{16\pi\kappa} \int d^4x \left[\left(\delta\sqrt{-g}\right)(R + \Lambda) + \sqrt{-g} R_{\mu\nu} \delta g^{\mu\nu} + \sqrt{-g} \delta R_{\mu\nu} g^{\mu\nu} \right] + \delta_g \mathcal{S}_M = 0. \tag{2.91}$$

Note that $\delta_g \Lambda = 0$. Now, let us separately analyse each term's contribution.

1. $\delta\sqrt{-g}$

Let's first analyse how the natural log of a determinant of a matrix A transforms (since the metric tensor is also a matrix and it would also transform analogously). Consider $\delta \ln |\det A| = \ln |\det (A + \delta A)| - \ln |\det (A)|$, using the property of ln, we get $\ln \frac{\det(A+\delta A)}{\det A}$, and solving this we get

$$\begin{aligned} \ln \tfrac{\det(A+\delta A)}{\det A} &= \ln \det \left(A^{-1}(A + \delta A)\right) \\ = \ln \det \left(1 + A^{-1}\delta A\right) &= \mathrm{Tr} \ln \left(1 + A^{-1}\delta A\right) \approx \mathrm{Tr}\; \mathrm{A}^{-1}\delta A. \end{aligned} \tag{2.92}$$

In all of these expressions, Tr stands for the trace of the matrix. In linear algebra, the trace of a $n \times n$ square matrix A is defined to be the sum of the elements on the main diagonal, i.e.,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$
$$\mathrm{Tr}\,(A) = a_{11} + a_{22} + \cdots + a_{nn} = \textstyle\sum_{j=1}^{n} a_{jj}. \tag{2.93}$$

Hence, in linear order we obtained $\delta \ln |\det A\ | \ \approx \mathrm{Tr}\left(A^{-1}\delta A\right)$. Applying this variation to the natural log of the square root of $-g$, we obtain the following

$$\begin{aligned} \delta \ln\sqrt{-g} \ &= \delta \ln\sqrt{|\det g_{\mu\nu}\ |} \ = -\tfrac{1}{2}\delta \ln |\det (g^{\mu\nu})\ | \ = -\tfrac{1}{2} g_{\mu\nu}\delta g^{\mu\nu}, \\ \delta \ln\sqrt{-g} \ &= \tfrac{1}{\sqrt{-g}}\delta\sqrt{-g}, \\ \delta\sqrt{-g} &= -\tfrac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}. \end{aligned} \tag{2.94}$$

2. $g^{\mu\nu}\delta R_{\mu\nu}\sqrt{-g}$

To find an expression for this we move to locally flat coordinates, i.e., the *local Minkowski reference frame* (LMRF). Observe: $\delta R^{\circ}{}_{\circ\circ\circ} \sim \partial_{\circ}\delta\Gamma^{\circ}{}_{\circ\circ} + \delta\Gamma^{\circ}{}_{\circ\circ}\Gamma^{\circ}{}_{\circ\circ} + \Gamma^{\circ}{}_{\circ\circ}\delta\Gamma^{\circ}{}_{\circ\circ}$. We know that in the LMRF, $\Gamma^{\circ}{}_{\circ\circ}(x) = 0$, and this simplifies the expression to $\delta R^{\circ}{}_{\circ\circ\circ} \sim \partial_{\circ}\delta\Gamma^{\circ}{}_{\circ\circ}$, so that

$$\delta R^{\lambda}{}_{\mu\alpha\nu}(x) = \delta\left(\Gamma^{\lambda}{}_{\mu\nu,\alpha} - \Gamma^{\lambda}{}_{\mu\alpha,\nu}\right) = \partial_{\alpha}\delta\Gamma^{\lambda}{}_{\mu\nu} - \partial_{\nu}\delta\Gamma^{\lambda}{}_{\mu\alpha,\nu}. \tag{2.95}$$

Since at this point $\Gamma^{\circ}{}_{\circ\circ}(x) = 0$, the ordinary derivative $\partial_{\circ}$ is the same as the covariant derivative $D_{\circ}$, making this change we obtain

$$\delta R^{\lambda}{}_{\mu\alpha\nu}(x) = D_{\alpha}\delta\Gamma^{\lambda}{}_{\mu\nu} - D_{\nu}\delta\Gamma^{\lambda}{}_{\mu\alpha}, \tag{2.96}$$

which is called the *Palatini identity.* This identity does not only hold in the locally flat coordinates, but also in general. By setting the upper index and the second lower index (i.e., $\lambda = \alpha$), we can write an expression for the variation of the Ricci tensor as follows

$$\delta R_{\mu\nu} = D_{\alpha}\delta\Gamma^{\alpha}{}_{\mu\nu} - D_{\nu}\delta\Gamma^{\alpha}{}_{\mu\alpha}. \tag{2.97}$$

Thus,

$$g^{\mu\nu}\delta R_{\mu\nu} = D_{\mu}\left(g^{\eta\rho}\Gamma^{\mu}{}_{\eta\rho} - g^{\eta\mu}\Gamma^{\rho}{}_{\eta\rho}\right) \equiv D_{\mu}\delta V^{\mu}, \tag{2.98}$$

where V^{μ} is a 4-vector. We can now rewrite the action integral for this term as

$$\int_{\zeta} d^4x\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} = \int_{\partial\zeta} d^4x\sqrt{-g}D_\mu\delta V^\mu, \tag{2.99}$$

where ζ is the spacetime manifold under consideration and $\partial\zeta$ is its boundary. To solve this integral we need to make use of *Gauss-Orstrogradsky theorem* also known as the *Divergence theorem*. The theorem states the following

$$\int d^nx\sqrt{-g}D_\alpha X^\alpha = \int d^{n-1}x\sqrt{-r}n_\alpha X^\alpha. \tag{2.100}$$

Upon comparison we find that $n = 4,\ \alpha = \mu,\ \&\ X^\alpha = \delta V^\mu$, so we write

$$\int d^4x\sqrt{-g}D_\mu V^\mu = \int d^3\chi\sqrt{-g^{(3)}}n_\mu\delta V^\mu = \oint d\Sigma_\mu\delta V^\mu, \tag{2.101}$$

where $d\Sigma_\mu = n_\mu\sqrt{-g^{(3)}}d^3\chi$ is a 4-vector normal to $\partial\zeta$, n_μ is the normal vector to the boundary, and $\left|g^{(3)}\right| = |\det g_{jk}\ |$ is the determinant of the induced 3-dimensional metric. This integral would vanish if the δV^λ vanishes on the boundary of the integration domain, as mentioned earlier (and as assumed while computing the Euler-Lagrange equations). Thus, $\delta V^\mu|_{\partial\zeta} = 0$, and we obtain

$$\int d^4x\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} = 0. \tag{2.102}$$

3. $\delta_g\mathcal{S}_M$

The matter action has to be an invariant under coordinate transformations, it takes the following form when varies with respect to the metric

$$\begin{aligned}\delta_g\mathcal{S}_M &= \delta_g \int \mathcal{L}d^4x\sqrt{-g} = \int d^4x\left[\tfrac{\partial\mathcal{L}}{\partial g^{\mu\nu}}\delta g^{\mu\nu}\sqrt{-g} + L\delta\sqrt{-g}\right]\\ &= \int d^4x\left[\tfrac{\partial\mathcal{L}}{\partial g^{\mu\nu}}\left(\delta g^{\mu\nu}\right) - \tfrac{1}{2}\sqrt{-g}g_{\mu\nu}L\left(\delta g^{\mu\nu}\right)\right]\sqrt{-g}\\ &= \int d^4x\left[\tfrac{\partial\mathcal{L}}{\partial g^{\mu\nu}} - \tfrac{1}{2}Lg_{\mu\nu}\right]\sqrt{-g}\delta g^{\mu\nu}.\end{aligned} \tag{2.103}$$

Now, we define the stress-energy tensor $T_{\mu\nu}$ as follows

$$\frac{\partial\mathcal{L}}{\partial g^{\mu\nu}} - \frac{1}{2}Lg_{\mu\nu} = \frac{1}{2}T_{\mu\nu}, \tag{2.104}$$

thus, the final form of the variation of the matter action is

$$\delta_g \mathcal{S}_M = \frac{1}{2}\int d^4x\sqrt{-g}T_{\mu\nu}\delta g^{\mu\nu} \tag{2.105}$$

Let's use all of the above results and rewrite the variation of the Einstein-Hilbert action integral.

$$\delta_g \mathcal{S}_{EH} = -\frac{1}{16\pi\kappa}\int d^4x\sqrt{-g}\delta g^{\mu\nu}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \frac{1}{2}g_{\mu\nu}\Lambda - 8\pi\kappa T_{\mu\nu}\right) = 0. \tag{2.106}$$

From this we obtain Einstein's equations.

$$R_{\mu\nu}\frac{1}{2}g_{\mu\nu}R - \frac{1}{2}g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{2.107}$$

2.7 Properties of the Einstein equations

Let's first analyse the fate of the equation in vacuum with no sources. This situation implies that $\Lambda = 0$ and $T_{\mu\nu} = 0$. Plugging these into the equation we get

$$R\mu\nu - \frac{1}{2}g_{\mu\nu}R = 0, \tag{2.108}$$

multiplying by $g^{\mu\nu}$ on either sides we get

$$\begin{aligned} R_{\mu\nu}g^{\mu\nu} - \tfrac{1}{2}\delta^{\mu\nu}_{\mu\nu}R = 0 \\ R_{\mu\nu}g^{\mu\nu} = 2R \overset{implies}{\Longrightarrow} R = 0, \end{aligned} \tag{2.109}$$

upon substituting this result in the vacuum, source-less version of the field equation we obtain

$$R_{\mu\nu} = 0. \tag{2.110}$$

This is called the *Ricci flatness condition*. It is important to note that the Ricci flatness condition does not imply vanishing curvature of spacetime, i.e., $R_{\mu\nu\alpha\beta} = 0$. There is something hidden in the field equation, it's trying to convey a key information. Let's see what this is. Take the covariant derivative of the equation with respect to the basis ν to get

$$D_\nu\left(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R - \tfrac{1}{2}\Lambda g_{\mu\nu}\right) = D_\nu\left(\tfrac{8\pi G}{c^4}T_{\mu\nu}\right),$$

$$D_\nu\left(R_{\mu\nu}\right) - \tfrac{1}{2}g_{\mu\nu}D_\nu\left(R\right) - \tfrac{1}{2}RD_\nu\left(g_{\mu\nu}\right) - \tfrac{1}{2}\Lambda D_\nu\left(g_{\mu\nu}\right) = \tfrac{8\pi G}{c^4}D_\nu\left(T_{\mu\nu}\right), \tag{2.111}$$

notice that the terms containing $D_\nu\left(g_{\mu\nu}\right)$ would vanish due to the local flatness condition. We can now manipulate the leftover terms to get

$$g_{\nu\nu}\left(D_\nu R^\nu_\mu\right) - \frac{1}{2}\left(D_\mu R\right)g^\mu_\nu g_{\mu\nu} = \frac{8\pi G}{c^4}\left(D^\nu T_{\mu\nu}\right)g_{\nu\nu}. \tag{2.112}$$

Dividing this equation by $g_{\nu\nu}$ on either sides we obtain the following

$$R^\nu{}_{\mu;\nu} - \frac{1}{2}\partial_\mu R = \frac{8\pi G}{c^4}D^\nu T_{\mu\nu}. \tag{2.113}$$

Observe the LHS, does it ring a bell? Yup, it's the first Bianchi identity, and since we already know that $R^\nu_{\mu;\nu} = \frac{1}{2}\partial_\mu R$, we can substitute this into the equation to obtain a very elegant result

$$D^\nu T_{\mu\nu} = 0. \tag{2.114}$$

We have just obtained back the conservation condition! This was the exact same condition we deduced in the earlier section. But wait, isn't something off here? We assumed that conservation holds and hence obtained the field equations, but what does it mean when the field equations themselves reproduce back the same condition, what are the equations trying to tell us? The answer is that the conservation of the stress-energy tensor is just a consequence of the Einstein field equations.

Example 2.1. Einstein-Hilbert action with a Cosmological constant
The Einstein-Hilbert action with a cosmological constant is given by

$$\mathcal{S} = \frac{1}{2\kappa}\int d^n x\sqrt{-g}\left(R - 2\Lambda\right), \tag{2.115}$$

where we have set $\kappa = 8\pi G$ and $c \equiv 1$. For different values of Λ, the solutions obey different asymptotic structures. Broadly these may be divided into three classes of solutions.
1. $\Lambda = 0$: This yields asymptotically flat solutions such as the Minkowski and Schwarzschild spacetimes.
2. $\Lambda > 0$: This yields asymptotically de Sitter (dS) solutions. Examples include the de Sitter and de Sitter-Schwarzschild metrics among others.
3. $\Lambda < 0$: This yields asymptotically Anti de Sitter (AdS) solutions. Examples include the BTZ black hole 3.57 and the other AdS metrics mentioned (see 1.228, 1.213, 6.31) among others.

Exercise 8

1. A scalar field Φ that is governed by the following action

$$\mathcal{S}[\Phi] = \frac{1}{c}\int d^4x\sqrt{-g}\left(\frac{1}{2}g_{\alpha\beta}\partial^\alpha\Phi\partial_\beta\Phi - V(\Phi)\right),$$

where $V(\Phi)$ is the potential of the scalar field.
a. Vary this action with respect to the metric tensor and obtain an expression for the stress-energy tensor of the scalar field.
b. Show that stress-energy tensor conservation yields the equation of motion of the scalar field.

2. Verify if varying the following action gives rise to the Klein-Gordon equation in curved spacetime as given in 1.239, for $n = 4$

$$\mathcal{S} = \frac{1}{2}\int d^4x\sqrt{-g}\; g^{\mu\nu}\left(D_\mu\varphi D_\nu\varphi + \frac{1}{6}R_{\mu\nu}\varphi^2\right).$$

Using the relation between the stress-energy tensor $T_{\mu\nu}$ and the action as given in (refer to 2.106), show that the variation of the action gives the following stress-energy tensor

$$T^{\mu\nu} = \frac{1}{2}g^{\mu\nu}\left(\frac{1}{6}\Box\varphi^2 + \frac{1}{6}R\varphi^2 - \varphi\Box\varphi\right) - \frac{1}{3}D^\mu D^\nu\varphi^2 + \varphi D^\mu D^\nu\varphi - \frac{1}{6}R^{\mu\nu}\varphi^2.$$

Finally, show that the trace of the stress-energy tensor vanishes.

3. The action that describes an electromagnetic field is given by

$$\mathcal{S}[A^\mu] = -\frac{1}{16\pi c}\int d^4x\sqrt{-g}\; F_{\mu\nu}F^{\mu\nu},$$

where $F_{\mu\nu} = 2D_{[\mu}A_{\nu]} = 2\partial_{[\mu}A_{\nu]}$.
a. Find the stress-energy tensor associated with the electromagnetic field.
b. What are the time-time and the time-space components of the stress energy tensor of the electromagnetic field in flat spacetime?
c. Show that this action is invariant under conformal transformation, $x^\mu \to x^\mu$, $A_\mu \to A_\mu$ and $g_{\mu\nu} \to e^{2\kappa(x)}g_{\mu\nu}$.

4. Nordstrøm devised a metric theory of gravity, before Einstein gave his, which relates the metric tensor to the stress-energy tensor by the following equations

$$W_{\mu\nu\alpha\beta} = 0, \quad R = \kappa g_{\mu\nu}T^{\mu\nu}, \tag{2.116}$$

where $W_{\mu\nu\alpha\beta}$ is the Weyl tensor. The vanishing of Weyl tells us that the metric

is conformally flat.
a. Show that $g_{\mu\nu} = e^{2\kappa(x)}\eta_{\mu\nu}$ is a solution by replacing this in the 2.116. Here $\eta_{\mu\nu}$ is the Minkowski metric.
b. The dimensionality of this spacetime is $n = 4$. Write down the line element and show that the Ricci scalar takes the form $R = -6e^{-2\kappa(x)}\left((\kappa'(x))^2 + \kappa''(x)\right)$.
c. For non-relativistic stress-energies, $T^{\mu}_{\mu} \approx T^0_0 = -\rho$. Use this to show that Nordstrøm's field equation reduces in the Newtonian limit to the gravitational field equations, and determine the value of κ.

5. Find the Einstein tensor of the induced metric on a timelike hypersurface to the BTZ metric 3.57, $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}R\gamma_{\mu\nu}$. Note that all the tensors are obtained using the induced metric $\gamma^{\mu\nu}$.

6. Show that the components of the Einstein tensor tensor vanishes identically in $n = 2$. Note that the Riemann tensor takes the form as given in 1.203.

7. The Brans-Dicke theory of gravity proposes the following action with respect to a field φ

$$\mathcal{S} = \frac{1}{2\kappa}\int d^4\sqrt{-g}\left(f(\varphi) + f_{,\varphi}(\varphi)\left(R - \varphi\right)\right) + \int d^4\mathcal{L}_M,$$

where $\mathcal{L}_M$ is the matter Lagrangian. This is an example of a scalar-tensor theory of gravity. Show that varying this action action with respect to φ gives

$$f_{,\varphi\varphi}\left(R - \varphi\right) = 0.$$

2.8 Noether's Theorem

In the framework of the Lagrangian theory, to each continuous group of transformations leaving the Lagrangian invariant there corresponds a quantity which is conserved. In particular, energy corresponds to time translations, linear momentum corresponds to space translations, and angular momentum corresponds to space rotations.

2.8.1 Symmetry and Conservation Laws

Symmetry plays an important role in Noether's theorem. A symmetry can be defined as an active coordinate transformation that does not change the value of the Lagrangian. Symmetry properties of the Lagrangian imply the existence of a conserved quantities. When the displacement of the system to a newly defined point in the configuration space, regardless of it's location in the configuration space, leaves the Lagrangian invariant, the transformation is called

an active one. Thus, if the Lagrangian does not contain explicitly a particular coordinate of displacement, then the corresponding canonical momentum is conserved and the absence of such an explicit dependence on the coordinate implies that the Lagrangian is invariant under the given transformation. Consider, for example, the Lagrangian $L = \dot{q}/2$, under a time independent coordinate transformation of $q \to q + \epsilon$, the velocity is unaffected and thus, the Lagrangian is invariant under the coordinate shift, i.e., $\delta \mathcal{L} = 0$.
In a more general example consider two arbitrary points in spacetime, say five minutes ago from where you were to where you are now. If we were to look at some quantity, described by a Lagrangian, which remains unchanged in the past five minutes, we would say that the Lagrangian is symmetric under spacetime translations. Let the Lagrangian exhibiting such a symmetry have the form $\mathcal{L}(x^\alpha, \dot{x}^\alpha)$. Then, the Euler-Lagrangian equations of this Lagrangian can be written as follows

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} = \frac{\partial \mathcal{L}}{\partial x^\alpha}, \tag{2.117}$$

and since the Lagrangian is invariant between the your position in the past five minutes, the time derivative of the Lagrangian would vanish, i.e.,

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} = 0. \tag{2.118}$$

Now, $\partial \mathcal{L}/\partial \dot{x}^\alpha = m\dot{x}^\alpha = p^\alpha$, is nothing but your 4-momentum. Since we have a time derivative, we observe that the equations yields

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} = \dot{p}^\alpha = 0. \tag{2.119}$$

Integration yields

$$p^\alpha = \mathcal{R}^\alpha, \tag{2.120}$$

where $\mathcal{R}^\alpha$ is some constant vector. We know from special relativity that the 4-momentum state is related to an energy E_p and a 3-dimensional momentum vector as follows

$$p^\alpha = (E_p, -\mathbf{p}), \tag{2.121}$$

i.e., if the Lagrangian is invariant as you changed your position in spacetime over the past five minutes, the energy and momentum are automatically conserved. Hence, Noether's theorem guarantees that continuous symmetry transformations give rise to conservation laws. We know that the time rate change of the Lagrangian is represented as the time rate change of the summation of the product of the momentum and the time derivative of the position, i.e.,

$\frac{d\mathcal{L}}{dt} = \frac{d}{dt}\sum_\alpha (p_\alpha \dot{q}_\alpha)$. Hence, the variation of the Lagrangian for the infinitesimal change in your coordinates in the past five minutes can be expressed as follows

$$\delta\mathcal{L} = \frac{d}{dt}\sum_\alpha (p_\alpha \delta\dot{q}_\alpha), \tag{2.122}$$

and since the Lagrangian is invariant

$$\frac{d}{dt}\sum_\alpha (p_\alpha \delta\dot{q}_\alpha) = 0. \tag{2.123}$$

Let the infinitesimal shift your coordinates to be

$$\delta q_\alpha = \mathcal{F}_\alpha(q)\delta, \tag{2.124}$$

i.e., for an infinitesimal rotation of Earth in the past five minutes, your new coordinates are

$$\begin{aligned} q_1 &\rightarrow q_1 + a\delta; \mathcal{F}_1 = a \\ q_2 &\rightarrow q_2 + b\delta; \mathcal{F}_2 = b \\ q_3 &\rightarrow q_3 + c\delta; \mathcal{F}_3 = c. \end{aligned} \tag{2.125}$$

Hence, replacing δq_α with $\mathcal{F}_\alpha(q)\delta$, where δ is a constant, we observe that a particular quantity has not changed in the past five minutes, i.e., it is conserved

$$\mathcal{Q} = \sum_\alpha p_\alpha \mathcal{F}_\alpha(q). \tag{2.126}$$

Modeling yourself to be a free-particle, your kinetic energy is of the form $T = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)$ and since your were unchanged by translations along x, y, and z axes[4],

$$\bar{T} = \frac{m}{2}\left(\left(\dot{x} + \dot{\delta}\right)^2 + \left(\dot{y} + \dot{\delta}\right)^2 + \left(\dot{z} + \dot{\delta}\right)^2\right) = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = T. \tag{2.127}$$

Hence, $\mathcal{L} = \bar{\mathcal{L}}$, and observing that $\mathcal{F}_\alpha(q) = 1$ here, the conserved quantity is

$$\mathcal{Q} = \sum_\alpha p_\alpha \mathcal{F}_\alpha(q) = p_\alpha, \tag{2.128}$$

i.e., linear momenta in x (when $\alpha = 1$), y (when $\alpha = 2$), and z (when $\alpha = 3$) directions are conserved.
The same can be demonstrated in an alternate manner. Consider the following

[4] i.e., $x \rightarrow x + \delta$, $y \rightarrow y + \delta$, and $z \rightarrow z + \delta$.

action functional

$$\mathcal{S} = m \int \sqrt{g_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}}, \tag{2.129}$$

where, the Lagrangian is $m\sqrt{g_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}}$. What transformation would leave this invariant that depends on $g_{\mu\nu}$? Assuming the Minkowski metric, i.e., $g_{\mu\nu} = \eta_{\mu\nu}$, then the Lagrangian is invariant under displacements $x^{\mu} \to x^{\mu} + a^{\mu}$ and we obtain the conserved quantity to be

$$\mathcal{Q} = \eta_{\mu\nu} a^{\mu} \frac{m\dot{x}^{\nu}}{\sqrt{\eta_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}}} = p_{\mu} a^{\mu}, \tag{2.130}$$

which holds for any and all a^{μ}, and hence, the coefficient of each component a^{μ} must be conserved. Therefore, the conserved quantities are

$$p_{\mu} = \eta_{\mu\nu} \frac{m\dot{x}^{\nu}}{\sqrt{\eta_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}}}, \tag{2.131}$$

which is nothing but the conservation of the 4-momentum of a test particle.

2.8.2 Mathematical Formulation of Noether's Theorem

Consider a functional of the following form

$$S[y] = \int_{\alpha}^{\beta} F(x, y, y')\, dx, \tag{2.132}$$

where x is an independent variable and $y = (y_1, y_2, y_3, ...)$ is a vector of n dependent variables. The functional has stationary paths defined by the Euler-Lagrange equations. Noether's theorem throws light upon how the value of this functional is affected by families of continuous transformations of the dependent and independent variables. Consider the following transformations defined in terms of a single parameter ϵ

$$\begin{aligned} \bar{x} &= \Lambda(x, y, y'; \epsilon) \\ \bar{y}_m &= \Xi_m(x, y, y'; \epsilon), \end{aligned} \tag{2.133}$$

for $m = 1, 2, ..., n$. Λ and Ξ are assumed to have continuous first derivatives with respect to all the variables, including ϵ. The transformations must reduce to identities when $\epsilon = 0$, i.e.,

$$\begin{aligned} \bar{x} &= \Lambda(x, y, y'; 0) \\ \bar{y}_m &= \Xi_m(x, y, y'; 0), \end{aligned} \tag{2.134}$$

for $m = 1, 2, ..., n$. Now, treating $\bar{x}$ and $\bar{y}_m$ as functions of the parameter ϵ and performing a Taylor expansion of them about $\epsilon = 0$ we get the following

$$\begin{aligned} \bar{x}(\epsilon) &= \bar{x}(0) + \left(\tfrac{\partial \Lambda}{\partial \epsilon}\right)_{\epsilon=0} (\epsilon - 0) + \mathcal{O}\left(\epsilon^2\right) \\ &= x + \epsilon \Lambda + \mathcal{O}\left(\epsilon^2\right), \end{aligned} \tag{2.135}$$

and,

$$\begin{aligned} \bar{y}_m(\epsilon) &= \bar{y}_m(0) + \left(\tfrac{\partial \Xi}{\partial \epsilon}\right)_{\epsilon=0} (\epsilon - 0) + \mathcal{O}\left(\epsilon^2\right) \\ &= y_m + \epsilon \Xi + \mathcal{O}\left(\epsilon^2\right), \end{aligned} \tag{2.136}$$

where $\Lambda(x, y, y') = \left(\frac{\partial \Lambda}{\partial \epsilon}\right)_{\epsilon=0}$ and $\Xi_m(x, y, y') = \left(\frac{\partial \Xi}{\partial \epsilon}\right)_{\epsilon=0}$, for $m = 1, 2, ..., n$. What Noether's theorem states is that when the action functional $\mathcal{S}[y]$ is invariant under the above transformations, i.e.,

$$\int_{\bar{\gamma}}^{\bar{\delta}} F(\bar{x}, \bar{y}, \bar{y}')\, d\bar{x} = \int_{\gamma}^{\delta} F(x, y, y')\, dx, \tag{2.137}$$

for all γ and δ such that $\alpha \leq \gamma < \delta \leq \beta$, where $\bar{\gamma} = \Lambda(\gamma, y(\gamma), y'(\gamma))$ and $\bar{\delta} = \Xi(\delta, y(\delta), y'(\delta))$, then for each stationary path of the action functional, the following equation holds good

$$\sum_{m=1}^{n} \frac{\partial F}{\partial y'_m} \Xi_m + \left(F - \sum_{m=1}^{n} y'_m \frac{\partial F}{\partial y'_m}\right) \Lambda = \mathbf{C}, \tag{2.138}$$

where C is a constant. Consider a test particle of mass m moving in a straight path in a time-independent potential $\Phi(x)$ with it's position at a time t given by the function $x(t)$. We know from Lagrangian mechanics that the path followed by the particle will be the path of least action, whose action functional is

$$\int_0^{\tau} \mathcal{L}(x, \dot{x}) dt = \int_0^{\tau} \left(\frac{1}{2} m \dot{x}^2 - \Phi(x)\right) dt. \tag{2.139}$$

Observe that the Lagrangian has no explicit time dependence since the potential field is time-independent. Hence, we might expect the functional to be invariant under translations in time, and thus Noether's theorem to hold. Consider the following time translation

$$\bar{t}(\epsilon) = t + \epsilon \Lambda + \mathcal{O}\left(\epsilon^2\right) = t + \epsilon, \tag{2.140}$$

and the following space translation

$$\bar{x}(\epsilon) = x + \epsilon \times 0 + \mathcal{O}\left(\epsilon^2\right) = x. \tag{2.141}$$

For the case of a simple time translation by an infinitesimal amount ϵ, $\Lambda = 1$

and observing the second equation where $\Xi = 0$ simply reflects the fact that we are only translating in the time direction. The invariance of the action under these transformations is expressed as follows

$$\int_0^{\bar{\tau}} \mathcal{L}\left(\bar{x}, \dot{\bar{x}}\right) d\bar{t} = \int_{\bar{0}-\bar{\epsilon}}^{\bar{\tau}-\bar{\epsilon}} \mathcal{L}\left(x, \dot{x}\right) dt = \int_0^{\tau} \mathcal{L}\left(x, \dot{x}\right) dt, \tag{2.142}$$

where the limits in the second integral follow from the change of the time variable from $\bar{\tau}$ to τ. Hence, Noether's theorem holds and reduces to the following with $\Lambda = 1$ and $\Xi = 0$

$$\mathcal{L} - \dot{x}\frac{\partial \mathcal{L}}{\partial \dot{x}} = \mathcal{C}, \tag{2.143}$$

where $\mathcal{C}$ is a constant. Evaluation yields

$$\begin{aligned} &\tfrac{1}{2}m\dot{x}^2 - \Phi(x) - \left(m\dot{x}^2\right) = 0 \\ &\tfrac{1}{2}m\dot{x}^2 + \Phi(x) = \mathcal{E} = \mathcal{C}, \end{aligned} \tag{2.144}$$

which is the conservation of energy.

2.8.3 Noether's Theorem in General Relativity

As discussed previously, Lagrangian corresponds to a local coordinate description to a function. Let x denote the independent variables x^α, where $\alpha = 1, ..., n$, let q denote the dependent variables q^b, where $b = 1, ..., m$, and let v denote the first derivatives of dependent variables v^b_α, $n \times m$ matrices. Then the Lagrangian L, a function of (x, q, v), is given by

$$\mathcal{L} = \mathcal{L}(x, q, v). \tag{2.145}$$

The *canonical momentum* is given by

$$p^\alpha_k = \frac{\partial \mathcal{L}}{v^k_\alpha}. \tag{2.146}$$

Let's now define *canonical stress* as follows

$$T^\alpha_\beta = p^\alpha_k v^k_\alpha - \mathcal{L}\delta^\alpha_\beta, \tag{2.147}$$

and let X^μ be a vector field generating a 1-parameter group of transformations of this domain leaving the Lagrangian invariant. Then, *Noether current* is defined as follows

$$J^\alpha = T^\alpha_\beta X^\beta. \tag{2.148}$$

Noether current is divergence-free, i.e.,

$$\partial_\alpha J^\alpha = 0. \tag{2.149}$$

Now, consider the Einstein-Hilbert action, sans the matter action and the cosmological constant, we have

$$\mathcal{L}_{EH} = -\frac{1}{16\pi\kappa} R, \tag{2.150}$$

where R is the Ricci scalar curvaturre and depends upon second derivatives of the metric tensor. This Lagrangian is the only one which gives rise to second order Euler–Lagrange equations (proven below).
Let the Lagrangian $\mathcal{L}$ be a function of $g^{\mu\nu}$, $g^{\mu\nu}_{,\alpha}$, $g^{\mu\nu}_{,\alpha\beta}$, $q_{(k)}$ which is an arbitrary space function, and $q_{(k),\xi}$. Then, by varying the action as follows

$$\delta \int \mathcal{L}_{EH} \sqrt{-g} d^4x, \tag{2.151}$$

we obtain differential equations equal in number to the functions[5]. The assumption that the Lagrangian is linear in $g^{\mu\nu}_{\alpha\beta}$ such that the coefficients of $g^{\mu\nu}_{\alpha\beta}$ depend only upon $g^{\mu\nu}$ is to be made so that we can replace the action with a more convenient form as follows

$$\int \mathcal{L}_{EH} \sqrt{-g} d^4x = \int \mathcal{L}^*{}_{EH} \sqrt{-g} d^4x + \zeta, \tag{2.152}$$

where ζ is an integral extended over the boundaries of the domain under consideration and $\mathcal{L}^*{}_{EH}$ depends only upon $g^{\mu\nu}$, $g^{\mu\nu}_{,\alpha}$, $q_{(k)}$, $q_{(k)\xi}$, but no longer on $g^{\mu\nu}_{\alpha\beta}$.By varying the above form we get a new variational form as follows

$$\delta \int \mathcal{L}^*{}_{EH} \sqrt{-g} d^4x. \tag{2.153}$$

By varying the action with respect to $g^{\mu\nu}$ we obtain,

$$\frac{\partial}{\partial x_\alpha} \frac{\partial \mathcal{L}^*}{\partial g^{\mu\nu}_\alpha} - \frac{\partial \mathcal{L}^*}{\partial g^{\mu\nu}} = 0, \tag{2.154}$$

and by varying it with $q_{(k)}$ we obtain the following equation

[5] $g^{\mu\nu}$ and $q_{(k)}$ are to be varied independently of each other such that at the boundaries of integration $\delta q_{(k)}$, $\delta g^{\mu\nu}$, and $\frac{\partial \delta g^{\mu\nu}}{\partial x_\alpha}$ all vanish

$$\frac{\partial}{\partial x_\alpha}\frac{\partial \mathcal{L}^*}{\partial q_{(k)\xi}} - \frac{\partial \mathcal{L}^*}{\partial q_{(k)}} = 0. \tag{2.155}$$

Noether's theorem depends on having a Lagrangian containing only the first derivatives of the unknown functions, whereas the Einstein-Hilbert Lagrangian contains second derivatives of the metric tensor. Thus, we cannot apply Noether's theorem directly to the gravitational Lagrangian. To make it easier, we introduce a new Lagrangian

$$\mathcal{L}^*{}_{EH} = -\frac{1}{16\pi\kappa}\mathcal{L}_{EH} + \partial_\omega \Lambda^\omega, \tag{2.156}$$

where,

$$\Lambda^\omega = -\frac{1}{16\pi\kappa}\sqrt{-g}\left(g^{\mu\nu}\Gamma^\omega_{\mu\nu} - g^{\mu\omega}\Gamma^\nu_{\mu\nu}\right), \tag{2.157}$$

and by doing this, we observe that $\mathcal{L}^*$ differs from the Einstein-Hilbert Lagrangian by a divergence, i.e.,

$$\mathcal{L}^*{}_{EH} = -\frac{1}{16\pi\kappa}\sqrt{-g}g^{\mu\nu}\left(\Gamma^\beta_{\mu\omega}\Gamma^\omega_{\nu\beta} - \Gamma^\beta_{\mu\nu}\Gamma^\omega_{\beta\omega}\right), \tag{2.158}$$

and hence, will yield the same field equations. Now, we define the canonical momentum as follows

$$p^{*\mu\omega\beta} = \frac{\partial \mathcal{L}^*{}_{EH}}{\partial v_{\mu\omega\beta}}, \tag{2.159}$$

and the canonical stress as follows

$$T^{*\mu}_\nu = p^{*\mu\omega\beta}v_{\mu\omega\beta} - \mathcal{L}^*{}_{EH}\delta^\mu_\nu. \tag{2.160}$$

Observe that the Lagrangian $\mathcal{L}^*{}_{EH}$ is invariant under translations such as $x^\mu \to x^\mu + \rho^\mu$ [6]. Now, the Noether current is found to be

$$J^{*\mu} = T^{*\mu}_\nu X^\nu, \tag{2.161}$$

and by Noether's theorem this current is divergence-free, i.e.,

$$\partial_\mu J^{*\mu} = 0, \tag{2.162}$$

and since c^μ are constants, the following conservation law holds true

---

[6] ρ^μ are constants and hence, a vector field generating a 1-parameter group of translations is defined as: $X = c^\alpha \frac{\partial}{\partial x^\alpha}$

$$\partial_\mu T^{*\mu}_{\nu} = 0. \tag{2.163}$$

Thus, the conservation of the stress-energy tensor arises organically from Noether's theorem.

2.8.4 Noether's Theorem and the Energy Momentum Tensor

We know very well that the Lagrangian is a function associated to the system and contains all necessary information about its dynamics. Let a system be described by the maps $\phi : \mathcal{M} \to \mathcal{R}$, then it's Lagrangian has the following form

$$\mathcal{L} : \left(R, T^*\mathcal{M}, g^{-1}\right) \to \mathcal{R}, \tag{2.164}$$

where g^{-1} is the inverse of the pseudo-Riemannian metric on $\mathcal{M}$. Thus the Lagrangian is $\mathcal{L} = \mathcal{L}\left(\phi, d\phi, g^{-1}\right)$. The action functional of such a Lagrangian has the form $\mathcal{S}(\phi) = \int_{\mathcal{M}} \mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg$. The action of the Lagrangian gives us the laws of evolution of the system. Thus, if we want to have a version of the conservation of energy then we should study the action $\mathcal{S}$ more geometrically. For the computation of $\mathcal{S}(\phi)$ we require a coordinate system x on $\mathcal{M}$ and need to construct diffeomorphisms on $\mathcal{M}$ since every diffeomorphism $f : \mathcal{M} \to \mathcal{M}$ induces a pullback coordinate system $y = f^*x$. Consider X, an arbitrary vector field. Since we seek for an object on M with local properties we may assume that X is compactly supported, i.e., $X = 0$ outside a compact region $U \subset M$. Let $\mathcal{F}_t$ be the associated flow of X. Now, each such diffeomorphism $\mathcal{F}_t$ defines a pullback coordinate system $y_t = \mathcal{F}_t^* x$ whose change of coordinates is $\mathcal{F}_t$. Then

$$\int_{\mathcal{M}} \mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg = \int_{F^{-1}(\mathcal{M})} \mathcal{F}_t^* \left(\mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg\right), \tag{2.165}$$

since $\mathcal{F}^{-1}(\mathcal{M}) = \mathcal{M}$,

$$\int_{\mathcal{M}} \mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg = \int_{\mathcal{M}} \mathcal{F}_t^* \left(\mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg\right). \tag{2.166}$$

This gives us

$$\begin{aligned} &\int_{\mathcal{M}} \frac{\mathcal{F}_t^*\left(\mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg\right) - \mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg}{t} = 0 \\ &\int_{\mathcal{M}} lt_{t\to 0} \left(\frac{\mathcal{F}_t^*\left(\mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg\right) - \mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg}{t}\right) = 0, \end{aligned} \tag{2.167}$$

and from the definition of Lie derivative we get

$$\int_{\mathcal{M}} \mathcal{L}_X \left(\mathcal{L}\left(\phi, d\phi, g^{-1}\right) dg\right). \tag{2.168}$$

Note that we were led to the above equation by only using the geometric structure of M. Upon application of the *Leibniz* and chain rule we obtain the following

$$\int_{\mathcal{M}} \left(\left(\frac{\partial \mathcal{L}}{\partial \phi} \mathcal{L}_X \phi + \frac{\partial \mathcal{L}}{\partial d\phi} \mathcal{L}_X d\phi + \frac{\partial \mathcal{L}}{\partial g^{-1}} \mathcal{L}_X g^{-1} \right) dg + \mathcal{L} \cdot \mathcal{L}_X dg \right). \tag{2.169}$$

Now, we make the following definitions and observations needed to solve the above integral:

a. Define *Cartlan's identity*, which is given by

$$\mathcal{L}_X d\phi = d(X\phi) = d(\mathcal{L}_X \phi). \tag{2.170}$$

b. Now, consider the vector field $\left(\frac{\partial \mathcal{L}}{\partial d\phi}\right)^k = \left(\frac{\partial \mathcal{L}}{\partial \phi_k}\right)$, where $\phi_k = \partial\phi/\partial x^k$. Then

$$\nabla \left(\left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) \mathcal{L}_X \phi \right) = d(\mathcal{L}_X \phi) \left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) + (\mathcal{L}_X \phi) \cdot \nabla \left(\left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) \right). \tag{2.171}$$

c. Application of the divergence theorem for $\left(\left(\frac{\partial \mathcal{L}}{\partial d\phi}\right) \mathcal{L}_X \phi\right)$ yields

$$\int_{\mathcal{M}} \nabla \left(\left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) \mathcal{L}_X \phi \right) dg = \int_{\partial \mathcal{M}} \left(\left(\left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) \mathcal{L}_X \phi \right) \cdot n \right) g d_{n-1} = 0, \tag{2.172}$$

since the boundary condition $X = 0$ implies that $\mathcal{L}_X \phi = X(\phi) = 0$ on ∂M.
d. Equations derived/defined in *a*, *b*, and *c* imply the following

$$\int_M \left(\left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) \mathcal{L}_X \phi \right) dg = \int_M (\mathcal{L}_X \phi) \left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) dg = - \int_M (\mathcal{L}_X \phi) \nabla \left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) dg. \tag{2.173}$$

Thus, using equation 2.173, we can rewrite the integral as follows

$$\int_{\mathcal{M}} \left(\left(\frac{\partial \mathcal{L}}{\partial \phi} \mathcal{L}_X \phi - (\mathcal{L}_X) \cdot \nabla \left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) + \frac{\partial \mathcal{L}}{\partial g^{-1}} \mathcal{L}_X g^{-1} \right) dg + \mathcal{L} \cdot \mathcal{L}_X dg \right) = 0, \tag{2.174}$$

which when simplified gives

$$\int_{\mathcal{M}} \left(\left[\partial \mathcal{L} \partial \phi - \nabla \left(\frac{\partial \mathcal{L}}{\partial d\phi} \right) \right] \mathcal{L}_X \phi + \frac{\partial \mathcal{L}}{\partial g^{-1}} \mathcal{L}_X g^{-1} \right) dg + \mathcal{L} \cdot \mathcal{L}_X dg \right) = 0. \tag{2.175}$$

In general, if $\phi : \mathcal{M} \to \mathcal{N}$ then we have the following Euler-Lagrange system of equations for every component of ϕ

$$\frac{\partial \mathcal{L}}{\partial \phi} - \nabla\left(\frac{\partial \mathcal{L}}{\partial d\phi}\right) = 0. \tag{2.176}$$

Thus, using this to simplify our integral we get

$$\int_{\mathcal{M}} \left(\left(\frac{\partial \mathcal{L}}{\partial g^{-1}}\mathcal{L}_X g^{-1}\right) dg + \mathcal{L}\cdot\mathcal{L}_X dg\right) = 0, \tag{2.177}$$

which can written alternatively as

$$\int_{\mathcal{M}} \left(\frac{\partial \mathcal{L}}{\partial g^{-1}}\mathcal{L}_X g^{-1} + \mathcal{L}\cdot\left(\frac{1}{2}g_{\mu\nu}(\mathcal{L}_X g)^{\mu\nu}\right)\right) dg = 0, \tag{2.178}$$

or,

$$\int_{\mathcal{M}} \left(\frac{\partial \mathcal{L}}{\partial g^{\mu\nu}}(\mathcal{L}_X g^{-1})^{\mu\nu} + \mathcal{L}\cdot\left(\frac{1}{2}g_{\mu\nu}(\mathcal{L}_X g)^{\mu\nu}\right)\right) dg = 0 \tag{2.179}$$

which finally yields

$$\int_{\mathcal{M}} \left((\mathcal{L}_X g^{-1})^{\mu\nu}\left[\frac{1}{2}g_{\mu\nu}(\mathcal{L}_X g)^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial g^{\mu\nu}}\right]\right) dg = 0. \tag{2.180}$$

Now, we define the stress-energy-momentum tensor as follows

$$T_{\mu\nu} = \left[\frac{1}{2}g_{\mu\nu}(\mathcal{L}_X g)^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial g^{\mu\nu}}\right]. \tag{2.181}$$

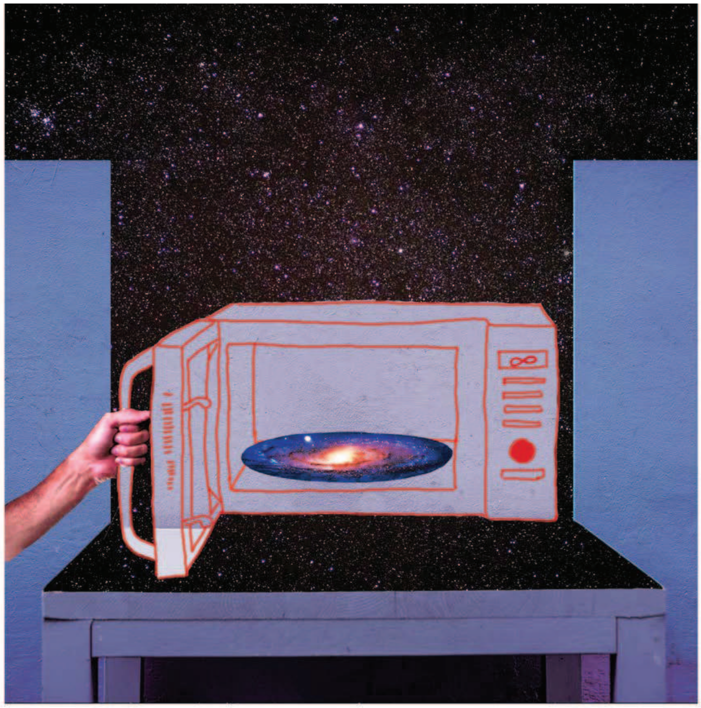
From observation, we see that $T_{\mu\nu}$ is symmetric[7].

The geometry of the space gave us equation 2.166 which was the key for the derivation of energy momentum tensor. It states that the average of those Lagrangians over the manifold $\mathcal{M}$ is the same. The equation

$$\mathcal{L}\left(\phi, d\phi, g^{-1}\right) = \mathcal{F}_t^*\left(\mathcal{L}\left(\phi, d\phi, g^{-1}\right)\right), \tag{2.182}$$

expresses the 1-parameter symmetries of the Lagrangian which eventually gave rise to the conservation of a quantity which was found to be the stress-energy tensor. Thus, in accordance to Noether's theorem, continuous symmetries give rise to conservation laws.

[7] Also, since the integral of the product of T and a derivative is zero then by integrating by parts we expect the integral of the product of the divergence of T and a vector field to be zero. Since the latter is going to be true for any vector field then we expect that the divergence of T vanishes. This is exactly what we were looking for and so it is the most natural candidate for the energy momentum tensor of the Lagrangian matter field ϕ



3

Lé Schwarzschild Solution

3.1 Introduction

Any metric whatsoever is a solution of the field equations as long as there are no restrictions imposed on the stress-energy tensor, since Einstein's equations then become just a definition of $T_{\mu\nu}$; thus, we are to first make assumptions about $T_{\mu\nu}$. Post this, we are to proceed by imposing symmetry conditions on the metric, by restricting the algebraic structure of the Riemann tensor, by adding field equations for the matter variables or by imposing initial and boundary conditions. The so called exact solutions have all been obtained by making some such restrictions. For a physical theory, we first mathematically analyse the set of differential equations and try finding as many exact solutions, or as complete a general solution, as possible. Next, comes the physical interpretations of these solutions which in the case of general relativity demands an analysis from a global perspective and the use of topological methods rather than just the purely local solution of the differential equations.

A metric would be referred to as an exact solutions if and only if its components could be expressed, in suitable coordinates, in terms of analytic functions (such as trigonometric functions, polynomials, etc.). Since general relativity is highly non-linear theory, it is not always easy to understand what qualitative features solutions might possess. In the initial years of general relativity, only a small number of exact solutions were proposed and discussed which had their origins in highly idealized physical problems, and possessed a very high degree of symmetry. Examples of these include the well-known spherical symmetric solutions of Schwarzschild (which shall be discussed in detail), Kerr, Reissner and Nordstrom, Tolman and Friedmann, Weyl, and the plane wave metrics. In the early days, relativists didn't possess high regard for the exact solutions, with the exception of cosmological and stellar models, because of the extreme weakness of the relativistic corrections to Newtonian gravity. Most of the problems that relativists tackled were marred by approximations methods, such as the weak field approximation. Probably one of the most important techniques in common use is the algebraic classification of the Weyl tensor into Petrov types and the understanding of the properties of algebraically special metrics. Another common technique, and the first to become popular, was the use of groups of motions, especially in the construction of cosmologies (more general than Friedmann's). Both the above discussed developments led to the use of invariantly-defined tetrad bases, rather than coordinate components. The null tetrad methods, and some ideas from the theory of group representations and algebraic geometry, gave rise to the spinor techniques which are now usually employed in the form given by Newman and Penrose. Using these methods, it was possible to obtain many new solutions whose growth still continues.

The Schwarzschild is the best known nontrivial exact solution of Einstein's field equations. The Schwarzschild metric describes the gravitational field in the vicinity of a gravitating object. This metric is named in honour of Karl Schwarzschild, who found the exact solution about a month after Einstein published his theory of general relativity, thus making this the first ever exact solution to Einstein's field equations. And, not only is it one of the simplest exact vacuum solutions, but it is also the most physically significant one. It is widely applied both in astrophysics and in considerations of orbital motions about the Sun or the Earth. It predicts the deviations from the Newtonian theory of gravity that are observed in the orbital motion in our solar system, and more importantly, in the deflection of light by the Sun. As we start to explore further into the derivation of this metric, one would observe that to achieve this metric almost in a month is no easy task. Yes, the calculations are taxing, but I can assure you is that it is worth the hard work. The specialty of the Schwarzschild metric is that it is a spherically symmetric solution of the Einstein field equations. The Schwarzschild metric is derived under the conditions that the Cosmological constant and the Energy-momentum tensor have null values, a solution of the field equations known as the vacuum solution. A null Energy-momentum tensor implies that neither are there sources nor any sinks of gravitational fields present in that spacetime apart from our gravitating object.

3.2 The Ricci Flatness Condition

To find the vacuum solution to Einstein's equations, set $\Lambda = 0$, and $T_{\mu\nu} = 0$, to get $R_{\mu\nu} = 0$. This relation is known as the Ricci flatness condition. This is the Einstein field equation in empty spacetime. It says that the Ricci tensor vanishes. We speak of the field equation in the singular, but in fact it consists of a set of equations according to the values taken by $\mu\nu$. It is of prime importance to note here that a vanishing Ricci tensor does not imply that the Riemann tensor vanishes too. Although we have the relation: $R_{\mu\nu}(x) = g^{\alpha\beta}(x) R_{\mu\nu\alpha\beta}(x)$, the Ricci tensor can vanish without the Riemann tensor having to do so. From the Riemann curvature tensor, we contract by setting $\alpha = \beta$, to obtain the Ricci tensor in terms of the Christoffel symbols as follows

$$\begin{aligned} R^{\alpha}_{\mu\beta\nu} &= D_{[\mu}D_{\beta]} = \left(\partial_{\beta}\Gamma^{\alpha}_{\mu\nu} + \Gamma^{\rho}_{\mu\nu}\Gamma^{\alpha}_{\beta\rho}\right) - \left(\partial_{\mu}\Gamma^{\alpha}_{\beta\nu} + \Gamma^{\rho}_{\beta\nu}\Gamma^{\alpha}_{\mu\rho}\right) \\ R_{\mu\nu} &= R^{\alpha}_{\mu\alpha\nu} = \left(\partial_{\alpha}\Gamma^{\alpha}_{\mu\nu} + \Gamma^{\rho}_{\mu\nu}\Gamma^{\alpha}_{\alpha\rho}\right) - \left(\partial_{\mu}\Gamma^{\alpha}_{\alpha\nu} + \Gamma^{\rho}_{\alpha\nu}\Gamma^{\alpha}_{\mu\rho}\right), \end{aligned} \tag{3.1}$$

As you can see from the above relations, when the Ricci tensor vanishes, the Riemann curvature tensor need not vanish. Moreover, where the Riemann curvature tensor is zero, spacetime would be flat. The metric we are trying to derive has the following form

$$ds^2 = \left(1 - \frac{r_s}{r}\right) dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 d\Omega^2 = g_{\mu\nu}(x)dx^{\mu}dx^{\nu}. \tag{3.2}$$

Before we start calculating let us analyse this metric. It may first be observed that the metric reduces to a Minkowskian one when $r \to \infty$. In the metric, the $r_s = \frac{2GM}{c^2}$, is the Schwarzschild radius. Since we are performing all calculations by setting the value of c ($c \equiv 1$), we can ignore the squared term in the Schwarzschild radius.

3.3 Singularities Already?

It can be observed that the metric degenerates when $r = 2GM$, i.e., as $r \to r_s$. This although may seem like a singularity, is a mere illusion created due to some cranky coordinates. This illusion fades away when better coordinate systems are used, for example, *Kruskal-Szekeres* coordinates, ingoing and outgoing *Eddington-Finkelstein* coordinates, *Lemaitre* coordinates, etc. If fact, it was Kruskal and Szekeres, who independently had discovered that this was a coordinate singularity that arose due to the use of the Schwarzschild coordinates. With the use of better coordinates, the metric becomes regular at $r = r_s$. However, the case of $r = 0$ is different, something terrible happens to the metric, physicists encounter their worst nightmare- infinity. As $r \to 0$, $\left(1 - \frac{r_s}{r}\right) \to -\infty$, and the most horrid part- we end up with a metric where all the terms are negative, implying that there are four positive eigen values and thus producing the metric signature-$(+, +, +, +)$. Hence, we have entered a region where there are four space directions and no time directions.
The case of $r \to 0$ is a true, physical singularity. To confirm this, we must turn to quantities that are independent of the choice of coordinates. One such is the *Kretschmann invariant*

$$K = R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} = \frac{12r_s^2}{r^6} = \frac{48G^2M^2}{c^4r^6} \tag{3.3}$$

As $r \to 0$, the curvature becomes infinite. Since curvature has dimensions of force, we say that the tidal forces at the point $r = 0$ is infinite and at that point, spacetime itself is not properly defined. It is important to note that this invariant is regular at $r = r_s$, i.e., the invariant corresponds to a finite tidal force. Hence, we conclude that the singularity at $r = r_s$ is not a real, physical singularity. There is another method in which we make a coordinate transform to such a metric tensor which is regular at this surface, in order to observe that the spacetime is regular at $r = r_s$. Such a coordinate transform is done using the *Tortoise coordinates*, which we shall explore in the next chapter.

The Schwarzschild metric represents the external gravitational field of a spherical gravitating object of mass m . Similar to the Newtonian picture, where the mass of the gravitating source of a stationary, asymptotically flat (i.e. the spacetime is flat at large distances from the source) gravitational field can be determined by integrating over a closed surface near spatial infinity, in general relativity, a similar integral has been introduced called the *Komar integral* which yields the *Komar mass* (see next chapter). The parameter m can be interpreted as the total mass, which is spherically distributed, inside the radius r. In the real world, the mass of all particles is necessarily positive. Thus, it is generally assumed that $m > 0$. However, as an exact mathematical solution of Einstein's equations, the Schwarzschild metric is also valid when m is negative. In such a case, the apparent singularity at $r = 2m$ would not appear, and the corresponding spacetime would not possess a similar interpretation. We shall discuss deeper questions pertaining to the mass of black holes in the latter part of the chapter.

3.4 Spherical Symmetry

Using the knowledge of intense mathematical foundations laid in the introductory chapter, we can reformulate concepts underlying the Schwarzschild metric very elegantly as follows: A three-dimensional Riemannian manifold $(\mathcal{M}, g)$ is said to be spherically symmetric if: The manifold is represented by one chart (V, z) with $z(V) = \mathbb{R}^3$, i.e., the image of the open set V, $z(V)$ exists in $\mathbb{R}^3$. An analogous way of conveying this is by placing a sphere $\mathcal{S}$ of $\mathbb{R}^3$ centred at some arbitrary point p. The pseudo-coordinates in $z(V)$, i.e., the spherical polar coordinates ρ, θ, and Φ are linked to the canonical coordinates x, y, and z of $\mathbb{R}^3$ by the following relations

$$x = \rho sin\theta sin\Phi, y = \rho sin\theta cos\Phi, z = \rho cos\theta. \tag{3.4}$$

In $z(V)$, g is represented by the metric of the following form

$$e^{u(\rho)} d\rho^2 + \kappa^2(\rho) d\Omega^2. \tag{3.5}$$

What these conditions imply is that $z(V)$ is foliated by a metric 2-spheres of constant ρ centred at the arbitrary point p. The areas of the spheres from the metric are $4\pi\kappa^2$. This metric is the general form of a metric invariant by rotations in R^3, centred at p. It is important to note here that the vanishing of $\kappa^2(0)$ does not imply the presence of a singularity in the metric, it just shows that spherical coordinates are not the best ones available at the point $\rho = 0$.

3.5 Constructing the Schwarzschild Metric

At large distances from the spherically symmetric gravitating object, spacetime is flat. Why? This is due to the fact that the influence of the gravitational field vanishes at large distances since it varies as $\approx r^{-2}$. Due to this, the ability of the tidal forces to curve spacetime at large distances from the gravitating object fades away thus resulting in a flat space. Note that when we talk about the gravitating object, we have assumed that there is no matter in the surroundings of our object, thus the vacuum field. We can represent this flat spacetime in terms of the line element as

$$ds^2 = dt^2 - \frac{1}{c^2}\left(dx^2 + dy^2 + dz^2\right). \tag{3.6}$$

Now, how would this metric change in the vicinity of our object? For this let us first write the basic form of the metric for a plane in polar coordinates. In flat space, the spatial distance between two points on a plane in polar coordinates is given by the equation

$$ds^2 = r^2 d\theta^2 + dr^2. \tag{3.7}$$

Now, let us change this metric. We first start by making the replacements, $sin\theta \to sinh\omega$, and $cos\theta \to cosh\omega$. Here, ω is the angle with which the hyperbola increases with respect to the origin (a timelike coordinate). Thus, we have changed from polar coordinates to hyperbolic coordinates. In this frame, the acceleration along a particular hyperbola is the same, however, the acceleration along different hyperbolae are different. An analogous relation can be drawn to that of circular motion here, similar to the acceleration remaining the same along a particular hyperbola, the acceleration of a particle moving around a circle is uniform, however, the acceleration around another concentric circle of a different radius who definitely not be the same. We now apply the following transformations

$$X = rcosh(\omega), T = rsinh(\omega), \tag{3.8}$$

such that,

$$X^2 - T^2 = r^2\left[cosh^2(\omega) - sinh^2(\omega)\right] = r^2. \tag{3.9}$$

Hence, producing the following metric

$$ds = r^2 d\omega^2 - dr^2. \tag{3.10}$$

Now, this is the metric in which our gravitating object lies, let us travel along a particular hyperbola and to determine the fate of the metric. Let the gravitating object, under consideration, be the super massive black hole located at the centre of our galaxy. Now, let us remove all the matter present outside this black hole ($T_{\mu\nu} = 0$) and for the moment assume that the value of the energy density of the vacuum of space is zero ($\Lambda = 0$). By performing these actions, we have established the vacuum conditions. From a small distance from a hyperbola that is present next to where Earth was, just a moment ago, we compute the metric. The black hole is almost 26,000 light years away from Earth. Placing the origin at the centre of the black hole, we re-define the position vector to be

$$r = R_{BH\to Hyp} + r' \tag{3.11}$$

where $R_{BH\to Hyp} = 26,000$ light years, is the distance between the black hole and the hyperbola trajectory which runs next to where Earth was a moment ago, and r' is the distance between the hyperbola and us. This distance is prone to vary since were nothing but mere particles floating in space but would never exceed that of $R_{BH\to Hyp}$, hence, $\frac{r'^2}{R^2_{BH\to Hyp}} \to 0$. Let us substitute this new relation into the metric and perform some manipulations,

$$\begin{aligned} ds^2 &= (R_{BH\to Hyp} + r')^2\, d\omega^2 - [d(R_{BH\to Hyp} + r')]^2 \\ ds^2 &= (R^2_{BH\to Hyp} + r'^2 + 2R_{BH\to Hyp} r')d\omega^2 - dr'^2 \\ &\approx \left(1 + \frac{2r'}{R_{BH\to Hyp}}\right) R^2_{BH\to Hyp} d\omega^2 - dr'^2. \end{aligned} \tag{3.12}$$

We know from previous chapters that proper acceleration, A, when the speed of light is set to unity is nothing but $\frac{1}{R}$. Hence, here, $A = \frac{1}{R_{BH\to Hyp}} = \mathbf{g}$. Define $R_{BH\to Hyp}(\omega) = t$.

$$ds^2 = (1 + 2r'\mathbf{g})\, dt^2 - dr'^2. \tag{3.13}$$

For a spherically symmetric gravitating object, we know from Newton's theory that the gravitational field is given by: $\mathbf{g} = -\frac{GM}{r'^2}$. Where M is the mass of the black hole and re-defining r' by stating that it is the distance at which we are placed next to the black hole (note that $r' > r_s$). We now get the final form of the metric to be

$$ds^2 = \left(1 - \frac{2GM}{r'}\right) dt^2 - dr'^2 = [1 + 2\Phi(r')]\, dt^2 - dr'^2. \qquad (3.14)$$

It is interesting to note that as $r' \to 2GM, d\tau^2 \to -dr'^2$, and the vanishing coefficient $\left(1 - \frac{2GM}{r'}\right)$, happens to be the black hole horizon. When the speed of light is not set to unity, and when $r' = x + y + z$, we obtain the metric of the form

$$ds^2 = \left(1 - \frac{2GM}{r'c^2}\right) dt^2 - \frac{1}{c^2}\left(dx^2 + dy^2 + dz^2\right). \qquad (3.15)$$

We now have a problem with the coordinate system, it's too hard to study central force problems in Cartesian coordinates. We move to another system which is tailor-made for studying the central force problem and one which acknowledges the spherical symmetry of our gravitating object-the spherical polar coordinates. Thus, we change the flat space metric by transformations to give

$$dx^2 + dy^2 + dz^2 = dr^2 + r^2\left(d\theta^2 + sin^2\theta d\Phi^2\right) = dr^2 + r^2 d\Omega^2. \qquad (3.16)$$

Substituting this into our metric, we obtain

$$ds^2 = \left(1 - \frac{2GM}{rc^2}\right) dt^2 - \frac{1}{c^2} dr^2 - \frac{1}{c^2} r^2 d\Omega^2. \qquad (3.17)$$

This whole struggle was to find the metric in the vicinity of our gravitating object. Let us recap- we started with the polar form of the metric describing a plane, we then transformed this metric to obtain a hyperbolic form, we then manipulated this form replacing the position vector with the net distance between us and the black hole, during simplification we realised the term that represented proper acceleration and that we could place a gravitating body with a gravitational field g in that spacetime. With such an object in place, we then realized that the first diagonal term of the metric represented the Newtonian potential, and since the object that we have placed is spherically symmetric, we have changed from Cartesian to spherical polar coordinates. So, this leaves us with only one question- is this metric correct? The answer- absolutely not! This is due to the reason that was mentioned on the paragraph about singularities. Let us discuss this point in detail and plan a surgery to obtain the correct metric. In the limit of $r \to 0$, the metric goes nuts

$$\begin{aligned} &ds^2 = \left(1 - \tfrac{2GM}{rc^2}\right) dt^2 - \tfrac{1}{c^2} dr^2 - \tfrac{1}{c^2} r^2 d\Omega^2 = -g_{\mu\nu}(x) dx^\mu dx^\nu \\ &lim_{r\to 0}\ ds, \quad \left(1 - \tfrac{2GM}{rc^2}\right) \to -\infty. \end{aligned} \qquad (3.18)$$

Hence, we obtain $|g_{\mu\nu}| = Diag\,(+\infty, A, B, C)$, where A, B,and C are positive constants. We ended up obtaining a metric where all the terms are negative, implying that there are four positive eigen values and thus producing the metric signature-$(+ \ + \ + +)$. Hence, we have entered a region where there are four space directions and no time directions. This is the reason why the metric derived is wrong. Let's perform a surgery to this metric such that we get the acceptable one. One might immediately suggest to add the term $\left(1 - \frac{2GM}{rc^2}\right)$ to the dr and $d\Omega$ components of the metric. But the correct metric derived from Einstein's field equations is of the form

$$ds^2 = \left(1 - \frac{2GM}{rc^2}\right) dt^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1} \frac{1}{c^2} dr^2 + \frac{1}{c^2} d\Omega^2. \tag{3.19}$$

Note that when dt^2 flips sign, dr^2 flips sign too. What this implies is that outside the Schwarzschild radius, there is no change in the metric and its signature is $(- \ + \ + +)$, i.e., one time direction and three space directions (r, θ, Φ). However, inside the Schwarzschild limit, space and time exchange their roles. The correction, $\left(1 - \frac{2GM}{rc^2}\right)^{-1}$, is quite negligible as long as we don't get too close where r is too small.

3.6 The Derivation

Here comes the easy part, the mathematical derivation of the Schwarzschild solution. Math is just a route (the best-known route) for us to utilize in order to get to the crux of the problem in hand. Without further ado, let's divulge into the math and derive the Schwarzschild solution. We make an ansatz (an educated guess) of a metric such that it preserves spherical symmetry, it has the following form

$$ds^2 = g_{tt} dt^2 + g_{tr} dt dr + g_{rt} dr dt + g_{rr} dr^2 + \zeta d\Omega^2. \tag{3.20}$$

Since $g_{tr} = g_{rt}$, we can re-write the metric in its final form as follows

$$ds^2 = g_{tt} dt^2 + 2 g_{tr} dt dr + g_{rr} dr^2 + \zeta d\Omega^2, \tag{3.21}$$

where $d\Omega^2 = d\theta^2 + sin^2\theta d\Phi^2, r \in [0, \infty), \Phi \in [0, 2\pi)$, and $\theta \in [0, \pi]$. Taking a time-slice, i.e., at a constant time $(dt = 0)$, space is sliced by concentric spheres, spheres whose radii are set up by $g_{rr}(r, t)$, and spheres whose areas are set up by $\zeta(r, t)$. The metric, when $dt = 0$ is given by

$$ds^2 = g_{rr}(r, t) dr^2 + \zeta(r, t) d\Omega^2. \tag{3.22}$$

This form of the metric is invariant under coordinate changes: $r = r(r', t^{'})$, and $t = t(r', t^{'})$. Now the metric tensor and $\zeta(r, t)$ would transform as follows

$$\begin{aligned} g'_{\alpha\beta}(y) &= g_{\mu\nu}(x)\frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial x^\nu}{\partial y^\beta} \\ \zeta'(y) &= \zeta[x(y)]. \end{aligned} \tag{3.23}$$

It is important to observe that $\zeta'(y)$ transforms as a scalar. Here $y^\alpha = (r,t)$, where α runs from 1 to 2. Using this freedom of choice of two functions, $r(r',t')$, and $t(r',t')$, we can choose two out of the four functions, $g_{tt}(r,t), g_{rt}(r,t), g_{rr}(r,t)$, and $\zeta(r,t)$. We fix the following standard notations for $g_{tt}(r,t)$, and $g_{rr}(r,t)$

$$g_{tt}(r,t) = e^{A(r,t)}, \quad g_{rr} = -e^{B(r,t)}. \tag{3.24}$$

Thus, we have arrived at the covariant form of our ansatz

$$ds^2 = e^{A(r,t)}dt^2 - e^{B(r,t)}dr^2 - r^2dr^2. \tag{3.25}$$

In the above metric, the two-dimensional surface with $r = const, t = const$ has the standard line element $dl^2 = r^2 d\Omega^2$ of the 2-sphere with the proper area $A = 4\pi r^2$. We are assuming, for the moment that $A > 0$, and $B > 0$, this implies $e^A > 1$, and $e^B > 1$ so that $t = const$ surfaces are spacelike and $r = const$ surfaces are timelike. From the metric, we observe the following

$$\begin{aligned} ||g_{\mu\nu}|| &= Diag\,(g_{00}, g_{11}, g_{22}, g_{33}) = Diag\left(e^A, -e^B, -r^2, -r^2 sin^2\theta\right) \\ ||g_{\mu\nu}|| &= Diag\left(g^{00}, g^{11}, g^{22}, g^{33}\right) = Diag\left(e^{-A}, -e^{-B}, -\tfrac{1}{r^2}, -\tfrac{1}{r^2 sin^2\theta}\right). \end{aligned} \tag{3.26}$$

Let's now calculate the Christoffel symbols for the metric. This can be a long and tiresome process, however, there are two methods of finding the Christoffel symbols- one is by direct evaluation, and the second is to obtain them from the variation of the geodesic equation. Although the latter is the preferred method, we shall do both the methods to highlight the elegance of the latter. Note the following conventions: $0 \to t, 1 \to r, 2 \to \theta, 3 \to \Phi$. We will be following the below mentioned method for the computation of the Christoffel symbols

$$\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu} = g^{\alpha\alpha}\Gamma_{\mu\nu\alpha} = g^{\alpha\alpha}\frac{1}{2}\left(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}\right). \tag{3.27}$$

The non-zero components of the Christoffel symbols are the following

$$\Gamma^1_{11} = g^{11}\Gamma_{111} = -\frac{e^{-B}}{2}\left(g_{11,1} + g_{11,1} - g_{11,1}\right) = -\frac{e^{-B}}{2}\frac{\partial e^{B(r,t)}}{\partial r} = \frac{B'}{2}$$

$$\Gamma^0_{10} = g^{00}\Gamma_{100} = -\tfrac{e^{-A}}{2}\left(g_{10,0} + g_{00,1} - g_{10,0}\right) = \tfrac{e^{-A}}{2}\tfrac{\partial e^{A(r,t)}}{\partial r} = \tfrac{A'}{2}$$

$$\Gamma^2_{33} = g^{22}\Gamma_{332} = -\frac{1}{2r^2}\left(g_{32,3} + g_{32,3} - g_{33,2}\right) = -\frac{1}{2r^2}\frac{\partial r^2 sin^2\theta}{\partial\theta} = -sin\theta cos\theta$$

$$\Gamma^0_{11} = g^{00}\Gamma_{110} = \frac{e^{-A}}{2}\left(g_{10,0} + g_{10,1} - g_{11,0}\right) = -\frac{e^{-A}}{2}\frac{\partial\left(-e^{A(r,t)}\right)}{\partial t} = \frac{\dot{B}}{2}e^{B-A}$$

$$\Gamma^1_{22} = g^{11}\Gamma_{221} = \frac{e^{-B}}{2}\left(g_{21,2} + g_{12,2} - g_{22,1}\right) = \frac{e^{-B}}{2}\frac{\partial\left(-r^2\right)}{\partial t} = -re^{-B}$$

$$\Gamma^1_{00} = g^{11}\Gamma_{001} = -\frac{e^{-B}}{2}\left(g_{01,0} + g_{10,0} - g_{00,1}\right) = \frac{e^{-B}}{2}\frac{\partial e^{A(r,t)}}{\partial r} = \frac{A'}{2}e^{A-B}$$

$$\Gamma^2_{12} = g^{22}\Gamma_{122} = -\frac{1}{2r^2}\left(g_{22,1} + g_{21,2} - g_{12,2}\right) = -\frac{1}{2r^2}\frac{\partial\left(-r^2\right)}{\partial r} = \frac{1}{r}$$

$$\Gamma^2_{12} = g^{22}\Gamma_{122} = -\frac{1}{2r^2}\left(g_{22,1} + g_{21,2} - g_{12,2}\right) = -\frac{1}{2r^2}\frac{\partial\left(-r^2\right)}{\partial r} = \frac{1}{r}$$

$$\Gamma^3_{13} = g^{33}\Gamma_{133} = \frac{-1}{2r^2 sin^2\theta}\left(g_{33,1} + g_{31,3} - g_{13,3}\right) = \frac{1}{2r^2 sin^2\theta}\frac{\partial\left(r^2 sin^2\theta\right)}{\partial r} = \frac{1}{r}$$

$$\Gamma^3_{23} = g^{33}\Gamma_{233} = \frac{-1}{2r^2 sin^2\theta}\left(g_{33,2} + g_{31,2} - g_{23,3}\right) = \frac{1}{2r^2 sin^2\theta}\frac{\partial\left(r^2 sin^2\theta\right)}{\partial\theta} = cot\theta$$

$$\Gamma^0_{00} = g^{00}\Gamma_{000} = \frac{e^{-A}}{2}\left(g_{00,0} + g_{00,1} - g_{00,0}\right) = \frac{e^{-A}}{2}\frac{\partial e^{A(r,t)}}{\partial t} = \frac{\dot{A}}{2}$$

$$\Gamma^1_{10} = g^{11}\Gamma_{101} = -\frac{e^{-B}}{2}\left(g_{01,1} + g_{11,0} - g_{10,1}\right) = -\frac{e^{-B}}{2}\frac{\partial\left(-e^{B(r,t)}\right)}{\partial t} = \frac{\dot{B}}{2}$$

$$\Gamma^1_{33} = g^{11}\Gamma_{331} = -\frac{e^{-B}}{2}\left(g_{31,3} + g_{13,3} - g_{33,1}\right) = \frac{e^{-B}}{2}\frac{\partial\left(-r^2 sin^2\theta\right)}{\partial r} = -rsin^2\theta e^{-B}. \tag{3.28}$$

The other components of the Christoffel symbols are zero. Now, let us compute the same in a different approach, let's use some physics! This method focuses on obtaining the Lagrangian from the metric and observing Christoffel symbols as part of the Euler-Lagrange equation.

We first obtain the Lagrangian from its definition, $S = \int ds = \int \mathcal{L} d\tau$, we get the Lagrangian to be

$$\Psi = \mathcal{L}^2 = \frac{ds^2}{d\tau^2} = e^{A(r,t)}\frac{dt^2}{d\tau^2} - e^{B(r,t)}\frac{dr^2}{d\tau^2} - r^2\frac{d\theta^2}{d\theta^2} - r^2 sin^2\theta\frac{d\Phi^2}{d\tau^2} = g_{\mu\nu}(x)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}$$

$$\Psi = e^{A(r,t)}\dot{t}^2 - e^{B(r,t)}\dot{r}^2 - r^2\dot{\theta}^2 - r^2 sin^2\theta\dot{\Phi}^2 = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu. \tag{3.29}$$

Now, we make use of the *Euler-Lagrange equation*, by matching the coefficients in the equation with those in the geodesic equation for each coordinate t, r, θ, Φ

$$
\begin{aligned}
\frac{d}{d\tau}\left(\frac{\partial\Psi}{\partial\dot{t}}\right) - \frac{\partial\Psi}{\partial t} &= \frac{d}{d\tau}\left(2\dot{t}e^{A}\right) - \left(\dot{A}e^{A}\dot{t}^{2} - \dot{B}e^{B}\dot{r}^{2}\right) \\
&= 2\left(e^{A}\ddot{t} + \left(\dot{A}e^{A}\dot{t} + A^{'}e^{A}\dot{r}\right)\dot{t}\right) - \left(\dot{A}e^{A}\dot{t}^{2} - \dot{B}e^{B}\dot{r}^{2}\right) \\
&= 2e^{A}\ddot{t} + 2\dot{A}e^{A}\dot{t}^{2} + 2A^{'}e^{A}\dot{r}\dot{t} - \dot{A}e^{A}\dot{r}\dot{t} + \dot{B}e^{B}\dot{r} \\
&= 2e^{A}\left(\ddot{t} + \frac{\dot{\mathbf{A}}}{\mathbf{2}}\dot{t}^{2} + \frac{\mathbf{A}^{'}}{\mathbf{2}}\dot{t}\dot{r} + \frac{\mathbf{A}^{'}}{\mathbf{2}}\dot{r}\dot{t} + \frac{\dot{\mathbf{B}}}{\mathbf{2}}\mathbf{e}^{\mathbf{B-A}}\dot{r}^{2}\right) = 0.
\end{aligned}
\tag{3.30}
$$

From this equation, we observe the following Christoffel symbols (highlighted in bold face)

$$
\Gamma^{0}_{00} = \frac{\dot{A}}{2}, \quad \Gamma^{0}_{01} = \Gamma^{0}_{10} = \frac{A^{'}}{2}, \quad \Gamma^{1}_{11} = \frac{\dot{B}}{2}e^{B-A} \tag{3.31}
$$

$$
\begin{aligned}
&\frac{d}{d\tau}\left(\frac{\partial\Psi}{\partial\dot{r}}\right) - \frac{\partial\Psi}{\partial r} = \frac{d}{d\tau}\left(-2\dot{r}e^{B}\right) - \left(A^{'}e^{A}\dot{t}^{2} - B^{'}e^{B}\dot{r}^{2} - 2r\dot{\theta}^{2} - 2rsin^{2}\theta\dot{\Phi}^{2}\right) \\
&= -2\left(e^{B}\ddot{r} + \left(\dot{B}e^{B}\dot{t} + B^{'}e^{B}\dot{r}\right)\dot{r}\right) - \left(A^{'}e^{A}\dot{t}^{2} - B^{'}e^{B}\dot{r}^{2} - 2r\dot{theta}^{2} - 2rsin^{2}\theta\dot{\Phi}^{2}\right) \\
&= -2e^{A}\ddot{r} - 2\dot{B}e^{B}\dot{t}\dot{r} - B^{'}e^{B}\dot{r}^{2} - A^{'}e^{A}\dot{t}^{2} + 2r\dot{\theta}^{2} + 2rsin^{2}\theta\dot{\Phi}^{2} \\
&= -2e^{A}\left(\ddot{r} + \frac{\dot{\mathbf{B}}}{\mathbf{2}}\dot{t}\dot{r} + \frac{\dot{\mathbf{B}}}{\mathbf{2}}\dot{r}\dot{t} + \frac{\mathbf{B}^{'}}{\mathbf{2}}\dot{r}^{2} + \frac{\mathbf{A}^{'}}{\mathbf{2}}\mathbf{e}^{\mathbf{A-B}}\dot{t}^{2} - \mathbf{re}^{-\mathbf{B}}\dot{\theta}^{2} - \mathbf{re}^{-\mathbf{B}}\mathbf{sin^{2}}\theta\dot{\Phi}^{2}\right) = 0.
\end{aligned}
\tag{3.32}
$$

From this equation, we observe the following Christoffel symbols (highlighted in bold face)

$$
\Gamma^{1}_{01} = \Gamma^{1}_{10} = \frac{\dot{B}}{2}, \; \Gamma^{1}_{11} = \frac{B^{'}}{2}, \; \Gamma^{1}_{00} = \frac{A^{'}}{2}e^{A-B}, \; \Gamma^{1}_{22} = -re^{-B}, \; \Gamma^{1}_{33} = -re^{-B}sin^{2}\theta \tag{3.33}
$$

$$
\begin{aligned}
\frac{d}{d\tau}\left(\frac{\partial\Psi}{\partial\dot{\theta}}\right) - \frac{\partial\Psi}{\partial\theta} &= \frac{d}{d\tau}\left(-2r^{2}\dot{\theta}\right) - \left(-2r^{2}sin\theta cos\theta\dot{\Phi}^{2}\right) \\
&= -2\left(2r\dot{r}\dot{\theta} + r^{2}\right)\ddot{\theta} + 2r^{2}sin\theta cos\theta\dot{\Phi}^{2} \\
&= -2r^{2}\left(\ddot{\theta} + \frac{\mathbf{1}}{\mathbf{r}}\dot{r}\dot{\theta} + \frac{\mathbf{1}}{\mathbf{r}}\dot{\theta}\dot{r} - \mathbf{sin}\theta\mathbf{cos}\theta\dot{\Phi}^{2}\right) = 0.
\end{aligned}
\tag{3.34}
$$

From this equation, we observe the following Christoffel symbols (highlighted in bold face)

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}, \quad \Gamma^2_{33} = -sin\theta cos\theta \tag{3.35}$$

$$\begin{aligned}
\tfrac{d}{d\tau}\left(\tfrac{\partial\Psi}{\partial\dot{\Phi}}\right) - \tfrac{\partial\Psi}{\partial\Phi} &= \tfrac{d}{d\tau}\left(-2r^2 sin^2\theta\dot{\Phi}\right) \\
&= -2\left(2r sin^2\theta\dot{r}\dot{\Phi} + 4r^2 sin\theta cos\theta\dot{\theta}\dot{\Phi} + 2r^2 sin^2\theta\ddot{\Phi}\right) \\
&= -4r^2 sin^2\theta\left(\ddot{\Phi} + \tfrac{\mathbf{1}}{\mathbf{r}}\dot{r}\dot{\Phi} + \mathbf{cot}\theta\dot{\theta}\dot{\Phi}\right) = 0.
\end{aligned} \tag{3.36}$$

From this equation, we observe the following Christoffel symbols (highlighted in bold face)

$$\Gamma^3_{13} = \frac{1}{r}, \quad \Gamma^3_{23} = cot\theta. \tag{3.37}$$

As you would have observed, obtaining the Christoffel symbols from the Euler-Lagrange equation is a simpler task, due to the main reason being that we obtain only the non-zero components, without any a priori information about the null components. We have 4 components and they need to be arranged in three slots of the Christoffel symbol (the way in which they are arranged matters). Let's lay out a $4\times4\times4$ cubic matrix, containing a total of 64 coefficients. Run a plane through the front face's diagonal and another plane through the top face's diagonal separately, what we are doing here is running through the thickness diagonal. By doing this we run through the front, bottom, back, and the top faces, and observe that there is a total of $4 \times 4 = 16$ coefficients. In addition to this, in the four faces we have travelled past there are a total of 6 coefficients above and below each diagonal thus resulting in a total of $4 \times 6 = 24$ coefficients above and 24 coefficients below the main diagonal. This would mean that we need to compute a total of $24 + 16 = 40$ independent Christoffel symbols just to realize that 15 of them have non-zero components. It is also interesting to note that in each of the cube's face there are 10 independent components, out of which 4 are consumed by general covariance, i.e., freedom to use arbitrary coordinates, and the remaining six correspond to 3 spatial rotations and 3 Lorentz boosts. Mathematically, if we wanted to sound smart, we can think that the independent components are the number of parameters of the *Lorentz-Poincare group*, i.e., the group of translations (4), spatial rotations (3), and velocity boosts (3), but let's not get into that. So, next time you compute the Christoffel symbols, please save the hard work for the next part.

What next part? Let's think about it. Thus far, since the diagonal elements of the metric tensor were known, we had computed something which was created from a combination of them, i.e., the Christoffel symbol(s). So, our next step is to compute something else which is constructed from a combination of the

Christoffel symbols. Three results pop up- the Riemann curvature tensor, the Ricci tensor, and the Ricci scalar, which one are we to pick? Turn to Einstein's field equations for the answer. In the equations, we have two quantities, the Ricci tensor and the Ricci scalar. Although it is true that we can compute the Ricci tensor and scalar from the Riemann curvature tensor, it is a cumbersome task, instead we directly compute the Ricci tensor. Now you may question me as to why I am taking the hard route of direct computation, is there no other way for avoiding the mess? Yes, there is: since we have 4 components that need to be arranged in the two slots of the Ricci tensor, we end with 16 different possibilities, out of which only 4 have non-zero components. The smarter way to proceed is to observe the null components of the Ricci tensor from symmetry considerations and then find the 4 non-zero components by evaluating the line interval on the unit sphere after a coordinate transformation. Similar to the approach in this section, we shall do both methods, not for wasting time, energy, space and information, but rather to understand the beauty and elegance of the latter method.

3.7 Method for Hard Workers

We solve the Einstein field equations separately for R_{00}, R_{11}, R_{22}, and R_{33}, and are left with the following painful-to-look-at equation

$$\begin{aligned}
R_{00} &= -\Gamma^1_{00,1} + \Gamma^0_{01}\Gamma^1_{00} - \Gamma^1_{00}\Gamma^1_{11} - \Gamma^1_{00}\Gamma^2_{12} - \Gamma^1_{00}\Gamma^3_{13},\\
R_{11} &= +\Gamma^0_{10,1} + \Gamma^0_{10}\Gamma^0_{01} - \Gamma^1_{11}\Gamma^0_{10} + \Gamma^2_{12,1} - \Gamma^1_{11}\Gamma^2_{12} + \Gamma^2_{12}\Gamma^2_{21} + \Gamma^3_{13,1}\\
&\quad -\Gamma^1_{11}\Gamma^3_{13} + \Gamma^3_{13}\Gamma^3_{31},\\
R_{22} &= -\Gamma^1_{22}\Gamma^0_{10} - \Gamma^1_{22,1} - \Gamma^1_{22}\Gamma^1_{11} + \Gamma^2_{21}\Gamma^1_{22} + \Gamma^3_{23,2} - \Gamma^1_{22}\Gamma^3_{13} + \Gamma^3_{23}\Gamma^3_{32},\\
R_{33} &= -\Gamma^1_{33}\Gamma^0_{10} - \Gamma^1_{33,1} - \Gamma^1_{33}\Gamma^1_{11} + \Gamma^3_{31}\Gamma^1_{33} - \Gamma^2_{33,2} - \Gamma^1_{33}\Gamma^2_{12} + \Gamma^2_{32}\Gamma^2_{33}.
\end{aligned} \tag{3.38}$$

Solving these equations by plugging in the appropriate Christoffel symbols, we obtain four equations, each which are to be solved to obtain the metric. Let's do this.

1. Calculation of R_{00}:

$$\begin{aligned}
R_{00} &= -\tfrac{1}{e^B}\left[\tfrac{(-e^B)'(e^A)'}{4(-e^B)} - \tfrac{(e^A)''}{2} - \tfrac{(e^A)'(e^A)'}{4(e^A)} - \tfrac{(e^A)'}{r}\right] \\
&= -e^{-B}\left[\tfrac{A'B'e^A}{4} - \tfrac{A''e^A}{2} - \tfrac{(A')^2 e^A}{4} - \tfrac{A'e^A}{r}\right] \\
&= e^{A-B}\left[\tfrac{A''}{2} + \tfrac{A'}{r} + \tfrac{(A')^2}{4} - \tfrac{A'B'}{4}\right].
\end{aligned} \tag{3.39}$$

2. Calculation of R_{11}:

$$\begin{aligned}
R_{11} &= \tfrac{(e^A)'(e^A)'}{4(e^A)} + \tfrac{(e^A)''}{2} - \tfrac{(e^A)'(-e^B)'}{4(-e^B)} - \tfrac{(-e^B)'(e^A)}{(-e^B r)} \\
&= e^A\left[\tfrac{(A')^2}{4} + \tfrac{A''}{2} - \tfrac{A'B'}{4} - \tfrac{B'}{r}\right].
\end{aligned} \tag{3.40}$$

3. Calculation of R_{22}:

$$\begin{aligned}
R_{22} &= \tfrac{(e^A)'r}{2(e^A)(-e^B)} - \tfrac{(-e^B)'r}{2(-e^B)^2} - \tfrac{1}{(-e^B)} - 1 \\
&= \tfrac{A'r-B'r}{2}e^{-B} - 1 + e^{-B}.
\end{aligned} \tag{3.41}$$

4. Calculation of R_{33}:

$$\begin{aligned}
R_{33} &= R_{22}sin^2\theta \\
&= \left[\tfrac{B'r-A'r}{2}e^{-B} - 1 + e^{-B}\right] sin^2\theta.
\end{aligned} \tag{3.42}$$

Alright, we have successfully found what we were looking for, now if I were to ask you the question as to why we are finding these equations, what would your answer be? Sometimes, we get too carried away with the math that we tend to forget the reason of their very creation/existence. An equation's beauty is highlighted by its physical significance, and an equation's importance is intensified when it serves a purpose. The purpose here is to find an equation that would describe the spacetime in the vicinity of the black hole. The foundation of the Schwarzschild solution lies in the assumptions made- null value of the energy-momentum tensor and the cosmological constant. Since there is no matter distribution outside the black hole, we arrived at a condition which is to be used for finding equations describing that spacetime- the *Ricci flatness condition.* This condition communicates to us that at any point in the vacuumed (i.e., zero matter distribution) spacetime the computation of the Ricci tensor yields a null result, i.e., $R_{\mu\nu} = 0$. Hence, let us impose this condition of the found equations and check for results. Before imposing the Ricci flatness condition blindly too all equations, let us carefully observe the equations obtained for bread crumbs in order to solve the final mystery of the metric.

Applying the Ricci flatness condition to R_{11} we observe the following

$$\begin{aligned} R_{11} &= e^A \left[\frac{(A')^2}{4} + \frac{A''}{2} - \frac{A'B'}{4} - \frac{B'}{r} \right] = 0 \\ \Longrightarrow \frac{B'}{r} &= \frac{(A')^2}{4} + \frac{A''}{2} - \frac{A'B'}{4}. \end{aligned} \tag{3.43}$$

Using this equation in the expression for R_{00} and imposing the Ricci flatness condition to the R_{00} component, we obtain

$$\begin{aligned} &R_{00} = 0 = e^{A-B} \left[\frac{A''}{2} + \frac{A'}{r} + \frac{(A')^2}{4} - \frac{A'B'}{4} \right] = e^{A-B} \left[\frac{A'+B'}{r} \right] \\ &\Longrightarrow A' + B' = 0, \end{aligned} \tag{3.44}$$

and from this, we see that

$$A = -B(r). \tag{3.45}$$

Now, applying the Ricci flatness condition to the R_{22} component, we observe

$$\begin{aligned} R_{22} = 0 &= \frac{B'r - A'r}{2} e^{-B} - 1 + e^{-B} \\ \left(1 - e^{-B}\right) &= \frac{B' - A'}{2} e^{-B} r, \end{aligned} \tag{3.46}$$

and since, from 3.43, $A' = -B'$, we have

$$\begin{aligned} \left(1 - e^{-B}\right) &= -\frac{2B'}{2} e^{-B} r \\ \Longrightarrow -\frac{dB}{(e^B - 1)} &- \int \frac{dr}{r} = 0, \end{aligned} \tag{3.47}$$

which is a first-order non-linear ODE whose solution is

$$\begin{aligned} B(r) &= -ln\left(\frac{C_1}{r} + 1\right) \\ \Longrightarrow e^{-B} &= e^A = \left(1 + \frac{C_1}{r}\right). \end{aligned} \tag{3.48}$$

Now, let us rewrite the metric using the above obtained result

$$ds^2 = \left(1 + \frac{C_1}{r}\right) dt^2 - \left(1 + \frac{C_1}{r}\right)^{-1} dr^2 - r^2 d\Omega^2. \tag{3.49}$$

From equation 1.151, the metric describing a gravitating object far away from the source was found to be: $ds^2 = (1 + 2r\mathbf{g})\, dt^2 - dr^2$. The acceleration due to gravity for a gravitating object is $\mathbf{g} = -\frac{GM}{r^2}$. Thus, the metric post this substitution reads

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - dr^2, \tag{3.50}$$

and upon comparison we find the constant C_1 to be

$$C_1 = -2GM. \tag{3.51}$$

It is to be noted that all our calculations were done setting the speed of light to unity, i.e., $c \equiv 1$. When the calculations are done sans this assumption, we obtain the constant to be $C_1 = 2GM/c^2$, and the final form of the metric reads

$$ds^2 = \left(1 - \frac{r_s}{r}\right) dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 d\Omega^2, \tag{3.52}$$

where r_s is the Schwarzschild radius. This concludes our derivation but raises a question- why did we compare our ansatz, whose purpose was to describe the geometry of a gravitating object in it's vicinity, to the metric describing the geometry, far away from the gravitating object. How can the metrics even be compared? We will learn in the next chapter that *Birkhoff's theorem* and *Israel's theorem* account for this comparison.

Exercise 9

1. Show that, by explicit calculations, the Kretschmann scalar of the Schwarzschild metric is $K = \frac{12r_s^2}{r^6}$.

2. Consider the Einstein field equations with a cosmological constant in vacuum,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 0. \tag{3.53}$$

a. Consider general spherically symmetric spacetime metric and show that it reduces to the following

$$ds^2 = -\left(1 - \frac{r_s}{r} + \frac{\Lambda}{3}r^2\right) dt^2 + \left(1 - \frac{r_s}{r} + \frac{\Lambda}{3}r^2\right)^{-1} dr^2 + r^2 d\Omega_2^2, \tag{3.54}$$

This is called the Schwarzschild-de Sitter metric and is the generalisation of the Schwarzschild solution which includes an arbitrary cosmological constant Λ. Notice that when $\Lambda = 0$, we see that the metric simplifies to the Schwarzschild case and when $m = 0$ (and hence, $r_s = 0$) we obtain the de Sitter metric which can be AdS or dS base on the sign of the cosmological constant.
b. Show that the Kretschmann scalar for this metric is $K = \frac{12r_s^2}{r^6} + \frac{8\Lambda^2}{3}$.

3. Consider the following transformations

$$\theta \to ir,\ \phi \to it,\ r \to z,\ t \to i\phi. \tag{3.55}$$

a. Apply these to the Schwarzschild metric and then perform an overall sign change. The resulting metric is called the metric of the type D solution which refers to the Petrov type D. Note that $sin^2(ir) = -sinh^2 r$.
b. Show that the Kretschmann scalar is the same as that of Schwarzschild with r replaced with z.
c. Will all the components of the Riemann tensor be equal to their Schwarzschild counterpart if the defined transformations are applied?

4. A Wick rotation is to make the substitution $t \to -i\tau$. This helps us to find a solution to a metric in the Minkowski space with a Lorentzian signature $(-++\ldots+)$ from a metric in the Euclidean space with signature $(++\ldots+)$.
a. Wick rotate the time coordinate in the global AdS_3 metric 1.228.
b. Parameterize the metric and show that it is related to the symmetry group $SL(2,\mathbb{C})$ which has the properties

$$SL(2,\mathbb{C}) \equiv \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,\ b,\ c,\ d \in \mathbb{C},\ ad - bc = 1 \right\}. \tag{3.56}$$

c. Simplify the metric by making the substitutions $r = lsinh\psi$ and $\bar{t} = l\tau$.

5. A rotating BTZ (Bañados-Teitelboim-Zanelli) black hole is described by the metric

$$ds^2 = -\frac{\left(r^2 - r_+^2\right)\left(r^2 - r_-^2\right)}{r^2 l^2} dt^2 + \frac{l^2 r^2}{\left(r^2 - r_+^2\right)\left(r^2 - r_-^2\right)} dr^2 + r^2 \left(d\phi - \frac{r_+ r_-}{l r^2} dt\right)^2 . \tag{3.57}$$

where r_+ and r_- are the black hole radii and l is the AdS radius. This is a black hole solution for $(2+1)$-dimensional gravity, i.e., two spatial dimensions (r and ϕ here) and one time dimension. Notice that this metric possesses off-diagonal components.
a. Simplify this metric by expressing it in terms of the mass $M = \frac{r_+^2 - r_-^2}{8Gl^2}$ and the angular momentum $J = \frac{2r_+ r_-}{4Gl}$.
b. Calculate the Ricci tensor components and show that the Ricci scalar takes the form $R = -6/l^2$.
c. Compare the Riemann tensor calculated to the one calculated via the expression 1.201 and fix the value of the cosmological constant. What type of a spacetime is the BTZ black hole in?
d. Calculate the Cotton tensor and comment on the conformal flatness of this metric.
e. Taking $r_- = 0$, show that the Kretschmann invariant for this metric is

$K = 12/l^4$. This is quite interesting since we obtain a constant as the answer and not a curvature singularity which blows up as some length parameter is taken to zero as in the case of the Schwarzshchild black hole. If there exists no singularity, why is the BTZ black hole a black hole? The answer is that in the BTZ black hole the singularity is due to the causal structure of the spacetime, not in its curvature. The type of singularity observed here is referred to as a causal singularity[1]

Example 3.1. A neat trick to have up ones sleeve is to realize that for diagonal metrics, the Christoffel symbols take a specific form. Let $g_{\alpha\beta}$ be a diagonal metric, then we have

$$\begin{aligned}
\Gamma^{\mu}_{\alpha\beta} &= g^{\mu\mu}\Gamma_{\alpha\beta\mu} = 0, \\
\Gamma^{\mu}_{\alpha\alpha} &= g^{\mu\mu}\Gamma_{\alpha\alpha\mu} = -\tfrac{1}{2}g^{\mu\mu}g_{\alpha\alpha,\mu} = -\tfrac{1}{2g_{\mu\mu}}\partial_\mu g_{\alpha\alpha}, \\
\Gamma^{\mu}_{\alpha\mu} &= g^{\mu\mu}\Gamma_{\alpha\mu\mu} = \tfrac{1}{2g_{\mu\mu}}\partial_\alpha g_{\mu\mu} = \partial_\alpha\left(\ln\sqrt{|g_{\mu\mu}|}\right), \\
\Gamma^{\mu}_{\mu\mu} &= g^{\mu\mu}\Gamma_{\mu\mu\mu} = \tfrac{1}{2g_{\mu\mu}}\partial_\mu g_{\mu\mu} = \partial_\mu\left(\ln\sqrt{|g_{\mu\mu}|}\right).
\end{aligned} \tag{3.58}$$

The advantages of this observation is twofold- firstly, we now know that all the mixed Christoffel symbols identically vanish since $\Gamma^i_{jk} = g^{ii}\Gamma_{jki}$ and Γ_{jki} contains (the derivative of) only non-diagonal components of the metric tensor which are all zero, and secondly, we now have expressions (for different classes such as all indices alike, base indices alike, and so on), to compute the remaining non-vanishing symbols. As an exercise rework the Christoffel symbols for the 4-sphere 1.202 and cross-check you results.

3.8 More on Isometries

3.8.1 Stationarity and Staticity of the Schwarzschild Metric

A *Killing Field* is a vector field the local flow of which preserves the metric. Equivalently, X satisfies the *Killing equation*,

$$\mathcal{L}_X g_{\mu\nu} = \nabla_{(\mu}X_{\nu)} = 0. \tag{3.59}$$

One among the many features of the Schwarzschild metric is that it's stationary with a Killing vector field $X = \partial_t$. A spacetime is defined to be *stationary* if there exists a Killing vector field X which approaches ∂_t in the asymptotically flat region, i.e., where $r \to \infty$ definitions) and generates a one parameter groups of isometries. A spacetime is called *static* if it is stationary and if the

[1] for more information see (33))

stationary Killing vector X is orthogonal to the hypersurface, i.e.,

$$X^{\beta} \wedge dX^{\beta} = 0, \tag{3.60}$$

where,

$$X^{\beta} = X_{\mu} dx^{\mu} = g_{\mu\nu} X^{\nu} dx^{\mu}. \tag{3.61}$$

Consider a general coordinate transformation, under which the metric tensor transforms as

$$\bar{g}^{\mu\nu}(\bar{x}) = g^{\alpha\beta}(x) \frac{\partial \bar{x}^{\mu}}{\partial x^{\alpha}} \frac{\partial \bar{x}^{\nu}}{\partial x^{\beta}}, \tag{3.62}$$

and the infinitesimal form of the transformation is to be found. If $\bar{x} = x^{\mu} + \epsilon^{\mu}(x)$, where $\epsilon^{\mu}(x)$ is a small vector field, then

$$\begin{aligned} \bar{g}^{\mu\nu}(\bar{x}) &\approx g^{\alpha\beta}(x) \left(\tfrac{\partial(\bar{x}^{\mu} + \epsilon^{\mu})}{\partial x^{\alpha}} \right) \left(\tfrac{\partial(\bar{x}^{\nu} + \epsilon^{\nu})}{\partial x^{\beta}} \right) \\ &\approx g^{\alpha\beta}(x) \left(\tfrac{\partial \bar{x}^{\mu}}{\partial x^{\alpha}} + \partial_{\alpha} \epsilon^{\mu} \right) \left(\tfrac{\partial \bar{x}^{\nu}}{\partial x^{\beta}} + \partial_{\beta} \epsilon^{\nu} \right), \end{aligned} \tag{3.63}$$

and we know that $\frac{\partial x^a}{\partial x^b} = \frac{dx^a}{dx^b} = \delta^a_b$, thus, in linear order

$$\begin{aligned} \bar{g}^{\mu\nu}(\bar{x}) &\approx g^{\alpha\beta}(x) \left(\delta^{\mu}_{\alpha} + \partial_{\alpha} \epsilon^{\mu} \right) \left(\delta^{\nu}_{\beta} + \partial_{\beta} \epsilon^{\nu} \right) \\ &\approx \left(g^{\alpha\beta} \delta^{\mu}_{\alpha} + g^{\alpha\beta} \delta_{\alpha} \epsilon^{\mu} \right) \left(\delta^{\nu}_{\beta} + \partial_{\beta} \epsilon^{\nu} \right) \\ &\approx g^{\mu\nu}(x) + g^{\mu} \partial \epsilon^{\nu} + g^{\nu} \partial \epsilon^{\mu} + \underbrace{\partial^{\nu} \epsilon^{\mu} \partial^{\mu} \epsilon^{\nu}}_{\approx 0}. \end{aligned} \tag{3.64}$$

Hence, we get

$$\bar{g}^{\mu\nu}(\bar{x}) \approx g^{\alpha\beta}(x) + \partial^{\mu} \epsilon^{\nu} + \partial^{\nu} \epsilon^{\mu} \equiv g^{\alpha\beta}(x) + \partial^{(\mu} \epsilon^{\nu)}. \tag{3.65}$$

Now, if for some transformation $\epsilon^{\mu} = K^{\mu}$, the metric tensor does not change, i.e.,

$$D^{(\mu} K^{\nu)} \equiv D^{\mu} K^{\nu} + D^{\nu} K^{\mu} = 0,$$

then the corresponding vector field K^{μ} is called a Killing vector and the transformations are called isometries.

3.8.2 Axisymmetric Spacetimes

A spacetime is called *axisymmetric* if there exists a Killing vector field Y, which generates a 1-parameter group of isometries, and which behaves like a rotation. This property is captured by requiring that all orbits 2π periodic, and that the set $\{Y = 0\}$, called the axis of rotation, is non-empty. In the extended Schwarzschild spacetime the set $\{r = 2m\}$ is a null hypersurface $\mathcal{E}$, the Schwarzschild event horizon. The stationary Killing vector $X = \partial_t$ extends to a Killing vector $\hat{X}$ in the extended spacetime which becomes tangent to and null on $\mathcal{E}$, except at the bifurcation sphere, where $\hat{X}$ vanishes.

3.8.3 Killing Vectors of a 2-Sphere

Consider the metric of a 2-sphere $ds^2 = d\omega_2^2$. After calculating the Christoffel symbols (which is left as an exercise), we write down the Killing equations $D_{(\mu}X_{\nu)} = 0$ where indices $\mu, \nu = 0, 1$. Here we have three Killing equations which read

$$
\begin{aligned}
&D_\theta K_\theta = \partial_\theta K_\theta - \Gamma^\alpha_{\nu\nu} K_\alpha = 0 \\
&\Rightarrow \partial_\theta K_\theta = 0, \\
&D_\phi K_\phi = \partial_\phi K_\phi - \Gamma^\alpha_{\phi\phi} K_\alpha = 0, \\
&\Rightarrow \partial_\phi K_\phi + K_\theta sin\theta cos\theta = 0, \\
&D_\theta K_\phi + D_\phi K_\theta = \partial_\theta K_\phi + \partial_\phi K_\theta - \Gamma^\alpha_{\theta\phi} K_\alpha - \Gamma^\alpha_{\phi\theta} K_\alpha = 0, \\
&\Rightarrow \partial_{(\theta} K_{\phi)} - K_\phi cot\theta = 0.
\end{aligned}
\tag{3.66}
$$

The two components of the Killing vector are functions of the coordinates, i.e., $K_\theta = K_\theta(\theta, \phi)$ and $K_\phi = K_\phi(\theta, \phi)$. Solving the first Killing equation, we obtain a constant that depends only on ϕ and hence, we can conclude that

$$K_\theta(\theta, \phi) = K_\theta(\phi). \tag{3.67}$$

We now differentiate the third Killing equation with respect to ϕ to obtain

$$\partial^2_\phi K_\phi + \partial_\theta \left(\partial_\phi K_\phi\right) - 2cot\theta \left(\partial_\phi K_\phi\right) = 0. \tag{3.68}$$

From the second Killing equation, we substitute for $\partial_\phi K_\phi$ in 3.68 to get

$$
\begin{aligned}
&\partial_\theta \left(-K_\theta sin\theta cos\theta\right) + \partial^2_\phi K_\theta - 2cot\theta \left(-K_\theta sin\theta cos\theta\right) = 0, \\
&K_\theta \left(-cos^2\theta + sin^2\theta\right) + \partial^2_\phi K_\theta + 2K_\theta cos^2\theta = 0, \\
&\Rightarrow \partial^2_\phi K_\theta + K_\theta = 0.
\end{aligned}
\tag{3.69}
$$

The solution to this differential equation is

$$K_\theta(\phi) = c_1 sin\phi + c_2 cos\phi, \tag{3.70}$$

where c_1 and c_2 are constants. Now we can insert this solution into the second Killing equation to find the solution to K_ϕ as follows

$$\partial_\phi K_\phi = -c_1 \; sin\phi \; sin\theta \; cos\theta - c_2 \; cos\phi \; sin\theta \; cos\theta, \tag{3.71}$$

which is a non-homogeneous differential equation whose solution is

$$K_\phi = c_1 \; cos\phi \; sin\theta \; cos\theta - c_2 \; sin\phi \; sin\theta \; cos\theta + f(\theta), \tag{3.72}$$

where $f(\theta)$ is a solution to the homogeneous differential equation $\partial_\phi K_\phi = 0$. Substituting 3.70 and 3.72 back into the third Killing equation we get

$$\begin{aligned} & f'(\theta) - 2f(\theta)cot\theta = 0, \\ & \partial_\theta \left(\frac{f_\theta}{sin^2\theta} \right) = 0, \\ & \Rightarrow f(\theta) = c_3 \;\; sin^2\theta. \end{aligned} \tag{3.73}$$

Thus, we have the following solutions for the Killing vectors

$$\begin{aligned} K_\theta &= c_1 \; sin\phi + c_2 \; cos\phi, \\ K_\phi &= c_1 \; cos\phi \; sin\theta \; cos\theta - c_2 \; sin\phi \; sin\theta \; cos\theta + c_3 \; sin^2\theta. \end{aligned} \tag{3.74}$$

To make the calculations nice, let's absorb the negative sign into the constant c_2 and define a new constant $\bar{c}_2$. Now, we have three independent constants c_1, $\bar{c}_2$ and c_3 for which we will have three independent Killing vectors K_1, K_2 and K_3.
a. Let $c_1 = 0$, $\bar{c}_2 = 1$, $c_3 = 0$

$$(K_1)_\theta = -cos\phi, \; (K_1)_\phi = sin\phi \; sin\theta \; cos\theta, \tag{3.75}$$

b. Let $c_1 = 1$, $\bar{c}_2 = 0$, $c_3 = 0$

$$(K_2)_\theta = sin\phi, \; (K_2)_\phi = cos\phi \; sin\theta \; cos\theta, \tag{3.76}$$

c. Let $c_1 = 0$, $\bar{c}_2 = 0$, $c_3 = 1$

$$(K_3)_\theta = 0, \; (K_3)_\phi = sin^2\theta. \tag{3.77}$$

We can now write down the Killing vectors in the following form

$$K_i = (K_i)^{\theta}\,\partial_\theta + (K_i)^{\phi}\,\partial_\phi. \tag{3.78}$$

Raising the indices using the metric tensor, we finally obtain the three Killing vectors

$$\begin{aligned} K_1 &= -cos\phi\;\partial_\theta + cot\theta\;sin\phi\;\partial_\phi, \\ K_2 &= sin\phi\;\partial_\theta + cot\theta\;cos\phi\;\partial_\phi, \\ K_3 &= \partial_\phi. \end{aligned} \tag{3.79}$$

If you lived on this 2-sphere, you have the freedom to do the following three things: you can walk in a particular direction, you can walk in a direction that is perpendicular to your first direction, and you can stand and spin in one place. All of these actions would not change the way you perceive the landscape or in other words these three symmetries would leave you invariant. The three Killing vectors we derived above correspond to these three symmetries.

3.8.4 Killing Vectors Algebra

For a general surface S in n-dimensional constant curvature spacetimes, we can have $n(n+1)/2$ number of Killing vectors. Consider now a surface in $\mathbb{R}^3$ where we have 3 Killing vectors which read

$$K_1 = \partial_x,\;\; K_2 = \partial_y,\;\; K_3 = y\partial_x - x\partial_y. \tag{3.80}$$

All these vectors form a bases of the infinitesimal isometry mappings. They are closed under a bracket $[K, M]$ called the commutator of the Killing vectors K and M which is defined as follows

$$[K, M] \equiv [a_i\partial_i b_j - b_i\partial_i a_j]\,\partial_j, \tag{3.81}$$

where $K = a_1\partial_x + a_2\partial_y$ and $M = b_1\partial_x + b_2\partial_y$. Under this bracket, the Killing vectors form a closed algebra which is shown below

$$
\begin{aligned}
[K_1, K_2] &= [\partial_x, \partial_y] = \partial_x\partial_y - \partial_y\partial_x = 0, \\
[K_1, K_3] &= [\partial_y, y\partial_x - x\partial_y] = \partial_x\left(y\partial_x - x\partial_y\right) - \left(y\partial_x - x\partial_x\right)\partial_x \\
&= y\partial_x^2 - \partial_y - x\partial_x\partial_y - y\partial_x^2 + x\partial_y\partial_x = -\partial_y, \\
[K_2, K_3] &= [\partial_y, y\partial_x - x\partial_y] = \partial_y\left(y\partial_x - x\partial_y\right) - \left(y\partial_x - x\partial_y\right)\partial_y \\
&= \partial_x + y\partial_y\partial_x - x\partial_y^2 - y\partial_x\partial_y + x\partial_y^2 = \partial_x, \\
&\Rightarrow [K_1, K_2] = 0,\ [K_2, K_3] = K_1,\ [K_3, K_1] = K_2.
\end{aligned}
\tag{3.82}
$$

3.8.5 Killing Vectors of Poincaré Half Plane

Consider the Poincaré half-plane metric 1.191. We have three Killing equations here which read

$$
\begin{aligned}
&D_x K_x = \partial_x K_x - \Gamma^{\alpha}_{\nu\nu} K_\alpha = 0 \\
&\Rightarrow \partial_x K_x - \tfrac{1}{y} K_y = 0, \\
&D_y K_y = \partial_y K_y - \Gamma^{\alpha}_{\nu\nu} K_\alpha = 0 \\
&\Rightarrow \partial_y K_y + \tfrac{1}{y} K_y = 0, \\
&D_x K_y + D_y K_x = \partial_x K_y + \partial_y K_x - \Gamma^{\alpha}_{xy} K_\alpha - \Gamma^{\alpha}_{yx} K_\alpha = 0 \\
&\Rightarrow \partial_x K_y + \partial_y K_x + \tfrac{2}{y} K_x = 0.
\end{aligned}
\tag{3.83}
$$

Solving the second Killing equation gives us

$$
K_y = \frac{1}{y} f'(x), \tag{3.84}
$$

where $f'(x)$ is the derivative of an arbitrary function of x.Using this in the first Killing equation gives us

$$
K_x = \frac{1}{y^2} f(x) + g(y). \tag{3.85}
$$

Now, substituting equations 3.84 and 3.85 into the third Killing equation gives us

$$
f''(x) + 2g(y) + yg'(y) = 0. \tag{3.86}
$$

This partial differential equation is separable and hence let $f''(x)$ be equal to a constant, say c_1. This yields

$$f''(x) = c_1 \Rightarrow f(x) = \frac{1}{2}c_1 x^2 + c_2 + xc_3. \tag{3.87}$$

and similarly, we let the differential equation for $g(y)$ to be equal to $-c_1$ so that adding this to the equation of $f(x)$ gives back the original form.

$$2g(y) + yg'(y) = -c_1 \Rightarrow g(y) = -\frac{1}{2}c_1 + \frac{1}{y^2}c_4. \tag{3.88}$$

Substituting these results back into 3.84 and 3.85 we obtain the Killing vectors which read

$$K_x = \frac{1}{2}\left(\frac{x^2}{y^2} - 1\right)c_1 + \frac{1}{y^2}\left(c_2 + xc_3 + c_4\right), \quad K_y = \frac{1}{y}\left(xc_1 + c_3\right) \tag{3.89}$$

Since the integration constants c_2 and c_4 serve the same purpose, we can either set one of them to zero or let $c_2 + c_4 = c_5$. Doing so and using the metric tensor to write the Killing vectors in the contravariant form as follows (with $c_1 = a$, $c_3 = b$ and $c_5 = c$)

$$K^x = \frac{1}{2}\left(x^2 - y^2\right)a + xb + c, \quad K^y = xya + by. \tag{3.90}$$

We now have three constants a, b and c for which we obtain three independent Killing vectors K_1, K_2 and K_3.
a. Let $a = 2$ (so that the factor cancels), $b = 0$, $c = 0$.

$$\begin{aligned} &(K_1)_x = \left(x^2 - y^2\right), \ (K_1)_y = 2xy. \\ &\Rightarrow K_1 = \left(x^2 - y^2\right)\partial_x + 2xy\partial_y. \end{aligned} \tag{3.91}$$

b. Let $a = 0$, $b = 1$, $c = 0$.

$$\begin{aligned} &(K_2)_x = x, \ (K_2)_y = y. \\ &\Rightarrow K_2 = x\partial_x + y\partial_y. \end{aligned} \tag{3.92}$$

c. Let $a = 0$, $b = 0$, $c = 1$.

$$\begin{aligned} &(K_3)_x = 1, \ (K_3)_y = 0. \\ &\Rightarrow K_3 = \partial_x. \end{aligned} \tag{3.93}$$

It can be checked that the Killing vectors form a closed algebra

$$[K_3, K_2] = K_3, \ [K_2, K_1] = K_1, \ [K_1, K_3] = -2K_2 \tag{3.94}$$

Exercise 10

1. Consider the Killing vectors for a 2-sphere 3.79. Show that they satisfy the Lie algebra for SO(3), i.e., the following commutator relations

$$[K_1, K_2] = K_3,$$
$$[K_2, K_3] = K_1, \qquad (3.95)$$
$$[K_3, K_1] = K_2.$$

2. Check that the Killing vectors of the Poincaré half-plane do form a closed algebra as shown in 3.94.

3. Find the the Killing vectors of the following
a. Hyperboloid $\mathbb{H}^2$ with metric $ds^2 = R^2\left(d\theta^2 + sinh^2\theta\; d\phi^2\right)$.
b. $ds^2 = -cosh^2\mu\; d\lambda^2 + d\mu^2$.
c. The Minkowski metric.
d. The Poincaré patch of AdS_3 1.199.
e. Metric of psuedospheres, i.e., surfaces of negative curvature, $ds^2 = K^2\left(d\vartheta^2 + e^{2\theta}d\varphi^2\right)$, where K is a constant.
f. The gravitational wave metric as given in 1.238.

4. The applicability of Killing vectors extends also to the infinite-dimensional dynamical systems, for example, those describing various fields. Given a Killing vector K and a conserved stress-energy tensor $T_{\mu\nu}$, we have the conserved current $J_\nu = K^\mu T_{\mu\nu}$. This implies the existence of the corresponding conserved charge. Show that, using the Killing equation and the fact that $T_{\mu\nu}$ is symmetric, $D_\nu J^\nu = 0$.

5. Show that is $(\mathcal{M}g)$ is Ricci flat and ξ is a Killing field, then ξ satisfies the Maxwell equations.

3.8.6 Conformal Killing Vectors

Conformal Killing vectors preserve the metric up to an overall factor of scaling. The condition for a vector K to be a conformal Killing vector is

$$\mathcal{L}_K g_{\alpha\beta} = \lambda g_{\alpha\beta} \qquad (3.96)$$

where λ is an arbitrary scalar function defined on the manifold. Contracting the indices we have

$$\lambda = \frac{2}{n} D_\alpha K^\alpha, \qquad (3.97)$$

and hence, the conformal equivalent of Killing's equation reads

$$D_{(\alpha}K_{\beta)} - \frac{1}{n}g_{\alpha\beta}D_\alpha K^\alpha = 0. \tag{3.98}$$

Now let's consider two conformally related metrics $\bar{g}_{\mu\nu} = e^{2\kappa(x)}g_{\mu\nu}$. The Christoffel symbols of the two metrics are related as follows

$$\begin{aligned}\bar{\Gamma}^\alpha_{\mu\nu} &= \bar{g}^{\alpha\alpha}\bar{\Gamma}^\alpha_{\mu\nu\alpha}\\ &= e^{2\kappa(x)}g^{\alpha\alpha}\tfrac{1}{2}\left(\bar{g}_{\alpha\mu,\nu} + \bar{g}_{\alpha\nu,\mu} - \bar{g}_{\mu\nu,\alpha}\right)\\ &= g^{\alpha\alpha}\tfrac{1}{2}\left(g_{\alpha\mu,\nu} + 2\kappa_{,\nu}g_{\alpha\mu} + g_{\alpha\nu,\mu} + 2\kappa_{,\mu}g_{\alpha\nu} - g_{\mu\nu,\alpha} - 2\kappa_\alpha g_{\mu\nu}\right)\\ &= \Gamma^\alpha_{\mu\nu} + \delta^\alpha_\mu\kappa_{,\nu} + \delta^\alpha_\nu\kappa_{,\mu} - g^{\alpha\rho}g_{\mu\nu}\kappa_{,\rho}.\end{aligned} \tag{3.99}$$

Now, the Killing equation reads

$$\begin{aligned}\bar{D}_{(\alpha}\bar{K}_{\beta)} &= \partial_{(\alpha}\bar{K}_{\beta)} - \bar{\Gamma}^\alpha_{(\mu\nu)}\bar{K}_\alpha\\ &= \partial_{(\alpha}\left(e^{2\kappa(x)}K_{\beta)}\right) - e^{2\kappa(x)}\Gamma^\alpha_{\mu\nu}K_\alpha - \left(\delta^\alpha_\mu\kappa_{,\nu} + \delta^\alpha_\nu\kappa_{,\mu} - g^{\alpha\rho}g_{\mu\nu}\kappa_{,\rho}\right)K_\alpha\\ &= e^{2\kappa(x)}\underbrace{D_{(\alpha}K_{\beta)}}_{=0} - \left(K_\mu\kappa_{,\nu} + K_\nu\kappa_{,\mu} - g_{\mu\nu}K^\rho\kappa_{,\rho}\right) = 0\\ \Rightarrow \bar{D}_{(\alpha}\bar{K}_{\beta)} &- \bar{g}_{\mu\nu}\bar{K}^\rho\kappa_{,\rho} = 0.\end{aligned} \tag{3.100}$$

Now, consider $D_\alpha K_\beta = D_\alpha\left(g_{\alpha\beta}K^\alpha\right) = g_{\alpha\beta}D_\alpha K^\alpha = 0$. This is implied since K^α is a Killing vector of $g_{\alpha\beta}$. Using this we have

$$\begin{aligned}g_{\alpha\beta}\bar{D}_\alpha\bar{K}^\alpha &= e^{2\kappa(x)}\left(\partial_\alpha K^\alpha + \Gamma^\alpha_{\mu\alpha}K^\mu\right) - g_{\alpha\beta}g_{\mu\nu}\bar{K}^\rho\kappa_{,\rho}\\ \Rightarrow g_{\mu\nu}\bar{K}^\rho\kappa_{,\rho} &= \tfrac{1}{n}\bar{D}_\alpha\bar{K}^\alpha,\end{aligned} \tag{3.101}$$

where the last step was obtained via contraction. Using this result we finally obtain the conformal equivalent of 3.98,

$$\bar{D}_{(\alpha}\bar{K}_{\beta)} - \frac{1}{n}\bar{D}_\alpha\bar{K}^\alpha = 0. \tag{3.102}$$

Now, let us consider the covariant derivative of 3.96 which reads

$$D_\gamma\left(D_\beta K_\alpha + D_\alpha K_\beta\right) - g_{\alpha\beta}\lambda_{,\gamma} = 0, \tag{3.103}$$

which is simplified to the following using 1.185 and the Killing equation

$$D_\alpha D_\beta K_\gamma + D_\beta D_\gamma K_\alpha + D_\gamma D_\alpha K_\beta = \frac{1}{2}\left(g_{\alpha\beta}\lambda_{,\gamma} + g_{\alpha\gamma}\lambda_{,\beta} + g_{\gamma\beta}\lambda_{,\alpha}\right). \quad (3.104)$$

Using the conformal Killing equation one more time $D_\gamma(D_\alpha K_\beta) = D_\gamma\left(D_\beta K_\alpha - g_{\alpha\beta}\lambda\right)$ enables us to make use of a familiar identity

$$D_\alpha D_\beta K_\gamma = \underbrace{D_\gamma D_\beta K_\alpha - D_\beta D_\gamma K_\alpha}_{=R^\rho_{\beta\alpha\gamma}K_\rho} + \frac{1}{2}\left(-g_{\alpha\beta}\lambda_{,\gamma} + g_{\alpha\gamma}\lambda_{,\beta} + g_{\gamma\beta}\lambda_{,\alpha}\right), \quad (3.105)$$

and hence, we obtain

$$D_\alpha D_\beta K_\gamma = R^\rho_{\beta\alpha\gamma}K_\rho + \frac{1}{2}\left(-g_{\alpha\beta}\lambda_{,\gamma} + g_{\alpha\gamma}\lambda_{,\beta} + g_{\gamma\beta}\lambda_{,\alpha}\right). \quad (3.106)$$

This equation implies that the second derivatives of the conformal Killing fields are determined by themselves and by the gradient of the scalar function defined on the manifold λ.

3.8.7 Conformal Killing Tensors

A conformal Killing tensor of rank m is defined as a totally symmetric tensor $K_{\nu_1\ldots\nu_m}$, that in the conformal frame obeys the conformal Killing tensor equation which reads as follows

$$D_{(\mu}K_{\nu_1\ldots\nu_m)} = mg_{(\mu\nu_1}\mathcal{K}_{\nu_2\ldots\nu_m)}, \quad (3.107)$$

where $\mathcal{K}_{\nu_2,\ldots,\nu_m}$ is a totally symmetric tensor of rank $m-1$ that can be found by taking the trace on both sides. To observe how this related to an ordinary Killing tensor, we can perform a conformal transformation. The covariant derivative in the conformal frame D is related to the ordinary covariant derivative $\hat{D}$ by a change of the Christoffel symbol,

$$\Gamma^\alpha_{\mu\nu} = \hat{\Gamma}^\alpha_{\mu\nu} + C^\alpha_{\mu\nu}, \quad (3.108)$$

where,

$$C^\alpha_{\mu\nu} = \omega^{-1}\left(\delta^\alpha_\mu\omega_{,\nu} + \delta^\alpha_\nu\omega_\mu - g_{\mu\nu}g^{\alpha\beta}\omega_{,\beta}\right), \quad (3.109)$$

where ω is the conformal factor in $\hat{g}_{\mu\nu} = \omega^2(x)g_{\mu\nu}$. We now have

$$\begin{aligned} D_\mu K_{\nu_1,\dots,\nu_m} &= \hat{D}_\mu K_{\nu_1\dots\nu_m} + C^\alpha_{\mu\nu_1} K_{\alpha\dots\nu_m} + \dots + C^\alpha_{\mu\nu_m} K_{\nu_1\dots\nu_m} \\ &= \hat{D}_\mu K_{\nu_1\dots\nu_m} + Q_{\mu\nu_1\dots\nu_m}, \end{aligned} \tag{3.110}$$

where $Q_{\mu\nu_1\dots\nu_m}$ is symmetric since $C^\alpha_{\mu\nu}$ is symmetric in lower indices. Now, we can define $Q_{\mu\nu_1\dots\nu_m} = mg_{(\mu\nu_1}\mathcal{K}_{\nu_2\dots\nu_m)}$ since it is consistent with the definition of $Q_{\mu\nu_1\dots\nu_m}$ when $\mathcal{K}_{\nu_2\dots\nu_m}$ is a totally symmetric tensor of rank $m-1$. Now using the property that $K_{\nu_1\dots\nu_m}$ is a Killing tensor, we obtain back the conformal Killing tensor equation 3.107 since the ordinary covariant derivative of the Killing tensor vanishes in accordance to the Killing equation, i.e., $\hat{D}_{(\mu}K_{\nu_1\dots\nu_m)} = 0$.

Exercise 11

1. Show that we require $\frac{1}{2}(n+1)(n+2)$ constants to calculate $D_\alpha D_\beta K_\gamma$.

2. Take the covariant derivative of 3.106 and show that for $n = 2$ the following equation holds

$$D_\eta D_\alpha D_\beta K_\gamma = \frac{1}{2} g_{\beta\zeta}\delta^{\xi\zeta}_{\gamma\eta} D_\eta (RK_\xi) + \frac{1}{2} D_\eta \left(-g_{\alpha\beta}\lambda_{,\gamma} + g_{\alpha\gamma}\lambda_{,\beta} + g_{\gamma\beta}\lambda_{,\alpha}\right) \tag{3.111}$$

Notice that both sides of 3.111 are covariant derivatives and in order to determine the scalar function λ and the Killing vector K_α we require an infinite number of constants. In other words there doesn't exist a finite basis of the conformal Killing fields exists in $n = 2$. This is a direct consequence of the fact that every 2-D metric is conformally flat and hence, there exists an infinite family of transformations preserving the conformally flatness.

3. Consider a Killing tensor $K_{\alpha\beta}$ which obeys the Killing equation $D_{(\mu}K_{\alpha\beta)} = 0$. Show that the product of two Killing vectors K_α and K_β is a Killing tensor. Similarly, prove that $K_{\nu_1}K_{\nu_2}\dots K_{\nu_m}$ is a Killing tensor if each of the K_{ν_i} $(i = 1,\dots,m)$ are Killing vectors. An elementary example of a Killing tensor is the metric tensor since $D_{(\alpha}g_{\mu\nu)} = 0$ holds.

4. Consider the Killing identity, $D_\mu D_\nu X_\beta = -R_{\gamma\mu\nu\beta}X^\gamma$. Show that $D_{(\mu}D_\nu X_{\beta)} = 0$ which implies that $D_\nu X_\beta$ is a Killing tensor. It is interesting to note that this holds when a Christoffel symbol built out of the Lie derivatives of metric tensors vanishes or in other words $\mathcal{L}_X \Gamma_{abc} = 0$ which is the statement of Affine collineation. Thus, we can conclude that a Killing symmetry ($\mathcal{L}_X g_{ab} = 0$) implies an Affine symmetry ($\mathcal{L}_X \Gamma^a_{bc} = 0$) which implies a Curvature symmetry ($\mathcal{L}_X R^a_{bcd} = 0$). This is called the theorem of inheritance of symmetries.

5. A rank p Killing tensor is a totally symmetric tensor $K_{\nu_1,\dots,\nu_p} = K_{(\nu_1,\dots,\nu_p)}$ that satisfies $D_{(\mu}K_{\nu_1,\dots,\nu_p)} = 0$. If $L^\mu \equiv \dot{x}^\mu$ is tangent to an affinely parametrised geodesic, show that the following quantity is constant along the geodesic

$$Q = L^{\nu_1 \ldots \nu_p} K_{\nu_1, \ldots, \nu_p}.$$

3.9 Killing Vectors and Isometries of The Schwarzschild Metric

We know that the Schwarzschild metric is time invariant and have no dependence on the angle ϕ. Hence, it's isometries include at least the transformations in time, i.e., $t \to t + \xi$, and rotations, i.e., $\phi \to \phi + \zeta$, for some constants ξ and ζ. The corresponding Killing vectors, $K^\mu = \left(K^t, K^r, K^\theta, K^\phi\right)$ take up the following form

$$K^\mu = (1, 0, 0, 0) \ \& \ K^\mu = (0, 0, 0, 1) \,. \tag{3.112}$$

3.9.1 Conserved Quantities in the Schwarzschild Metric

Let a test particle move along a world line $x^\mu(s)$ with a 4-velocity given by $v^\mu(s) = \frac{dx^\mu}{ds}$. It's derivative is given as

$$\frac{d}{ds}\left(K^\mu v_\mu\right) = \frac{d}{ds}\left(K^\mu v_\mu \times \frac{dx^\nu}{dx^\nu}\right) = \frac{d}{dx^\nu}\left(K^\mu v_\mu\right)\frac{dx^\nu}{ds} = \partial_\nu\left(K^\mu v_\mu\right)\frac{dx^\nu}{ds}, \tag{3.113}$$

which upon computation yields

$$\partial_\nu\left(K^\mu v_\mu\right)\frac{dx^\nu}{ds} = K^\mu v_\mu D_\nu v_\mu + v^\nu v^\mu D_\nu K_\mu. \tag{3.114}$$

Now, for a particle moving along a geodesic, $v^\nu D_\nu v_\mu = 0$ and if K^μ is the killing vector, then $D_{(\mu} K_{\nu)} = 0$. Thus, $\frac{d(K^\mu v_\mu)}{ds} = 0$, i.e., the corresponding quantity is conserved.

$$(K^\mu v_\mu) = Const, \tag{3.115}$$

for the motion along a geodesic. Now, considet the Schwarzschild spacetime and the killing vector $K^\mu = (1, 0, 0, 0)$. The conserved quantity is given as follows

$$(K^\mu v_\mu) = v_0 = g_{00} v^0 = \left(1 - \frac{r_s}{r}\right)\frac{dt}{ds} \equiv \frac{E}{m}, \tag{3.116}$$

and similarly, for the killing vector $K^\mu = (0, 0, 0, 1)$, we obtain the following conserved quantity

$$r^2 sin^2\theta \, \frac{d\phi}{ds} \equiv \frac{L}{m}. \tag{3.117}$$

Let us choose the plane to be at $x_3 = 0$, i.e., at $\theta = \pi/2$. The the conserved quantity becomes

$$r^2 \frac{d\phi}{ds} \equiv \frac{L}{m}. \tag{3.118}$$

Notice that the quantity $r^2 \frac{d\phi}{ds}$ is nothing but the area swept by the radius vector of the orbiting test particle in a given time interval. Thus, the equation above is nothing but Kepler's second law!

Furthermore, the 4-velocity must obey (with $\theta = \pi/2$)

$$g_{\mu\nu} v^\mu v^\nu = \frac{ds^2}{ds^2} = \left(1 - \frac{r_s}{r}\right)\left(\frac{dt}{ds}\right)^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\left(\frac{dr}{ds}\right)^2 - r^2\left(\frac{d\phi}{ds}\right)^2 = 1. \tag{3.119}$$

Using the two conservation laws, we find that the world line of a massive particle in the Schwarzschild spacetime obeys

$$\left(\frac{dr}{ds}\right)^2 = \left(\frac{E}{m}\right)^2 - \left(1 - \frac{r_s}{r}\right)\left(1 + \frac{L^2}{m^2 r^2}\right) \tag{3.120}$$

Physical Meaning of *E* and *L*

Consider the Newtonian and non-relativistic limits, i.e., $r_s << r$ and $dr/dt << 1$ respectively. Then $dr/ds \approx dr/dt$, and

$$\frac{E^2 - m^2}{2m} \approx \frac{m}{2}\left(\frac{dr}{dt}\right)^2 + \frac{L^2}{2mr^2} - \frac{mr_s}{2r}. \tag{3.121}$$

If $E^2 - m^2 = (E - m)(E + m) \approx 2m\mathcal{E}$, where $\mathcal{E} = (E - m) << m$ is the non-relativistic total energy, and L is the angular momentum, then the equation defines the trajectory of a massive particle present in the Newtonian gravitational field. Thus, the conservation of energy (E) follows from the invariance under time translations and the conservation of angular momentum (L) follows from the invariance under rotations.

3.10 Orbits in the Schwarzschild Metric

3.10.1 Radial Plummet or Crash

With zero angular momentum, i.e., $L = 0$, the orbit equation takes up the form

$$\left(\frac{dr}{ds}\right)^2 = \left(\frac{E}{m}\right)^2 - \left(1 - \frac{r_s}{r}\right). \tag{3.122}$$

Now, let's assume that the particle starts its free fall at infinity with null velocity, i.e., $dr/ds \to 0$ as $r \to \infty$. Then,

$$\left(\frac{dr}{ds}\right)^2 = \left(\frac{E}{m}\right)^2 - 1 = 0 \Longrightarrow E = m, \tag{3.123}$$

and the equation simplifies to

$$\left(\frac{dr}{ds}\right)^2 = \frac{r_s}{r}. \tag{3.124}$$

Now, the propertime of the particle's free fall from a radius $r = R$ to the horizon $r = r_s$ is[2]

$$s = -\int_R^{r_s} \sqrt{\frac{r}{r_s}} dr = \frac{2}{3} r_s \left(\left(\frac{R}{r_s}\right)^{\frac{3}{2}} - 1\right). \tag{3.125}$$

Thus, it takes a finite propertime for a particle to cross the Schwarzschild black hole horizon. It follows from the energy conservation that $dt/ds = \left(1 - \frac{r_s}{r}\right)^{-1}$ and thus, the ratio of dr/ds to dt/ds is given as

$$\frac{dr}{dt} = -\sqrt{\frac{r_s}{r}} \left(1 - \frac{r_s}{r}\right). \tag{3.126}$$

The time necessary for a particle to fall from a radius $r = R(>> r_s)$ to a radius $r = r_s + \epsilon$ in the vicinity of the horizon ($\epsilon << r_s$) is

$$t\left(R \to r_s + \epsilon\right) = \mathcal{T} = -\int_R^{r_s+\epsilon} \left(\frac{r}{r_s}\right) \frac{r}{r - r_s} dr \approx r_s \; ln \frac{R}{\epsilon}, \tag{3.127}$$

and hence, as $\epsilon \to 0$, $t \to \infty$, i.e., the particle cannot approach the horizon within finite time as measured by an observer fixed over the black hole. This can also be observed from another perspective. Consider the action functional

[2] note that the negative sign of the integral is due to the fact that for the trajectory of the fall, $dr < 0$ & $ds > 0$

of the Schwarzschild metric[3]

$$\begin{aligned}\mathcal{S} &= -m \int \sqrt{\left(1-\tfrac{r_s}{r}\right)\tfrac{dt^2}{dt^2}\left(1-\tfrac{r_s}{r}\right)^{-1}\tfrac{dr^2}{dt^2}}dt \\ &= -m \int \sqrt{\left(1-\tfrac{r_s}{r}\right)-\left(1-\tfrac{r_s}{r}\right)^{-1}\dot{r}^2}dt = -m\int \mathcal{L}dt.\end{aligned} \tag{3.128}$$

Now, the Hamiltonian, $\mathcal{H}$, given by $\mathcal{H} = \sum p\dot{r} - \mathcal{L} = p_r\dot{r} - \mathcal{L} = \left(\frac{\partial\mathcal{L}}{\partial\dot{r}}\dot{r}\right) - \mathcal{L}$, takes up the following form

$$\mathcal{H} = \frac{m}{\mathcal{L}}\left(1-\frac{r_s}{r}\right) = \mathcal{E}, \tag{3.129}$$

which upon squaring on either side and solving for $\dot{r}$ we obtain

$$\dot{r} = \left(1-\frac{r_s}{r}\right)\sqrt{1-\frac{m^2}{E^2}\left(1-\frac{r_s}{r}\right)}. \tag{3.130}$$

Hence, as the particle falls towards the horizon, it's velocity goes to zero as it approaches the Schwarzschild limit. The object's velocity diminishes along the trajectory instead of accelerating! It asymptotically gets nearer and nearer to the horizon, but never gets there.

3.10.2 Circular Orbit

Consider a particle travelling in a circular orbit around the black hole. It's orbit, with time looks like a helix and the projection of the helical trajectory on a surface results in a circular orbit. Let $\left(1-\frac{r_s}{r}\right) = \mathcal{F}(r)$ and $\left(1-\frac{r_s}{r}\right)^{-1} = \mathcal{G}(r)$ and since we restrict our observation to a surface, $d\phi = 0$. The action functional reads

$$\mathcal{S} = -m\int \sqrt{\mathcal{F}(r) - \mathcal{G}(r)\dot{r}^2 - r^2\dot{\theta}^2}dt = -m\int \mathcal{L}dt. \tag{3.131}$$

Now, for a circular orbit, the energy (E) and the angular momentum (L) are conserved. Thus, $L = \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = p_\theta$, and a straightforward computation yields

$$p_\theta = m\Xi\left(r,\dot{r},\dot{\theta}\right),\ \Xi = \frac{r^2\dot{\theta}}{\mathcal{L}}, \tag{3.132}$$

and the momentum associated with the radial coordinate r reads

$$p_r = \frac{m\mathcal{G}\dot{r}}{\mathcal{L}}. \tag{3.133}$$

It is important to note here that Ξ is conserved since the angular momentum,

[3] with $c \equiv 1$

L is conserved and the function $\varXi$ is called *reduced angular momentum*. But, the momentum associated with the radial coordinate, i.e., p_r is not conserved since r is not conserved under any law. Now, we write the Hamiltonian as follows

$$\begin{aligned}\mathcal{H} &= \left(p_\theta \dot{\theta} + p_r \dot{r}\right) - \mathcal{L} \\ &= \tfrac{m\dot{\theta}\varXi}{\mathcal{L}} + \tfrac{m\mathcal{G}\dot{r}}{\mathcal{L}} - \mathcal{L} \\ &= \tfrac{m\mathcal{F}(r)}{\mathcal{L}} = E,\end{aligned} \tag{3.134}$$

and for circular orbits, $\dot{r} = 0$ thus,

$$E = \frac{m\mathcal{F}(r)}{\sqrt{\mathcal{F}(r) - r^2\dot{\theta}^2}}, L = \frac{mr^2\dot{\theta}^2}{\sqrt{\mathcal{F}(r) - r^2\dot{\theta}^2}}. \tag{3.135}$$

Let's find an expression for $\dot{\theta}$ from the equation of angular momentum.

$$\begin{aligned}L &= \tfrac{mr^2\dot{\theta}^2}{\sqrt{\mathcal{F}(r) - r^2\dot{\theta}^2}} = m\varXi \\ \dot{\theta}^2 &\left(r^2 + \varXi^2\right) = \mathcal{F}(r)\tfrac{\varXi^2}{r^2} \\ \dot{\theta} &= \tfrac{\varXi}{r}\sqrt{\tfrac{\mathcal{F}(r)}{r^2+\varXi^2}} = f(r, \varXi)\end{aligned} \tag{3.136}$$

Substituting this into the expression for energy, we obtain

$$E = m\sqrt{\frac{\mathcal{F}(r)\left(r^2 + \varXi^2\right)}{r^2}} = g(r, \varXi). \tag{3.137}$$

3.10.3 Orbiting Photon

Consider an orbiting photon with fixed energy and angular momentum. Since $\varXi$ is very large, $r^2\varXi^2 >> r^4$. The energy of this photon is deduced to be

$$E = \frac{m}{r^2}\sqrt{\mathcal{F}(r)\left(r^2\varXi^2\right)} = \frac{mr\varXi}{r^2}\sqrt{\mathcal{F}(r)} = \frac{L}{r}\sqrt{\mathcal{F}(r)}. \tag{3.138}$$

Let $A(r) = \frac{\sqrt{\mathcal{F}(r)}}{r} = \frac{1}{r}\sqrt{1 - \frac{r_s}{r}}$. The orbit of the photon will be positioned where the function $A(r)$ is stationary, i.e., either maximum or minimum. Let us study this in separate cases.
1. Case I:
At large distances, $A(r) = \frac{1}{r}$ and $E = \frac{L}{r}$.

2, Case II:

At closer distances, $A(r) = 0$ and $E = 0$.

Photon Orbit

We find the maxima of the function $A(r)$ to be

$$A^{'}(r) = \frac{3r_s - 2r}{2r^3\sqrt{1-\frac{r_s}{r}}} = 0$$
$$r = \tfrac{3}{2}r_s. \tag{3.139}$$

This radius at which the photon orbits is called the *photon sphere* (see figure 3.1). Thus, it is implied that photon spheres can only exist in the space surrounding an extremely compact object such as a neutron star or a black hole. This unstable orbit is independent of the angular momentum. Any particle (with mass) inside the photon sphere, moving in an angular direction, would wind up inspiraling into the singularity.

No signal from the star's surface can escape to infinity once the surface has passed through $r = r_s$. For the external observer, the surface never actually reaches $r = r_s$, but as $r \to r_s$ the redshift of light leaving the surface increases exponentially fast and the star effectively disappears from view within a time $\approx GM/c^3$. The late time appearance is dominated by photons escaping from the unstable photon orbit at $r = 1.5r_s$.

3.10.4 The Schwarzschild Potential

Using the two laws of the conserved quantities of the Schwarzschild metric, we found that the world line of a massive particle in the Schwarzschild spacetime obeyed a certain equation. Let us now rewrite the equations in terms of the effective potential, $V^2(r)$ as

$$\left(\frac{dr}{ds}\right) = \frac{1}{m}\sqrt{E^2 - V^2(r)}, \quad V^2(r) = m^2\left(1 - \frac{r_s}{r}\right)\left(1 + \frac{L^2}{m^2 r^2}\right). \tag{3.140}$$

Let us define $h = L/m$. Hence, in terms of h the equation takes the following form

$$\left(\frac{dr}{ds}\right) = \frac{1}{m}\sqrt{E^2 - \left[m^2\left(1 - \frac{2GM}{r}\right)\left(1 + \frac{h^2}{r^2}\right]^2\right)}, \tag{3.141}$$

where $r_s = 2GM$. Now, to understand the nature of the relativistic orbits, let us determine the maxima and minima of this effective potential in terms of dimensionless variables $w \equiv GM/c^2 r$ and $\hat{L} = L/Mm = Lc/GMm$. The

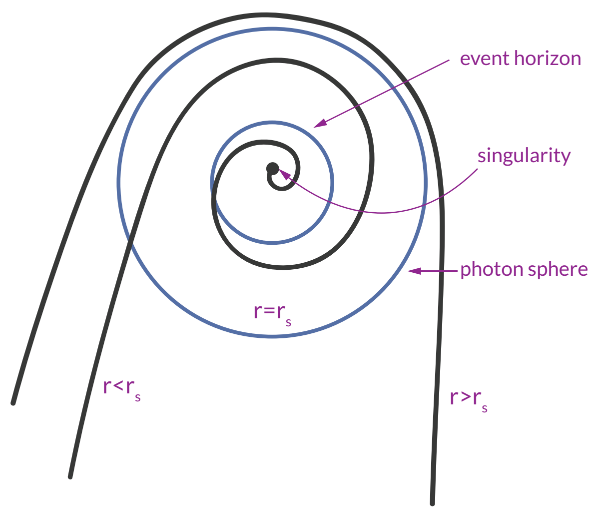


Fig. 3.1. Photon sphere and various orbits.

maxima of the function $V^2(r)$ occurs at

$$w_m \equiv \frac{1 \pm \sqrt{1 - \frac{12}{\hat{L}}}}{6}, \tag{3.142}$$

with the maximum potential being

$$V_m^2(\hat{L}) = m^2 \left(1 - 2w_m\right) \left(1 + \hat{L}^2 w_m^2\right). \tag{3.143}$$

Observe that for $\hat{L} > \sqrt{12}$, i.e., for $L > 2\sqrt{3}GMm$, the effective potential has one maximum and one minimum. There is a unison of the two extrema for $L = 2\sqrt{3}GMm$ and the function becomes monotonic for $L < 2\sqrt{3}GMm$. The maximum potential is reached when $L = 4GMm$. Several important aspects of the motion can be deduced by plotting a graph of $V(r)/m$ against rc^2/GM[4] for different values of L.

[4] = r/M in natural units, i.e, when $c \equiv G \equiv 1$

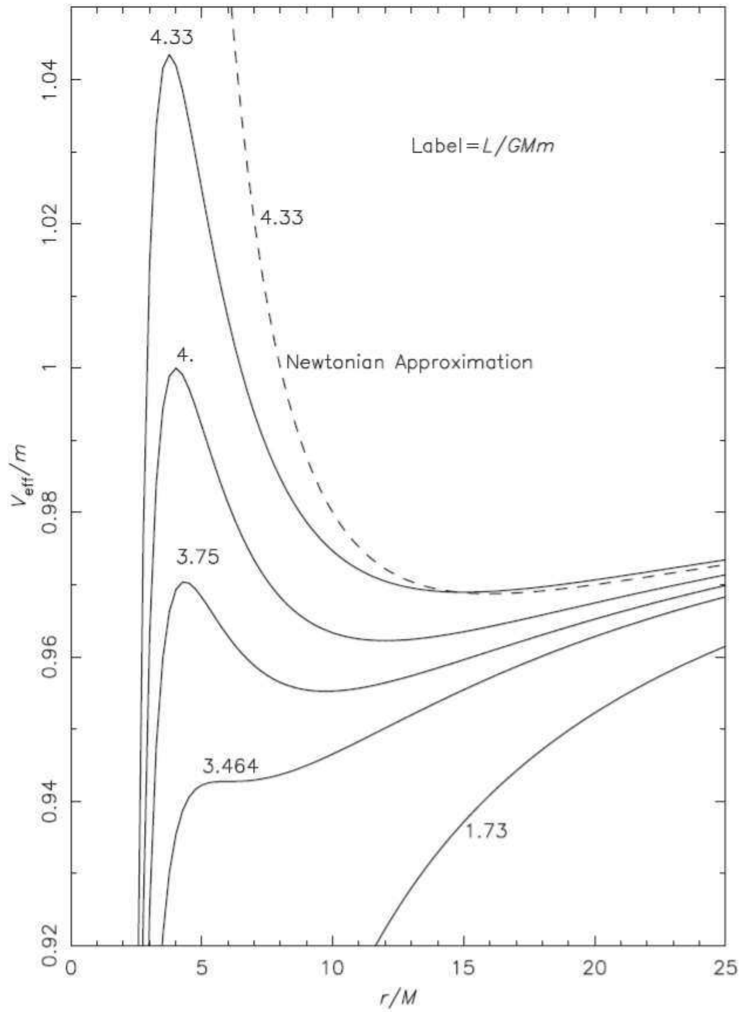


Fig. 3.2. The plot shows the effective potential for an object in the Schwarzschild metric for varying r/M

Notice that for $h = 0$, the radial plummet is just as it is for the Newtonian case because the potential is the same. For a given value of L and E the nature of the orbit will be governed by the turning points inn r, determined by the equation $V^2(r) = E^2$. If $L > 4GMm$, the the value of V_{max} is greater than m. Also, $V_{max} \to m$ as $r \to \infty$ for all values of L. If the energy E of an object is lower that m and $L > 4GMm$, then there will be two turning points. The object will orbit the central body with a perihelion and an aphelion undergoing precession (similar to elliptical orbits in Newtonian gravity).

For $m < E < V_{max}(L)$, there will be only one turning point. The object will approach the central mass from infinity, reach a radius of closest approach and travel back to infinity (similar to hyperbolic orbits in the Newtonian gravity).

For $E = V_{max}(L)$, the orbit will be circular at some fixed radius $\bar{r}$ determined by the condition $V'(\bar{r}) = 0$, $V(\bar{r}) = E$. Solving these equations simultaneously, we find that the the radii of circular orbits and their energies are given by

$$\frac{\bar{r}}{2GM} = \frac{L^2}{4G^2M^2m^2}\left(1 \pm \sqrt{1 - \frac{12G^2M^2m^2}{L^2}}\right), \quad E^2 = \frac{L^2}{GM\bar{r}}\left(1 - \frac{2GM}{\bar{r}}\right)^2 . \tag{3.144}$$

The upper and lower signs refer to the stable and the unstable orbit respectively. The stable orbit closest to the center has parameters $\bar{r} = 6GM$, $L = 2\sqrt{3}GMm$ and $E = m\sqrt{8/9} \approx 0.943m$. When an object falls into the black hole from the stable circular orbit closest to the center, it can release a fraction 0.057 of it's energy in radiation. An interesting case is that when $E > V_{max}(L)$, the object falls to the center. This behaviour to Newtonian gravity in which an object with non-zero angular momentum can never reach $r = 0$.

3.11 Null Hypersurfaces

Let's remind ourselves the definition of a normal of a hypersurface whose equation is $\Sigma \equiv r - const = 0$, given as $n_\mu = \Sigma_{,\mu} = (0, 1, 0, 0)$. From the Schwarzschild metric and the definition of a normal of a hypersurface, in the case of $r = const$ surface, we have

$$n_\mu n_\nu g^{\mu\nu} = g^{rr} = 1 - \frac{r_s}{r}. \tag{3.145}$$

Thus, the surfaces $r = const$ are spacelike if $r < r_s$, null if $r = r_s$, and timelike if $r > r_s$. The null hypersurface $r = r_s$ separates the regions of space where $r = const$ are timelike hypersurfaces from regions where $r = const$ are spacelike hypersurfaces; this implies that a particle crossing a null hypersurface can never comeback. Hence, the null hypersurface $r = r_s$ is called a horizon. The region with $r > r_s$, where the $r = const$ hypersurfaces are timelike; the $r \to \infty$ limit, where the metric becomes at, is in this region, so we can consider this region as the exterior of the Schwarzschild black hole. The region with $r < r_s$, where the $r = const$ hypersurfaces are spacelike; an object which falls inside the horizon and enter in this region can only continue falling to decreasing values of r, until it reaches the curvature singularity $r = 0$; this region is then considered the interior of the black bole.

3.11.1 Null Geodesic Generators

Now, consider a surface $\mathcal{S}$ given by equation $\mathcal{S}(x^\mu) = 0$. It is well known that the normal n_μ is in the direction $\partial_\mu \mathcal{S}$ to the surface. Consider another vector t^μ such that t^μ is orthogonal to the normal, i.e., $t^\mu n_\mu = 0$. Let the vector $t^\mu = dx^\mu/d\lambda$ for some curve $x^\mu(\lambda)$ on that surface, then the orthogonality condition implies that the vector t^μ is tangent to the normal and that $\mathcal{S}$ does not change along the curve, i.e., $(dx^\mu/d\lambda)\partial_\alpha \mathcal{S} = 0$. When the norm of $\partial_\mu \mathcal{S}$ is non-zero, we have normalize this vector so that its norm is ± 1. For a null hypersurface Σ, the normal to the hypersurface is also to it. This is satisfied by the the normal itself when it is null, i.e., $n_\mu n^\mu = 0$. When the normal is null we will henceforth use the symbol l instead. The normal l is tangent to the null curves $x^\mu(\lambda)$ in Σ: $l^\mu = dx^\mu/d\lambda$. In fact, the integral curves of l are null geodesics on the surface $\mathcal{S}$ in the hypersurface σ. Let $l^\mu = f g^{\mu\nu}\partial_\nu \mathcal{S}$, where f is an arbitrary function. We have

$$
\begin{aligned}
l \cdot \nabla l^\nu &= l^\mu \nabla_\mu \left(f g^{\nu\rho}\partial_\rho \mathcal{S}\right) \\
&= l^\mu \left(\nabla_\mu f\right) g^{\nu\rho}\partial_\rho \mathcal{S} + f l^\mu g^{\nu\rho}\nabla_\mu \partial_\rho \mathcal{S} \\
&= \left(l \cdot \nabla f\right) f^{-1} l^\nu + f l^\mu g^{\nu\rho}\nabla_\mu \partial_\rho \mathcal{S}.
\end{aligned}
\tag{3.146}
$$

The second term reduces as follows

$$
\begin{aligned}
f l^\mu g^{\nu\rho}\nabla_\mu \partial_\rho \mathcal{S} &= f l^\mu g^{\nu\rho}\nabla_\rho \partial_\mu \mathcal{S} \\
&= f l^\mu g^{\nu\rho}\nabla_\mu \left(f^{-1} l_\mu\right) \\
&= f g^{\nu\rho}\nabla_\mu \left(f^{-1} l_\mu\right) \underbrace{l^2}_{=0} + l^\mu g^{\nu\rho}\nabla_\rho l_\mu \\
&= \tfrac{1}{2} g^{\nu\rho}\nabla_\rho (l^2) \\
&\propto l^\nu .
\end{aligned}
\tag{3.147}
$$

In the last line, note that $l^2 = 0$ on Σ does not necessarily imply $\nabla_\rho(l^2) = 0$, because l^2 can be non-zero outside of Σ and hence its derivative can be non-zero. However, any non-zero contribution to $\nabla_\rho(l^2)$ must be proportional to the normal to Σ: $\nabla_\rho(l^2) \propto l_\rho$. Together with (3.81) this implies that $l \cdot^\nu \propto l^\nu$ and therefore the integral curves of l^μ are geodesics. If the integral curves of l^μ are not affinely parametrised, then we can always find some function $\zeta(x)$ such that $\bar{l}^\mu = \zeta(x) l^\mu$ has affinely parametrised integral curves, i.e., $\bar{l} \cdot \nabla \bar{l}^\mu = 0$. These curves are called the *null geodesic generators* of the null hypersurface.

Exercise 12

1. Consider the following metric which is a solution to the Einstein equation with $\Lambda > 0$

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2 d\omega_2^2,$$

a. show that $f(r) = \left(1 - \frac{r^2}{l^2}\right)$, where $l^2 = \frac{3}{\Lambda}$.
b. Let τ denote the proper time of a particle. The action can be written as follows

$$\mathcal{S} = \int d\tau \left(-f(r)^2 \dot{t}^2 + f(r)^{-2} \dot{r}^2 + r^2 \left(\dot{\theta}^2 + sin^2\theta\ \dot{\phi}^2\right)\right),$$

where $\dot{x}^\mu = dx^\mu/d\tau$. Show that there are two conserved quantities, the angular momentum $L = \frac{1}{2}\frac{d\mathcal{L}}{d\dot{\phi}} = r^2 sin^2\dot{\phi}$ and the energy $E = \frac{1}{2}\frac{d\mathcal{L}}{d\dot{t}} = f(r)\dot{t}$.
c. The equations of motion arising from the action should be supplemented with some constraint that informs us whether we're dealing with a massive or massless particle. For a massive particle, the constraint ensures that the trajectory is timelike, i.e, $-f(r)^2\dot{t}^2 + f(r)^{-2}\dot{r}^2 + r^2\left(\dot{\theta}^2 + sin^2\theta\ \dot{\phi}^2\right) = -1$. Restricting to geodesics that lie in the $\theta = \pi/2$ plane, show that the effective potential is given by $V_{eff}(r) = \left(1 + \frac{L^2}{r^2}\right) f(r)^2$.

2. A particle which is in a circular orbit around a black hole is perturbed in such a way that it's angular momentum is unchanged, but the energy is slightly increased so that there is a small velocity component outwards. Describe and sketch the resulting behavior, for initial radii $3M$, $4M$, $5M$ and $6M$.

3. Consider the metric

$$ds^2 = -dt^2 + \left(1 + \alpha r^2\right) dr^2 + r^2 d\Omega_2^2,$$

where α is a positive constant.
a. Consider the null geodesics on the equatorial plane ($\theta = \pi/2$) and show that they satisfy

$$\left(\frac{dr}{d\phi}\right)^2 = r^2\left(1 + \alpha r^2\right)\left(\beta r^2 - 1\right),$$

where β is a constant.
b. Show that, by integrating this equation, the paths of light rays are ellipse.

4. Find the geodesic equations for a wormhole which is described by the metric 5.54.

5. The Friedmann-Lemaître-Robertson–Walker (FLRW) has the following metric

$$ds^2 = -dt^2 + a(t)^2\left(\frac{1}{1 - kr^2} dr^2 + r^2 d\Omega_2^2\right), \tag{3.148}$$

where $a(t)$ is a scaling parameter called the expansion factor and $k \in \{1, 0, -1\}$. This metric describes a universe that is spatially homogeneous and isotropic at each instant of time.
a. Show that the Ricci scalar takes the form, $R = 6a^{-2}\left(k + \dot{a}^2 + a\ddot{a}\right)$.
b. Show that the Kretschmann scalar is $K = 4a^{-4}\left(29 + 18\dot{a}^2 + 3\dot{a}^4 + 3a^2\ddot{a}^2\right)$.
c. Show that the Weyl tensor vanishes for this metric.
d. The vanishing of the Weyl tensor implies that there exists a coordinate system in which this metric (for all k) is conformal to the Minkowski metric. Show that the spatially flat FLRW metric with $k = 0$ can be expressed as $g_{\mu\nu} = a(\eta)^2\eta_{\mu\nu}$, where η is the conformal time coordinate defined as $\eta = \int a(t)^{-1}dt$ and $\eta_{\mu\nu}$ is the Minkowski metric.

6. In the FLRW metric 3.148, consider the case with $k = 0$, when a fluid system that is described by the stress energy tensor

$$T_{\mu\nu} = \left(\rho(t), -a(t)^2p(t), -a(t)^2p(t), -a(t)^2p(t)\right),$$

where ρ is mass density and p is pressure.
a. Find the Einstein tensor and obtain the following Friedmann equations for such a system (with $c \equiv 1$)

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho, \ \frac{2\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 = -8\pi Gp.$$

b. Show that these two Friedmann equations lead to the following

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3p\right).$$

This equation tells us that $\ddot{a} > 0$, i.e., the universe will undergo accelerated expansion, only when $(\rho + 3p) < 0$. Note that for normal matter, $(\rho + 3p) > 0$ and hence, $\ddot{a} < 0$ which implies that the universe will have a decelerating expansion.
c. Apply the stress-energy conservation $D_\mu T^{\mu\nu} = 0$ and show that it yields

$$\dot{\rho} + 3\frac{\dot{a}}{a}\left(\rho + p\right) = 0.$$

d. Solve this equation which governs the evolution of $\rho(t)$ in a Friedmann universe for a fluid with an equation of state $p = w\rho$, where w is a constant, and show that $\rho(t) \propto a(t)^{-3(w+1)}$. Hint: Set $H = \dot{a}/a$ and solve the equation, this is called the Hubble parameter.
e. Consider the Einstein field equation with a cosmological constant. Derive the Friedmann equations for such a case and write down an explicit expression for the scaling factor for a relativistic fluid with $w = 1/3$.

7. Consider the Kerr metric in Boyer-Lindquist coordinates which describes the geometry around a rotating black hole,

$$\begin{aligned} ds^2 = &-\left(1 - \frac{r_s r}{\Sigma}\right) dt^2 - \frac{2r_s a r sin^2\theta}{\Sigma}\, dt d\phi + \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2 \\ &+ \left(r^2 + a^2 + \frac{r_s a^2 r sin^2\theta}{\Sigma}\right) sin^2\theta d\phi^2, \end{aligned} \tag{3.149}$$

where $\Sigma = r^2 + a^2 cos^2\theta$, $\Delta = r^2 - r_s r + a^2$, and the Kerr parameter $a = J/M$, i.e., the ratio of the angular frequency to the mass of the black hole.
a. What happens to the metric when there is no angular frequency, i.e., $a = 0$.
b. Consider the g_{rr} component, this is singular when $\Delta = 0$. Show that this has two solutions $r_\pm$ and also check if $r_+ \to r_s$ and $r_- \to 0$ as $a \to 0$. This implies a presence of an event horizon at r_+. The expression of r_+ indicates that the angular momenta of Kerr black holes are limited by the square of their mass. Kerr black holes with the largest possible angular momentum $J = M^2$ are called extremal black holes. Matter, principally from an accretion disk around the black holes, spirals into the black hole and thus adds its angular momentum to it[5].
c. Consider the g_{tt} component, the condition $g_{tt} = 0$ has two solutions r_i andr_o, find these. The outer solution r_o defines the boundary of the ergoregion which surrounds the event horizon.
d. Show that the normal vector to the surface $r = r_+$ is null.
e. Rewrite 3.149 for $\theta = \pi/2$ and find out the effective potential at the equatorial plane felt by a test particle moving in the spacetime.

[5] This implies that many black holes would naturally tend towards becoming extremal.

4

Singularities and their Eliminations

4.1 Singularities and The Cosmic Censorship Hypothesis

It was previously discussed that there are two singularities present in the Schwarzschild metric, the unphysical coordinate one at $r = r_s$, and the physical one at $r = 0$. After falling through the event horizon, the particle continues to move to smaller values of r, and will experience larger tidal distortions. At the hypothetical point $r = 0$ however, the particle would experience an infinite tidal force. This physical singularity is a harmless one, harmless since it does not affect the black hole exterior as its hidden behind an event horizon. But, what if a singularity has nowhere to hide? The singularity at $r = 0$ in the past is called the *white hole singularity.* What's a white hole? For now, let's put it this way: A white hole is an object which is the exact opposite of a black hole, a black hole *sucks* while a white hole *pukes.* Before proceeding it is important to note here that a black hole doesn't suck matter, it ain't a cosmic vacuum cleaner! If our Sun were to be replaced by a black hole, the gravitational field would almost remain the same and the Earth would continue to orbit around the new incognito Sun, it is only beyond the Schwarzschild radius that objects can't escape out of the black hole via the event horizon. Hence, I have exercised my poetic-licence when I attributed black holes to suck. For the sake of symmetry, every black hole has a white hole counterpart such that the total energy is conserved. A question may arise now whether all the particles that go into a black hole come out through the white hole which leads to a deeper question revolving around whether the big bang really did happen at a point. The latter question would be addressed later. Well, according to the idea presented above, they might but not in one piece.

When, Mr. Absolute Zero falls into a black hole, he would get *spaghettified* (since the gravity acting at his feet is greater that the gravity at his head), he would be ripped into pieces and the pieces would get ripped into smaller ones, until Mr. Absolute Zero is just an accumulation of the most basic particles that form matter. These particles would then have to travel through a throat-like bridge (known as the *Einstein-Rosen bridge*) to the other side and then be spit out by the white hole. What if Mr. Absolute Zero wants to get to the other side in one piece? He then has to travel faster than the speed of light (which is clearly impossible with today's technology) & if he does then he would experience the mysterious new universe.

A white hole singularity is known as a *naked singularity*. The *cosmic censorship hypothesis* proposed by Roger Penrose states that these naked singularities are illegal. It is basically a mathematical conjecture, i.e., it is as yet neither proven nor refuted. This conjecture states that singularities are never naked, i.e., they can never exist outside of black holes. Einstein's field equations can yield surplus unphysical solutions like the case of having naked singularities (not covered by event horizons). Naked singularities are points where Einstein's classical theory of gravity breaks down and quantum gravity takes over. In 2006, there emerged a new conjecture which states that gravity is always the weakest force in any universe, this was the *Weak Gravity Conjecture*. A simple way to check this is to calculate the and compare the coulombic and the gravitational force between two electrons separated by a unit distance. The gravitational force happens to be of the order of 10^{-71} N, while the coulombic force is of the order of 10^{-28} N. Now think of this- What if the quantum particles that exist in the universe gravitationally collapse into a black hole due to a high energy electric field? Such a black hole formed would have a naked singularity and thus, we can establish that there does exist a connection between the two conjectures.

A physicist would believe that in real situations black holes are results of collapse processes, he would expect that during these processes any quantity which is physically measurable remains regular and that the formation of a black hole by the collapse of matter is not a time-symmetric process which leads to a black hole without the associated white hole. In fact, physicists John Preskill and Kip Thorne debated Hawking that observable naked singularities could exist and this gave rise to a bet which was made in 1991. However, after supercomputer simulations indicated that naked singularities could exist Hawking conceded the bet on the 5^{th} of February, 1997. Hawking presented his colleagues with 'adequate raiment's to shield their nakedness from the vulgar view'. The goods he presented consisted of two T-shirts, which was apparently inscribed with "an appropriate message" from Hawking!

Here's an interesting question to ask: When does the singularity of a black hole form? The answer is simple but deep. When a star collapses to form a black hole, the horizon is first formed before the singularity, since if the events occurred in reverse, we would have a naked singularity and these were declared outlaws by judge Penrose, right? Now the time inside the event horizon is frozen with respect to Mr. Zero Entropy who is an outside observer, thus Mr. Zero Entropy outside would never observe the singularity, but in the reference frame of the collapsing star, the singularity forms in approximately milliseconds after the horizon's formation. Hence, with respect to Mr. Zero Entropy, the singularity never formed and he would assume the black hole formed to be singularity-less!

Another way to define a singularity is by stating that it (singularity) is a condition in which geodesics are incomplete. For example, if Mr. Absolute Zero is dropped into a black hole, his world-line terminates at the singularity. He is not just destroyed due to being spaghettified, but also possesses no future world-lines. It is important to note here that the black hole singularity is a type of curvature singularity, i.e., as Mr. Absolute Zero approaches the singularity, the curvature of spacetime diverges to infinity as measured by a curvature invariant such as the Ricci scalar. There is another type of singularity called the conical singularity. This one like the tip-of-the-cone singularity, once geodesics hit the tip there is arbitrariness in the direction it would proceed in[1]. Hence, geodesics are incomplete. The singularities involving geodesic incompleteness are not coordinate singularities which were discussed above. Now, let's tackle the question regarding the big bang-did the big bang occur at a point? First off, the singularities that we discuss in general relativity are not points in spacetime; it's like the hole in the topology of a manifold. A simple answer to this question is no, the big bang did not happen at a point, instead it happened everywhere in the universe at the same time. Consequences of the above answer includes the fact that the universe has no centre, the big bang did not happen at a central point in the universe that it is expanding from. This can be supported with a simple example- imagine the universe was a balloon on the surface of which are galaxies, if Mr. Absolute Zero was on one of the galaxies he would observe the other galaxies drifting away from him as the balloon (i.e. universe) expands. From this observation, Mr. Absolute Zero would conclude that he is on the centre of the universe, but on another universe, his cousin Mr.Zero Entropy would make a similar observation and this would lead him to believe that he and not his cousin is on the centre of the universe. Every person in every galaxy on the balloon would think that they are the centre of the universe (literally!). This is the reason why there is no centre for the universe's expansion. Another consequence of the answer is that the universe isn't expanding into anything, there is no space outside into which it can expand.

4.2 Birkhoff's Theorem

4.2.1 An Introduction to Birkhoff's Theorem

From the Schwarzschild metric, we know that at large distances from our spherically symmetric gravitating object, i.e., as $r \to \infty$, the metric translates

[1] for examples see the section on embeddings

to flat spacetime. Such metrics abiding by the above rule are referred to as *asymptomatically flat*. Thus, an asymptomatically flat, spherically symmetric metric in vacuum ($\Lambda = 0, T_{\mu\nu} = 0$) is static (i.e., time independent and diagonal). This is Birkhoff's theorem. When we say static, we mean that the coefficients of the Schwarzschild metric are independent of t; such a metric is said to be stationary. Also, the metric does not have any non-diagonal components, such as $dtd\Phi, d\theta dr$, etc. A stationary metric having only diagonal components is said to be static. Using this new-found knowledge of spherical symmetry in the previous chapter, Birkhoff's theorem can be restated as follows

Theorem 4.1. *A smooth spherically symmetric metric solution of the vacuum equations of Einstein field equations is necessarily static.*

Observe the elegance of the theorem, every word of the theorem now has an embedded mathematical note, thus making the theorem a physical symphony! Such is the power of math. Let's understand the power of this theorem, to do that lets create a black hole!

4.2.2 Shell Theorem

Consider a spherically symmetric shell of incoming radiation, incoming with the speed of light and which carries energy and momentum. Newtonian theory suggests that on the interior of this shell, there is no gravitational field, while on the outside one would see the gravitational field as it would be if all the mass was concentrated at a point. This is the famous *Shell theorem* of Newton and this holds true even if the shell is moving. Moreover, the gravitational field inside such a shell varies linearly with distance from its centre. To check this, consider the following experiment: let Mr. Absolute Zero to be placed inside the shell at a distance r from the centre, now since there is no net gravitational force exerted by the shell on any particle inside we can ignore all the shells of greater radii, in accordance to the shell theorem. Thus, the remaining mass m is proportional to r^3 and from Gauss law ($\int g d\mathbf{S} = -4\pi G m$), we get the gravitational field to be proportional to $\frac{m}{r^2}$ and using the proportionality of m, we conclude that the gravitational field is proportional to $\frac{r^3}{r^2} = r$ (hence the linear relation). Now, let's analyse this shell the GR way. Birkhoff's theorem states that a spherically symmetric solution is static, and a shell, more precisely a vacuum shell corresponds to the radial branch of the Schwarzschild solution in some radial interval $r \in [r_1, r_2]$. Here, since there is no mass M, at the centre of the shell (corresponding to the region $r \in [0, r_2]$), the Schwarzschild radius is zero ($r_s = \frac{2GM}{rc^2} = 0$). Hence, the metric reduces to the following form

$$ds^2_{shell} = c^2 dt^2 - \frac{1}{c^2}(dr^2 - r^2 d\Omega^2). \tag{4.1}$$

This is nothing but the flat Minkowski metric in spherical polar coordinates. This implies that the gravitational field must vanish inside this spherically symmetric shell. This is in perfect unison with what the shell theorem of Newtonian theory suggests. An interesting consequence of the Birkhoff theorem is that no radial changes in a spherical star (expansion, pulsation, contraction, etc.) can affect its external gravitational field (reason for this statement is provided in the experiment performed in the upcoming paragraph): all spherically symmetric vacuum gravitational fields are indistinguishable for $r > r_s$. This implies that a spherically pulsating star cannot emit gravitational waves, nor can it gravitationally radiate away its mass. This can be viewed in another way; any spherically symmetric perturbation and the subsequent collapse of an equilibrium configuration cannot affect the external gravitational field as long as exact spherical symmetry is maintained. Analogous to how Maxwell's laws prohibit monopole electromagnetic waves, Einstein's laws prohibit monopole gravitational waves! There is absolutely no way for any gravitational influence of the radial collapse to propagate outward.

4.2.3 Gedankenerfahrung

What can be the explanation for why spherical symmetry collapse would not lead to gravitational waves? Let us understand this intuitively via a *Gedankenerfahrung* (thought experiment). Consider the collapse of a static, solid and non-rotating star instead of that of a vacuumed shell. The star is modelled to be an ideal spherically symmetric ball filled with matter surrounded by vacuum. It is important to note the assumption that spherical symmetry of the ball is never to be violated with time. If the symmetry is violated by perturbations, they will grow in time due to tidal forces. Also, any violation of the spherical symmetry of the star would lead to creation of time dependent gravitational fields in the vacuum assumed, i.e., to a creation of gravitational waves. Initially the star wouldn't have been static due to some thermonuclear processes going on inside it, which was the reason for the internal pressure of the star. But, say at a time $t = 0$ the entire fuel was used up and the star turns out to be momentarily static. We are to further assume that there is a homogeneous distribution of pressure-less material (called *dust* by physicists) inside the ball. Such a massive star would create a spherically symmetric gravitational field that is asymptotically flat. Hence, from Birkhoff's theorem, we conclude that the metric outside the star is the Schwarzschild metric. Now, we are to do something amazing, electrically charge the ball such that the charge is homogeneously distributed over its volume (i.e., uniform volume charge density). Alright, since we have laid the foundations, let's answer the question. Imagine the ball starts to contract rapidly, but while doing so the spherical symmetry and homogeneity of it is intact. The vital observation during such a process is to see that independent of the radius of the ball, the coulomb field outside the ball remains unchanged. Let's prove this before proceeding the experiment any further. From Gauss theorem, we have

$$\oint_A \mathbf{E} d\mathbf{S} = \frac{1}{\epsilon_0} \oint_V \rho dV, \quad E(r) = \begin{cases} \frac{\rho}{3\epsilon_0} r = \frac{1}{4\pi\epsilon_0}\frac{Q}{R^3} r, & R > r \\ \frac{\rho}{3\epsilon_0} r = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^2}, & R < r \end{cases} \tag{4.2}$$

Note that when the ball shrinks from radius $R \to r'$, the coulomb field on the interior would change, but not the exterior since the electric field for $R < r$ is independent of the radius of the ball, R and hence, independent of any evolution of R with time. Thus, the electric field is time independent and becomes an electrostatic field. In this case we would observe that the magnetic field outside the ball is vanishing ($\nabla \times E = -\frac{\partial \mathbf{B}}{\partial t} = 0$) and such a motion of charge does not create an electromagnetic radiation. Resuming the experiment, we observe that all momenta are zero with respect to the centre, but to have radiation we need a time varying dipole moment. Similarly, for gravitational radiation we need a time varying quadruple moment which we don't observe in our experiment. Thus, there is no creation of gravitational waves (which transport energy as gravitational radiation) due to spherical symmetry collapse.

For a very elegant proof of Birkhoff's theorem, consider the time translation: $t \to t + \delta t$, where δt is a small fraction of time and is a constant. Applying this time translation to the Schwarzschild solution we find that we obtain back the original line element. Thus, as was discussed in the previous chapter, the metric is invariant under time translations (since g_{tt} and g_{rr} do not depend on time), and this implies that the spherically symmetric solution of Einstein's vacuum field equation is necessarily static, which is what the Birkhoff theorem states. Considering the field of a pulsating star, what this theorem implies is that since spherically symmetry is maintained constantly and although the mass is pulsating (i.e., mass remains at m but its radius keeps varying), the exterior metric remains static (i.e., the star has the same external field of a star, of the same mass, at rest). This also implies that spherically symmetric bodies cannot produce gravitational waves. It turns out that it takes a spherically asymmetric situation for the production of gravitational waves.

4.2.4 Birkhoff's Theorem: A Mathematical Proof

Now we move on to the actual mathematical proof of Birkhoff's theorem. All that talk about spherical symmetry implies that spacetime can be foliated by 2-spheres, each with a line element

$$ds^2 = r^2 d\theta^2 + sin^2\theta d\Phi^2 \tag{4.3}$$

Let's set r to be our third coordinate and refer to the rest of the coordinates as t. Spherical symmetry of the geometry further implies that we can choose the r and t directions to be orthogonal to the spheres, then the complete form the line element is

$$ds^2 = -\alpha dt^2 + \beta dr^2 + 2\gamma dtdr + r^2 d\theta^2 + sin^2\theta d\Phi^2. \tag{4.4}$$

where α, β, and γ are functions of r and t only. For a constant θ and Φ , we can express the metric in the (t, r)-plane as follows

$$ds^2 = -\alpha dt^2 + \beta dr^2 + 2\gamma dtdr = -\left(\alpha dt - \frac{\gamma}{\alpha}dr\right)^2 + \left(\beta^2 - \frac{\gamma^2}{\alpha^2}\right). \tag{4.5}$$

Now, we can use the uniqueness property of differential equations- assuming that the differential equation $\frac{dt}{dr} = \frac{\gamma}{\alpha^2}$ has a solution $t = \zeta(r)$, then

$$d(t - \zeta(r)) = dt - \frac{\gamma}{\alpha^2}dr = \frac{\alpha dt - \frac{\gamma}{\alpha}dr}{\alpha}. \tag{4.6}$$

And thus, we can set $t^{'} = t - \zeta(r)$ and $\rho = \alpha$. Thus, any differential in two dimensions is a multiple of an exact differential, and we have

$$\alpha dt - \frac{\gamma}{\alpha}dr = \rho dt^{'}. \tag{4.7}$$

We can now write the line element in terms of the newly defined quantities, ρ and $t^{'}$ as

$$ds^2 = -\rho^2 dt^{'2} + \chi^2 dr^2, \tag{4.8}$$

where $\chi^2 = \beta^2 - \frac{\gamma^2}{\alpha^2}$. Note that χ^2 is positive as the metric signature is invariant. This also implies that ρ and χ are now functions of $t^{'}$ and r. We henceforth drop the primes; this argument shows that we could have simply assumed that $\gamma = 0$ (thus, $\rho = \alpha$ and $\chi = \beta$). The step of the proof is to compute the Einstein tensor for the original line element (with $\gamma = 0$), $ds^2 = -\alpha dt^2 + \beta dr^2 + r^2 d\theta^2 + sin^2\theta d\Phi^2$

$$G^t_t = -\frac{1}{r^2\beta^2}\left(\beta^2 - 1 + \frac{2r}{\beta}\frac{\partial\beta}{\partial r}\right)$$

$$G^t_r = -\frac{2}{\alpha^2\beta r}\frac{\partial\beta}{\partial t}$$

$$G^r_r = -\frac{1}{r^2\beta^2}\left(\beta^2 - 1 + \frac{2r}{\alpha}\frac{\partial\alpha}{\partial r}\right)$$

$$G^\theta_\theta = G^\Phi_\Phi = \frac{1}{r^2\beta^2}\left(-\frac{r}{\beta}\frac{\partial\beta}{\partial r} + \frac{r}{\alpha}\frac{\partial\alpha}{\partial r} + \frac{r^2}{\alpha^2}\frac{\partial^2\alpha}{\partial r^2} - \frac{r^2}{\beta}\frac{\partial^2\beta}{\partial t^2} - \frac{r^2}{\alpha\beta}\frac{\partial\beta}{\partial r}\frac{\partial\alpha}{\partial r} + \frac{r^2}{\alpha\beta}\frac{\partial\beta}{\partial t}\frac{\partial\alpha}{\partial t}\right). \tag{4.9}$$

Setting $G^t_r = 0$, we observe that β must be a function of r alone. Similarly, setting $G_t^t - G_r^r = 0$, we obtain the following

$$\frac{1}{\beta}\frac{\partial\beta}{\partial r} = -\frac{1}{\alpha}\frac{\partial\alpha}{\partial r}. \tag{4.10}$$

This implies that $\beta = \frac{\Lambda}{\alpha}$, where $\Lambda(= f(t))$ acts as an integration constant. By setting $dt^{'} = \frac{dt}{\Lambda}$, we can absorb this integration constant into the definition of t. We can therefore safely assume that $\Lambda = 1$, so that

$$\beta = \frac{1}{\alpha}, \tag{4.11}$$

and the line element take the form

$$ds^2 = -\alpha dt^2 + \frac{1}{\alpha^2}dr^2 + r^2 d\theta^2 + sin^2\theta d\Phi^2, \tag{4.12}$$

where α is a function of r alone. Setting $G_t^t = 0$ and using 4.10 and 4.11 we get

$$\frac{1}{\alpha}\frac{\partial\alpha}{\partial r} = \frac{1}{\beta}\frac{\partial\beta}{\partial r} = \frac{\beta^2 - 1}{2r} = \frac{1-\alpha^2}{2r\alpha^2}. \tag{4.13}$$

Upon separation we yield a differential equation in α and r, and hence solve for it as

$$\begin{aligned} &\textstyle\int \frac{\alpha d\alpha}{1-\alpha^2} = \int \frac{dr}{2r} \\ &-ln(1-\alpha^2) = ln(cr) \\ &\alpha^2 = 1 - \frac{1}{cr}. \end{aligned} \tag{4.14}$$

where c is an integration constant. The final step of the proof is to insert $G_t^t = 0$ into $G_\theta^\theta = G_\Phi^\Phi$ and using 4.10 along with the final relationship

$$\frac{\partial}{\partial r}\left(\frac{1}{\alpha}\frac{\partial\alpha}{\partial r}\right) = \frac{1}{\alpha}\frac{\partial^2\alpha}{\partial r^2} - \frac{1}{\alpha}\left(\frac{\partial\alpha}{\partial r}\right)^2. \tag{4.15}$$

which shows that these components vanish identically, thus completing the proof.

4.3 Israel's Theorem

Now to Israel's theorem which is just the converse of Birkhoff's theorem, i.e., it states that given an asymptomatically Minkowski spacetime, the Schwarzschild solution is the only valid static spacetime solution for Einstein's field equations. It is important to note that not every spherically symmetric spacetime will have a metric like the Schwarzschild one, but it can be made to look like the latter by suitable coordinate transformations. The Schwarzschild geometry is the unique static vacuum gravitational field in general relativity, this can be said to be a sort of uniqueness theorem. It is very easy to create asymptomatically flat spacetimes, for example, if a doughnut (a very heavy one) is placed in spacetime, it would create a very complicated stationary spacetime metric which is asymptotically flat. The black hole uniqueness theorem proves that we cannot create such a metric when we have only vacuum, it has to be either the Minkowski metric or the Schwarzschild one. Israel's theorem was the start of black hole uniqueness theorems. This theorem can also be framed in another way- If the black hole is static, then it must be spherically symmetric whose geometry can only be described by the Schwarzschild solution. What this would imply is that in the absence of angular momentum, the gravitational collapse of a star (whose mass is above a certain threshold) must result in a Schwarzschild black hole. This statement raises more questions than it answers. This statement would also imply that the star would collapse into a Schwarzschild black hole irrespective of its initial shape. What if the star, prior to its collapse, had a non-spherical shape? Richard Price, proposed in 1972 that any non-spherical protrusion must be radiated completely away as gravitational waves by an object collapsing to a black hole. This is known as *Price's theorem* . These protrusions are higher order multipole moments and are either radiated away, either out to infinity as gravitational waves or to the black hole. After all such radiations fade away, the black hole settles down in a spherical shape. An analogue to elucidate this theorem would be strumming a chord on a guitar. Say we strum the chord of $C\#$ (C sharp), its sound is heard as long as the strings vibrate. The vibrations create sound waves, which are Fourier transformed by our ears to be heard and the sound persists until the vibrations dampen off. Similar to that of a string, the horizon of a black hole vibrates, sending off gravitational waves. These gravitational waves carry away the energy of the horizon's deformation, and as the vibrations dampen, the horizon settles to a spherical shape. In this analogy, the role of our ears currently being played by the gravitational wave detectors such as LIGO, VIRGO, etc. These detectors act as our cosmic ears, performing multiple Fourier transforms that enables us to hear and differentiate the many violent astrophysical events.

4.4 The Structure of Isometry Groups of Asymptotically Flat Spacetimes

[2] A prerequisite for studying stationary spacetimes is the understanding of the structure of the isometry groups[3] which can arise, together with their actions. For the theorem that follows we do not assume anything about the nature of the Killing vectors or of the matter present; it is therefore convenient to use a notion of asymptotic flatness which uses at the outset 4-dimensional coordinates. A metric on Ω will be said to be asymptotically flat if there exist $\zeta > 0$ and $k \geq 0$ such that

$$|g_{\mu\nu} - \eta_{\mu\nu}| + r\,|\partial_\alpha g_{\mu\nu}| + ... + r^k\,|\partial_{\alpha_1}...\partial_{\alpha_k} g_{\mu\nu}| \leq Cr^{-\zeta}, \tag{4.16}$$

for some constant C ($\eta_{\mu\nu}$ here is the Minkowski metric). Ω will be called a *boost-type* domain if

$$\Omega = \{(t, \mathbf{x}) \in R \times R^3 : |\mathbf{x}| \geq R, |t| \leq \Xi r + C\}, \tag{4.17}$$

for some constants $\Xi > 0$ and $C \in R$. Let ϕ_t denote the flow of a Killing vector field X. $(\mathcal{M}, g_{\mu\nu})$ will be said to be stationary-rotating and if ϕ_t satisfies[4]

$$\phi_{2\pi}\left(x^\mu\right) = x^\mu + A^\mu + \mathcal{O}\left(r^{-\delta}\right), \quad \delta > 0 \tag{4.18}$$

in the asymptotically flat end, where A^μ is a timelike vector of Minkowski spacetime (in particular $A^\mu \neq 0$). We can think of $\partial/\partial\phi + a\partial/\partial t, a \neq 0$ as a model for the behavior involved.

4.4.1 Asymptotically Flat Stationary Metrics

A spacetime $(\mathcal{M}, g)$ will be said to possess an asymptotically flat end if M contains a spacelike hypersurface M_{ext} diffeomorphic to $\mathbb{R}^n$ $B(R)$, where $B(R)$ is a coordinate ball of radius R, with the following properties: there exists a constant $\zeta > 0$ such that, in local coordinates on M_{ext} obtained from $\mathbb{R}^n$ $B(R)$, the metric g induced by g on M_{ext}, and the extrinsic curvature tensor K of M_{ext}, satisfy the fall-off conditions, for some $k > 1$,

$$g_{\mu\nu} - \delta_{\mu\nu} = \mathcal{O}_k\left(r^{-\zeta}\right), \quad K_{\mu\nu} = \mathcal{O}_{k-1}\left(r^{-1-\zeta}\right), \tag{4.19}$$

[2] Necessary for the mathematically inclined, optional for the physics mind, and avoidable if you haven't done chapter 1

[3] see chapter 1

[4] also,if the matrix of partial derivatives of X_μ asymptotically approaches a rotation matrix in $\mathcal{S}_{ext}$

where we write $\mathcal{F} = \mathcal{O}_k\left(r^{-\zeta}\right)$ if $\mathcal{F}$ satisfies

$$\partial_{k_1}...\partial_{k_p}\mathcal{F} = \mathcal{O}\left(r^{\zeta-p}\right), \quad 0 \leq p \leq k. \tag{4.20}$$

For simplicity it is assumed that the spacetime is vacuum. Along any spacelike hypersurface $\mathcal{S}$, a Killing vector field X of $(\mathcal{M}, g)$ can be decomposed as follows

$$X = nN + Y, \tag{4.21}$$

where Y is tangent to $\mathcal{S}$, and n is the unit future-directed normal to M_{ext}. The fields N and Y are called *Killing initial data.* The vacuum field equations, together with the Killing equations imply the following set of equations on $\mathcal{S}$

$$\begin{aligned} &D_\mu K_\nu + D_\nu Y_\mu = 2NK_{\mu\nu}, \\ &R_{\mu\nu}(g) + K^k_k K_{\mu\nu} - 2K_{\mu k}K^k_\nu - N^{-1}\left(\mathcal{L}_Y K_{\mu\nu} + D_\mu D_\nu N\right) = 0, \end{aligned} \tag{4.22}$$

where $R_{\mu\nu}(g)$ is the Ricci tensor of g. The above equations are called Killing initial data equations. Under the boundary conditions, an analysis of these equations provides detailed information about the asymptotic behavior of (N, Y). In particular we can prove that if the asymptotic region $\mathcal{S}_{ext}$ is part of initial data set $(\mathcal{S}, g, K)$, we can then choose adapted coordinates so that the metric can be, locally, written as follows

$$g = -\mathcal{V}^2\left(dt + \underbrace{\phi_\mu dx^\mu}_{=\phi}\right)^2 + \underbrace{g_{\mu\nu}dx^\mu dx^\nu}_{=g}, \tag{4.23}$$

with

$$\begin{aligned} &\partial_t\mathcal{V} = \partial_t\phi = \partial_t g = 0 \\ &g_{\mu\nu} - \delta_{\mu\nu} = \mathcal{O}_k\left(r^{-\zeta}\right), \ \phi_\mu = \mathcal{O}_k\left(r^{-\zeta}\right), \ \mathcal{V} - 1 = \mathcal{O}_k\left(r^{-\zeta}\right). \end{aligned} \tag{4.24}$$

Methods known in principle show that, in this gauge, all metric functions have a full asymptotic expansion in terms of powers of $ln|r|$ and inverse powers of r. In the new coordinates we can take

$$\zeta = n - 2. \tag{4.25}$$

By inspection of the equations we can further infer that the leading order corrections in the metric can be written in the Schwarzschild form[5].

[5] this is dealt in greater depth in the next chapter

4.5 Spherical Collapse and Collapse with Small Non-Spherical Perturbations

So how does a realistic collapse look like? The entire process can be stated in one line: In a star, *instability* causes an *implosion* which leads to the creation of a *horizon* and a *singularity*.

When the star has exhausted all of its nuclear fuel, it starts contracting inwards at a slow pace. Eventually it begins to squeeze its electrons or photons, which sustain the pressure, onto its atomic nuclei; this "softens" the equation of state leading to an instability being induced. This sets up a chain reaction and in less than a second, the instability develops into a full-scale implosion. The star's surface falls through its gravitational radius, for an idealized spherical case, thus forming the horizon. From the star's interior frame, within a short proper time interval after passing through the horizon, a singularity is reached. The singularity is a point of zero radius which possesses infinite density and infinite tidal gravitational forces.

What would be the result if small non-spherical perturbations are introduced during the star's collapse? Richard H. Price performed calculations which suggested that during the collapse, all things that can be radiated away are completely radiated in part to "infinity" and in part "down the black hole", such that the final field is characterised by its conserved quantities. Let's study these non-spherical perturbations. Such a collapse would lead to perturbations in the star's density, angular momentum, and electromagnetic field. Let's analyse these perturbations individually.

4.5.1 Perturbations of density

When the star begins to collapse, it possesses a small non-spherical protrusion in its density distribution. This protrusion grows larger as collapse proceeds, and this growing protrusion radiates gravitational waves. This protrusion on the star remains on it as the star plunges via the horizon during its collapse, thus creating a deformed horizon. It is key to note that the radiated gravitation waves have two flavours, waves of short wavelength and waves of long wavelength. The former ($\lambda \ll M$), emitted from near the horizon ($r-2M \leq M$), are partly propagated to infinity and are partly backscattered by the background Schwarzschild curvature of spacetime. These backscattered waves propagate their way into the horizon formed during collapse of the star. The latter flavour, i.e., waves of long wavelength ($\lambda \gg M$) emitted from near horizon are fully backscattered by the curvature of spacetime. These waves, however, don't reach far (no further than $r \approx 3M$) and eventually end up propagating down the black hole. If Mr. Absolute Zero were examining the protrusion throughout the collapse of the star, he can never learn of the existence of the final protrusion, he can do so only by examining the deformation, i.e. the quadruple moment, in the final gravitational field. This final deformation in field propagates with the speed of light, in the form of gravitational waves. Thus, the final external field is perfectly spherical, protrusion-free and possesses a Schwarzschild geometry!

4.5.2 Perturbations in angular momentum

During the collapse process of a star, it possesses a small spin, or nonzero intrinsic momentum S. S is conserved throughout the collapse[6]. What kind of a geometric object is S? It is defined by measurements made at far distances from the source, where, with receding distance, spacetime is becoming asymptotically flat. Thus, asymptotic flatness is key to definability of S and also M (the mass of the star, because more the mass more the curvature and longer distance in spacetime for the metric to become asymptotically flat), and at far away distances from the source (, i.e. at weak curvature) the intrinsic angular momentum (and also the 4-momentum) will reveal themselves by their imprints on the spacetime geometry. This imprinted S cannot propagate outward from near horizon due to its conservation law. Hence, the final external field is that of an undeformed, slowly rotating black hole

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 - \left(\frac{4Ssin\theta}{r^2}\right) rsin\theta d\Phi dt. \tag{4.26}$$

The first part of the metric describes the usual Schwarzschild geometry we all are aware of, while the latter part describes the rotational imprint (polar axis oriented along S).

4.5.3 Perturbations in electromagnetic field

Since a star has an internal charge distribution, it possesses an electric field, it also possesses a magnetic field which are generated by currents in its interior, and its intensely hot matter emits electromagnetic radiation. The electric monopole moment of the star is conserved (other quantities vary), i.e., the number of Maxwell tubes is equal to the product of the solid angle (4π) and charge e. The total flux never changes and remains constant prior to the collapse, during the collapse, and into the black hole stage. The final external electromagnetic field is a spherically symmetric Coulomb field; and the final spacetime geometry is that of Reissner and Nordstrøm, given by

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 d\Omega^2. \tag{4.27}$$

A collapsing spherical star with an arbitrary non-spherical scalar charge distribution, generates an external scalar field Ψ, and its vacuum field equation

[6] A small, negligible change due to angular momentum is carried off by waves; that change is proportional to the square of the amplitude of the perturbations in star, i.e., to S^2

is $\partial^\alpha \partial_\alpha \Psi = \Psi^{\alpha}_{;\alpha} = 0$. An assumption that the spin field S is small and has a minuscule perturbation on the star's external, Schwarzschild geometry. Let's attribute an integer s to the fields (such that $s = 0$ is a scalar field; $s = 1$ is a vector field; etc.), and an integer l to the different poles (such that $l = 0$ for monopole; $l = 1$ for dipole; $l = 2$ for quadruple; etc.). During the collapse process, all multipole fields with $l < s$ are conserved. The scalar field Ψ $(s = 0)$ conserves nothing. From the above passages we know that the star has an electromagnetic field (vector field, thus, $s = 1$) that conserves only monopole parts, the gravitational field (tensor field, thus, $s = 2$) that conserves its monopole part (with imprint equal to mass), and its dipole parts (with imprints on the spacetime metric measuring angular momentum). Radiation is possible only for $l \geq s$, i.e. scalar waves can have any multipolarity, electromagnetic waves must be dipole and higher, and gravitational waves must be quadruple and higher. Price's theorem states that, for near spherical star collapses (to form black holes), all things that can be radiated, i.e. all the multipoles $l \geq s$, get radiated away completely in part to infinity and in part down the black hole. Thus, the final field is characterised completely by its conserved quantities, i.e. multipole moments with $l < s$.

4.6 Black holes get bald too

According to *Price's theorem*, all protrusions are radiated away completely. This provides us an explanation of the mechanism which makes black holes *hairless*. This act of black holes becoming bald is called as the *no hair theorem*. Let's try and understand this theorem via a thought experiment.

The horizon of the black hole is a region of infinite redshift. Position two observers, Mr. Absolute Zero and his brother Mr. Zero Entropy at coordinates r_{AZ} and r_{ZE}. We observe that Mr. Absolute Zero emits two signals towards his brother. There would be a coordinate time separation equal to δt between the emissions. Now, we assume that the metric is time independent so that the time separation between the arrival of the signals to Mr. Zero Entropy is also δt. We already know the relation between the proper time and time as: $d\tau = \sqrt{g_{00}} dt$. For δ changes, the proper time separations of the signals would have the following relation

$$\delta t = \frac{\delta \tau_{AZ}}{\sqrt{g_{00}(r_{AZ}}} = \frac{\delta \tau_{ZE}}{\sqrt{g_{00}(r_{ZE}}}. \tag{4.28}$$

Since frequency is inversely related to clock rate, we can approximately say the following is true

$$\omega_{ZE} = \omega_{AZ} \frac{\sqrt{g_{00}(r_{AZ})}}{\sqrt{g_{00}(r_{ZE})}}. \tag{4.29}$$

From the above equation we observe that when the position of Mr. Absolute

Zero is slowly changed to the Schwarzschild radius, i.e., as $r_{AZ} \rightarrow r_S$, the frequency of the signal received by his brother tends to zero, i.e., $\omega_{ZE} \rightarrow 0$. This implies that the signal Mr. Zero Entropy receives is infinitely redshifted. This behaviour is independent of the where Mr. Zero Entropy is positioned (r_{ZE}). Hence, if Mr. Absolute Zero emits black body radiation behind the horizon of the black hole, his brother would only see a stationary field due to the infinite redshift. Now apply the same analogy to that of a collapsing star. When the source of all the gravitational radiations (the protrusions) approaches the Schwarzschild radius, the radiation would experience a greater magnitude of redshift. Hence, observers, such as ourselves, would observe stationary fields independent of the initial of the frequency of the initial radiations. During the collapse, all that can be radiated away is radiated away so that distant observes detect only stationary fields. But wait! I had previously stated that the protrusions give rise to multipole moments that are emitted away as gravitational waves, but what moments are capable of providing stationary fields (since they are what we detect)? To have gravitational waves, we need a time rate of change of multipole momenta (usually quadruple moment). In order to have electromagnetic radiations, there needs to exist a time rate of change of dipole moment. Similarly, only the monopole electric moment (charge), the monopole gravitational moment (mass) and dipole gravitational moment (angular momenta) are able enough to provide exclusive stationary fields, hence we can conclude that they are the ones that must remain after the collapse process has ended. This implication of the above lines is that a black hole, in the presence of some electromagnetic fields, can possess only three parameters-coulombic charge, mass, and angular momentum. The black hole cannot carry any "scalar hairs". This is the no hair theorem. If you still crave for a simplified explanation, we can say that the no hair theorem explains as to why a black hole wants to be spherical with no strings attached!

As stated earlier, in accordance to Price's theorem, the influences of the black hole's mass, angular momentum (spin), and charge are the only things that remain after the end of a gravitational collapse. All of the other higher momenta would be carried away as radiations. What this means is that we cannot perform any experiment in which measurements of the final three parameters of the hole would reveal any features of the star it imploded to form it, except for the star's mass, charge, and spin. The black hole makes the intricate details of its humble origins anonymous. However, the firm and ultimate proof that a black hole has no hair (except for the scalar hairs) was not given by Price. Prince's theorem is restricted to star implosions that had very small deviations from spherical symmetry and whose spin was very slow. The proof determining the fate of a rapidly spinning, highly deformed, imploding star was done by Brandon Carter, Stephen Hawking and Werner Israel, but that's a topic for another day.

4.7 That mass though...

Electromagnetic radiation cannot escape a black hole since it travels at the speed of light. Similarly, gravitational radiation cannot escape a black hole either, because it too travels at the speed of light. A black hole, however can have an electric charge, which means there is an electric field around it (as described in the paragraphs above), there is nothing to be alarmed at here due to the fact that a static electric field is not the same as electromagnetic radiation. Similarly, black holes have mass, so they have a gravitational field around them, there is nothing to be shocked at here too because a gravitational field is different from gravitational radiation. The fundamental difference between a standard gravitational field and a gravitational wave is quite subtle-the latter consists of propagating ripples which carry energy in the form of gravitational radiation. With this said, let's ask ourselves an interesting question-gravitational attraction that we would feel next to a black hole carries information about the amount of mass within the black hole. We know that no information can escape from a black hole, but then how do we reason as to why the information of the black hole's mass escapes? The answer is that the information doesn't have to escape from inside of the horizon, because the information about its mass is not inside but on the horizon. This statement raises more questions than it answers one and the main question is- then where is the mass of the black hole located (or stored)? Yes, the information is on the horizon, but where's the mass? Before you shout out, "at the singularity!", allow me to remind you that singularity is not a point but rather a hole in the topology of a manifold[7].

4.7.1 Mass in General Relativity

Now, let's answer the question. I did state that singularities are not points on the topological manifold because if they were points, then they could well be explained using Euclidean geometry by representing their reference frame using a coordinate chart in the Cartesian coordinates of (t, x, y, z), Singularities are something much more complicated- they are holes in topological manifold, more elegantly, they are similar to having punctures in the tyres of our car. Not all terrains contain rocks and materials which are sharp and heavy enough to poke holes in our tyres (our tyres were assumed to be very elastic remember?), only a few can. Analysing all the rocks (stars) we find that only a few cause punctures and all of these can be differentiated from the rest using a mass limit. Hence, the ones which puncture our manifold are black-holes (or rather are stars who are heavy enough that their collapse yields a black-hole), whose information is not stored in the threads but rather stored on boundaries of the punctures (information is on the event horizons of the black holes). So where is the mass? Well, since you know the actual theory now, you must ask yourself,

[7] understand what this statement means before answering the question, via the simple analogy presented in the first chapter. See Manifolds: A Pedestrian Approach

"how can mass be stored in punctures?", well that can't be right! Punctures are mere holes on the tyres and assigning mass to holes makes no sense. The answer is that we don't know, this is one of the many puzzling questions to which we are still to find answers but certainly will one day.

Mass in General Relativity has multiple definitions and can be confusing. It is almost impossible to find a generalized definition for a system's total mass, for the reason observe Einstein's field equations, the gravitational field energy is not a part of the energy-momentum tensor. The energy-momentum tensor, $T_{\mu\nu}$ represents the energy due to matter and electromagnetic fields, but this does not include any contribution from the gravitational energy. Now we find ourselves in a catch 22 situation because we can argue that gravitational energy does not act as the source of gravity, but since Einstein's equations are non-linear, this would imply that gravitational waves interact with each other and hence, we can argue that gravitational waves interact with each other and hence, we can argue that gravitational energy is a source of gravity. However, it is possible to define mass for a stationary spacetime. This is called the *Komar mass*. We shall not go into the details but will try and understand this using our analogy. Let us consider the puncture in the topological manifold created by a specific spherically symmetric rock. We here make a crucial assumption that all the rocks on all the terrains are spherically symmetric. Hence, on application of Israel's theorem we find that the only metric that can explain the rock's (star's) geometry is the Schwarzschild metric (of course there is a possibility that this spherically symmetric metric need not resemble the Schwarzschild one, but one can always use coordinate transformations to achieve a Schwarzschild-like look). Since the Schwarzschild metric is a stationary one, it satisfies the condition for using Komar mass. We can find with ease, by setting $t = const, d\Phi = 0$, that the radial acceleration that is required to hold a test mass stationary at a Schwarzschild coordinate of r is-$a = \frac{m}{r^2\sqrt{1-\frac{2GM}{rc^2}}} = \frac{m}{r^2\sqrt{g_{tt}}}$.

Now, the motion of the test mass as the tyre rotates and translates through terrains, is given by the geodesic equation. Hence, we can see that the acceleration, in covariant form can be given as-$\dot{x}^\mu = -\Gamma^\mu_{\alpha\beta}u^\alpha u^\beta$, and the Christoffel symbol has the dimensions of potential, i.e., $\left[\Gamma^\mu_{\alpha\beta}\right] = \frac{1}{r} \approx \Phi(r)$. Now, the potential can be expressed as the negative integral of force, i.e., $\Phi(r) = -\int F dr$, and finally since the force has dimensions of inverse length squared which is also the dimension of the Riemann tensor (also the Ricci tensor since the metric tensor is dimensionless) we can arrive at the conclusion that the mass is proportional to the following integral- $m \propto \int_V R_{\mu\nu}u^\mu u^\nu$. Furthermore, we can find that $R_{\mu\nu} = -\frac{8\pi G}{c^4}T$ by setting $\mu = \nu$. Thus, the mass integral is-$m \propto \int_V \left(2T_{\mu\nu} - Tg_{\mu\nu}\right)u^\mu u^\nu dV$. If we had done some hard work and derived the equation, it would have the following form

$$m = \int_V \sqrt{g_{tt}} \left(2T_{\mu\nu} - Tg_{\mu\nu}\right) u^\mu u^\nu dV \tag{4.30}$$

this is the Komar mass integral. The spacetime in which this is defined is a stationary one, i.e., it exhibits time translational symmetry. There is another class of situations referred to as asymptotically flat, in which we can study the mathematical model of mass carried away from massive gravitating systems in the form of gravitational waves. Similar to Komar mass, we can define (asymptotically flat class) the *ADM* and the *Bondi* masses. For an isolated gravitating system, whose spacetime is asymptotically flat and which emits gravitational waves, there is a precise amount of total mass/energy-momentum and of its loss through gravitational radiation, referred to as the *Bondi-Sachs conservation law.*

The deeper we divulge into relativity, the more we discover that we haven't completely understood anything. You may go on to ask that if there is such a huge confusion of the mass of a black hole, how are we to ever find the mass of the universe. Does the universe have a mass? Well we don't completely know the extent or the 'edge' of the universe to confine it and find the mass... so we can't ask that question. Then what is the mass of the observable universe? Well, since the observable universe cannot be considered to be an isolated system, it is neither asymptotically flat nor stationary and hence none of the 'definitions' of mass in General Relativity apply to it. Now, if we considered the universe to be closed, then would the closed universe have a mass? Well, I'm sorry to disappoint you yet again because the answer is a huge no. As John Wheeler put it, "There is no such thing as the energy (or angular momentum or charge) of a closed universe, according to general relativity, and this for a simple reason. To weigh something, one needs a platform on which to stand to do the weighing..."

4.8 Elimination of Singularities

We already know by now that there is only one physical singularity and that is when $r \to 0$. As we have already observed, to see to that the above case is a true singularity we are to turn to quantities that are independent of the choice of coordinates. We chose the Kretschmann invariant and note that it is indeed regular at $r = r_s$. Another method to see that the spacetime described by the Schwarzschild metric is regular at $r = r_s$ is to make such a coordinate transformation to such a metric tensor which is regular at this surface so that it would eliminate this unphysical singularity. The Schwarzschild metric can be re-written as

$$ds^2 = \left(1 - \frac{r_s}{r}\right)\left[dt^2 - \left(1 - \frac{r_s}{r}\right)^2 dr^2\right] - r^2 d\Omega^2 \tag{4.31}$$

Our aim is to change the radial coordinate which would map the horizon to negative infinity such that the resulting coordinate system covers the region $r > \frac{2MG}{c^2}$, exclusively. Hence, we define new coordinates called the *tortoise coordinates.*

$$r_* = r + r_s ln\left|\frac{r}{r_s} - 1\right|, dr_* = \left(1 - \frac{r_s}{r}\right)^{-1} dr \tag{4.32}$$

Plugging this into the re-written form of the Schwarzschild metric, we obtain a new metric which has the following form

$$ds^2 = \left(1 - \frac{r_s}{r}\right)\left[dt^2 - dr_*^2\right] - r^2 d\Omega^2 \tag{4.33}$$

Notice that the radial-time part of the metric takes up a form called conformally flat. Any space can be called conformally flat if it can be brought to the following form

$$ds^2 = H(x) dx^\mu dx^\nu \eta_{\mu\nu} \tag{4.34}$$

Where $\eta_{\mu\nu}$ is the Minkowski metric. When we fix θ and Φ, the radial-time, two-dimensional space is conformally flat. It is important to note that in these coordinates, as $r \to \infty$, $r_* \to \infty$ and as $r \to r_s$, i.e. at the horizon, $r_* \to -\infty$. This is a good coordinate system but not the best, hence, let's try and find a better system which would cover the entire Schwarzschild spacetime completely. We have to discover coordinates for performing the mentioned task, but where do we start? This is the wrong question to ask. The question we must ask ourselves is how do we start this? Firstly, we need coordinates that would cover the entire spacetime manifold. To do this we need to observe the entire spacetime manifold, there is a desperate need for "the perfect view". On obtaining such a view, we can go ahead and try and measure its boundaries and try and represent the spacetime in terms of a cooked-up but correct coordinates. Yeah, I know what you're asking, to view the entire manifold we need to observe the infinities, how do we even imagine such a think? Hold the thought, we will come back to it. Although this manifold trip might sound like an enjoyable one, to get this view we need to trek it using math. The mathematically picture is much more elegant. We will make use of our car analogy and try and make sense of the math by drawing relations.
Consider the same car which represents our entire spacetime manifold with four tyres that record all the information of the terrains it has travelled upon in its threads, which are infinite in number. Now remember that the deformations on the tyres are created due to the stones present in the terrain and that not all puncture the tyres since a specific mass and sharpness is required by the stones to poke a hole in the tyre. When the tyres get punctured, nothing happens to the car, it's perfectly fine and resumes its travels with a specific speed. Before your imaginations run wild let me intrude by stating that you are the driver and not anything else! Now think of this, your car need not

be the only one in town, there are multiple cars, few similar to yours and a few not. Now you want to fix the punctures but your economic condition prevent you from doing so. You now start wondering how the costlier, high-end, luxury cars manage their tyres so well and hence observe the behaviour of a puncture in the tyre of the costlier car. Since I trust you (because well you are a future physicist after all!), I hand you my super expensive luxury car. You try to simulate a puncture by driving over terrains you had visited earlier but fail to obtain a puncture of similar geometry. A brilliant idea strikes you and you cut out a patch of your car's tyre and a paste it to my car's tyre, where you had cut out a hole of similar dimensions. You paste the patch using a special adhesive. You successfully simulate a puncture and observe the details of the hole created. The puncture is more pronounced now because its geometry is not in any way influenced by your car's tyre material (which is more prone to puncture than mine). You realize that you can now study and access the information of the terrain (which is stored in the threads and the punctures) in greater detail when parts of your tyre are glued to mine. You immediately replace your tyres with mine and glue patches of your tyre and simulate exactly similar type of punctures and observe them in greater detail. This concludes our analogy, now let's view this mathematically. As mentioned previously the punctures are the topological holes in the spacetime manifold, i.e., black holes. Now recollect that your car is a 4-D Lorentzian manifold and mine being a high-end luxury car represents a manifold of a different dimension. The detailed observation of the punctures was possible only when the tyre material was glued to mine, this mathematically translates to us embedding the spacetimes in higher dimensions. The special adhesive that we used is called a hyperplane in math-land. In conclusion, we can summarize our thought experiment by stating that to find better coordinates that cover the entire spacetime (of Minkowskian flavour), you had to embed the spacetime as a hyperplane into a higher dimensional Minkowski space in order to find coordinates that cover the hyperplane completely, and hence, cover the entire original curved spacetime. We know very well that at every point of our 4-D spacetime, its metric, being a symmetric 2-tensor would have 10 independent components (Since the metric tensor has 4 diagonal components and 6 other symmetrical components to both the left and right sides of the diagonal. Total number of independent components $= 6 + 4 = 10$). From this we can subtract 4 degrees of freedom in accordance to the four coordinate transformations, $\bar{x}^{\mu}(x)$, resulting in 6 independent degrees of freedom at each point. Hence, we have just embedded a 4-D spacetime locally as a hyperplane into a $(4 + 6)$ dimensional Minkowski spacetime. Or simply put, you have glued your 4-D car's tyre onto my $(6 + 4)$-D car's tyre using a special adhesive in order to understand better the geometry of the puncture and cook-up a coordinate a system which would cover the entire boundary of the adhesive.

It turns out that the curved spacetime has extra, available symmetries, then one can embed it into a flat space of a dimension less than ten. The Schwarzschild spacetime (which is symmetric) can be embedded into a 6-D flat space. This embedding is done with using the *Kruskal-Szekeres coordinates*. Before proceeding to this, let's reflect on the above few lines. There won't be any gravity in the 6-D spacetime because the metric is globally Minkowskian, i.e., $R_{\mu\nu\alpha\beta} = 0$ everywhere, identically. In order to understand the deep implications and the overall elegance in this method let us try to show as to why it is possible to embed the Schwarzschild metric in higher dimensions. To check this, we must embed the metric in a 6-D flat spacetime and prove that the embedded metric reproduces the geometry of the 4-D Schwarzschild metric. Hence, thinking of the Schwarzschild metric as an induced metric on a four-dimensional hyperplane embedded in a flat six-dimensional spacetime with the line element given as

$$ds_6^2 = +d\Psi_1^2 - d\Psi_2^2 - d\Psi_3^2 - d\Psi_4^2 - d\Psi_5^2 - d\Psi_6^2 \tag{4.35}$$

where,

$$\begin{aligned}
\Psi_1 &= 4GM\sqrt{1 - \tfrac{2GM}{r}} sinh\left(\tfrac{t}{4GM}\right) \\
\Psi_2 &= 4GM\sqrt{1 - \tfrac{2GM}{r}} cosh\left(\tfrac{t}{4GM}\right) \\
\Psi_3 &= \pm\int \left[\tfrac{2GM}{r} + \left(\tfrac{2GM}{r}\right)^2 + \left(\tfrac{2GM}{r}\right)^3 + O\left[\left(\tfrac{2GM}{r}\right)^4\right)\right] \\
\Psi_4 &= r sin\theta cos\Psi, \quad \Psi_5 = r sin\theta sin\Psi, \quad \Psi_6 = r cos\theta
\end{aligned} \tag{4.36}$$

By taking the derivatives (which is left for you to compute), we would end up with a metric. The crucial part of the proof is to observe the fate of the Ψ_3 term. Notice that this term is nothing but the Taylor expansion of $\left(1 - \frac{2GM}{r}\right)^{-1}$ at $x = 0$, where $x = \frac{2GM}{r}$. Finally, we obtain and observe the following

$$ds_6^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 - r^2 d\Omega^2 = ds_4^2 \tag{4.37}$$

Thus, we have just shown that the metric of the embedded metric in 6-D flat space is nothing but the 4-D Schwarzschild metric. This type of embedding is known as *Fronsdal embedding.* Now arises the question as to how we have embedded it. The embedding had to be performed in such a way that the metric reproduces itself and as mentioned earlier this is done using the Kruskal-Szekeres coordinates. Embedding diagrams provide an effective pathway for visualizing the geometry of a spherically symmetric metric such as the Schwarzschild metric. We shall first study the details of such a diagram before understanding and divulging into the math surrounding the Kruskal coordinates, by doing this you would be able to appreciate the ingenious coordinate transformations and obtain a holistic understanding.

Exercise 13

1. The Reissner-Nordstrøm metric is a solution to Einstein's field equations that describes the spacetime around a spherically symmetric non-rotating body with mass M and an electric charge Q. The line element takes the following form

$$ds^2 = -\left(1 - \frac{r_s}{r} + \frac{r_Q^2}{r^2}\right) dt^2 + \left(1 - \frac{r_s}{r} + \frac{r_Q^2}{r^2}\right) dr^2 + r^2 d\Omega_2^2, \tag{4.38}$$

where $r_Q^2 \equiv \frac{GQ^2}{4\pi\epsilon c^4}$ and $r_s = \frac{2GM}{c^2}$.
a. Show that the Ricci scalar vanishes identically.
b. Show that the Kretschmann scalar is $K = \frac{4}{r^8}\left(14r_Q^4 - 12r_Q^2 r r_s + 3r^2 r_s^2\right)$.
c. Show that the metric is only singular at r if $Q > M$.
d. Another singular point is at $r = r_\pm$ for the case $Q < M$. Show that $r_\pm = M \pm \sqrt{M^2 - Q^2}$ with $G \equiv c \equiv 1$ and $\frac{1}{4\pi\epsilon} \equiv 1$.
e. Define the tortoise coordinate r_* by

$$dr_* = \left(1 - \frac{r_s}{r} + \frac{r_Q^2}{r^2}\right)^{-1} dr,$$

and solve to obtain an expression for r_*.

2. In the Reissner-Nordstrøm metric 4.38, setting constants $G \equiv c \equiv k \equiv 1$ (where $k = \frac{1}{4\pi\epsilon}$ consider the special case that the charge and mass are equal, $Q = M$. This special case describes the extremal Reissner-Nordstrøm geometry.
a. What happens to the inner and outer horizons r_+ and r_-?
b. Show that the Ricci scalar is $R = \frac{2M^2(M-2r)^2}{(M-r^4)r^2}$.
c. Consider the case with an imaginary charge $Q^2 < 0$. Although this is unphysical, the resulting metric is well-defined. Check if the singularity, instead of being gravitationally repulsive, becomes gravitationally attractive (this involves some tiresome calculations).

3. In AdS_3 space, there is $SL(2,\mathbb{R})_L \times SL(2,\mathbb{R})_R$ symmetry. The generators of $SL(2,\mathbb{R})_L$ read

$$L_0 = i\partial_u,\ L_{\pm 1} = ie^{\pm iu}\left(coth(2\psi)\ \partial_u - \frac{1}{sinh2\psi}\ \partial_v \mp \frac{i}{2}\ \partial_\psi\right), \tag{4.39}$$

where $u = t + \phi$ and $v = t - \phi$. $SL(2,\mathbb{R})_R$ is generated by $\bar{L}_{0,\pm 1}$ which can be obtained by simply interchanging u and v.
a. Rewrite the global AdS_3 metric in terms of coordinates $\hat{\phi} = lcosh\psi - it$, $\gamma = tanh\psi\ e^{iu}$, and $\hat{\gamma} = tanh\psi\ e^{iv}$.
b. Rewrite the $SL(2,\mathbb{R})_L$ and $SL(2,\mathbb{R})_R$ generators in terms of the new coordinates.

c. Check that the following forms a closed algebra

$$[L_0, L_{\pm 1}] = \mp L_{\pm 1}, \; [L_1, L_{-1}] = 2L_0. \tag{4.40}$$

d. Find the Killing vectors for the AdS_3 metric (either in global coordinates as in 1.228 or in Poincaré coordinates as in 1.199) and check if the Killing vectors obey the $SL(2, \mathbb{R})$ algebra shown above.

5

Embedding Diagrams and Extensions of the Schwarzschild Metric

5.1 Embedding Diagrams and Their Machinery

We shall first address the question that you would have asked in the previous chapter, i.e., to view the entire manifold we need to observe the infinities, how do we even imagine such a thing? How do we view something from the outside if we don't know where it ends? After reading the previous section you would propose embedding as an approach to such a problem, which is correct. Embedding forms the crux of the idea of an extrinsic view of topology. This is true simply due to the reason that we cannot view something from the outside unless it can be confined to some larger, or higher-dimensional space. Suppose there was a population of intelligent creatures living on our car (topological manifold), and if they wanted to observe and study the geometry of the punctures (topological punctures, i.e., black holes) the only way they could achieve such a task is to cut portions of the tyres and glue it to the tyre of a different car (my car, a higher-dimensional topological manifold) that they can observe. It is impossible for them to probe the geometrical dimensions of the car they live in similar to how it is impossible for us to find the edge of the universe. Hence, embedding refers to how a topological object such as a manifold is positioned in space. Taking a time-slice of the Schwarzschild metric, i.e., $t = const$, the two-dimensional geometry of the $\theta = \pi/2$ surface is given by the following metric (metric signature followed is $(- + ++)$)

$$ds^2 = \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Phi^2 = \frac{dr^2}{\zeta(r)} + r^2 d\Phi^2 \tag{5.1}$$

Our aim is to visualize the geometry by building a two-dimensional surface embedded in a three-dimensional flat space. It is important to note that when $M(r) = 0$, then the function $\zeta(r) = 1$ and the metric assumes a flat geometry. This can be thought of as a flat plane, $z = 0$, in cylindrical coordinates. Of course, the metric describes the geometry of a gravitating object of non-zero mass, i.e., $M(r) \neq 0$. Hence, we can assume the metric to be describing a surface of revolution given be $z = z(r)$, with the Φ component of the cylindrical coordinates having the range $0 < \Phi \leq 2\pi$. The metric of such a surface embedded in three dimensions is described by the following metric

$$ds^2 = dz^2 + dr^2 + r^2 d\Omega^2 = \left[\left(\frac{dz}{dr}\right)^2 + 1\right] dr^2 + r^2 d\Omega^2. \tag{5.2}$$

On comparing the above metric to the Schwarzschild time slice, we observe that the surface equation is related to the function $\zeta(r)$ by the following relation

$$\left(\frac{dz}{dr}\right)^2 + 1 = \zeta(r)^{-1}. \tag{5.3}$$

Notice that $\frac{dz}{dr}$ is continuous, we will use this later for obtaining the embedding diagram. We can now integrate this equation to obtain the expression that describes the surface of revolution.

$$z(r) = \int_0^r \sqrt{\frac{1-\zeta(x)}{\zeta(x)}} dx = \int_0^r \left[\frac{x}{2GM(r)} - 1\right]^{-\frac{1}{2}} dx. \tag{5.4}$$

For $r > r_s$, i.e., outside the gravitating object we observe that the above integral takes the following form

$$z(r) = \sqrt{8GM(r - 2GM)} + C \tag{5.5}$$

We can observe that at a large radius, $z(r) \propto r^{\frac{1}{2}}$. Near the centre of the star, we can approximate $M(r)$ as $\frac{4\pi}{3}\rho r^3$ and integrate the equation after introducing a variable $\xi = \sqrt{\frac{3c^2}{8\pi G\rho}}$. Upon integration, we obtain an equation which hints that the surface is a segment of sphere of radius ξ near the centre of the star, given by

$$(\xi - z(r))^2 + r^2 = \xi^2. \tag{5.6}$$

The above equation is valid for $r \gg \xi$ and this result is exact for any star of a constant density and is approximately correct near the origin of any other relativistic gravitational object model. Upon stitching the geometry represented by the above equation with that represented by the equation for $r > R$, we obtain an embedding diagram that represents the curvature around a spherically symmetric gravitational source. It is important to note that z and r are monotonically increasing functions of each other. This implies that the embedded surface always opens upward and outward like a bowl. The geometry never has a neck and it never flattens out except asymptotically at $r = \infty$. Remember that $\frac{dz}{dr}$ was continuous? This hints that the interior and exterior geometries in the embedding diagram will join smoothly. For any relativistic density and pressure, the geometry would be similar.The time-slice of the Schwarzschild metric is nothing but the quartic surface defined by the equation

$$x^2 + y^2 = \left(\frac{z^2}{8M} + 2M\right)^2, \tag{5.7}$$

embedded in a three-dimensional Euclidean space with cartesian coordinates (x, y, z). To see this, consider the following coordinate transformations: $x = rcos\Phi, y = rsin\Phi$, and $z = \sqrt{(8M(r-2M))}$. Applying these transformations to the three-dimensional Euclidean metric: $ds^2 = dx^2 + dy^2 + dz^2$, which I leave for you as an exercise, you would get back the time-sliced Schwarzschild metric. To analyse the geometry of the embedding, we can perform two parametric plots, one with $(rcos\Phi, rsin\Phi, \sqrt{(8(r-2))}$, and another with $(rcos\Phi, rsin\Phi, -\sqrt{(8(r-2))}$ shown in figure 5.1. Both have similar ranges with $r \in [0, 10]$ and $\Phi \in [0, 2\pi]$ (note that mass M has been set to unity for simplicity).

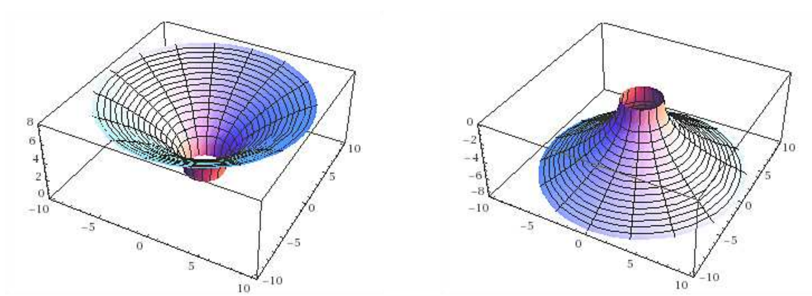


Fig. 5.1. Parametric plots of the upper half (L) and the lower half (R)

While gluing these diagrams together, observe that the two halves have been attached together at the circle $r = 2M = r_s$. The two halves correspond to two separate regions. This is known as the *Einstein-Rosen bridge*, which is one among the many examples of a *wormhole.* No observer can ever cross this bridge, this can be clearly verified from the *Kruskal diagram* (which is in the subsequent sections). If you wish to cross the wormhole from the upper half to the lower half or vice-versa, then your trajectory must be spacelike somewhere.

5.2 Embedding in N-Dimensions

Let's represent the Schwarzschild metric in $(n+1)$ dimensions with a $(-+++)$ metric signature,

$$ds^2 = -\left(1 - \frac{r_s}{r^{n-2}}\right) dt^2 + \left(1 - \frac{r_s}{r^{n-2}}\right)^{-1} dr^2 + r^2 d\Omega^2. \tag{5.8}$$

Note that here $d\Omega^2$ is the round unit metric on a sphere S^{n-1}. Following the embedding machinery as described in the previous section, we shall embed this metric (it's time slice) in a $(n+1)$-dimensional Euclidean space as follows

$$ds^2 = dz^2 + (dx^1)^2 + (dx^2)^2 + ... + (dx^n)^2 = \left[\left(\tfrac{dz}{dr} + 1\right)\right]^2 dr^2 + r^2 d\Omega^2. \tag{5.9}$$

This equation will coincide with the spacelike component of the time-sliced $(n+1)$-dimensional Schwarzschild metric. Thus, a smooth function is obtained.

$$\frac{dz}{dr} = \pm\sqrt{\frac{r_s}{r^{n-2} - r_s}}. \tag{5.10}$$

Note that this function can be explicitly integrated in 3 and 4 dimensions. Let us analyse the cases separately.

5.2.1 $(n = 3)$-dimensions

The smooth function takes the following form in $(n = 3)$ dimensions

$$\frac{dz}{dr} = \pm\sqrt{\frac{r_s}{r - r_s}} \tag{5.11}$$

Integration yields the following result

$$z(r) = \pm 2r_s\left(\sqrt{\frac{r_s}{r - r_s}}\right)^{-1} + C = \pm 2\sqrt{2m}\sqrt{r - 2m} + C, \tag{5.12}$$

where $r_s = 2m$. You know the reason for the positive sign- it represents the exterior of a black hole. What about the negative sign? It actually corresponds to the other side of the Einstein-Rosen bridge, which is an asymptotically flat region. Solving for $r(z)$ (and setting the integration constant) we obtain

$$r = 2m + \frac{z^2}{8m}. \tag{5.13}$$

The geometry of this is that of a paraboloid. Since this was first denoted by Flamm, its called the *Flamm paraboloid*. This embedding in $(n = 3)$ is visualized in figure 5.2.

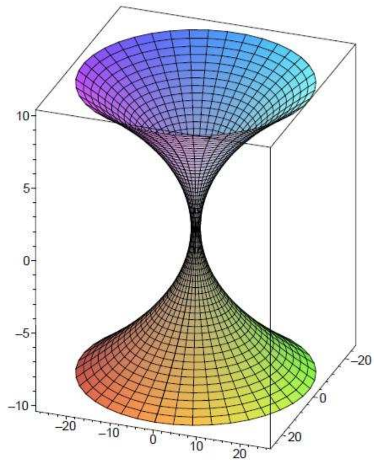


Fig. 5.2. Embedding of an $n = 3$ dimensional Schwarzschild black hole into a four-dimensional Euclidean space. The two halves of the black hole are glued by the Einstein-Rosen bridge $r = 2m$

5.2.2 $(n = 4)$-dimensions

The smooth function takes the following form in $(n = 4)$ dimensions

$$\frac{dz}{dr} = \pm\sqrt{\frac{r_s}{r^2 - r_s}}. \tag{5.14}$$

Integration yields the following result

$$z(r) = \pm\sqrt{r_s}ln\left(\sqrt{r^2 - r_s} + r\right) + C = \pm\sqrt{2m}\; ln\left(\sqrt{r^2 - 2m} + r\right) + C. \tag{5.15}$$

Solving for $r(z)$ (and setting the integration constant) we obtain

$$r = \sqrt{2m}\; cosh\left(\frac{z}{2m}\right). \tag{5.16}$$

This embedding in $(n = 4)$ is visualized in figure 5.3.

5.2.3 $(n \geq 5)$-dimensions

Things get a bit complicated in $n \geq 5$ dimensions as $z(r)$ is expressed in terms of elliptic functions. The qualitative behaviour in this dimension, or for that matter in $n \geq 5$ dimensions is quite different, here $z(r)$ asymptotically diminishes to a finite value as $r \to \infty$. The smooth function takes the following form

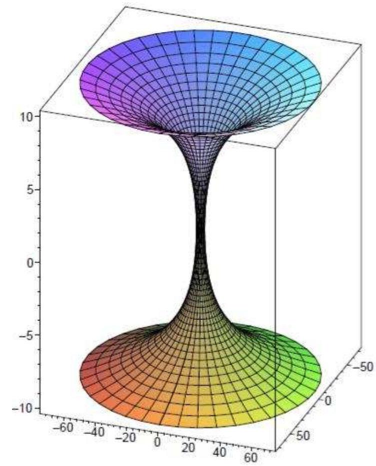


Fig. 5.3. Embedding of an $n = 4$ dimensional Schwarzschild black hole into a five-dimensional Euclidean space. The two halves of the black hole are glued by the Einstein-Rosen bridge $r = \sqrt{2m}$

in $(n = 5)$ dimensions

$$\frac{dz}{dr} = \pm\sqrt{\frac{r_s}{r^3 - r_s}}. \tag{5.17}$$

Integration yields the following result (to simplify calculations let us make the assumption that $m = 1$)

$$z(r) = \pm\frac{2i}{3^{\frac{1}{4}}}\sqrt{(-1)^{\frac{5}{6}}(r-1)\frac{r^2+r+1}{r^3-1}}F\left(sin^{-1}\left(\frac{\sqrt{-ir-(-1)^{\frac{5}{6}}}}{3^{\frac{5}{6}}}\right)\right)(-1)^{\frac{1}{3}}, \tag{5.18}$$

where $F(x|m)$ is the elliptic integral of the first kind with parameter $m = k^2$. The form of the function in $(n = 6)$ dimensions is the following

$$\frac{dz}{dr} = \pm\sqrt{\frac{r_s}{r^4 - r_s}}. \tag{5.19}$$

Integration yields the following result

$$z(r) = \pm F\left(sin^{-1}(r)| - 1\right) + C. \tag{5.20}$$

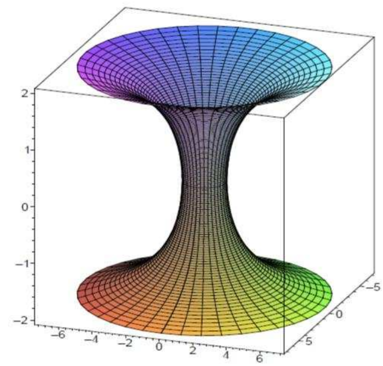


Fig. 5.4. The glued plot which is actually the isometric embedding of the space-geometry of a $(5+1)-$ dimensional Schwarzschild black hole into six-dimensional Euclidean space.

5.3 Embedding of the Schwarzschild Metric in Six-Dimensional Space

As we mentioned in the earlier, the minimal dimension N of the flat space in which the Schwarzschild metric can be embedded is equal to six; let us revisit the embedding but first lets rewrite the line interval of the Schwarzschild metric ($G \equiv c \equiv 1$), we have (metric signature followed is $(-+++)$)

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{5.21}$$

where $r_s(= 2m)$ is the Schwarzschild radius. In the following sections, we discuss few variants of embeddings.

5.3.1 Kasner Embedding

$$ds^2 = d\Psi_1^2 + d\Psi_2^2 - d\Psi_3^2 - d\Psi_4^2 - d\Psi_5^2 - d\Psi_6^2$$

$$\Psi_1 = \left(1 - \tfrac{r_s}{r}\right)^{1/2} cos(t),\ \Psi_2 = \left(1 - \tfrac{r_s}{r}\right)^{1/2} sin(t),\ \Psi_3 = f(r), \tag{5.22}$$

$$\Psi_4 = r sin\theta sin\Phi,\ \Psi_5 = r sin\theta cos\Phi,\ \Psi_6 = r cos\theta$$

Where $\left(f(r)'\right)^2 = \frac{2mr^3+m^2}{r^3(r-2m)}$. Historically, this was the first Schwarzschild metric embedding. This covers only the region $r > r_s$ and has a conical singularity at $r = r_s$. This embedding, however, is not asymptotically flat. Let's try and work this out. After computing the derivatives and their squares (which I leave to you as an exercise), you should land up with the following metric

(with $r_s = 2m$)

$$ds^2 = -\left(1 - \frac{2m}{r}\right) dt^2 + \left(1 - \frac{(m-1)m}{r^4}\right)\left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \quad (5.23)$$

Our aim is to visualize the geometry by building a two-dimensional surface embedded in a three-dimensional flat space. Let's consider the time-slice of this metric at $\theta = \frac{\pi}{2}$, and proceed with the embedding operations.

$$\begin{aligned} ds^2 &= \left(1 - \tfrac{(m-1)m}{r^4}\right)\left(1 - \tfrac{2m}{r}\right)^{-1} dr^2 + r^2 d\Phi^2 = \tfrac{dr^2}{\zeta(r)} + r^2 d\Phi^2 \\ &= \left[\left(\tfrac{dz}{dr}\right)^2 + 1\right] dr^2 + r^2 d\Omega^2 \end{aligned} \quad (5.24)$$

We can now integrate to obtain an expression for $z(r)$ as follows

$$z(r)|_{m\equiv 1} = \int_0^r \sqrt{\frac{1-\zeta(x)}{\zeta(x)}} dx = \sqrt{2} r \sqrt{\frac{1}{r} - 1} - \frac{1}{\sqrt{2}} tan^{-1}\left(\frac{\sqrt{\frac{1}{r} - 1}(2r-1)}{2(r-1)}\right) + C \quad (5.25)$$

Apply the transformations- $x = rcos\Phi, y = rsin\Phi$, and z as in expression, to the three-dimensional Euclidean metric- $ds^2 = dx^2 + dy^2 + dz^2$, and plot the parametric plots for (x, y, z) and $(x, y, -z)$. Observe the plots, shown in figure 5.5, of the upper and lower halves and compare them with the previous embedding diagrams.

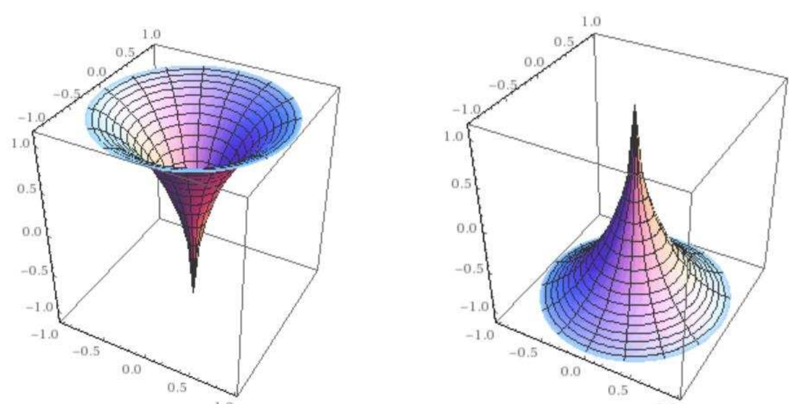


Fig. 5.5. Kasner Embedding: Parametric plots of the upper half (L) and the lower half (R)

5.3.2 Fronsdal Embedding

$$ds^2 = d\Psi_1^2 - d\Psi_2^2 - d\Psi_3^2 - d\Psi_4^2 - d\Psi_5^2 - d\Psi_6^2$$

$$r > r_s: \ \Psi_1 = \left(1 - \tfrac{r_s}{r}\right)^{1/2} sinh(t), \ \Psi_2 = \left(1 - \tfrac{r_s}{r}\right)^{1/2} cosh(t), \ \Psi_3 = \kappa(r),$$

$$r < r_s: \ \Psi_1 = \left(1 - \tfrac{r_s}{r}\right)^{1/2} cosh(t), \ \Psi_2 = \left(1 - \tfrac{r_s}{r}\right)^{1/2} sinh(t), \ \Psi_3 = \kappa(r),$$

$$\Psi_4 = rsin\theta sin\Phi, \ \Psi_5 = rsin\theta cos\Phi, \ \Psi_6 = rcos\theta, \tag{5.26}$$

where $\kappa(r) = \int \sqrt{\frac{2mr^3+m^2}{r^2(r-2m)}}dr = \int \sqrt{\left(\frac{2m}{r}\right) + \left(\frac{2m}{r}\right)^2 + \left(\frac{2m}{r}\right)^3}dr$

5.3.3 Kruskal-Szekeres Coordinates

This was the type of embedding explained earlier, in the previous section. The outcome of such an embedding was ds_4^2, and using the knowledge of tortoise coordinate, we obtained the equation

$$ds^2 = \left(1 - \frac{2m}{r}\right)\left[dt^2 - dr_*^2\right] - r^2 d\Omega^2. \tag{5.27}$$

Now, let us define lightlike coordinates: $u = t + r_*$ and $v = t - r_*$. Add these two equations to obtain an expression for t. Differentiate and square the expression to get: $4dt^2 = du^2 + dv^2 + 2dudv$, and observing that $dt^2 - dr_*^2 = dudv$, we obtain the following metric

$$ds^2 = \left(1 - \frac{r_s}{r(u,v)}\right) dudv - r^2(u,v)d\Omega^2. \tag{5.28}$$

From these lightlike coordinates we can also find an expression for r_*. Subtracting the two equations we get,

$$r_* = \frac{v-u}{2}. \tag{5.29}$$

We already know the form of r_* from the previous chapter. The coordinate singularity of the metric, i.e., the one that occurred as $r \to r_s$, is now replaced at $r_* \to -\infty$, or $v - u \to -\infty$. Define the following transformations

$$U = -2r_s e^{-\frac{u}{2r_s}} = -2r_s e^{-\frac{t-r_*}{2r_s}} = -2r_s e^{-\frac{t-r}{2r_s}} \left(\frac{r}{r_s - 1}\right)^{\frac{1}{2}},$$
$$V = 2r_s e^{\frac{v}{2r_s}} = -2r_s e^{\frac{t+r_*}{2r_s}} = -2r_s e^{\frac{t+r}{2r_s}} \left(\frac{r}{r_s - 1}\right)^{\frac{1}{2}}. \tag{5.30}$$

Taking the derivatives of the defined transformations we obtain expressions for du and dv, which we can substitute into the metric as follows

$$du = e^{-\frac{u}{2r_s}} \left(\frac{r}{r_s-1}\right)^{-\frac{1}{2}} dU, \quad dv = e^{\frac{v}{2r_s}} \left(\frac{r}{r_s-1}\right)^{-\frac{1}{2}} dV$$

$$ds^2 = e^{\frac{v-u}{2r_s}} \left(1-\frac{r}{r_s}\right)\left(\frac{r}{r_s-1}\right)^{-1} dUdV - r^2(U,V)d\Omega^2 \tag{5.31}$$

$$ds^2 = \frac{r_s}{r(U,V)} e^{-\frac{r(U,V)}{r_s}} dUdV - r^2(U,V)d\Omega^2.$$

Here, $r(U,V)$ is an implicit function given by

$$\left(\frac{r(U,V)}{r_s} - 1\right) e^{-\frac{r(U,V)}{r_s}} = -\frac{UV}{4r_s^2}. \tag{5.32}$$

We can write $r(U,V)$ explicitly by making use of the *Lambert W-function* as follows

$$r(U,V) = r_s \left[1 - W\frac{UV}{e}\right]. \tag{5.33}$$

We can also replace u and v by spacelike and timelike coordinates, U and V, called the *Kruskal-Szekeres coordinates* defined by

$$r > r_s : U = \left(\frac{r}{r_s} - 1\right)^{\frac{1}{2}} e^{\frac{r}{2r_s}} cosh\left(\frac{t}{2r_s}\right), \quad V = \left(\frac{r}{r_s} - 1\right)^{\frac{1}{2}} e^{\frac{r}{2r_s}} sinh\left(\frac{t}{2r_s}\right).$$

$$r < r_s : U = \left(\frac{r}{r_s} - 1\right)^{\frac{1}{2}} e^{\frac{r}{2r_s}} sinh\left(\frac{t}{2r_s}\right), \quad V = \left(\frac{r}{r_s} - 1\right)^{\frac{1}{2}} e^{\frac{r}{2r_s}} cosh\left(\frac{t}{2r_s}\right). \tag{5.34}$$

Taking the derivatives of the coordinates (for $r > r_s$) and substituting it into the Schwarzschild line element we obtain

$$dU^2 = \left(\frac{r}{r_s} - 1\right) \frac{e^{\frac{r}{r_s}}}{4r_s^2} sinh^2\left(\frac{t}{2r_s}\right) dt^2 + \frac{r^2 e^{\frac{r}{r_s}}}{4r_s^4} \left(\frac{r}{r_s} - 1\right)^{-1} cosh^2\left(\frac{t}{2r_s}\right) dr^2$$
$$+ \frac{r e^{\frac{r}{r_s}}}{2r_s^3} sinh\left(\frac{t}{2r_s}\right) cosh\left(\frac{t}{2r_s}\right) drdt \tag{5.35}$$

$$dV^2 = \left(\frac{r}{r_s} - 1\right) \frac{e^{\frac{r}{r_s}}}{4r_s^2} cosh^2\left(\frac{t}{2r_s}\right) dt^2 + \frac{r^2 e^{\frac{r}{r_s}}}{4r_s^4} \left(\frac{r}{r_s} - 1\right)^{-1} sinh^2\left(\frac{t}{2r_s}\right) dr^2$$
$$+ \frac{r e^{\frac{r}{r_s}}}{2r_s^3} sinh\left(\frac{t}{2r_s}\right) cosh\left(\frac{t}{2r_s}\right) drdt \tag{5.36}$$

Subtracting equation 5.36 from 5.35, we observe the following

$$dU^2 - dV^2 = \left(\frac{r}{r_s} - 1\right) \frac{e^{\frac{r}{r_s}}}{4r_s^2} cosh^2\left(\frac{t}{2r_s}\right) dt^2 + \frac{r^2 e^{\frac{r}{r_s}}}{4r_s^4} \left(\frac{r}{r_s} - 1\right)^{-1} sinh^2\left(\frac{t}{2r_s}\right) dr^2$$
$$= \frac{e^{\frac{r}{r_s}}}{4r_s^4} \left[-\frac{r}{r_s}\left(1 - \frac{r_s}{r}\right) dt^2 + \frac{r}{r_s}\left(1 - \frac{r_s}{r}\right)^{-1} dr^2\right]$$
$$= \frac{e^{\frac{r}{r_s}}}{4r_s^4} \frac{r}{r_s} \left[-\left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2\right] = \frac{e^{\frac{r}{r_s}}}{4r_s^4} \frac{r}{r_s} ds^2 \tag{5.37}$$

Let's rewrite the above equation in terms of the line element ds^2 (with non-zero angle element $d\Omega$) as follows

$$ds^2 = \frac{4r_s^3}{r} e^{-\frac{r}{r_s}} \left(dU^2 - dV^2\right) - r^2 d\Omega^2 = \frac{32m^3G^3}{rc^6} e^{-\frac{r}{r_s}} \left(dU^2 - dV^2\right) - r^2 d\Omega^2 \tag{5.38}$$

In natural units (i.e., $c \equiv G \equiv 1$), we define the Schwarzschild line element in the Kruskal-Szekeres coordinates as follows

$$ds^2 = \frac{32m^3}{r} e^{-\frac{r}{r_s}} \left(dU^2 - dV^2\right) - r^2 d\Omega^2 \tag{5.39}$$

where r is an implicit function of U and V defined by $\left(\frac{r}{r_s} - 1\right) e^{-\frac{r}{r_s}} = U^2 - V^2$. Notice that the metric is now regular at $r = r_s$, and it contains only one physical singularity at $r = 0$. The property associated to the *Kretschmann invariant* being finite at $r = r_s$ implies that the Schwarzschild spacetime is extensible, i.e., it can be embedded in a larger spacetime whose manifold is not covered by the Schwarzschild coordinates with $r > r_s$. Thus, we can conclude that it is indeed possible to embed the Schwarzschild spacetime and both of its extensions (i.e., the *Eddington-Finkelstein* extension and the Eddington-Finkelstein white hole extension of $r < r_s$) in a larger spacetime which contains an additional copy of the Schwarzschild spacetime. The manifold description of the Fronsdal embedding is closely related to the use of Kruskal-Szekeres coordinates. It is to be noted that the metric signature of $(+ - + - --)$ is used to construct the embedding in a finite region of a manifold with the Schwarzschild metric, but this embedding would contain singularities.

Properties of the Kruskal metric

1. $UV = -4r_s^2 \left(\frac{r(U,V)}{r_s} - 1 \right) e^{-\frac{r(U,V)}{r_s}}$.
Hence, for constant r, we have hperbolae and for conatant t, we have straight lines passing through the origin. The hyperbolae degenerate as $r \to r_s$ and we obtain straight lines defined by $UV = 0$.

2. From $UV = 0$, we obtain: $U = 0$ and $V = 0$. U and V are lightlike coordinates since the equations $dU = 0$ and $dV = 0$ describe light rays. To visualize the global manifold we define the following lightlike coordinate

$$U = T - R, \; V = T + R. \tag{5.40}$$

In the Kruskal diagram shown in figure 5.6, outgoing light rays move along curves $U = const$ and ingoing light rays move along curves $V = 0$.

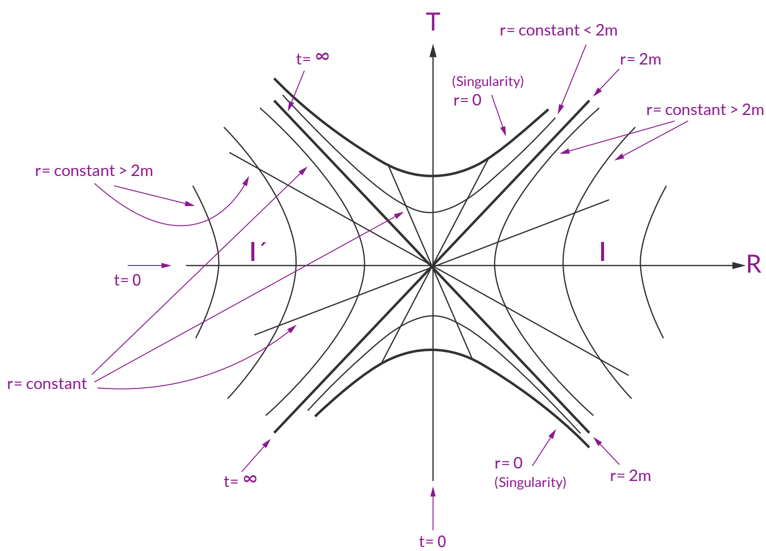


Fig. 5.6. Kruskal diagram

Note that the Schwarzschild coordinates cover only the first quadrant and the metric is singular along the hyperbola, $r = 0$. Observation tells that the metric describing a gravitating object is not *static* in the Kruskal-Szekeres coordinates. The metric in is explicitly dependent upon the time coordinate T. As shown in section 5.2.1, the two-dimensional surface embedded in three-dimensional Euclidean space, having this geometry is generated by the rotating curve (about the z-axis)

$$z = \sqrt{8m(r - 2m)}. \tag{5.41}$$

The resulting geometry is the *Flamm paraboloid* and the region connecting the two asymptotically flat regions is called the *Einstein-Rosen bridge*. This is visualized in figure 5.7.

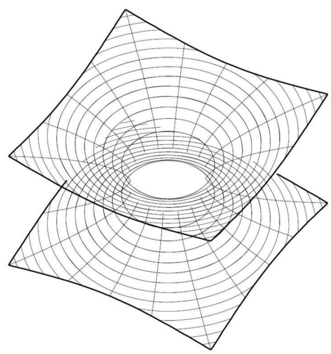


Fig. 5.7. This is the surface defined by equation 5.41 with $t = const$ and $\theta = \pi/2$. For large r the geometry of this surface becomes approximately flat. The z-axis runs vertically through the middle of the throat, i.e., the Einstein-Rosen bridge, the r-coordinate is the distance from that axis.

5.3.4 Fujitani-Ikeda-Matsumoto Embedding

$$\begin{aligned}
&ds^2 = d\Psi_1^2 + d\Psi_2^2 - d\Psi_3^2 - d\Psi_4^2 - d\Psi_5^2 - d\Psi_6^2 \\
&\Psi_1 = t\left(1 - \tfrac{r_s}{r}\right)^{\frac{1}{2}}, \quad \Psi_2 = \tfrac{1}{\sqrt{2}\alpha}\left(\tfrac{\alpha^2 t^2}{2} - 1\right)\left(1 - \tfrac{r_s}{r}\right)^{\frac{1}{2}} + \tfrac{h(r)}{\sqrt{2}}, \\
&\Psi_3 = \tfrac{1}{\sqrt{2}\alpha}\left(\tfrac{\alpha^2 t^2}{2} + 1\right)\left(1 - \tfrac{r_s}{r}\right)^{\frac{1}{2}} + \tfrac{h(r)}{\sqrt{2}}, \quad \Psi_4 = r sin\theta sin\Phi, \\
&\Psi_5 = r sin\theta cos\Phi, \quad \Psi_6 = r cos\theta,
\end{aligned} \tag{5.42}$$

where

$$h(r) = \frac{\alpha r(2r + 3r_s)}{4}\sqrt{1 - \frac{r_s}{r}} + \frac{3\alpha r_s^2}{8} ln\left(\frac{2r}{r_s}\left(1 + \sqrt{1 - \frac{r_s}{r}}\right) - 1\right), \tag{5.43}$$

and in all the above equations $\alpha = \sqrt{2}$. It is to be noted that there are two possibilities of metric signatures we can follow: $(++----)$ and $(--+---)$, and lets use the former for the embedding. Computing the derivatives and their squares we land up with the following metric (expressed in signature $(-+++)$)

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} \left(\frac{4r^4 - 2r^3 r_s + r^2 r_s^2 t^2 - r_s^2}{4r^4})\right) dr^2 + r^2 d\Omega^2. \tag{5.44}$$

The embedded metric resembles the Schwarzschild metric except for the additional term which is a function of t and r, i.e., $f(r,t) = \left(\frac{4r^4 - 2r^3 r_s + r^2 r_s^2 t^2 - r_s^2}{4r^4})\right)$. Again, follow the embedding machinery explained previously and consider the time-slice at $\theta = \frac{\pi}{2}$ (with $r_s = 2m$).

$$\begin{aligned} ds^2 &= \left(1 - \frac{r_s}{r}\right)^{-1} \left(\frac{4r^4 - 2r^3 r_s + - r_s^2}{4r^4})\right) dr^2 + r^2 d\Phi^2 \\ &= \frac{dr^2}{\zeta(r)} + r^2 d\Phi^2 = \left[\left(\frac{dz}{dr}\right)^2 + 1\right] dr^2 + r^2 d\Phi^2. \end{aligned} \tag{5.45}$$

Integration yields the following result

$$z(r) = \int_0^r \sqrt{\frac{1 - \zeta(x)}{\zeta(x)}} dx = -m \sum_{\omega : -\omega^4 + m\omega^3 + 2m = 0} \frac{m\ ln(r - \omega) - 2\omega^3\ ln(r - \omega)}{-4\omega^3 + 3m\omega^2}. \tag{5.46}$$

Taking a series expansion (generalized *Puiseux expansion*) we obtain the following expression (with non-zero m)

$$\begin{aligned} z(r) &= -2m\ ln(r) + \frac{2m}{r} + \frac{3m^3 + m^2(2m^2 - 1)}{3r^3} + \frac{m^3(2m^2 + 3)}{r^4} \\ &+ \frac{m^4(2m^2 + 7)}{5r^5} + O\left(\left(\frac{1}{r}\right)^6\right). \end{aligned} \tag{5.47}$$

For simplicity, we can set $m \equiv 1$ and obtain the parametric plots. This type of embedding covers only the region $r > r_s$ and has a conical singularity at $r = r_s$. It is important to note that if the signature $(- - + - --)$ is used to construct an embedding of the Schwarzschild metric at $r < r_s$. Observe the parametric plots of the upper and lower half, shown in figure 5.8, by making appropriate substitutions as explained in the previous sections.

Example 5.1. Consider the Barriola-Vilenkin metric which describes the gravitational field of a global monopole,

$$ds^2 = -dt^2 + dr^2 + k^2 r^2 d\Omega^2 \tag{5.48}$$

To obtain the embedding metric we first scale the metric of the embedded surface by $r \to kr$ so that (see 5.2) now reads

$$ds^2 = \left[1 + \frac{1}{k^2}\left(\frac{dz}{dr}\right)^2\right] k^2 dr^2 + k^2 r^2 d\phi^2. \tag{5.49}$$

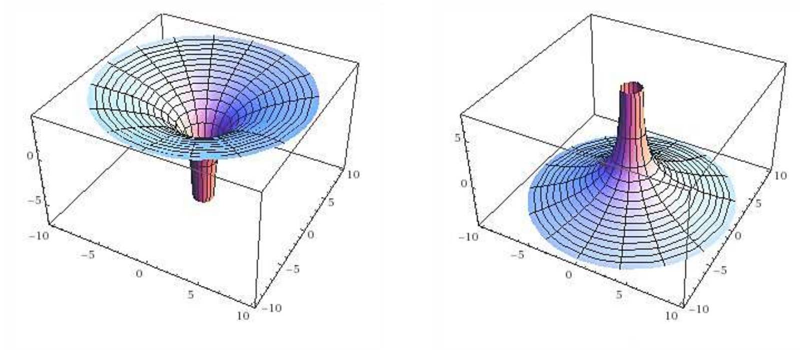


Fig. 5.8. Fujitani-Ikeda-Matsumoto Embedding: Parametric plots of the upper half (L) and the lower half (R)

We can now compare this metric to the time-slice of the Barriola-Vilekin metric at $\theta = \pi/2$ to obtain the embedding function as follows (for $k < 1$)

$$\begin{aligned} &\left[1 + \tfrac{1}{k^2}\left(\tfrac{dz}{dr}\right)^2\right] k^2 = 1, \\ &z(r) = \sqrt{1-k^2}r. \end{aligned} \tag{5.50}$$

Exercise 14

1. Consider the metric for a cosmic string the Schwarzschild spacetime whose line element reads

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 r^2 \left(d\theta^2 + \beta^2 sin^2\theta \; d\phi^2\right), \tag{5.51}$$

where β is called the string parameter. Show that the embedding function for $\beta^2 < 1$ this metric takes the form

$$z(r) = r\left(1 - \frac{r_s}{r}\right)\sqrt{\frac{r}{r - r_s} - \beta^2} - \frac{r_s}{2\sqrt{1-\beta^2}} \; ln \frac{\sqrt{\frac{r}{r-r_s} - \beta^2} - \sqrt{1-\beta^2}}{\sqrt{\frac{r}{r-r_s} + \beta^2} - \sqrt{1-\beta^2}}, \tag{5.52}$$

and check if it reduces to the embedding function of Schwarzschild metric when $\beta^2 = 1$.

2. Show that the embedding function for the Janis-Newman-Winicour metric 6.33 is

$$z(r) = 2\sqrt{rr_s}\; F_1\left(-\tfrac{1}{2}; \tfrac{\gamma+1}{2}, -\tfrac{1}{2}; \tfrac{1}{2}, \tfrac{r_s}{r}, \tfrac{r_s(\gamma+1)^2}{4r\gamma^2}\right)$$
$$-\tfrac{2\pi\gamma}{\gamma+1}\; {}_2F_1\left(-\tfrac{1}{2}, \tfrac{\gamma+1}{2}; 1; \tfrac{4\gamma}{(\gamma+1)^2}\right), \tag{5.53}$$

where F_1 is the Appell-F_1 function and ${}_2F_1$ is the hypergeometric function. Why doesn't this reduce to the Schwarzschild case when $\gamma = 1$ although the metric does?

3. Consider the simplest metric which describes a wormhole

$$ds^2 = -dt^2 + dl^2 + \left(b_0^2 + l^2\right) d\Omega_2^2, \tag{5.54}$$

where b_0 is the throat radius of the wormhole, l is a radial coordinate, and $r^2 = b_0^2 + l^2$. This was first given by Morris and Thorne.
a. Show that the Kretschmann scalar for this metric is $K = \frac{12b_0^2}{\left(b_0^2+l^2\right)^4}$.
b. Find the embedding function for this metric and plot the same for a wormhole with a unit throat radius.

5.4 Extensions of the Schwarzschild Metric

Let's address another problem, we know very well now that the metric is singular at $r = r_s$ and $r = 0$ and we must tear them off the manifold defined by the coordinates (t, r, θ, Φ). The manifold would get disconnected into regions $0 < r < r_s$ and $r_s < r < \infty$ if we pluck off the surface defined by $r = r_s$. Observing the region $r > r_s$ we realize that this is nothing but the external field. Is there a larger manifold $\bar{M}$ into which $\mathcal{M}$ can be immersed? Is the Schwarzschild spacetime extensible? We have proven this without even realizing it. Consider the manifold M with a Schwarzschild metric g. To prove that $(\mathcal{M}, g)$ can be extended we introduce a new coordinate defined by

$$dr_* = \left(1 - \frac{r_s}{r}\right)^{-1} dr \rightarrow r_* = r + r_s ln(r - r_s). \tag{5.55}$$

Remember this? This was the Tortoise coordinate. Now, since you recollect you would predict that the next move would be to introduce the light-like coordinates $v = t + r_*$ and $u = t - r_*$ and then use the ingoing coordinates (v, r, θ, ϕ) to obtain the *ingoing Eddington-Finkelstein* metric $\bar{g}$ defined by (see figure 5.9

$$ds^2 = \left(1 - \frac{r_s}{r}\right) dv^2 - 2dvdr - r^2 d\Omega^2. \tag{5.56}$$

Well, you are spot on! Notice that the manifold M over the external field region is $r_s < r < \infty$, but the ingoing Eddington-Finkelstein metric $\bar{g}$ is non-singular and hence the manifold is not disconnected. Since the manifold is no longer singular at $r = r_s$, we can safely say that this metric is valid in a manifold $\bar{\mathcal{M}}$, which is larger and for which $0 < r < \infty$.

Now, consider a point p and the open set V it belongs to existing in $(\mathcal{M}, g)$, where g is the Schwarzschild metric defined by coordinates $x(p) = (r, t, \theta, \Phi)$, and its image point in the open set U existing in $(\bar{\mathcal{M}}, \bar{g})$, where $\bar{g}$ is the ingoing Eddington-Finkelstein metric defined by coordinates $\bar{x}(q) = (v, r, \theta, \Phi)$. There exists a map z from $\mathcal{M}$ to $\bar{\mathcal{M}}$ such that it is a one-one C^k map and its inverse z^{-1} is a C^k map from $\bar{M}$ to $\mathcal{M}$, i.e., the coordinates of the image point $z(p) = \bar{x}(p)$ in an open set U are k-times continuously differentiable functions of the coordinates $x(p)$. Thus, there exists a map z between $\mathcal{M}$ and $\bar{\mathcal{M}}$ which is a C^k diffeomorphism, and we conclude that the region of $(\bar{\mathcal{M}}, \bar{g})$ for which $0 < r < r_s$ is isometric to the region of the Schwarzschild metric for which $0 < r < r_s$. Thus, changing of coordinates implies moving to a different manifold!

The different manifold here is $\bar{\mathcal{M}}$ to which we have extended the Schwarzschild metric such that it is no longer singular at $r = r_s$ and this region where $r = r_s$ on $\bar{M}$ is called a null surface. Notice that as $r \to 0$, the Kretschmann invariant diverges as $R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} \to \frac{1}{r^6}$, therefore $r = 0$ is the real deal and is called a physical singularity. The pair $(\bar{M}, \bar{g})$ cannot be extended in a C^0 manner, i.e., cannot be extended continuously across $r = 0$. Note that the same extension also works for the outgoing coordinates defined by (u, r, θ, Φ), where $u = t - r_*$. Using these coordinates, we obtain the *outgoing Eddington-Finkelstein* metric g' defined by (see figure 5.10)

$$ds^2 = \left(1 - \frac{r_s}{r}\right) du^2 + 2dudr - r^2 d\Omega^2. \tag{5.57}$$

Again, we notice that the manifold $\mathcal{M}$ over the external field region is $r_s < r < \infty$, but the outgoing Eddington-Finkelstein metric g' is non-singular and hence the manifold is not disconnected. Since the manifold is no longer singular at $r = r_s$, we can safely say that this metric is valid in a manifold larger manifold $\mathcal{M}'$ for which $0 < r < \infty$. Also notice that the new region $0 < r < r_s$ of the manifold $\mathcal{M}'$ equipped with the outgoing Eddington-Finkelstein metric defined by the coordinates (u, r, θ, Φ) is isometric to the region $0 < r < rs$ of the Schwarzschild metric. There is something amazing happening here and it lies right in front of you. Try and predict the metric's fate under a time reversal, $t \to -t$. Upon performing this exercise, you would land up with the following metric

$$ds^2 = \left(1 - \frac{r_s}{r}\right) du^2 - 2dudr + r^2 d\Omega^2, \tag{5.58}$$

which is similar to the ingoing metric. What could be the connection? The answer is that the isometry reverses the direction of time. The surface defined by $r = r_s$ is a null surface in $\mathcal{M}'$, and permits only the past-directed time-like

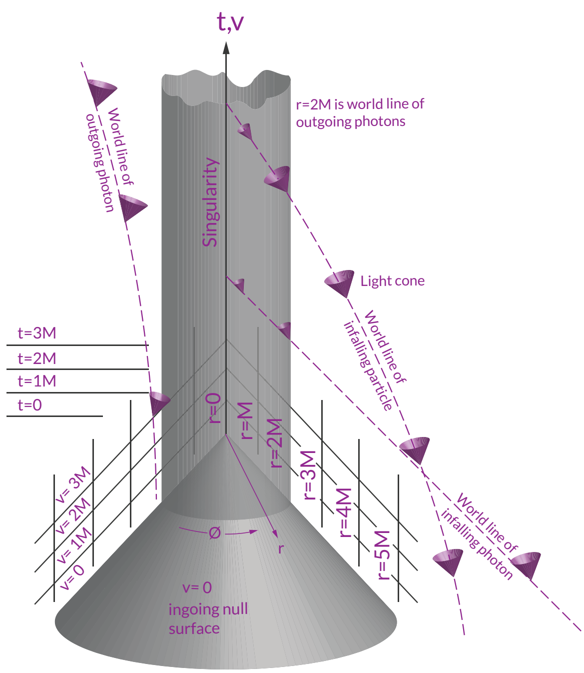


Fig. 5.9. Ingoing Eddington-Finkelstein coordinates (with $\theta = \pi/2$). Surfaces of constant v, being ingoing null surfaces, are plotted on a 45^0 slant, just as they would be in flat spacetime. Here, $r = r_s = 2M$

curves cross from $r > r_s$ to $r < r_s$. We can make both extensions $(\bar{\mathcal{M}}, \bar{g})$ and $(\mathcal{M}', g')$ simultaneously. We came to this conclusion in the chapter on singularities where we stated that it is possible to embed the Schwarzschild spacetime and both of its extensions in a larger spacetime which contains an additional copy of the Schwarzschild spacetime. Translating this to the language of manifolds, we say that there exists a larger manifold $\mathcal{M}^*$ with the metric g^* into which both the extensions $(\bar{\mathcal{M}}, \bar{g})$ and $(\mathcal{M}', g')$ can be embedded isometrically. This is done such that it coincides on the region $r > r_s$ which is isometric to the Schwarzschild spacetime $(\mathcal{M}, g)$. This embedding is done using the Kruskal-Szekeres coordinates to obtain the *Kruskal extension* $(\mathcal{M}^*, g^*)$, where g^* is defined using the coordinates (U, V, θ, Φ). Since Kretschmann scalar $R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$ diverges as r^{-6} when r approaches zero, we can conclude that the metric cannot be extended across the set $r = 0$, at least in the class of C^2 metrics.

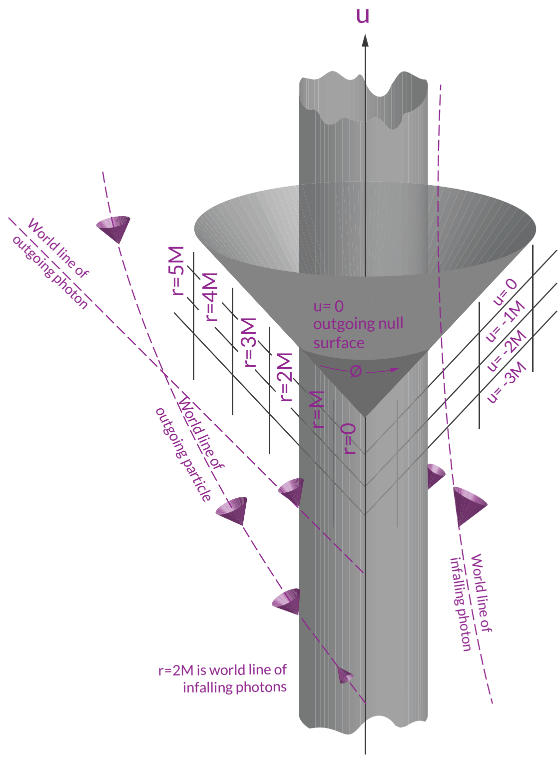


Fig. 5.10. Outgoing Eddington-Finkelstein coordinates (with $\theta = \pi/2$). Surfaces of constant u, being outgoing null surfaces, are plotted on a 45^0 slant, just as they would be in flat spacetime. Here, $r = r_s = 2M$

Exercise 15

1. Define ingoing coordinate $u = t + r_*$ and rewrite the Reissner-Nordstrøm metric 4.38.

2. A much better coordinate system for describing the Kerr metric 3.149 near the surfaces $r = r_\pm$ is Eddington–Finkelstein coordinates (v, r, θ, φ), where r and θ are the same as before whereas v and $\bar{\varphi}$ are related to Boyer–Lindquist coordinates via the following formulae

$$dv = dt + \frac{r^2 + a^2}{\Delta} dr, \;\; d\varphi = d\phi + \frac{a}{\Delta} dr.$$

Rewrite the Kerr metric in the above defined coordinates and check that the singularities are absent.

5.5 Killing Horizons

5.5.1 A Pedestrian Approach

Consider Mr. Absolute Zero to be travelling along the X-axis in a trajectory $X = f(\tau)$, $T = g(\tau)$, where f and g are specified functions and τ is the proper time in the clock carried by him. Let the point p be some event with Minkowski coordinates (T, X) to which Mr. Absolute Zero assigns the coordinates (t, x). We know from Special Relativity that the two coordinates are related by

$$\begin{aligned} X - cT &= f(t - x/c) - ch(t - x/c) \\ X + cT &= f(t + x/c) + ch(t + x/c) \end{aligned} \tag{5.59}$$

Let us now apply this to Mr. Absolute Zero travelling along the X-axis with a uniform acceleration g. The equation of motion

$$\frac{d}{dT}\left(\frac{v}{\sqrt{1 - v^2/c^2}}\right) = g. \tag{5.60}$$

The proper time τ shown by Mr. Absolute Zero's watch while he is being uniformly accelerated can be related to the coordinate time by the standard result

$$\tau = \int_0^T \sqrt{1 - v^2} dT = \frac{1}{g} sinh^{-1}(gT), \tag{5.61}$$

and using this result, we can express his trajectory in a parameterized form (in terms of the proper time) as follows

$$gX = cosh(g\tau) \equiv g\ f(\tau),\ dT = sinh(g\tau) \equiv g\ h(\tau). \tag{5.62}$$

Thus, the coordinate relations with $(c \equiv 1)$ become

$$X - T = g^{-1} e^{-g(t-x)},\ X + T = g^{-1} e^{g(t+x)}, \tag{5.63}$$

which yields

$$X = g^{-1} e^{gx} cosh(gt),\ T = g^{-1} e^{gx} sinh(gt). \tag{5.64}$$

This provides the transformation between the inertial coordinate system and that of Mr. Absolute Zero. Observe that the above transformations are nonlinear and hence won't preserve the line element ds^2. The coordinate frame based on (t, x) is called the *Rindler frame*. Now, using

$$dT^2 - dX^2 = e^{2gx}\left(dt^2 - dx^2\right), \tag{5.65}$$

we get

$$ds^2 = -dt^2 + dX^2 + dY^2 + dZ^2 = e^{2gx}\left(-dt^2 + dx^2\right) + dy^2 + dz^2. \quad (5.66)$$

Using the coordinate transformation $(1 + gx') = e^{gx}$, where we change to a new space coordinate x', we get

$$ds^2 = -\left(1 + \frac{gx'}{c^2}\right)^2 c^2dt^2 + dx'^2 + dy'^2 + dz'^2. \quad (5.67)$$

An alternative form of the *Rindler metric* is

$$ds^2 = -\left(\frac{2gw}{c^2}\right) c^2dt^2 + \left(\frac{2gw}{c^2}\right)^{-1} dw^2 + dy^2 + dz^2, \quad (5.68)$$

which is obtained by making a coordinate transformation $1 + gx' = \sqrt{2gw}$. Now, consider the generalized Schwarzschild metric

$$ds^2 = -\mathcal{F}(r) + \mathcal{F}^{-1}(r)dr^2 + r^2d\Omega^2, \quad (5.69)$$

with the condition that $\mathcal{F}(r)$ has a simple zero at $r = r_k$ with $\mathcal{F}'(r) \neq 0$. Our aim is to observe that such metrics will have a horizon at $r = r_k$ and will also share many of the physical features in the Schwarzschild metric. In this case, the metric near $r = r_k$ take the following form

$$ds^2 \approx -\mathcal{F}'(r_k)(r - r_k)dt^2 + \left(\mathcal{F}'(r_k)(r - r_k)\right)^{-1} dr^2 + d\mathcal{L}^2, \quad (5.70)$$

where $d\mathcal{L}^2$ denotes the metric on the $t = const$, $r = const$ surface. Introducing the variable $\chi \equiv \frac{1}{2}\mathcal{F}'(r_k)$ and a new coordinate $w \equiv (r - r_k)$ in place of r, the metric takes the form

$$ds^2 \approx -2\chi w dt^2 + (2\chi w)^{-1} dw^2 + d\mathcal{L}^2. \quad (5.71)$$

Notice that the above form is precisely the modified metric of the Rindler frame with the horizon now being located at $w = 0$. We see that, when the flat spacetime is described in the Rindler coordinates, the metric is singular at $w = 0$; but since the underlying spacetime is flat, we know that the geometry has no real singularity at $w = 0$ and the peculiar behaviour of the metric must be due to the bad choice of coordinates. This is a clear indication that, for metrics of the above form, the spacetime geometry at $r = r_k$ is well-defined.
Consider the Schwarzschild spacetime, given in terms of the Kruskal-Szekers coordinates

$$ds^2 = -\frac{32m^3}{r}e^{-\frac{r}{r_s}}\left(dU^2 - dV^2\right) + r^2 d\Omega^2 \equiv -\mathcal{C}^2(U,V)\left(dU^2 - dV^2\right) + d\mathcal{L}^2, \tag{5.72}$$

where

$$U = -2r_s e^{-\frac{u}{2r_s}}, \; V = 2r_s e^{\frac{v}{2r_s}}, \tag{5.73}$$

with $u = t + r_*$ and $v = t - r_*$, where r_* is the tortoise coordinate. In the metric, $\mathcal{C}^2(U,V)$ is the conformal factor. The above metric is a generalization of the general spherically symmetric metric in Schwarzschild spacetime, with χ replacing[1] $\frac{1}{2r_s}$. The Killing vector, $\boldsymbol{\xi} = \partial_t$ representing time translation invariance in the Schwarzschild like coordinates can be expressed in the Kruskal-Szekeres coordinates as

$$\boldsymbol{\xi} = \partial_t \times \frac{\partial X^\alpha}{\partial X^\alpha} = \frac{\partial X^\alpha}{\partial t}\frac{\partial}{\partial X^\alpha} = \chi\left(V\frac{\partial}{\partial V} - U\frac{\partial}{\partial U}\right), \tag{5.74}$$

and the norm of the Killing vector is given as follows

$$\xi^2 = g_{UV}\xi^U\xi^V = \mathcal{C}^2\chi^2 UV. \tag{5.75}$$

From the nature of the transformations, we know that the product UV is negative in in the right wedge and positive in the future light cone. It vanishes on the horizon $\mathcal{H}$, where $UV = 0$ (which implies $U = V = 0$), and given by $U = T - R = 0$ and $V = T + R = 0$. The conformal factor, $\mathcal{C}$ is finite at the horizon. It follows that the norm of the Killling vector ξ^2 vanishes vanishes on the horizon and switches sign there. Take the Schwarzschild metric for instance, we know from the previous chapters that the first Killing vector $K^\mu = (1,0,0,0)$ and to find the covariant forms, we simply need to lower with the metric. In Schwarzschild we have

$$\boldsymbol{\xi} = K_\mu = g_{\mu\nu}K^\nu = \left\{-\left(1 - \frac{r_s}{r}\right), 0, 0, 0\right\}, \tag{5.76}$$

and thus, the norm is

$$\xi^2 = -\left(1 - \frac{r_s}{r}\right). \tag{5.77}$$

The killing vector ξ itself becomes

$$\boldsymbol{\xi}|_{\mathcal{H}} = (\partial_t)_{\mathcal{H}} = \chi V \partial_V = \frac{1}{2r_s}2r_s e^{\frac{v}{2r_s}}\frac{\partial}{\partial\left(2r_s e^{\frac{v}{2r_s}}\right)} = \partial_v. \tag{5.78}$$

on the future horizon $U = 0$. Therefore, the Killing vector ξ is both normal

[1] since $\chi = \frac{1}{2}\mathcal{F}'(r) = \frac{1}{2}lt_{r\to r_s}\left(1 - \frac{r_s}{r}\right) = \frac{1}{2r_s}$

and tangential to the horizon surface which, of course, is possible only because the horizon is a null surface. Similar conclusions apply on the past horizon $V = 0$. Now, given a Killing vector ξ, its integral curves are called the orbits of ξ and are illustrated in figure 5.11. Observe that the orbits of ξ are hyperbolas in the right and left wedges; they degenerate to straight lines on the horizons with the origin $U = 0$, $V = 0$ being a fixed point. Just like any other event in the UV plane, the origin also represents a 2-sphere with coordinates θ and ϕ on it and is called a *bifurcation* 2-*sphere*.

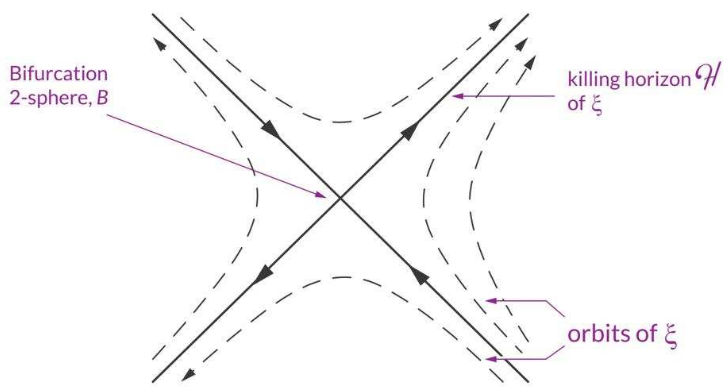


Fig. 5.11. A birfurcate Killing horizon

A *Killing horizon* $\mathcal{H}$ is a type of null surface which arises if a Killing vector field ξ is normal to the null surface. Let l be normal to the null hypersurface Σ and affinely parametrised such that $l \cdot \nabla l = 0$. Then, for some function f, $\xi = fl$ on Σ. It follows that ξ satisfies

$$\begin{aligned} \xi^\mu \nabla_\mu \xi^\nu &= f l^\mu \nabla_\mu (f l^\mu) \\ &= f l^\mu l^\nu \nabla_\mu f + f^2 l^\mu \nabla_\mu l^\nu \\ &= f l^\nu l^\mu \nabla_\mu f \\ &= \kappa \xi^\nu , \end{aligned} \tag{5.79}$$

on Σ. κ is called *surface gravity*. It takes its name from the fact that κ is constant over the horizon and equals the force that an observer at infinity would have to exert in order to keep a unit mass at the horizon.

Killing Horizons in the Schwarzschild Spacetime

Performing a coordinate change to ingoing Eddington-Finkelstein coordinates causes the metric to take the form $(-+++)$

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dv^2 + dr\ dv + dv\ dr + r^2 d\Omega^2. \tag{5.80}$$

We know that the Killing vector has the following form

$$\xi^\mu = K^\mu = (1, 0, 0, 0),\ \&\ \xi_\mu = K_\mu = g_{\mu\nu} K^\nu = \left\{-\left(1 - \frac{r_s}{r}\right), 1, 0, 0\right\}. \tag{5.81}$$

Thus, $\xi^\mu \nabla_\mu \xi^\nu = \kappa \xi^\nu$ gives the following differential equation

$$\frac{1}{2} \frac{\partial}{\partial r} \left(1 - \frac{r_s}{r}\right) = \kappa, \tag{5.82}$$

and upon solving this we get $\kappa = \frac{1}{2r_s}$. Observe that at the horizon, we obtained χ from the Schwarzschild metric to be $\chi = \frac{1}{2}\mathcal{F}'(r) = \frac{1}{2r_s} = \kappa$, this is nothing but surface gravity.

5.5.2 A Mathematical Approach

A *Killing horizon* is a *null-hypersurface* defined by the vanishing of the norm of a vanishing Killing vector field. Mathematically, we can define it as follows: A null hypersurface which coincides with a connected component of the following set

$$\mathcal{A}_X = \mathcal{A}(X) = \{g(X, X) = 0, X \neq 0\}, \tag{5.83}$$

where X is a Killing vector, with X tangent to $\mathcal{A}$, is called a Killing horizon associated to X[2].

Consider a spacelike submanifold S of co-dimension 2 in a spacetime $(\mathcal{M}, g)$, and suppose that there exists a Killing vector field X which vanishes on S. Then, the 1-parameter group of isometries, $\phi_t[X]$, generated by X leaves S invariant and, along S, the tangent maps, $\phi_t[X]_*$, induce isometries of TM to itself. At every point $q \in S$ there exist precisely two null directions- vector-space of $n_\pm$ in the tangent space of M, i.e., $n_\pm \subset T_q M$, where $n_\pm$ are two distinct null future-directed vectors normal to S. Now, since every geodesic is uniquely determined by its initial point and its initial direction, we can conclude that the null geodesics through the point q are mapped to themselves by the flow of X which implies that X is tangent to those geodesics. There exist two null hypersurfaces $\mathcal{A}_\pm$ threaded by those null geodesics, intersecting at S. Let $\mathcal{A}_{\pm+}$ be the connected components of the hypersurface sans the set of null

[2] Here it is implicitly assumed that the hypersurface is embedded

Killing vectors, i.e., $\mathcal{A}_\pm \setminus \{X = 0\}$ lying to the future of S and accumulating at S. Similarly, let $\mathcal{A}_{\pm-}$ be the connected components of $\mathcal{A}_\pm \setminus \{X = 0\}$ lying to the past of S and accumulating at S. Then, the $\mathcal{A}_{\pm\pm}$ are Killing horizons which, together with S, form a *bifurcate Killing horizon* with bifurcation surface S.

5.5.3 Killing Pre-horizons

Let X be a Killing vector, then every connected[3] null hypersurface $\mathcal{A}_0 \subset \mathcal{A}_X$, with $\mathcal{A}_X$[4], with the property that X is tangent to $\mathcal{A}_0$, is called a *Killing pre-horizon*. This is not to be confused with a Killing horizon since a horizon is necessarily embedded while a pre-horizon is not allowed to embedded. Hence, every Killing horizon is a Killing pre-horizon, but not every Killing pre-horizon is a Killing horizon.

5.6 Penrose-Carter Diagrams and the Idea of Pinning Down Infinities

The mathematical idea of a *Penrose-Carter diagram* is to select a relevant 2-dimensional part of a spacetime and make it's *stereographic projection* on a compact space. What this idea translates to in English is- we use a coordinate transformation on the spacetime $(\mathcal{M}, g)$ to pin an infinity to a finite coordinate distance, so that we can draw the entire spacetime on a sheet of paper. A 2-dimensional metric, being a symmetric 2×2 matrix, has three components among which two can be fixed via transformations of two coordinates. Thus, any 2-dimensional metric on $R^{1,1}$ can be transformed to the following form

$$g_{\alpha\beta} = \omega^2(x)\eta_{\alpha\beta}, \ \alpha = 1, 2 \tag{5.84}$$

where ω is a non-zero differentiable function, and $\omega^2(x)$ is a spacetime dependent function called *conformal factor*.

5.6.1 Causal and Conformal Relations

Since spacetime is locally Lorentzian, any two events in a sufficiently small neighbourhood can be joined by lines that are everywhere either spacelike, timelike or null.

a. In the case in which the events are separated by a timelike line: One event occurs before the other. Such events are said to be *causally related*. The first event is contained within the past light cone of the second, which it can causally influence. The second event is contained inside the future light cone of the first.

[3] not necessarily embedded

[4] $\mathcal{A}_X = \mathcal{A}(X) = \{g(X, X) = 0, X \neq 0\}$

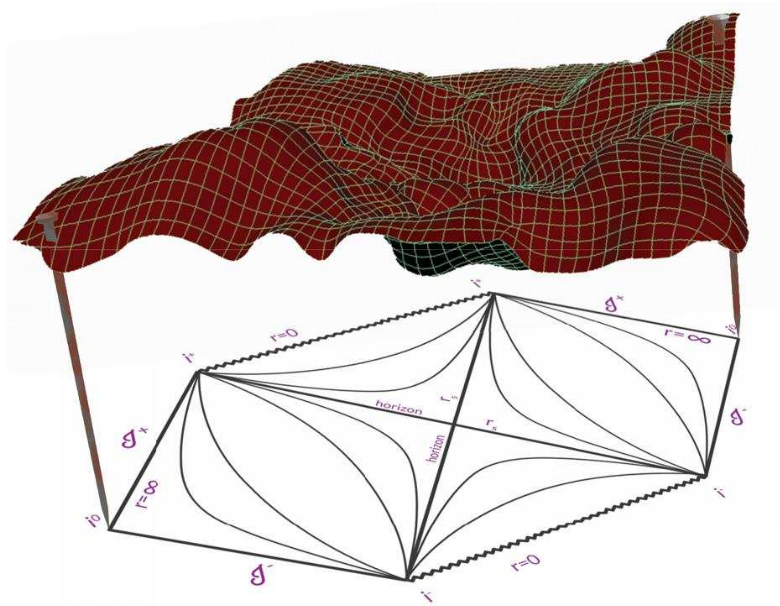


Fig. 5.12. The idea of pinning down infinities via a Penrose-Carter diagram. Here i^0 is a spacelike infinity while $\mathcal{I}^{\pm}$ are null infinities.

b. In the case in which the events are separated by a spacelike line are causally unrelated and information which cannot propagate faster than the speed of light, cannot travel between them.

In some spacetimes, infinite world-lines exist which remain permanently outside each other's light cones. Observers on these world-lines could never be aware of each other's existence! Distinct regions of spacetime that contain families of such world-lines are said to be *causally separated* and the presence of these regions are to be recognized in spacetimes that contain them.

The idea behind a Penrose-Carter diagram, as previously seen, is that angles and lightlike world-lines are preserved under conformal transformations. The metric will typically diverge as we approach the infinities, i.e., the edges of the finite diagram. To fix this, we perform a conformal transformation on g to obtain a new metric $\bar{g}$ that is regular on the edges. Then $(\mathcal{M}, \bar{g})$ is a good representation of the original spacetime $(\mathcal{M}, g)$ insofar that it has exactly the same causal structure. A remarkable property of the conformal transformation is that the Weyl tensor for the original and conformal spacetimes are identical, i.e., $W_{\alpha\beta\gamma\delta} = W_{\alpha\beta\gamma\delta}$. Curvature tensors, however, are not preserved under conformal transformations, i.e., for example $\bar{R}\alpha\beta\gamma\delta \neq R_{\alpha\beta\gamma\delta}$. Thus, it is generally not possible for both the original and the conformal spacetimes to be vacuum, or to correspond to the same type of source.

For any spacetime with metric g and manifold M, a related spacetime with metric $\bar{g}$ and manifold $\bar{M}$ is defined by the conformal transformation $\bar{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}$, where the conformal factor Ω can be, in general, an arbitrary function. Tensor indices in the original and conformally related spacetimes are lowered and raised using the corresponding metric, with $\bar{g}^{\mu\nu} = \Omega^{-2} g^{\mu\nu}$. This conformal transformation is particularly useful if it maps the asymptotic regions at infinite proper distance in the original spacetime to finite regions in the conformal spacetime. Taking the conformal factor Ω to be positive in M, the asymptomatic boundary of M maps to the hypersurface in $\bar{M}$ on which $\Omega = 0$. This boundary is referred to as *conformal infinity* and denoted as $\mathcal{I}$ (pronounced ′scri′). Conformal diagrams take a particularly simple form when describing spherically symmetric spacetimes. This enables the spacetime to be visualized in a 2-dimensional picture in which every point represents a typical point on a 2-sphere at some time. For spacetimes with less symmetry, it is still possible to construct conformal diagrams for specific sections, but to visualize their complete causal structure in a suitable higher-dimensional conformal picture is usually much more difficult.

5.6.2 Spatial vs Future Null Infinities

Mr. Zero Entropy and Mr. Absolute Zero are in charge of calculating the amount of mass carried away by gravitational and electromagnetic waves during a supernova explosion. They first measure the asymptomatic form of the metric not just at spatial infinity but at future null infinity. Mr. Absolute Zero measures the mass before the explosion m_B by examining the asymptomatic form of g_{00} at spatial infinity, the following are the results he reported back to base

$$g_{00} = -1 + \frac{2m_B}{r} + \mathcal{O}\left(\frac{1}{r^2}\right), \tag{5.85}$$

as $r \to \infty, t = const$. Mr. Zero Entropy waits in the spaceship is at a fixed radius r until the radiation has flowed completely past their point. He measures the mass after explosion m_A. The following are the results he reported back to base

$$g_{00} = -1 + \frac{2m_A}{r} + \mathcal{O}\left(\frac{1}{r^2}\right), \tag{5.86}$$

as $r \to \infty$ with $t - r = const$. Thus, the measurement of m_A is done by examining the asymptomatic form of g_{00} not at spatial infinity, but at future null infinity.

5.6.3 Conformal Mapping of Minkowski Spacetime

The simplest spacetime is that which is flat everywhere. Such a spacetime contains no matter and no gravitational field. It is known as Minkowski space, and is the spacetime of special relativity. It typically occurs as a weak-field limit of many solutions of general relativity, and may also appear as the asymptotic limit of the gravitational field of bounded sources. For all these reasons it is most important that its structure be clearly understood. As is well known, in a 3 + 1-dimensional spacetime, the maximum number of symmetries is 10. The Minkowski metric has this precise number of isometries, which can be considered to correspond to four translations, three spatial rotations and three special Lorentz boosts. Such a maximally symmetric spacetime necessarily has constant curvature, which is zero in this case. Thus, in any coordinate representation in which it may not be immediately recognisable, Minkowski space (or part of it) can always be uniquely identified by the fact that its curvature tensor vanishes identically. the Minkowski metric can be expressed in the Cartesian form as follows

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2, \tag{5.87}$$

in which the coordinates t, x, y, z cover their full natural ranges $(-\infty, \infty)$. In spherical polar coordinates, the metric takes up the following form

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2. \tag{5.88}$$

But in this form, the metric has apparent singularities when $r = 0$, when $sin\theta = 0$, and when $r \to \infty$. However, these are simply coordinate singularities which correspond, respectively, to the origin, the axis of spherical polar coordinates and to spatial infinity. These apparent singularities thus have no physical significance. In this case of a flat Minkowski space, the conformal relation is obtained by considering the (t, x) part of spacetime, and applying the following coordinate transformations

$$\begin{aligned} t + x &= tan\left(\tfrac{\psi+\xi}{2}\right) \\ t - x &= tan\left(\tfrac{\psi-\xi}{2}\right). \end{aligned} \tag{5.89}$$

Thus, if $t, x \in (-\infty, \infty)$, then $\psi, \xi \in [-\pi, \pi]$. We have just pinned down the infinities to finite values $-\pi$ and π via appropriate functions. Now,

$$dt^2 - dx^2 = \frac{1}{\left[2cos\left(\frac{\psi+\xi}{2}\right) cos\left(\frac{\psi-\xi}{2}\right)\right]^2}\left(d\psi^2 - d\xi^2\right). \tag{5.90}$$

Comparing this to the standard form of conformal transformation, we find that the conformal factor is

$$\Omega = \frac{1}{\left[2cos\left(\frac{\psi+\xi}{2}\right) cos\left(\frac{\psi-\xi}{2}\right)\right]^2}, \tag{5.91}$$

and it blows up at $|\psi \pm \xi| = \pi$, which makes the boundary of the compact (ψ, ξ) spacetime infinitely far away from it's internal point. This allows us to map the compact (ψ, ξ) spacetime onto the non-compact (t, x) spacetime. The boundary of the physical Minkowski spacetime is then given by the points where $\Omega = 0$, i.e., $\psi + \xi = \pi$ and $\psi - \xi = -pi$. These points corresponds to infinities in the physical spacetime. For the section on which $\theta = \pi/2$, the entire Minkowski space is conformal to the region between the two null cones. Such null cones are shown in figure 5.13, and Minkowski space can thus be seen to be conformal to the region between them.

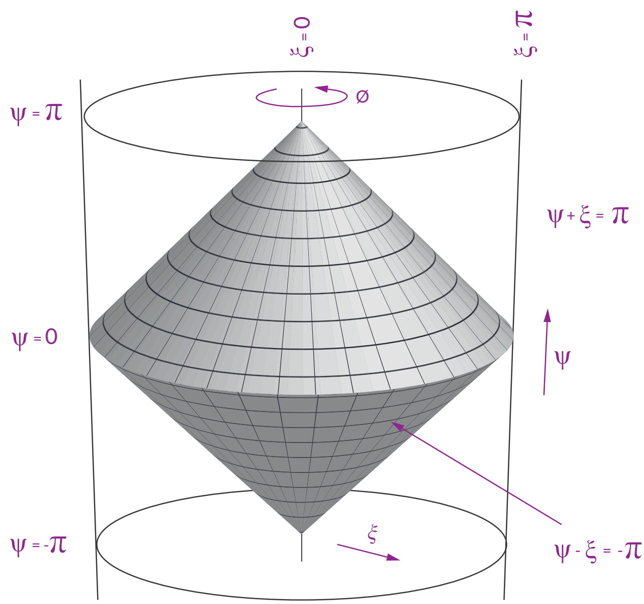


Fig. 5.13. The section $\theta = \pi/2$ of Minkowski space is conformal to the region between the two null cones $\psi + \xi = \pi$ and $\psi - \xi = \pi$ that is within the timelike cylinder of radius π which represents the *Einstein static universe*

This immediately illustrates the important fact that the conformal infinity of Minkowski space is null, and denoted as $\mathcal{I}$. Also, since the radius of the 2-spheres vanishes at $\xi = \pi$, it is demonstrated that the point at spatial infinity i^0 is indeed just a point. $\xi = 0$ represents the origin of the spherical coordinates

$r = 0$, and $\psi = 0$ corresponds to $t = 0$. Table 5.1 gives the definitions of all the pinned infinities.

Label	Name	Definition
$\mathcal{I}^-$	denotes the past null infinity	$(\psi - \xi = -\pi,\ 0 < \xi < \pi)$
$\mathcal{I}^+$	denotes the future null infinity	$(\psi + \xi = \pi,\ 0 < \xi < \pi)$
i^-	denotes the past timelike infinity	$(\psi = -\pi, \xi = 0)$
i^0	denotes the spatial infinity	$(\psi = 0, \xi = 0)$
i^+	denotes the future timelike infinity	$(\psi = \pi, \xi = 0)$

Table 5.1. The symbols associated to and the conditions of the pinned infinities

The metric of the new compact (ψ, ξ) space, in polar coordinates has the following form

$$ds^2 = \frac{-d\psi^2 + d\xi^2}{4cos^2\left(\frac{\psi+\xi}{2}\right) cos^2\left(\frac{\psi-\xi}{2}\right)} + r^2(\psi, \xi)\left(d\theta^2 + sin^2 d\phi^2\right). \qquad (5.92)$$

Note that if $dt^2 - dx^2 = 0$, then $d\psi^2 - d\xi^2 = 0$. Thus, the conformal factor is irrelevant in the study of the properties of lightlike world-lines which obey $ds^2 = 0$.

5.6.4 Conformal Mapping of Schwarzschild Spacetime

Consider the Schwarzschild metric in the Kruskal-Szekeres coordinates

$$ds^2 = \frac{32m^3}{r} e^{-\frac{r}{r_s}} \left(dU^2 - dV^2\right) - r^2\left(d\theta^2 + sin^2 d\phi^2\right). \qquad (5.93)$$

As seen previously, these coordinates are already in the conformally flat form hence we need only to make their ranges compact. This can be done using the following coordinate transformations

$$U \pm V = tan\left(\frac{\psi \pm \xi}{2}\right). \qquad (5.94)$$

These transform the metric into the form

$$ds^2 = \frac{32m^3}{r} e^{-\frac{r}{r_s}} \frac{d\psi^2 - d\xi^2}{4cos^2\left(\frac{\psi+\xi}{2}\right) cos^2\left(\frac{\psi-\xi}{2}\right)} - r^2\left(d\theta^2 + sin^2 d\phi^2\right). \qquad (5.95)$$

the global structure of the Schwarzschild spacetime is illustrated in figure 5.14. It shows all possible regions of the complete analytically extended manifold. In particular, exterior to the horizons there exist two causally separated static regions $r > r_s$ that are asymptotically Minkowski like. The figure also shows the spacelike character of the initial and final curvature singularities $r = 0$.

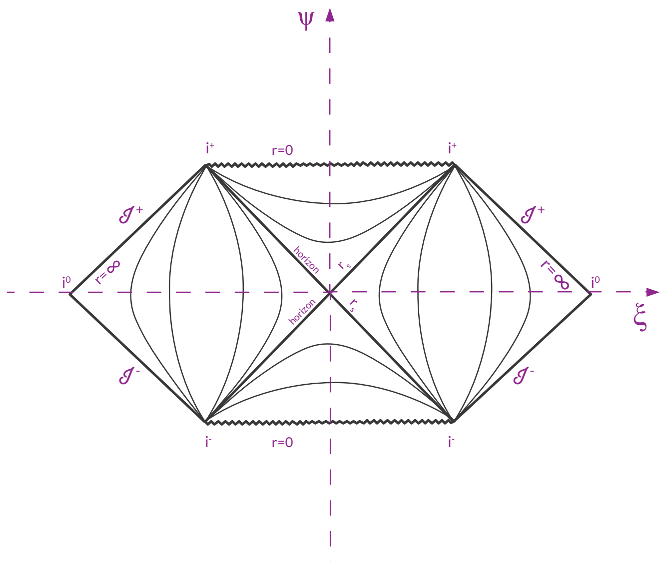


Fig. 5.14. The figure shows the Penrose diagram for the complete Schwarzschild spacetime. Here, the θ and ϕ coordinates are suppressed so that each point represents a 2-sphere of radius r. All lines shown are hypersurfaces on which r is a constant.

Figure 5.15 illustrates the character of the spacelike infinity i^0, null infinity $\mathcal{I}^{\pm}$, and the horizons that surround the initial and final curvature singularities. The initial and final curvature singularities are also referred to as the *white hole* and the *black hole*.

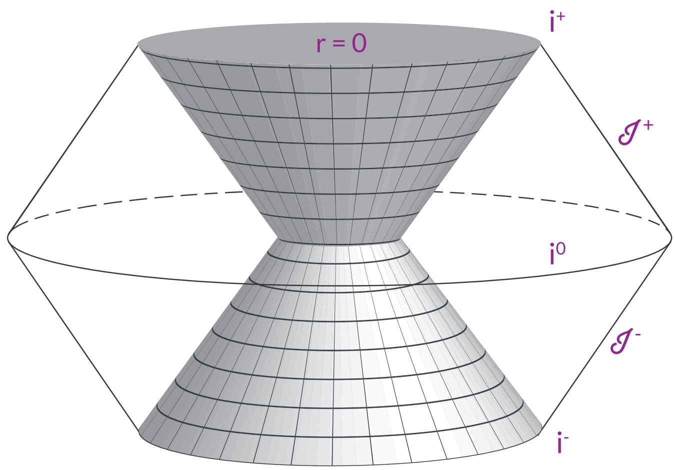


Fig. 5.15. The exterior region of the Schwarzschild spacetime with the angular coordinate ϕ reintroduced. The complete spacetime contains two such regions that are *causally separated* (see subsection 5.6.1) from each other. The horizons at $r = r_s$ are drawn at 45^0 to reflect the fact that they are null, but their area at all times remains constant.

5.6.5 Why Curvature Singularities?

Due to the immense physical significance of the Schwarzschild solution as describing the spacetime exterior to a massive spherical object or black hole, we have concentrated on the case in which $m > 0$. This solution with a negative parameter m, i.e., $m < 0$ may also be considered[5]. In such a case, there is no horizon and the curvature singularity at $r = 0$ is timelike, globally naked and unstable. The global nakedness of the singularity tells us that the curvature singularity could be literally seen at every point in the spacetime. This case, however, does not correspond to any known physical situation. The Penrose and conformal diagrams for this solution are illustrated in figure 5.16.

[5] the assumption that $r \in (0, \infty)$ still holds

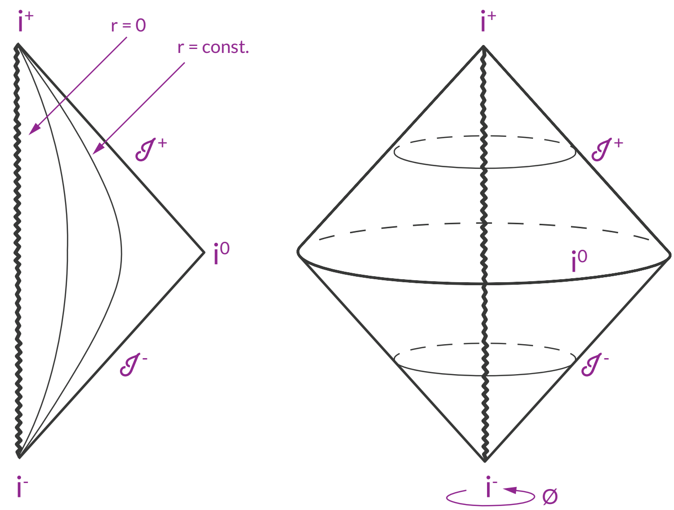


Fig. 5.16. Penrose and conformal diagrams for the Schwarzschild spacetime for the case when $m < 0$. Notice the globally naked timelike singularity at $r = 0$.

5.7 Higher Dimensions

A conventional coordinate system for the Schwarzschild metric is the *Isotropic coordinates* in which we introduce a new radial coordinate $\bar{r}$, which is implicitly defined by the following formula

$$r = \bar{r}\left(1 + \frac{m}{2\bar{r}}\right)^2, \tag{5.96}$$

with some computations we obtain,

$$g_m = \left(1 + \frac{m}{2|x|}\right)^4 \left(\sum_{\alpha=1}^{3} (dx^\alpha)^2\right) - \left(\frac{1 - \frac{m}{2|x|}}{1 + \frac{m}{2|x|}}\right)^2 dt^2, \tag{5.97}$$

where x^α are coordinates on R^3 with $|x| = \bar{r}$. Those coordinates show explicitly that the space-part of the metric is conformally flat.

5.7.1 Painlevé - Gullstrand Coordinates

The Schwarzschild spacetime has the agog property of possessing flat spacelike hypersurfaces. They appear when introducing the *Painlevé-Gullstrand* coordinates as follows:
Starting from the standard Schwarzschild metric we introduce a new time τ via the following equation

$$t = \tau - 2r\sqrt{\frac{2m}{r}} + 4m\ tanh^{-1}\left(\sqrt{\frac{2m}{r}}\right), \tag{5.98}$$

such that

$$dt = d\tau - \frac{\sqrt{\frac{2m}{r}}}{\frac{2m}{r}}dr. \tag{5.99}$$

This yields the following metric

$$ds^2 = -\left(1 - \frac{2m}{r}\right)d\tau^2 + 2\sqrt{\frac{2m}{r}}drd\tau + dr^2 + r^2 d\Omega^2, \tag{5.100}$$

or, alternatively in standard Cartesian coordinates

$$ds^2 = -\left(1 - \frac{2m}{r}\right)d\tau^2 + 2\sqrt{\frac{2m}{r}}drd\tau + dx^2 + dy^2 + dz^2. \tag{5.101}$$

5.7.2 Wave Coordinates and their Asymptotic Behaviour

Wave coordinates are another set of coordinates that act as an effective tool for PDE analysis of spacetimes. In spherical coordinates associated to wave coordinates $(t, \hat{x}, \hat{y}, \hat{z})$, with radius function $\hat{r} = \sqrt{\hat{x}^2 + \hat{y}^2 + \hat{z}^2}$, the Schwarzschild metric takes the following form [6]

$$ds^2 = -\frac{\hat{r} - m}{\hat{r} + m}dt^2 + \frac{\hat{r} + m}{\hat{r} - m}d\hat{r}^2 + (\hat{r} + m)^2 d\Omega^2. \tag{5.102}$$

Now, consider the Schwarzschild metric in dimensions greater than 3, i.e., $n \geq 3$,

$$ds_n^2 = -\left(1 - \frac{2m}{r^{n-2}}\right)dt^2 + \left(1 - \frac{2m}{r^{n-2}}\right)^{-1}dr^2 + r^2 d\Omega^2, \tag{5.103}$$

[6] this is obtained by replacing r in the Schwarzschild metric with $\hat{r} = r - m$

and consider a general spherically symmetric static metric of the the following form

$$ds^2 = -e^{2\alpha}dt^2 + e^{2\beta}d\hat{r}^2 + e^{2\gamma}\hat{r}^2 d\Omega^2, \tag{5.104}$$

where α, β, and γ depend only upon r. Now let us define ϕ and ψ such that

$$\phi = e^{\alpha+\beta+(n-3)\gamma}, \quad \psi = e^{\alpha+\beta+(n-1)\gamma}\left(e^{-2\beta} - e^{-2\gamma}\right). \tag{5.105}$$

Now, we proceed to perform all our calculations in a coordinate system in which the vector (x, y, z) is aligned along the x-axis, $(x, y, z) = (r, 0, 0)$. Then the metric in a spacetime dimension of $n+1$ reads

$$g = \begin{pmatrix} -e^{2\alpha} & 0 & 0 & 0\cdots & 0 \\ 0 & e^{2\beta} & 0 & 0\cdots & 0 \\ 0 & 0 & e^{2\gamma} & 0\cdots & 0 \\ 0 & 0 & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & e^{2\gamma} \end{pmatrix}, \tag{5.106}$$

which implies the following

$$det\ g = |g| = e^{2(\alpha+\beta)+2(n-1)\gamma}, \tag{5.107}$$

still at $(x, y, z) = (r, 0, 0)$. Spherical symmetry implies that this equality holds everywhere. Now, we have

$$\begin{aligned} det\ g\ \partial_g \partial^g x &= \partial_\xi \left(\sqrt{|g|}g^{\nu\mu}\right) \\ &= \partial_\xi \left(e^{-2\gamma}\delta^{\mu\nu} + \left(e^{-2\beta} - e^{-2\gamma}\right)\tfrac{x^\mu x^\nu}{r^2}\right) \\ &= \partial_\xi \left(\underbrace{e^{\alpha+\beta+(n-3)\gamma}}_{=\phi}\delta^{\mu\nu} + \underbrace{e^{\alpha+\beta+(n-1)\gamma}\left(e^{-2\beta} - e^{-2\gamma}\right)}_{=\psi}\tfrac{x^\mu x^\nu}{r^2}\right) \\ &= \left(\phi' + \psi' + \tfrac{n-1}{r}\psi\right)\tfrac{x^\mu}{r}, \end{aligned} \tag{5.108}$$

this is called the *harmonicity condition*. It follows from this condition that

$$0 = \frac{d(\phi+\psi)}{d\hat{r}} + \frac{n-1}{\hat{r}}\psi, \tag{5.109}$$

equivalently,

$$\frac{d\left[\hat{r}^{n-1}\left(\phi+\psi\right)\right]}{d\hat{r}}=\left(n-1\right)\hat{r}^{n-2}\phi. \tag{5.110}$$

Now, comparing the above equation with the standard metric defined, we find

$$e^{\alpha}=\sqrt{1-\frac{2m}{r^{n-2}}},\quad e^{\beta}=\frac{dr}{r}e^{-\alpha},\quad e^{\gamma}=\frac{r}{\hat{r}}. \tag{5.111}$$

Since $\phi+\psi=e^{\alpha-\beta+(n-1)\gamma}$ and $\phi=\frac{dr}{d\hat{r}}\left(\frac{r}{\hat{r}}\right)^{n-3}$, we finally obtain [7]

$$\frac{d}{dr}\left[r^{n-1}\left(1-\frac{2m}{r^{n-2}}\right)\frac{dr}{d\hat{r}}\right]=(n-1)\hat{r}r^{n-3}. \tag{5.113}$$

The characteristic exponents are 1 and $n-1$ so that, after matching a few leading coefficients, the standard theory of such equations provides solutions with the behaviour

$$\hat{r}=r-\frac{m}{(n-2)\,r^{n-3}}+\begin{cases}\frac{m^2}{4}r^{-2}\,ln(r)+\mathcal{O}\left(r^{-5}\,ln(r)\right), & n=4\\ \mathcal{O}\left(r^{5-2n}\right) & ,n\geq 5.\end{cases} \tag{5.114}$$

Thus, somewhat surprisingly, we find logarithms of r in an asymptotic expansion of $\hat{r}$ in dimension $n=4$. However, for $n\geq 5$ there is a complete expansion of $\hat{r}$ in terms of inverse powers of r, without any logarithmic terms in those dimensions.

5.7.3 Schwarzschild-Tangherlini Metric

In spacetime dimension $(n+1)$, the metrics take the form

$$ds^2=-\mathcal{V}^2(r,t)dt^2+\mathcal{V}^{-2}dr^2+r^2d\Omega^2, \tag{5.115}$$

with

$$\mathcal{V}^2=1-\frac{2m}{r^{n-2}}, \tag{5.116}$$

[7] which upon introduction of $x=1/r$ becomes

$$\frac{d}{dr}\left[x^{3-n}\left(1-2mx^{n-2}\right)\frac{d\hat{r}}{dx}\right]=(n-1)\hat{r}x^{1-n}, \tag{5.112}$$

this is called an equation with a *Fuchsian singularity* at $x=0$

where the mass m is called *Arnowitt-Deser-Misner* (or ADM) *mass* in spacetime dimension four [8]. Making the assumption that $m > 0$, a maximal analytic extension can be constructed by a simple modification of the calculations presented above, leading to a spacetime with a global structure.
Consider a metric of the following form

$$ds^2 = -\mathcal{R}dt^2 + \mathcal{R}^{-1}dr^2 + \underbrace{h_{\mu\nu}dx^\mu dx^\nu}_{=h}, \tag{5.117}$$

with $\mathcal{R} = \mathcal{R}(r)$, where $h = h_{\mu\nu}\left(t, r, x^\gamma dx^\mu dx^\nu\right)$ is a family of Riemannian metrics on an $(n2)$-dimensional manifold which possibly depend on t and r[9]. It is assumed that $\mathcal{R}$ is defined for r in a neighborhood of $r = r_0$, at which $\mathcal{R}$ vanishes, with a simple zero[10] there. Equivalently

$$\mathcal{R}\left(r_0\right) = 0, \quad \mathcal{R}^{'}\left(r_0\right) \neq 0. \tag{5.118}$$

Let us define the following

$$\begin{aligned} &u = t - f(r), \; v(r) = t + f(r), \; f = \tfrac{1}{\mathcal{R}} \\ &\hat{u} = -e^{-cu}, \; \hat{v} = e^{cv}, \end{aligned} \tag{5.119}$$

we are led to the following form of the metric

$$ds^2 = -\frac{\mathcal{R}}{c^2}e^{-2cf(r)}d\hat{u}d\hat{v} + h. \tag{5.120}$$

Since $\mathcal{R}$ has a simple zero, it factorizes as follows

$$\mathcal{R} = \left(r - r_0\right)\mathcal{H}(r), \tag{5.121}$$

for a function $\mathcal{H}$ which has no zeros in a neighborhood of r_0. This follows immediately from the following formula

$$\mathcal{R}{-}\mathcal{R}\left(r - r_0\right) = \int_0^1 \frac{d\mathcal{R}\left(t\left(r - r_0\right) + r_0\right)}{dt}dt = \left(r - r_0\right)\int_0^1 \mathrm{R}^{'}\left(t\left(r - r_0\right) + r_0\right)dt. \tag{5.122}$$

[8] and is proportional to that mass in higher dimensions

[9] It is convenient to write $\mathcal{R}$ for $\mathcal{V}^2$, as the sign of $\mathcal{R}$ did not play any role; similarly the metric h was irrelevant for the calculations performed in the previous section(s).

[10] If f is a function that is meromorphic in a neighbourhood of a point z_0 of the complex plane, then there exists an integer n such that $f(z)\left(z - z_0\right)^n$. Simple zero is one of order $|n| = 1$

Now, in the following equation

$$
\begin{aligned}
\frac{1}{\mathcal{R}(r)} &= \frac{1}{(r-r_0)\mathcal{H}(r_0)} + \frac{1}{\mathcal{R}(r)} - \frac{1}{(r-r_0)\mathcal{H}(r_0)} \\
&= \frac{1}{(r-r_0)\mathcal{H}(r_0)} + \frac{\mathcal{H}(r_0)-\mathcal{H}(r)}{\mathcal{H}(r)\mathcal{H}(r_0)(r-r_0)},
\end{aligned}
\tag{5.123}
$$

an analysis of $\mathcal{H}(r) - \mathcal{H}(r_0)$ followed by integration yields

$$
f(r) = \frac{1}{\mathcal{R}'(r_0)} ln\,(r - r_0) + \hat{f}(r) \tag{5.124}
$$

for some function $\hat{f}$ which is smooth near r_0. Substituting this into the metric with

$$
c = \frac{\mathcal{R}'(r_0)}{2}, \tag{5.125}
$$

we get

$$
ds^2 = \mp \frac{4\mathcal{H}(r)}{\left(\mathcal{R}'(r_0)\right)^2} e^{-\hat{f}(r)\mathcal{R}'(r_0)} d\hat{u}d\hat{v} + h, \tag{5.126}
$$

with a negative sign if we started in the region with $r > r_0$, and positive otherwise. The function r is again implicitly defined by[11]

$$
\hat{u}\hat{v} = \mp (r - r_0)\, e^{\hat{f}(r)\mathcal{R}'(r_0)}. \tag{5.127}
$$

The function f defined previously, for a $(4+1)$-dimensional *Schwarzschild-Tangherlini* solution can be calculated to be the following

$$
f = r + \sqrt{2m}\; ln\left(\frac{r - \sqrt{2m}}{r + \sqrt{2m}}\right), \tag{5.128}
$$

which results in the following metric

$$
ds^2 = -\frac{8m\left(r + \sqrt{2m}\right)^2}{r^2} e^{-\frac{r}{2m}} d\hat{u}d\hat{v} + d\Omega^2. \tag{5.129}
$$

---

[11] Notice that the RHS has a derivative which equals $\mp e^{\frac{\hat{f}(r_0}{\mathcal{R}'(r_0)}} \neq 0$ at r_0, and hence this equation defines a smooth function $r = r(\hat{u}\hat{v})$ for r near r_0 by the implicit function theorem. The above discussion applies to $\mathcal{R}$ which are of C^k differentiability class, with some losses of differentiability. Also, the above equation also provides an extension of C^{k2} differentiability class, which leads to the restriction $k \geq 2$. However, the implicit function argument just given requires h to be differentiable, so we need in fact $k \geq 3$ for a cogent analysis.

The isotropic coordinates in higher dimensions lead to the following form of the Schwarzschild-Tangherlini metric

$$ds^2 = \left(1 + \frac{m}{2|x|^{n-2}}\right)^{\frac{4}{n-2}} \left(\sum_{\alpha=1}^{n} (dx^\alpha)^2\right) - \left(\frac{1 - \frac{m}{2|x|^{n-2}}}{1 + \frac{m}{2|x|^{n-2}}}\right) dt^2, \qquad (5.130)$$

where the radial coordinate $|x|$ has the following relation with r

$$r = \left(1 + \frac{m}{2|x|^{n-2}}\right)^{\frac{2}{n-2}} |x|. \qquad (5.131)$$

An explicit formulation of the same metric, expressed as functions of elementary coordinates, may be given as follows

$$ds^2 = -8m\left(dxdy + \frac{y^2}{xy + 2m}dx^2\right) - (xy + 2m)^2 d\Omega^2, \qquad (5.132)$$

where the coordinates (x, y) are related to standard Schwarzschild coordinates (t, r) are follows

$$\begin{aligned} r &= xy + 2m, \\ t &= xy + 2m\left(1 = ln\left|\tfrac{y}{x}\right|\right), \\ |x| &= \sqrt{|r - 2m|}e^{\frac{r-t}{4m}}, \\ |y| &= \sqrt{|r - 2m|}e^{\frac{t-r}{4m}}. \end{aligned} \qquad (5.133)$$

In higher dimensions there exists an explicit, manifestly globally regular form of the metric, in spacetime dimension $n + 1$

$$\begin{aligned} ds^2 &= -2\frac{q^2\left(-(r)^{-n+2}2^{n+1}m^{n+1}+4m^2((n+1)(2m-r)+3r-4m)\right)}{m(2m-r)^2}dW^2 \\ &+8mdWdz + r^2 d\Omega_{n-1}^2, \end{aligned} \qquad (5.134)$$

where $r \geq 0$ is the following function

$$r(W, z) = 2m + (n - 2)Wz, \qquad (5.135)$$

and $d\Omega_{n-1}^2$ is the metric of a unit $n - 1$ sphere.

6

Advanced Topics

6.1 The Vielbein Spin Connection

A neat trick to have in one's pocket is the knowledge of the Vielbein spin connection or the tetrad formalism. This is specifically useful in removing *waste*, i.e., terms which are null. This formalism is very useful when we have a diagonal metric where the non-null components of the Christoffel symbol and Riemann tensor can be computed with ease. It can also be employed with non-diagonal metrics although the calculations are a bit length but still less when compared to the usual method. In the following paragraphs, I will be providing a brief outline of this method, enough to do computations. Consider a metric $ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta$. We introduce a Vielbein which is nothing but a basis of 1-forms

$$e^a = e^a_\alpha dx^\alpha, \tag{6.1}$$

where the components e^a_α possess the property

$$g_{\alpha\beta} = \eta_{ab}e^a_\alpha e^b_\beta. \tag{6.2}$$

The indices a and b are called tangent space indices and η_{ab} is the flat metric. It is to be noted that the metric η_{ab} will possess the same signature of $g_{\alpha\beta}$. The choice of vielbeins is arbitrary but we can specifically choose an orthogonal-type transformation to obtain another valid Vielbein $\bar{e}^a$ which is related to the original Vielbein by

$$\bar{e}^a = \Lambda^a_b e^b, \tag{6.3}$$

where Λ^a_b satisfies the condition

$$\eta_{ab}\Lambda^a_c\Lambda^b_d = \eta_{cd}. \tag{6.4}$$

We can now define the connection 1-form as follows

$$\omega^a_b = \omega^a_{b\alpha} dx^\alpha \tag{6.5}$$

which can be regarded as a non-degenerate $n \times n$ matrix with 1-form entries. The connection 1-form is related to the connection coefficient, in local coordinate basis, by the following expression

$$\omega^a_b = \Gamma^a_{cb} dx^c. \tag{6.6}$$

The torsion 2-form is given by

$$T^a = de^a + \omega^a_b \wedge e^b. \tag{6.7}$$

Now, if the connection is torsion-free (which it usually is), then from 6.6 it can be seen to satisfy the following

$$\begin{aligned}\left(\omega^a_b \wedge e^b\right)_{\alpha\beta} &= \Gamma^a_{cb}\left(e^c_\alpha e^b_\beta - e^c_\beta e^b_\alpha\right) = 2\Gamma^a_{cb} e^c_{[\alpha} e^b_{\beta]} \\ &= -2D_{[\alpha} e^a_{\beta]} \\ &= -\left(de^a\right)_{\alpha\beta}\end{aligned} \tag{6.8}$$

This is called Cartan's first structural equation. The covariant derivative commutes with the process of contracting the tangent space indices a, b with η_{ab} provided that we have

$$D\eta_{ab} \equiv d\eta_{ab} - \omega^c_a \eta_{cb} - \omega^c_b \eta_{ac} = 0. \tag{6.9}$$

Realizing that the components of η_{ab} are constants, we obtain the equation of metric compatibility which reads

$$\omega_{ab} = -\omega_{ba}, \tag{6.10}$$

where $\omega_{ab} \equiv \eta_{ac}\omega^c_b$. Cartan's first structural equation 6.8 and the metric compatibility equation 6.10 help determine the connection 1-form ω^a_b uniquely. By definition, the exterior derivatives of the Vielbeins e^a are given by

$$de^a = -\frac{1}{2} C^a_{bc} e^b \wedge e^c. \tag{6.11}$$

Form this definition of the exterior derivative and Cartan's first structure equation, we find that the solution to ω_{ab} reads as follows

$$\omega_{ab} = \frac{1}{2}\left(C_{abc} + C_{acb} - C_{bca}\right), \tag{6.12}$$

where $C_{abc} \equiv \eta_{cd} C^d_{ab}$. A curvature 2-form is defined as follows indexcurvature 2-form

$$\Omega^a_b = d\omega^a_b + \omega^a_c \wedge \omega^c_b. \tag{6.13}$$

Taking the exterior derivative of connection 1-form 6.6 we obtain

$$d\omega^a_b = \Gamma^a_{cb,n} dx^n \wedge dx^c = \frac{1}{2}\left(\Gamma^a_{cb,n} - \Gamma^a_{cn,b}\right) dx^n \wedge dx^c \tag{6.14}$$

and the wedge of two connection 1-forms read

$$\begin{aligned} \omega^a_c \wedge \omega^c_b &= \Gamma^a_{cn} \Gamma^c_{bm} dx^n \wedge dx^m \\ &= \tfrac{1}{2}\left(\Gamma^a_{cn}\Gamma^c_{bm} - \Gamma^a_{cm}\Gamma^c_{bn}\right) dx^n \wedge dx^m. \end{aligned} \tag{6.15}$$

Substituting equations 6.15 and 6.14 in 6.13, we obtain

$$\begin{aligned} \Omega^a_b &= \tfrac{1}{2}\left(\Gamma^a_{cb,n} - \Gamma^a_{cn,b} + \Gamma^a_{cn}\Gamma^c_{bm} - \Gamma^a_{cm}\Gamma^c_{bn}\right) dx^n \wedge dx^m \\ &= \tfrac{1}{2} R^a_{bnm} dx^n \wedge dx^m \end{aligned} \tag{6.16}$$

Let E^n_a denote the inverse of the Vielbein e^a_n, satisfying $E^n_b e^a_n = \delta^a_b$. The Reimann tensor components can now be defined entirely in the tangent basis as

$$R^a_{bcd} \equiv E^n_c E^m_d R^a_{bnm}, \tag{6.17}$$

and hence, in terms of the tangent indices, we have

$$\Omega^a_b = \frac{1}{2} R^a_{bcd} e^c \wedge e^d, \tag{6.18}$$

which is called Cartan's second equation of structure. Consider a spherically symmetric metric with line element

$$ds^2 = -e^{2\mu(t,r)} dt^2 + e^{2\nu(t,r)} dr^2 + r^2 d\Omega_2^2, \tag{6.19}$$

We first introduce the orthonormal Vielbein basis[1]

[1] usually the script are hated, i.e., $e^{\hat{t}}$ should be used instead of e^t we are using here to denote that they are coordinate basis. We can afford to be a snobbish here.

$$e^t = e^{\mu} dt,\ e^r = e^{\nu} dr,\ e^{\theta} = r d\theta,\ e^{\phi} = r sin\theta d\phi, \tag{6.20}$$

whose exterior derivative reads

$$\begin{aligned}
de^t &= \dot{\mu} e^{\mu}\ dt \wedge dr + \mu' e^{\mu}\ dr \wedge dt = -\mu' e^{-\nu}\ e^t \wedge e^r \\
de^r &= \dot{\nu} e^{\nu}\ dt \wedge dr = -\dot{\nu} e^{-\mu}\ e^r \wedge e^t, \\
de^{\theta} &= dr \wedge d\theta = -\tfrac{1}{r} e^{-\nu}\ e^{\theta} \wedge e^r, \\
de^{\phi} &= sin\theta\ dr \wedge d\phi + r cos\theta\ d\theta \wedge d\phi = -\tfrac{1}{r} e^{-\nu}\ e^{\phi} \wedge e^r - \tfrac{1}{r} cot\theta\ e^{\phi} \wedge e^{\theta}.
\end{aligned} \tag{6.21}$$

Since we have $de^a = -\omega^a_b \wedge e^b$ to hold for a torsion-less connection, we can use these equations to directly read-off the connection 1-forms as follows

$$\begin{aligned}
&\omega^t_r = \mu' e^{-\nu}\ e^t,\ \omega^r_t = \dot{\nu} e^{-\mu}\ e^r, \\
&\omega^{\theta}_r = \tfrac{1}{r} e^{-\nu}\ e^{\theta},\ \omega^{\phi}_r = \tfrac{1}{r} e^{-\nu}\ e^{\phi},\ \omega^{\phi}_{\theta} = \tfrac{1}{r} cot\theta\ e^{\phi}.
\end{aligned} \tag{6.22}$$

With the connection 1-forms found, we can make use of equation 6.13 to curvature 2-forms, but before we do that let us write down the expressions for each 2-form and see what components of the wedge product we need to evaluate

$$\begin{aligned}
\Omega^t_r &= d\omega^t_r + \omega^t_t \wedge \omega^t_r + \omega^t_r \wedge \omega^r_r + \omega^t_{\theta} \wedge \omega^{\theta}_r + \omega^t_{\phi} \wedge \omega^{\phi}_r = d\omega^t_r, \\
\Omega^r_t &= d\omega^r_t + \omega^r_r \wedge \omega^r_t + \omega^r_t \wedge \omega^t_t + \omega^r_{\theta} \wedge \omega^{\theta}_t + \omega^r_{\phi} \wedge \omega^{\phi}_t = d\omega^r_t, \\
\Omega^{\theta}_r &= d\omega^{\theta}_r + \omega^{\theta}_{\theta} \wedge + \omega^{\theta}_r \omega^{\theta}_r \wedge \omega^r_r + \omega^{\theta}_{\phi} \wedge \omega^{\phi}_r + \omega^{\theta}_t \wedge \omega^t_r = d\omega^{\theta}_r + \omega^{\theta}_{\phi} \wedge \omega^{\phi}_r \\
&= d\omega^{\theta}_r - \underbrace{\omega^{\phi}_{\theta} \wedge \omega^{\phi}_r}_{\propto\ e^{\phi} \wedge e^{\phi} = 0}, \\
\Omega^{\phi}_r &= d\omega^{\phi}_r + \omega^{\phi}_t \wedge \omega^t_r + \omega^{\phi}_{\theta} \wedge \omega^{\theta}_r + \omega^{\phi}_{\phi} \wedge \omega^{\phi}_r + \omega^{\phi}_r \wedge \omega^r_r = d\omega^{\phi}_r + \omega^{\phi}_{\theta} \wedge e^{\theta}_r, \\
\Omega^{\phi}_{\theta} &= d\omega^{\phi}_{\theta} + \omega^{\phi}_{\phi} \wedge \omega^{\phi}_{\theta} + \omega^{\phi}_r \wedge \omega^r_{\theta} + \omega^{\phi}_t \wedge \omega^t_{\phi} + \omega^{\phi}_{\theta} \wedge \omega^{\theta}_{\theta} = d\omega^{\phi}_{\theta} - \omega^{\phi}_r \wedge \omega^{\theta}_r.
\end{aligned} \tag{6.23}$$

Upon careful computations, we would be able to obtain the following expressions for the curvature 2-form

$$
\begin{aligned}
\Omega^t_r &= (\mu'\nu' - \mu'' - \nu'')\, e^{-2\nu}\ e^t \wedge e^r, \\
\Omega^r_t &= (\dot\mu\dot\nu - \ddot\mu - \ddot\nu)\, e^{-2\mu}\ e^r \wedge e^t, \\
\Omega^\theta_r &= \tfrac{\nu'}{r} e^{-2\nu}\ e^\theta \wedge e^r + \tfrac{\dot\nu}{r} e^{-\mu-\nu}\ e^\theta \wedge e^t, \\
\Omega^\phi_r &= \tfrac{\nu'}{r} e^{-2\nu}\ e^\phi \wedge e^r + \tfrac{\dot\nu}{r} e^{\mu-\nu}\ e^\phi \wedge e^t, \\
\Omega^\phi_\theta &= \tfrac{1-e^{-2\nu}}{r^2}\ e^\phi \wedge e^\theta.
\end{aligned}
\tag{6.24}
$$

We can now compare these expressions to Cartan's second structure equation 6.18 and directly read off the Riemann tensors

$$
\begin{aligned}
R^t_{rtr} &= (\mu'\nu' - \mu'' - \nu'')\, e^{-2\nu},\ R^r_{trt} = (\dot\mu\dot\nu - \ddot\mu - \ddot\nu)\, e^{-2\mu}, \\
R^\theta_{r\theta r} &= \tfrac{\nu'}{r} e^{-2\nu},\ R^\theta_{t\theta t} = \tfrac{\dot\nu}{r} e^{-\mu-\nu}, \\
R^\phi_{r\phi r} &= \tfrac{\nu'}{r} e^{-2\nu},\ R^\phi_{r\phi t} = \tfrac{\dot\nu}{r} e^{-\mu-\nu},\ R^\phi_{\theta\phi\theta} = \tfrac{1-e^{-2\nu}}{r^2}.
\end{aligned}
\tag{6.25}
$$

These components of the Riemann tensor are nothing but Ricci tensor components in disguise. Consider R^t_{rtr}, we can write the covariant description as $R_{trtr} = \eta_{tt} R^t_{rtr}$ and finally the Ricci tensor as $R_{rr} = \eta^{tt} R_{trtr} = R^t_{rtr}$. It is to be noted here that we are raising and lowering indices using the Minkowski metric, η_{ab}, since the indices are not holonomic. For a diagonal metric, as we have here, the holonomic components of a $(1,1)$-tensor coincide with the non-holonomic components. This is the power of the curvature 2-form method. No need to calculate Christoffel symbols and more importantly, there is no need of blind calculations since the only components we solve for are the non-null components.

6.1.1 Curvature 2-form Bianchi Identity

For every tensor valued p-form Ψ of type (n,m) there exists a unique tensor valued $(p+1)$-form $D\Psi$, which is also of type (n,m) and has the following components with respect to a basis $\{e^a\}$

$$
(D\Psi)^{a_1\ldots a_2}_{b_1\ldots b_m} = d\Psi^{a_1\ldots a_2}_{b_1\ldots b_m} + \omega^{a_1}_{c_1} \wedge \Psi^{c_1 a_2 \ldots a_n}_{b_1\ldots b_m} + \ldots + \omega^{a_n}_{c_n} \wedge \Psi^{a_1\ldots c_n}_{b_1\ldots b_m} - \omega^{c_1}_{b_1} \wedge \Psi^{a_1\ldots a_n}_{c_1 b_2\ldots b_m} \cdots
$$
$$
-\omega^{c_m}_{b_m} \wedge \Psi^{a_1\ldots a_n}_{b_1\ldots c_m}. \tag{6.26}
$$

Now, consider the form $D\Omega^a_b$ which can be written as

$$
D\Omega^a_b = d\Omega^a_b + \omega^a_c \wedge \Omega^c_b - \omega^c_b \wedge \Omega^a_c \tag{6.27}
$$

Using 6.13, we can write the exterior derivative of the curvature 2-form as follows

$$d\omega^a_b = d\left(d\omega^a_b + \omega^a_c \wedge \omega^c_b\right) = d\left(d\omega^a_b\right) + d\omega^a_c \wedge \omega^c_b - \omega^c_b \wedge d\omega^a_c, \tag{6.28}$$

where $d\left(d\omega^a_b\right) = 0$. Substituting this into 6.27 we obtain the identity

$$D\Omega^a_b = 0. \tag{6.29}$$

This corresponds to the differential Bianchi identity.

Exercise 16

1. Show that the Schwarzschild metric has the following structure coefficients

$$C^t_{tr} = \frac{r_s}{2r^2\sqrt{1-\frac{r_s}{r}}},\ C^\theta_{r\theta} = C^\phi_{r\phi} = -\frac{1}{r}\sqrt{1-\frac{r_s}{r}},\ C^\phi_{\theta\phi} = \frac{cot\theta}{r}. \tag{6.30}$$

2. The AdS_4 metric in global coordinates has the line element

$$ds^2 = -\left(1+\frac{r^2}{l^2}\right)dt^2 + \left(1+\frac{r^2}{l^2}\right)^{-1}dr^2 + r^2 d\Omega^2_2, \tag{6.31}$$

where l is called the AdS radius. You can solve this by first introducing the following orthonormal Vielbein basis

$$e^0 = \left(1+\frac{r^2}{l^2}\right)^{\frac{1}{2}}dt,\ e^1 = \left(1+\frac{r^2}{l^2}\right)^{-\frac{1}{2}}dr,\ e^2 = rd\theta,\ e^3 = rsin\theta d\phi. \tag{6.32}$$

a. Find the structure coefficients and use them to find the connection 1-forms.
b. Find the curvature 2-forms and hence, the Riemann tensor components for this metric.

3. Consider a spacetime in spherical coordinates $(t, r, \vartheta, \varphi)$ described by the following line element

$$ds^2 = -\alpha^\gamma c^2 dt^2 + \alpha^{-\gamma}dr^2 + r^2\alpha^{-\gamma+1}\left(d\vartheta^2 + sin^2\vartheta\ d\varphi^2\right), \tag{6.33}$$

where $\alpha = 1 - \frac{r_s}{\gamma r}$. This is called the Janis-Newman-Winicour spacetime.
a. Define a Veilblein basis and show that the structure constants of this metric are as follows

$$\begin{aligned} C^t_{tr} &= \frac{r_s}{2r^2}\alpha^{\frac{(\gamma-2)}{2}},\ C^\varphi_{\vartheta\varphi} = -\frac{cot\vartheta}{r}\alpha^{\frac{(\gamma-1)}{2}}, \\ C^\vartheta_{r\vartheta} &= C^\varphi_{r\varphi} = -\frac{2\gamma r - r_s(\gamma+1)}{2\gamma r^2}\alpha^{\frac{(\gamma-2)}{2}}. \end{aligned} \tag{6.34}$$

b. Using the structure coefficients compute the connection 1-forms following 6.12.
c. Show that the only non-zero component of the Ricci tensor is $R_{rr} = -\frac{r_s(\gamma^2-1)}{2\gamma^2 r^4 \alpha^2}$ and find the Ricci scalar.

d. Find the Kretschmann scalar and show that it matches the Kretschmann scalar of the Schwarzschild case when we set $\gamma = 1$. This is true since when $\gamma = 1$, the line element reduces to that of Schwarzschild.

4. The curvature 2-form method works well even for non-diagonal metrics. Consider the TaubNUT metric which describes a spinning black hole. The line element, in Boyer-Lindquist like spherical coordinates $(t, r, \vartheta, \varphi)$ reads

$$ds^2 = -\frac{\Delta}{\Sigma}\left(dt + 2l cos\vartheta \; d\varphi\right)^2 + \Sigma\left(\frac{dr^2}{\Delta} + d\vartheta^2 + sin^2\vartheta \; d\varphi^2\right), \tag{6.35}$$

where $\Sigma = r^2 + l^2$ and $\Delta = r^2 - 2Mr - l^2$. We have made use of natural units here. M here is the mass of the black hole and l characterizes the magnetic monopole strength of the black hole.
a. Find the connection 1-forms using the following orthonormal Vielblein basis

$$e^0 = \sqrt{\frac{\Delta}{\Sigma}}\left(dt + 2l cos\vartheta \; d\varphi\right), \; e^1 = \sqrt{\frac{\Sigma}{\Delta}} \; dr, \; e^2 = \sqrt{\Sigma} \; d\vartheta, \; e^3 = \sqrt{\Sigma} sin\vartheta \; d\varphi.$$

b. Find the Curvature 2-forms and show that the Ricci scalar is $R = 0$.

6.2 Further Discussions on The Einstein Equations

Consider the *principle part*, i.e., the part containing the highest derivatives of the metric (which is the 2^{nd} derivative) is

$$P\left\{R_{\mu\nu}\right\} = \frac{1}{2} g^{\alpha\beta}\left\{\partial_\mu \partial_\alpha g_{\beta\nu} + \partial_\nu \partial_\alpha g_{\beta\mu} - \partial_\mu \partial_\nu g_{\alpha\beta} - \partial_\alpha \partial_\beta g_{\mu\nu}\right\} \tag{6.36}$$

The *character* of the Einstein equations as reflected in their symbol, which is defined by replacing in the principal part $\partial_\mu \partial_\nu g_{\alpha\beta}$ by $\xi_\mu \xi_\nu \dot{g}_{\alpha\beta}$, where ξ_μ are the components of a covector and $\dot{g}_{\alpha\beta}$ (the components of a possible variation of g). Then, for a given background metric metric g, ρ_ξ at a point $p \in \mathcal{M}$ and the covector $\xi \in \; ^*T_p\mathcal{M}$ are given by

$$\left(\rho_\xi \cdot \dot{g}\right)_{\alpha\beta} = \frac{1}{2} g^{\alpha\beta}\left(\xi_\mu \xi_\alpha \dot{g}_{\beta\nu} + \xi_\nu \xi_\alpha \dot{g}_{\beta\mu} - \xi_\mu \xi_\nu \dot{g}_{\alpha\beta} - \xi_\alpha \xi_\beta \dot{g}_{\mu\nu}\right), \tag{6.37}$$

let $\left(i_\xi \dot{g}\right)_\nu = g^{\alpha\beta} \xi_\alpha \dot{g}_{\beta\nu}$, $(\xi, \xi) = g^{\alpha\beta} \xi_\alpha \xi_\beta$, $(\xi \otimes \lambda)_{\mu\nu} = \xi_\mu \lambda_\nu$, and $g^{\alpha\beta} \dot{g}_{\alpha\beta} = Tr \; \dot{g}$, then

$$\{\rho_\xi \cdot \dot{g}) = \frac{1}{2}\left(\xi \otimes i_\xi \dot{g} + i_\xi \dot{g} \otimes \xi - Tr\ \dot{g}\xi \otimes \xi - (\xi,\xi)\,\dot{g}\right\}. \tag{6.38}$$

6.2.1 The Symbol of a system of Euler-Lagrange Equations

Let x denote the independent variables x^α, where $\alpha = 1, ..., n$, let q denote the dependent variables q^b, where $b = 1, ..., m$, and let v denote the first derivatives of dependent variables v^b_α, $n \times m$ matrices. Then the Lagrangian $\mathcal{L}$, a function of (x, q, v), is given by

$$\mathcal{L} = \mathcal{L}\,(x, q, v)\,, \tag{6.39}$$

now, a set of functions $u^b(x) : b = 1, ..., m$ is a solution of the Euler–Lagrange equations, if $q^b = u^b(x)$, and $v^b_\alpha = \frac{\partial v^b}{\partial x^\alpha}(x)$ gives the following equation

$$\frac{\partial}{\partial x^\alpha}\left(\frac{\partial \mathcal{L}}{\partial v^b_\alpha}(x, u(x), \partial u(x))\right) - \frac{\partial \mathcal{L}}{\partial q^b}\,(x, u(x), \partial u(x)) = 0. \tag{6.40}$$

We can condense the Euler-Lagrange equations by defining $p^\alpha_b = \frac{\partial \mathcal{L}}{\partial v^b_\alpha}$, and $g_b = \frac{\partial \mathcal{L}}{\partial q^b}$, to give[2]

$$\frac{\partial p^\alpha_b}{\partial x^\alpha} = g_b. \tag{6.41}$$

Now, the principle part of the Euler-Lagrange equations is given by

$$h^{\alpha\beta}_{bc}\frac{\partial^2 u^c}{\partial x^\alpha \partial x^\beta}(x, u(x), \partial u(x)), \tag{6.42}$$

where $h^{\alpha\beta}_{bc} = \frac{\partial^2 \mathcal{L}}{\partial v^b_\alpha \partial v^c_\beta}(x, q, v)$. The equations of variation are linearized equations, satisfied by a variation via solutions as follows: If $\dot{u}^b$ denotes the variations of the functions u^b, the principle part of the linearized equations is as follows

$$h^{\alpha\beta}_{bc}(x, u(x), \partial u(x))\frac{\partial^2 \dot{u}^c}{\partial x^\alpha \partial x^\beta}. \tag{6.43}$$

Considering oscillatory solutions of the equations of variations, such as

[2] where q, v, p, g are analogous to position, velocity, momentum and force respectively, in classical mechanics.

$$\dot{u}^b = \dot{k}^b e^{i\theta}, \tag{6.44}$$

and substituting (θ/δ) for θ in the *high-frequency limit* of $\delta \to 0$, we get

$$h^{\alpha\beta}_{bc}(x, u(x), \partial u(x))\dot{k}^c \frac{\partial\theta}{\partial x^\alpha}\frac{\partial\theta}{\partial x^\beta} = 0. \tag{6.45}$$

Notice that the LHS of the above equation, after replacing α with μ, represents the symbol $\rho_\xi \cdot \dot{k}$, where $\xi_\mu = \frac{\partial\theta}{\partial x^\mu}$. Generalizing, the symbol of the Euler-Lagrange is given by

$$(\rho_\xi \cdot \dot{u})^b = h^{\alpha\beta}_{bc}\xi_\alpha\xi_\beta\dot{u}^c = \zeta_{bc}(\xi)\dot{u}^c, \tag{6.46}$$

where $\zeta_{bc}(\xi) = h^{\alpha\beta}_{bc}\xi_\alpha\xi_\beta$ is an $n \times m$ matrix whose entries are homogeneous quadratic polynomials in ξ. The meaning of the above operations can be understood from a global perspective as follows: $x^\alpha, \alpha = 1, ..., n$ are local coordinates on an n-dimensional manifold $\mathcal{M}$ and x denotes an arbitrary point on $\mathcal{M}$, while $q^b, b = 1, ..., m$ are local coordinates on an m-dimensional manifold $\mathcal{N}$ and q denotes an arbitrary point on $\mathcal{N}$. The unknown u is the mapping $u : \mathcal{M} \to \mathcal{N}$ and the functions $u^b(x), b = 1, ..., m$ describe this mapping in terms of the given local coordinates.
If $\mathcal{M}$ is an n-dimensional manifold, the characteristic subset, $C^*_x \in {}^*T_x\mathcal{M}$ is defined as follows

$$C^*_x = \xi \neq 0 \in {}^*T_x\mathcal{M} : \textit{null space of} \rho_\xi \neq 0; \; = \xi \neq 0 \in {}^*T_x\mathcal{M} : Det\zeta(\xi) = 0. \tag{6.47}$$

Hence, $\xi \in C^*_x$ if and only if $\xi \neq 0$ and the null space of ρ_ξ is non-trivial. A very basic example of an Euler-Lagrange equation with a non-empty characteristic is the linear wave equation

$$\partial_\alpha\partial^\alpha u = g^{\mu\nu}\nabla_\mu(\partial_\nu u) = 0, \tag{6.48}$$

which arises from the following Lagrangian

$$\mathcal{L} = \frac{1}{2}g^{\mu\nu}v_\mu v_\nu. \tag{6.49}$$

The symbol is $\rho_\xi \cdot \dot{u} = (g^{\mu\nu}\xi_\mu\xi_\nu)\,\dot{u}$ and the characteristic is given as follows

$$C^*_x = \{\xi \neq 0 \in {}^*T_x\mathcal{M} : (\xi, \xi) = g^{\mu\nu}\xi_\mu\xi_\nu = 0\}, \tag{6.50}$$

i.e., C^*_x is the null cone in the cotangent space ${}^*T_x\mathcal{M}$ associated to the metric g.

6.2.2 Back to symbol for the Einstein Equations

By setting $\dot{g} = \lambda \otimes \xi + \xi \otimes \xi$ for an arbitrary covector $\lambda \in {}^{*}T_x\mathcal{M}$. Then,

$$i_\xi \dot{g} = (\lambda, \xi)\, \xi + (\xi, \xi)\lambda = g^{\mu\nu}\lambda_\mu \xi_\nu + (\xi, \xi)\lambda, \tag{6.51}$$

and,

$$Tr\ \dot{g} = 2(\lambda, \xi). \tag{6.52}$$

We observe that

$$\rho_\xi \cdot \dot{g} = 0. \tag{6.53}$$

Thus, the null space of ρ_ξ is a non-trivial covector ξ. This degeneracy is due to the fact that the equations are generally covariant, i.e., if g is a solution of the Einstein equations and f is a diffeomorphism of the manifold onto itself, then the pullback, defined by $f * g$, is also a solution. If X is a vector field on $\mathcal{M}$, then X generates a 1-parameter group $\{f_l\}$ of diffeomorphisms of $\mathcal{M}$ and

$$\mathcal{L}_X g = \frac{d}{dl} f_l^* g|_{l=0}, \tag{6.54}$$

the Lie derivative (with respect to X of g), is a solution of the linearized equations. From the introductory chapter, we know from section 1.4.2 that the Lie derivative is defined as follows

$$\left(\mathcal{L}_X g\right)(Y, Z) = g\left(\nabla_Y X, Z\right) + g\left(\nabla_Z X, Y\right). \tag{6.55}$$

Setting $Y = e_\mu$ and $Z = e_\nu$, where $e_\mu : \mu = 0, 1, 2, 3$ is an arbitrary frame,

$$\left(\mathcal{L}_X g\right)_{\mu\nu} = \nabla_\mu X_\nu + \nabla_\nu X_\mu, \tag{6.56}$$

where $X_\mu = g_{\mu\alpha}X^\alpha$, and the symbol of the Lie derivative id given by

$$\dot{g}_{\mu\nu} = \xi_\mu \lambda_\nu + \xi_\nu \lambda_\mu, \tag{6.57}$$

where $\lambda_\mu = \dot{X}_\mu$. In General Relativity we must factor out solutions of the form $\mathcal{L}_X g = \dot{g}$ for any vector field X. Thus, consider the equivalence relation

$$\dot{g}_1 \sim \dot{g}_2 \iff \dot{g}_2 = \dot{g}_1 + \lambda \otimes \xi + \xi \otimes \xi, \lambda \in {}^{*}T_x\mathcal{M}, \tag{6.58}$$

which gives the quotient space[3] $\mathcal{Q}$, of dimension $4-1=3$. Consider the null space ρ_ξ (with ρ_ξ defined on $\mathcal{Q}$). We can now distinguish two cases according as to whether the covector ξ satisfies $(\xi,\xi)\neq 0$ or $(\xi,\xi)=0$.

Case I: $(\xi,\xi)\neq 0$. If ξ is not null, then $\rho_\xi\cdot\dot{g}=0$ implies that

$$\dot{g}=\lambda\otimes\xi+\xi\otimes\xi, \tag{6.59}$$

where $\lambda=\frac{\left(i_\xi\dot{g}-\frac{1}{2}Tr\ \dot{g}\xi\right)}{(\xi,\xi)}$. Thus, ρ_ξ has only trivial null spaces on $\mathcal{Q}$.

Case II: $(\xi,\xi)=0$. If ξ is null, we can choose ξ in the same component of the null cone $\mathcal{N}_x^*\in\ {}^*T_x\mathcal{M}$ such that $(\xi,\xi)=-2$ and there is then a unique representative $\dot{g}$ in each equivalence class $\{\dot{g}\}\in\mathcal{Q}$ such that

$$i_{\bar{\xi}}\dot{g}=0. \tag{6.60}$$

Thus,

$$\rho_\xi\cdot\dot{g}=0\iff\xi\otimes i_\xi\dot{g}+i_\xi\dot{g}\otimes\xi-\xi\otimes\xi Tr\ \dot{g}=0. \tag{6.61}$$

Upon taking the inner product with $\bar{\xi}$, we observe that $(i_\xi\dot{g},\xi)=\left(i_{\bar{\xi}}\dot{g},\xi\right)=0$, hence

$$-2i_\xi\dot{g}+2\xi Tr\ \dot{g}=0. \tag{6.62}$$

Again, taking the inner product with $\bar{\xi}$ yields

$$-4Tr\ \dot{g}=0\Longrightarrow Tr\ \dot{g}=0, \tag{6.63}$$

and substituting this gives

$$i_\xi\dot{g}=0. \tag{6.64}$$

Therefore, if $\xi\in\mathcal{N}_x^*$, then we can conclude that the null space of ρ_ξ can be identified with the space of trace-free quadratic forms on the 2-dimensional

[3] The quotient space (also called factor spaces) $\mathcal{Q}$ of a topological space $\mathcal{M}$ and an equivalence relation $\sim$ on $\mathcal{M}$ is the set of equivalence classes of points in M (under the equivalence relation $\sim$) together with the following topology given to subsets of $\mathcal{Q}$: a subset U of $\mathcal{M}$ is called open if and only if $\cup_{b\in U}b$ is open in $\mathcal{M}$. This can be stated in terms of maps as follows: if $q:\mathcal{M}\to\mathcal{Q}$ denotes the map that sends each point to its equivalence class in $\mathcal{Q}$, the topology on $\mathcal{Q}$ can be specified by prescribing that a subset of $\mathcal{Q}$ is open if and only if $q^{-1}[\textit{the set}]$ is open.

spacelike plane Σ, the g-orthogonal complement of the linear span of ξ and $\bar{\xi}$. This is the space of gravitational degrees of freedom at a point (two polarizations) [4].

6.3 The Minkowskian Approximation

Consider weak gravitational fields, say, the gravitational field from a far away gravitational wave producing source. In this region, any change in the matter distribution, i.e., in $T_{\mu\nu}$, will induce a change in the gravitational eld, which will be recorded as a change in metric which is expressed as follows

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{6.65}$$

where, $\eta_{\mu\nu} = Diag(-1, 1, 1, 1)$, and $h_{\mu\nu}$ is a tensor describing the variations induced in the spacetime metric. It can be thought of as a small perturbation in the otherwise flat spacetime. We now proceed with the *linearization approach* which assumes that the tensor $h_{\mu\nu}$ is small ($|h_{\mu\nu}| << 1$), so that we need only keep terms linear in $h_{\mu\nu}$ in calculations. Upon calculation, the Einstein equations in vacuum have the following form

$$R_{\mu\nu} \approx -\frac{1}{2}\eta^{\alpha\beta}\partial^2_{\alpha\beta}h_{\mu\nu} + \frac{1}{2}\partial_\mu\xi_\mu + \frac{1}{2}\partial_\nu\xi_\mu = 0, \tag{6.66}$$

where

$$\xi_\mu = \partial_\alpha h^\alpha_{mu} - \frac{1}{2}\partial_\mu h^\alpha_\alpha. \tag{6.67}$$

When $\xi_\mu = 0$, the linear eld equations in vacuum have the form

$$\eta^{\alpha\beta}\partial^2_{\alpha\beta}h_{\mu\nu} = 0. \tag{6.68}$$

These are a system of waves equations for h, the perturbation. They can be expressed alternatively as

$$\left[-\frac{\partial^2}{\partial t^2} + \nabla^2\right] h_{\mu\nu} \equiv \partial_\lambda\partial^\lambda h_{\mu\nu} = 0, \tag{6.69}$$

which is the 3-dimensional wave equation. The equations $\xi_\mu = 0$ are interpreted as *polarization conditions* satisfied by gravitational waves. The simplest solution to the wave equation is a plane wave solution of the form

$$h_{\mu\nu} = S_{\mu\nu}e^{ik_\lambda x^\lambda}, \tag{6.70}$$

[4] Conversely, $i_{\bar{\xi}}\dot{g} = i_\xi\dot{g} = 0$ and $Tr\ \dot{g} = 0$ implies that $\dot{g}$ lies in the null space of ρ_ξ

where $S_{\mu\nu}$ is a constant symmetric tensor called the polarization tensor. This is where the information about the amplitude and the polarization of the waves is encoded. k_λ is a constant vector called the *wave vector* which determines the propagation direction of the wave and its frequency. It is to be noted that only the real part of the wave solution is used in physical applications.

6.4 Hilbert's Gauge Condition

Consider again the weak gravitational field with a small perturbation. With the linearization approach we can calculate the Christoffel symbols to be (the raising and lowering of indices done by $\eta_{\mu\nu}$)

$$\Gamma^\gamma_{\alpha\beta} = \frac{1}{2}\eta^{\gamma\lambda}\left(\partial_\alpha h_{\lambda\beta} + \partial_\beta h_{\lambda\alpha} - \partial_\lambda h_{\beta\alpha}\right) = \frac{1}{2}\left(\partial_\alpha h^\gamma_\beta + \partial_\beta h^\gamma_\alpha - \partial^\gamma h_{\beta\alpha}\right). \quad (6.71)$$

The curvature tensor is found to be

$$\begin{aligned} R^\gamma_{\beta\alpha\lambda} &= \tfrac{1}{2}\partial_\alpha\left(\partial_\lambda h^\gamma_\beta + \partial_\beta h^\gamma_\lambda - \partial^\gamma h_{\beta\lambda}\right) - \tfrac{1}{2}\partial_\lambda\left(\partial_\alpha h^\gamma_\beta + \partial_\beta h^\gamma_\alpha - \partial^\gamma h_{\beta\alpha}\right) \\ &= \tfrac{1}{2}\left(\partial_\alpha\partial_\beta h^\gamma_\lambda + \partial_\lambda\partial^\gamma h_{\beta\alpha} - \partial_\alpha\partial^\lambda h_{\beta\lambda} - \partial_\lambda\partial_\beta h^\lambda_\alpha\right). \end{aligned} \quad (6.72)$$

We can now contract the above equation on two of the indices by multiplying η^λ_γ throughout the equation to obtain the Ricci tensor and scalar

$$\begin{aligned} R_{\beta\alpha} &= \tfrac{1}{2}\left(\partial_{\alpha\partial\beta}h + \partial_\zeta\partial^\zeta h_{\beta\alpha} - \partial_\alpha\partial_\lambda h^\lambda_\beta - \partial_\lambda\partial_\beta h^\lambda_\alpha\right), \\ R &= R^\beta_\beta = \eta^{\beta\alpha}R_{\beta\alpha} = -\partial_\zeta\partial^\zeta h_{\beta\alpha} + \partial_\lambda\partial_\beta h^{\beta\lambda}. \end{aligned} \quad (6.73)$$

We can now use these equation to obtain the Einstein tensor, $G_{\beta\alpha}$, and write Einstein's equations in the following form

$$\partial_\alpha\partial_\beta h + \partial_\zeta\partial^\zeta h_{\beta\alpha} - \partial_\alpha\partial_\lambda h^\lambda_\beta - \partial_\lambda\partial_\beta h^\lambda_\alpha - \eta_{\beta\alpha}\left(\partial_\zeta\partial^\zeta h_{\beta\alpha} + \partial_\lambda\partial_\beta h^{\beta\lambda}\right) = -16\pi\kappa T_{\beta\alpha}. \quad (6.74)$$

It is convenient to change the variables $h_{\beta\alpha}$ to $\bar{h}_{\beta\alpha}$, whose relation is defined as follows

$$\bar{h}_{\beta\alpha} = h_{\beta\alpha} - \frac{1}{2}\eta_{\beta\alpha}h. \quad (6.75)$$

Rewriting the field equations in terms of the newly defined variable produces

$$\partial_\zeta\partial^\zeta\bar{h} + \eta_{\beta\alpha}\partial_\lambda\partial_\gamma\bar{h}^{\lambda\gamma} - \partial_\alpha\partial_\lambda\bar{h}^\lambda_\beta - \partial_\beta\partial_\lambda\bar{h}^\lambda_\alpha = -16\pi\kappa T_{\beta\alpha} \quad (6.76)$$

Notice that this equation is invariant under the following gauge transformation

$$h'_{\beta\alpha} = h_{\beta\alpha} - \partial_\beta \xi_\alpha - \partial_\alpha \xi_\beta, \tag{6.77}$$

which leaves the curvature tensor $R^\gamma_{\beta\alpha\lambda}$ invariant and can be expressed as

$$R_{\mu\beta\nu\alpha} \approx g_{\mu\delta} \left(\partial_\nu \Gamma^\delta_{\beta\alpha} - \partial_\alpha \Gamma^\delta_{\beta\nu} \right) \approx 2 \left(\partial_\nu \Gamma_{\mu\beta\alpha} - \partial_\alpha \Gamma_{\mu\beta\nu} \right). \tag{6.78}$$

By explicit substitution, the curvature tensor takes the following form

$$R_{\mu\beta\nu\alpha} - \xi_{\mu[,\alpha,\nu],\beta} + \xi_{\beta[,\alpha,\nu],\mu} - \xi_{\alpha[,\mu,\nu],\beta} + \xi_{\alpha[,\beta,\nu],\mu}. \tag{6.79}$$

Observe that in the above equation the partial derivatives commute with each other, the additional terms cancel out and the curvature remains invariant the gauge transformation. Further note that under a coordinate shift, $x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)$, the metric changes by $-\partial_{(\mu}\xi_{\nu)}$ [5]. Similarly, the change in the newly defined variable $\bar{h}^{\beta\alpha}$ is

$$\bar{h}'^{\beta\lambda} = \bar{h}^{\beta\lambda} - \partial^\beta \xi^\lambda - \partial^\lambda \xi^\beta + \eta^{\beta\lambda} \partial_\gamma \xi^\gamma. \tag{6.80}$$

We can now choose four functions ξ_α to put four conditions on the tensor field $h_{\beta\alpha}$ as follows

$$\partial_\lambda h'^{\beta\lambda} = 0, \tag{6.81}$$

which is called the *Hilbert's gauge condition* (or sometimes the *de Donder's gauge conditions* or the *Fock's gauge conditions* or just simply the *Einstein's gauge conditions*). When applied to the obtained Einstein's equation in terms of the newly defined variable, we get

$$\partial_\zeta \partial^\zeta \bar{h}^{\beta\alpha} = -16\pi\kappa T^{\beta\alpha}, \tag{6.82}$$

which in vacuum with no sources becomes $\partial_\zeta \partial^\zeta \bar{h}^{\beta\alpha} \equiv \left[-\frac{\partial^2}{\partial t^2} + \nabla^2 \bar{h}^{\beta\alpha} \right] = 0$. Each component of the perturbation satisfies the wave equation. Hence, all the components of the metric just move down the axis like waves with the speed of light.

[5] this is true in a flat space time. However, in an arbitrary curved spacetime with metric tensor defined as $g'_{\alpha\beta} = g_{\alpha\beta} + \delta g_{\alpha\beta}$, the change in the metric is expressed as $\delta g_{\alpha\beta} = -D_{(\beta}\xi_{\alpha)}$

6.5 Plane Waves

Let us consider the vacuum solutions to $\partial_\zeta \partial^\zeta \bar{h}^{\mu\nu} = -16\pi\kappa T^{\mu\nu}$. The study of this is of prime importance, at least to *homo sapiens* since the detection of these gravitational waves is taking place far away from its sources. In vacuum, Einstein's equations are expressed as $\partial_\zeta \partial^\zeta \bar{h}^{\mu\nu} = 0$, with supplementary gauge conditions called the Hilbert's gauge conditions, $\bar{h}^{\mu\nu}_{,\nu}$. The solution to these equations (for distances far from the source) can be expressed as a superposition of plane waves

$$\bar{h}_{\mu\nu} = \mathbf{Re}\left(H_{\mu\nu} e^{ik_\lambda x^\lambda}\right), \tag{6.83}$$

with additional constraints

$$k^\nu H_{\mu\nu} = 0, \tag{6.84}$$

that implies that the waves are transverse. Here k^ν, as explained previously, is the real wavevector and $H_{\mu\nu}$ is a constant symmetric complex matrix describing the polarization of the wave. The wave equation reduces to

$$k_\rho k^\rho = 0, \tag{6.85}$$

i.e., the solution describes a wave with frequency $\omega \equiv k^0 = \sqrt{k_x^2 + k_y^2 + k_z^2}$ which propagates with the speed of light. Consider the null vector defined by $k^\rho = \omega(1,0,0,1)$. Now, $e^{ik_\rho x^\rho} = e^{-i\omega\left(t - \frac{z}{c}\right)}$ describes a wave of frequency ω propagating in along the z-axis with the speed of light. $H_{\rho 1} = H_{\rho 2} = 0$ and the transverse condition reduces to

$$H_{\rho 0} + H_{\rho 3} = 0. \tag{6.86}$$

It is to be noted that the Hilbert's gauge condition does not eliminate all gauge freedom. Consider the gauge transformation $h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\xi_{\nu)}$, imposing the gauge condition $\partial^\nu \bar{h}_{\mu\nu} = 0$, we observe

$$\partial^\nu \bar{h}_{\mu\nu} \to \partial^\nu \bar{h}_{\mu\nu} + \partial^\nu \partial_\nu \xi_\mu, \tag{6.87}$$

this would preserve the gauge condition if and only if ξ_μ obeys the following wave equation

$$\partial^\nu \partial_\nu \xi_\mu = 0. \tag{6.88}$$

Thus, we can conclude that there is a residual gauge freedom which we can use to simplify the solution. Consider a simple solution to the above wave equation

$$\xi_\mu(x) = A_\mu e^{ik_\zeta x^\zeta}, \tag{6.89}$$

now using $\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \eta_{\mu\nu}\partial^\zeta \xi_\zeta$, we observe that the residual gauge freedom in this particular case is

$$H_{\mu\nu} \to H_{\mu\nu} + i\left(k_\mu A_\nu + k_\nu A_\mu - \eta_{\mu\nu} k^\zeta A_\zeta\right), \tag{6.90}$$

which can be used to achieve the longitudinal gauge

$$H_{0\mu} = 0. \tag{6.91}$$

But notice that this still doesn't determine A_μ uniquely, which is fixed when we use the gauge condition to set

$$H^\mu_\mu = 0. \tag{6.92}$$

This is known as the *trace-free condition*. Now, The longitudinal gauge condition combined with the transversality condition yields $H_{3\mu} = 0$. Upon using the trace free conditions we get

$$H_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & H_+ & H_\times & 0 \\ 0 & H_\times & H_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{6.93}$$

where the solution is specified by the two constants H_+ and $H_\times$ corresponding to two independent polarizations. Thus, gravitational waves are transverse and have two possible polarizations. This is what we mean when we say that the gravitational field has two degrees of freedom per spacetime point[6].

6.6 Petrov-Penrose Classification

Since the *Weyl tensor* is a measure of curvature of a pseudo-Riemannian manifold, a spacetime is said to be *conformally flat* if its Weyl tensor vanishes, i.e., if $W_{\alpha\beta\gamma\delta} = 0$. Generally, gravitational fields are classified in accordance to the *Petrov-Penrose classification* of their corresponding Weyl tensor. This is an algebraic classification based on the idea that the curvature tensor can be thought of as a 6×6 matrix and the reduction of these matrix naturally results in general categories of curvature tensors.

[6] The same conclusion was obtained in a highly mathematical fashion in section 6.2

6.6.1 Matrix Representation of the Curvature Tensor

From the symmetries of the *Riemann curvature tensor*, we can write it as $R^{\alpha\beta}_{\gamma\delta}$ and associate an index $\mathcal{I} = 1, 2, ..., 6$ with each pairs 01, 02, 03, 23, 31, 12 of independent values that $\alpha\beta$ and $\gamma\delta$ can take. The curvature tensor can be expressed as a 6×6, $\mathcal{M}^{\mathcal{I}}_{\mathcal{K}}$ matrix as given below

$$\left(\begin{array}{ccc|ccc} R^{01}_{01} & R^{01}_{02} & R^{01}_{03} & R^{01}_{23} & R^{01}_{31} & R^{01}_{12} \\ R^{02}_{01} & R^{02}_{02} & R^{02}_{03} & R^{02}_{23} & R^{02}_{31} & R^{02}_{12} \\ R^{03}_{01} & R^{03}_{02} & R^{03}_{03} & R^{03}_{23} & R^{03}_{31} & R^{03}_{12} \\ \hline R^{23}_{01} & R^{23}_{02} & R^{23}_{03} & R^{23}_{23} & R^{23}_{31} & R^{23}_{12} \\ R^{31}_{01} & R^{31}_{02} & R^{31}_{03} & R^{31}_{23} & R^{31}_{31} & R^{31}_{12} \\ R^{12}_{01} & R^{12}_{02} & R^{12}_{03} & R^{12}_{23} & R^{12}_{31} & R^{12}_{12} \end{array}\right) \quad (6.94)$$

$\mathcal{M}^{\mathcal{I}}_{\mathcal{K}}$ can be alternatively written in it's more enlightening form as follows

$$\mathcal{M}^{\mathcal{I}}_{\mathcal{K}} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ -\mathcal{B}^T & \mathcal{C} \end{pmatrix}, \quad (6.95)$$

where $\mathcal{A}, \mathcal{B}, \mathcal{C}$, are 3×3 matrices. Notice that the trace of matrix $\mathcal{B}$ is null. This can be shown by first lowering the index and making use of the property of the *Levi-Civita symbol* as follows

$$Tr\ \mathcal{B} = R^{01}_{23} + R^{02}_{31} + R^{03}_{12} = \epsilon^{011} R_{123} + \epsilon^{022} R_{231} + \epsilon^{033} R_{312}. \quad (6.96)$$

In the above equation, ϵ is called the the three-dimensional standard for of the Levi-Civita symbol (see figure 6.1). The following equation accounts for it's property

$$\epsilon^{ijk} = \begin{cases} +1, & if\ (i,j,k)\ is\ (1,2,3), (2,3,1),\ or\ (3,1,2) \\ -1, & if\ (i,j,k)\ is\ (3,2,1), (1,3,2),\ or\ (2,1,3) \\ 0, & if\ i = j,\ or,\ j = k,\ or\ k = i \end{cases} \quad (6.97)$$

Now, from the above properties, that when $j = k$ in the three-dimensional standard for of the Levi-Civita symbol, $\epsilon^{ijk} = \epsilon^{ikk} = \epsilon^{ijj} = 0$, i.e., $\epsilon^{011} = \epsilon^{022} = \epsilon^{033} = 0$. Thus, we obtain the result

$$Tr\ \mathcal{B} = 0. \quad (6.98)$$

Also, notice that the matrices $\mathcal{A}$ and $\mathcal{C}$ are equal to their transposes, i.e., $\mathcal{A} = \mathcal{A}^T$ and $\mathcal{C} = \mathcal{C}^T$. The structure of the matrix represented in equation 6.94 is based on separating the components of the Riemann curvature tensor

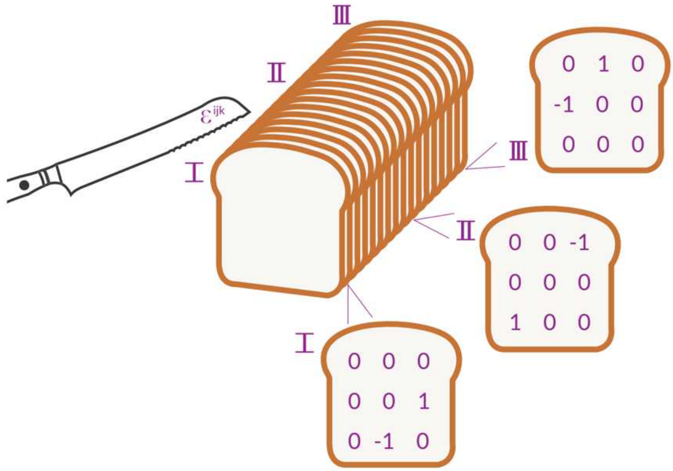


Fig. 6.1. Visualizing the three dimesional Levi-Civita symbol to be equivalent to the three slices of a three dimensional loaf of bread. Similarly, the four dimensional Levi Civita we encountered in section 2.3 would be four slices off a four dimensional loaf.

into three distinct sets, $R_{0\alpha 0\beta}$, $R_{0\beta\gamma\delta}$, and $R_{\gamma\delta\mu\nu}$. Observe that the first set is a 3×3 matrix in the indices α and β and as for the other two, they are to be fixed by removal of antisymmetricity they possess. Thus, we introduce the following 3×3 matrices

$$\Psi_{\alpha\beta} = R_{0\alpha 0\beta},\ \Sigma_{\alpha\beta} = \frac{1}{2}\epsilon_{\alpha\gamma\delta} R^{\gamma\delta}_{0\beta},\ \Lambda_{\alpha\beta} = \frac{1}{4}\epsilon_{\alpha\gamma\delta}\epsilon_{\beta\mu\nu} R^{\gamma\delta\mu\nu}, \tag{6.99}$$

where ϵ_{abc} is a three-dimensional Levi-Civita tensor. These matrices yield the following relations under the *Ricci flatness* condition, $R_{XY} = 0$

$$\Psi_{\alpha\alpha} = 0,\ \Sigma_{\alpha\beta} = \Sigma_{\beta\alpha},\ \Psi_{\alpha\beta} = -\Lambda_{\alpha\beta}. \tag{6.100}$$

According to the definitions given above we have the matrix $\Psi_{\alpha\beta}$ to have the following form

$$\Psi_{11} = R_{0101},\ \Psi_{12} = R_{0102},\ \Psi_{0103} = R_{0103},\ \ldots$$

$$\Longrightarrow \Psi_{\alpha\beta} = \begin{pmatrix} R_{0101} & R_{0101} & R_{0103} \\ R_{0201} & R_{0202} & R_{0203} \\ R_{0301} & R_{0302} & R_{0303} \end{pmatrix} = \begin{pmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} \\ \Psi_{21} & \Psi_{22} & \Psi_{23} \\ \Psi_{31} & \Psi_{32} & \Psi_{33} \end{pmatrix} \tag{6.101}$$

Comparing this matrix to the 6×6 form obtained previously, we find that $\Psi_{\alpha\beta}$

is comprised of the components of first quarter of the matrix (after lowering their index). Thus, $\Psi_{\alpha\beta} = \mathcal{A}$. Now, to the matrix $\Sigma_{\alpha\beta}$. Observe that in the components of the $\Sigma_{\alpha\beta}$

$$\Sigma_{11} = \frac{1}{2}\epsilon_{123}R^{23}_{01},\ \Sigma_{12} = \frac{1}{2}\epsilon_{123}R^{23}_{02},\ \Sigma_{13} = \frac{1}{2}\epsilon_{123}R^{23}_{03},\ ..., \tag{6.102}$$

the factor 1/2 is removed by the symmetry of the matrix, i.e., since $\Sigma_{\alpha\beta} = \Sigma_{\beta\alpha}$, $\Sigma_{12} = \Sigma_{21}$, ..., and hence

$$\Sigma_{(12)} = 2\Sigma_{12} = \Sigma_{12} + \Sigma_{21} = \underbrace{\epsilon_{123}}_{=1} R^{23}_{02} = R^{23}_{02}. \tag{6.103}$$

Similarly, we can calculate the other components to obtain the following matrix

$$\Sigma_{\alpha\beta} = \begin{pmatrix} R^{23}_{01} & R^{23}_{02} & R^{23}_{03} \\ R^{31}_{01} & R^{31}_{03} & R^{31}_{03} \\ R^{12}_{01} & R^{12}_{02} & R^{12}_{03} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}. \tag{6.104}$$

Comparing this matrix to the 6×6 form obtained previously, we find that $\Sigma_{\alpha\beta}$ is comprised of the components of third quarter of the matrix (, i.e. the first half of the second row). Thus, $\Sigma_{\alpha\beta} = -\mathcal{B}^T$. In matrix $\Lambda_{\alpha\beta}$, notice that there is symmetry in the indices and also among matrix components due to the *block symmetry* of the curvature tensor. the following are the components of the matrix $\Lambda_{\alpha\beta}$

$$\begin{aligned} &\Lambda_{11} = \tfrac{1}{4}\epsilon_{123}\epsilon_{123}R^{1323},\ \Lambda_{12} = \tfrac{1}{4}\epsilon_{123}\epsilon_{231}R^{2331},\ \Lambda_{13} = \tfrac{1}{4}\epsilon_{312}\epsilon_{312}R^{1212}, \\ &\Lambda_{21} = \tfrac{1}{4}\epsilon_{213}\epsilon_{123}R^{1323},\ ... \end{aligned} \tag{6.105}$$

We know that $\Lambda_{\alpha\beta}$ is a symmetric matrix thus, components such as $\Lambda_{12} = \Lambda_{21} \Longrightarrow \Lambda_{(12)} = 2\Lambda_{12}$, and this eliminates the factor (1/2). Now, to account for the remaining (1/2), consider the matrix components $a_{12} = \Sigma_{12}$ and $a_{21} = \Sigma_{21}$ (using index a to avoid confusion), in which there exists a block symmetry[7] between Riemann curvature tensor components, $R^{2331} = R^{1323}$. This implies that

$$\begin{aligned} &a_{12} = \Sigma_{12} = \tfrac{1}{2}\epsilon_{123}\epsilon_{231}R^{2331} = \tfrac{1}{2}\epsilon_{213}\epsilon_{123}R^{1323} = \Sigma_{21} = a_{21} \\ &\Longrightarrow 2a_{(12)} = a_{12} + a_{21} = \underbrace{\epsilon_{123}}_{=1}\underbrace{\epsilon_{231}}_{=1} R^{2331} = R^{2331}. \end{aligned} \tag{6.106}$$

Similarly, we can calculate the other components to obtain the following matrix

$$\Lambda_{\alpha\beta} = \begin{pmatrix} R^{2323} & R^{2331} & R^{2312} \\ R^{3123} & R^{3131} & R^{3112} \\ R^{1223} & R^{1231} & R^{1212} \end{pmatrix} = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{pmatrix} \tag{6.107}$$

[7] $R^{\alpha\beta\gamma\delta} = R^{\delta\gamma\alpha\beta} = R^{\beta\alpha\gamma\delta}$

Comparing this matrix to the 6×6 form obtained previously, we find that $\Lambda_{\alpha\beta}$ is comprised of the components of fourth quarter of the matrix (, i.e. the second half of the second row). Thus, $\Lambda_{\alpha\beta} = \mathcal{C}$.

Let $\Omega_{\alpha\beta}$ be a symmetric complex tensor defined as follows

$$\Omega_{\alpha\beta} = \frac{1}{2}\left(\Psi_{\alpha\beta} + 2i\Sigma_{\alpha\beta} - \Lambda_{\alpha\beta}\right) = \frac{1}{2}\left(\Psi_{\alpha\beta} + 2i\Sigma_{\alpha\beta} + \Psi_{\alpha\beta}\right) = \Psi_{\alpha\beta} + i\Sigma_{\alpha\beta} \tag{6.108}$$

The classification of the Riemann curvature tensor can be reduced to a simple eigen value problem. Consider the eigen value equation $\Omega_{\alpha\beta}k_\beta = \lambda k_\alpha$, in which the complex eigenvalues $\lambda = \lambda_R + i\lambda_I$ satisfy the condition $\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)} = 0$ since $\Omega^\alpha_\alpha = 0$. The matrix's classification is now dependent on the number of independent eigenvectors and leads to six different cases, called Petrov Types I, II, D, III, N, and O.

6.6.2 Petrov Type I

In this case there are three independent eigenvectors and upon diagonalizing $\Omega_{\alpha\beta}$ and separating it's real and imaginary parts we obtain the real part to be

$$\Psi_{\alpha\beta} = Diag\left(\lambda_R^{(1)}, \lambda_R^{(2)}, -\left(\lambda_R^{(1)} + \lambda_R^{(2)}\right)\right), \tag{6.109}$$

and the imaginary part to be

$$\Sigma_{\alpha\beta} = Diag\left(\lambda_I^{(1)}, \lambda_I^{(2)}, -\left(\lambda_I^{(1)} + \lambda_I^{(2)}\right)\right). \tag{6.110}$$

Upon computation we can show that the eigenvalues can be expressed in terms of the following scalars

$$\begin{aligned} \mathcal{I}_1 &= \tfrac{1}{48}\left(R_{abcd}R^{abcd} - i\ R_{abcd}{}^{*}R^{abcd}\right) \\ \mathcal{I}_2 &= \tfrac{1}{96}\left(R_{abcd}R^{cdef}R^{gh}_{ef} + i\ R_{abcd}R^{cdef*}R^{gh}_{ef}\right), \end{aligned} \tag{6.111}$$

where ${}^*R_{abcd} = \frac{1}{2}\epsilon_{abef}R^{ef}_{cd}$ is the dual of the Riemann curvature tensor. Also, notice that the real part of the scalar $\mathcal{I}_1$, divided by a factor of 48, is nothing but the *Kretschmann invariant* we have encountered in previous chapters, i.e.,

$$\frac{1}{48}\mathbf{Re}\left(\mathcal{I}_1\right) = R_{abcd}R^{abcd} = \frac{12r_s^2}{r^6} = K. \tag{6.112}$$

Calculating the scalars using the diagonalized matrices, $\Psi_{\alpha\beta}$ and $\Sigma_{\alpha\beta}$, we obtain

$$\mathcal{I}_1 = \frac{1}{3}\left(\lambda^{(1)2} + \lambda^{(2)2} + \lambda^{(3)2}\right),\ \mathcal{I}_2 = \frac{1}{3}\lambda^{(1)}\lambda^{(2)}\left(\lambda^{(1)} + \lambda^{(2)}\right). \tag{6.113}$$

These formulae enable us to calculate $\lambda^{(1)}$, $\lambda^{(2)}$ starting from the values of the Riemann curvature tensor $R_{\alpha\beta\gamma\delta}$ in any reference frame.

Petrov Type D

The special case in which $\lambda^{(1)} = \lambda^{(2)}$ is called Type D.

6.6.3 Petrov Type II

There are two independent eigenvectors such that the square of one of them is then equal to zero (i.e., $k_\alpha k^\alpha$ vanishes for one of them). This implies that we can choose this vector to lie in the $x^1 - x^2$ plane so that $k_2 = ik_1$ and $k_3 = 0$. The eigenvalues equations now read $\Omega_{11} + i\Omega_{12} = \lambda$, $\Omega_{22} + i\Omega_{12} = \lambda$ so that we can write

$$\Omega_{11} = \lambda - i\mu,\ \Omega_{22} = \lambda + i\mu,\ \Omega_{12} = \mu. \tag{6.114}$$

The complex quantity $\lambda = \lambda_R + i\lambda_I$ is a scalar and cannot be changed. But the quantity μ can be given any non-zero value by a suitable complex rotation; hence, we can assume it to be real. We obtain the following matrices

$$\Psi_{\alpha\beta} = \begin{pmatrix} \lambda_R & \mu & 0 \\ \mu & \lambda_R & 0 \\ 0 & 0 & -2\lambda_R \end{pmatrix}$$
$$\Sigma_{\alpha\beta} = \begin{pmatrix} \lambda_R - \mu & 0 & 0 \\ \mu & \lambda_R + \mu & 0 \\ 0 & 0 & -2\mu \end{pmatrix}. \tag{6.115}$$

In this case there are just two invariants λ_R and λ_I so in accordance to equation 6.113, we have $\mathcal{I}_1 = \lambda^2$ and $\mathcal{I}_2 = \lambda^3$ so that $\mathcal{I}_1^3 = \mathcal{I}_2^2$.

Petrov Type N

The special case when $\lambda = 0$ corresponds to a situation in which both the curvature invariances vanish and this type is called Type N.

6.6.4 Petrov Type III

In this case there is just one eigenvector with $k_\alpha k^\alpha = 0$ and all other eigenvalues are identically zero. The eigenvalue equation, $\Omega_{\alpha\beta} n_\beta = \lambda k_\alpha$, has the following solution

$$\Omega_{11} = \Omega_{22} = \Omega_{12} = 0,\ \Omega_{13} = \mu,\ \Omega_{23} = i\mu, \tag{6.116}$$

so that

$$\Psi_{\alpha\beta} = \begin{pmatrix} 0 & 0 & \mu \\ 0 & 0 & 0 \\ \mu & 0 & 0 \end{pmatrix}$$
$$\Sigma_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \mu \\ 0 & \mu & 0 \end{pmatrix}. \tag{6.117}$$

In this case the curvature tensor has no non-zero invariants even though the spacetime is curved.

6.6.5 Discussion on the Petrov types

The matrix $\Omega_{\alpha\beta}$ can be expressed in terms of $\mathcal{A}$ and $\mathcal{B}$ as

$$\Omega_{\alpha\beta} = \Psi_{\alpha\beta} + i\Sigma_{\alpha\beta} = \mathcal{A} - i\mathcal{B}^T. \tag{6.118}$$

Now, let $\mathcal{Q} = \Omega_{\alpha\beta}$. Thus, the eigenvalue equation can alternatively be written as follows

$$\mathcal{Q}\mathbf{k} = \lambda\mathbf{k}, \tag{6.119}$$

i.e., in which we have determined the eigenvectors $\mathbf{k}$ and the eigenvalues λ of the complex symmetric and traceless 3×3 matrix $\mathcal{Q}$: from the four-dimensional Lorentz frame we have passed to a three-dimensional complex space with Euclidean metric. This eigenvalue problem led to a characteristic equation $det\ \mathcal{Q} - \lambda I) = 0$. The results are tabulated in the table 6.1. The algebraic type of the matrix $\mathcal{Q}$ provides an invariant characterization of the gravitational field at a given point p; these characteristics are independent of the coordinate system at p.

Petrov Type O

We know that the form the *Weyl tensor* takes up in $n \geq 3$-dimension is

$$W_{\alpha\beta\mu\nu} = R_{\alpha\beta\mu\nu} - \frac{2}{n-2}\left(g_{\alpha[\nu}R_{\mu]\beta} + g_{\beta[\mu}R_{\nu]\alpha}\right) + \frac{2}{(n-1)(n-2)}Rg_{\alpha[\mu}g_{\nu]\beta}. \tag{6.120}$$

Petrov Types	Eigenvalues	Matrix Criterion
I	$\lambda_1 \neq \lambda_2 \neq \lambda_3$	$(\mathcal{Q} - \lambda_1 I)(\mathcal{Q} - \lambda_2 I)(\mathcal{Q} - \lambda_3 I) = 0$
D	$\lambda_2 = -2\lambda_1$	$(\mathcal{Q} - \lambda_1 I)(\mathcal{Q} - \lambda_2 I) = 0$
II	$\lambda_2 = -2\lambda_1$	$(\mathcal{Q} - \lambda_1 I)^2 (\mathcal{Q} - \lambda I) = 0$
N	$\lambda_1 = \lambda_2 = \lambda_3 = 0$	$\mathcal{Q}^2 = 0$
III	$\lambda_1 = \lambda_2 = \lambda_3 = 0$	$\mathcal{Q}^3 = 0$
O	$\lambda_1 = \lambda_2 = \lambda_3 = 0$	$\mathcal{Q} = 0$

Table 6.1. The table shows the matrix criteria for distinct Petrov Types

Clearly, the Weyl tensor is composed of the Riemann and the Ricci tensors and the curvature scalar. Now, the matrix criteria for Petrov Type O reads

$$\begin{aligned}
&\mathcal{Q} = 0 \\
&\Longrightarrow \Omega_{\alpha\beta} = \Psi_{\alpha\beta} + i\Sigma_{\alpha\beta} = 0 \\
&\Longrightarrow \mathbf{Re}\left(\Omega_{\alpha\beta}\right) = \Psi_{\alpha\beta} == \begin{pmatrix} R_{0101} & R_{0101} & R_{0103} \\ R_{0201} & R_{0202} & R_{0203} \\ R_{0301} & R_{0302} & R_{0303} \end{pmatrix} = 0 \qquad (6.121) \\
&\Longrightarrow R_{\alpha\beta\gamma\delta} = 0 \Longrightarrow R_{\alpha\beta} = 0 \Longrightarrow R = 0.
\end{aligned}$$

The vanishing of the Riemann tensor, the Ricci tensor and the curvature scalar imply a vanishing Weyl tensor[8] which in turn implies that the space is *conformally flat*. Thus, Type O regions are conformally flat regions which are associated with places where the Weyl tensor vanishes identically (see section 1.8.5).

6.7 Causality Theory

The study of a black hole present in a spacetime is done in two distinct regions- the black hole's interior and it's exterior. These regions are distinct since they are distinguished by the property that all external observers are causally separated (see subsection 5.6.1) from events that go on in the inside. Let us device a thought experiment to understand some of the deep ideas associated with black holes. Let Mr. Absolute Zero have a radio device that constantly sends signals back to the spaceship where Mr. Zero Entropy tracks his location. Once

[8] But the converse depends on the dimension. In dimension 1, every metric is flat, and the Riemann, Ricci, and scalar curvatures are always zero. In dimension 2, if the scalar curvature is zero, the metric is flat. In dimension 3, if the Ricci curvature is zero, the metric is flat, but there are non-flat metrics with zero scalar curvature. In dimensions 4 and up, there are plenty of examples of non-flat metrics with both Ricci and scalar curvatures equal to zero

Mr Absolute Zero enters the black hole, he can no longer send signals back to the spaceship. He would disappear off the radar as if he went into stealth mode!
Let's try to associate mathematical terms to simplify this concept. Let Mr. Absolute Zero sending light signals to the mother ship be event Z. We can associate a *null cone* to characterize the set of all points (a.k.a. *events*) that can be reached from the event Z by future-directed null curves (in general they can be either timelike or null). Such a set is called the *causal future* of Z and is denoted by $\mathcal{J}^+(Z)$. Similarly, associate a null cone to the set of all events that can be (or was reached) from Z by past-directed null curves (again, in general it can be timelike or null). Such a set is called the *causal past* of Z and is denoted by $\mathcal{J}^-(Z)$. Now, rather than just localizing our definition to an event (or a point) of Mr. Absolute Zero's adventures, let's associate similar definitions to the set of all such events S_Z, or for the sake of simplicity , just S. If S is the set of all events in which Mr. Absolute Zero has sent light signals, then $\mathcal{J}^+(S)$ is the union of the causal futures of all the events Z contained in S. Similarly, if S is the set of all events in which Mr. Absolute Zero had sent light signals, then $\mathcal{J}^-(S)$ is the union of the causal pasts of all the events Z contained in S.

A spacetime is said to contain a black hole if there exists null geodesics that never reach future null infinity $\mathcal{I}^+$. These geodesics originate from the interior of the black hole which is a region characterised by the very fact that all future-directed curves originating from it fail to reach $\mathcal{I}^+$. Thus, from this we can conclude that events lying within the black hole interior cannot be in the causal past of $\mathcal{I}^+$. Using all this logic we can now mathematically define a black hole. The black hole region $\mathcal{B}$ of the spacetime manifold $\mathcal{M}$ is the set of all events Z that do not belong to the causal past of future null infinity, i.e.,

$$\mathcal{B} = \mathcal{M} - \mathcal{J}^- \left(\mathcal{I}^+\right) , \tag{6.122}$$

and the *event horizon* $\mathcal{H}$ is then defined to be the boundary of the black hole, i.e.,

$$\mathcal{H} = \partial B = \partial \left[\mathcal{J}^- \left(\mathcal{I}^+\right)\right] . \tag{6.123}$$

$\partial \left[\mathcal{J}^- \left(\mathcal{I}^+\right)\right]$ is sometimes alternatively written as $\dot{\mathcal{J}}^+ \left(\mathcal{I}^+\right)$. Generally, $\dot{\mathcal{J}}^+(S)$ is called the *boundary* of the causal future $\mathcal{J}^+(S)$ and similarly, $\dot{\mathcal{J}}^-(S)$ is called the boundary of the causal past $\mathcal{J}^-(S)$. Table 6.2 lists all the mathematical labels defined.

Label	Name
$\mathcal{I}^-(Z)$	The past null infinity of event Z or the chronological past
$\mathcal{I}^+(Z)$	The future null infinity of event Z or the chronological future
$\mathcal{J}^-(Z)$	The causal past of event Z
$\mathcal{J}^+(Z)$	The causal future of event Z
$\mathcal{J}^-(S)$	The causal past of set S
$\mathcal{J}^+(S)$	The causal future of set S
$\dot{\mathcal{J}}^-(S)$	The boundary of the causal past, $\mathcal{J}^-(S)$
$\dot{\mathcal{J}}^+(S)$	The boundary of the causal future,$\mathcal{J}^+(S)$
$\dot{\mathcal{J}}^-\left(\mathcal{I}^+\right)$	The totality (or union) of all future horizons
$\dot{\mathcal{J}}^+\left(\mathcal{I}^-\right)$	The totality (or union) of all past horizons
i^-	The past timelike infinity of event Z
i^0	The spatial infinity of event Z
i^+	The future timelike infinity of event Z
$\mathcal{C}(s)$	Causal curve that is nowhere spacelike

Table 6.2. Causality labels

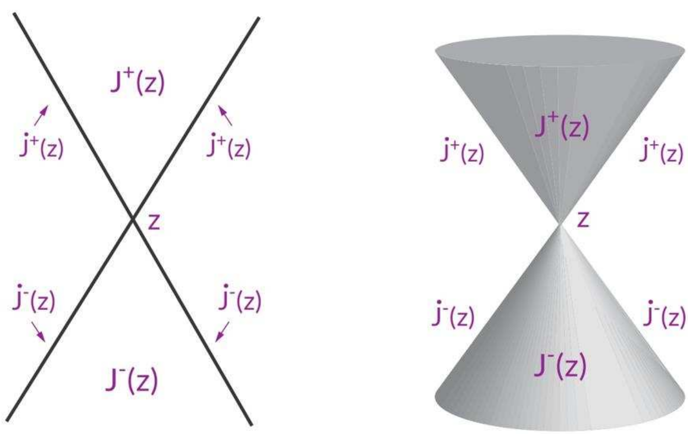


Fig. 6.2. Causal labels associated to Z.

6.7.1 Gedankenerfahrung: Mr.Absolute Zero and the Global Structure of Horizons

Consider another scenario, one in which Mr. Absolute Zero receives light signals from his spaceship as opposed to him transmitting them. In such a reversal of events, the light signals sent to Mr.Absolute Zero will not only have a spacelike future i^0 but also have a future null infinity $\mathcal{I}^+$ since light continues to travel along null geodesics. In Mr.Absolute Zero's frame of reference, when he is at a point Z in space, the signals he receives causally precede Z and hence belong to his causal past $\mathcal{J}^-(Z)$. Imagine now that Mr.Absolute Zero has an infinite number of clones, all invited to join the experiment from various parallel universes. Let's say that all the Absolute Zeros travel at the speed of light, which they commence in unison with a specific light signal they all receive simultaneously from the spaceship (i.e., the spaceship deploys an infinite number of light signal to the infinite number of Absolute Zero's simultaneously). The unison of all the future boundaries that all the Absolute Zero's reach, i.e., the boundary of the domain $\mathcal{J}^-(\mathcal{I}^+)$ is $\dot{\mathcal{J}}^-(\mathcal{I}^+)$. Now, after an infinite distance, if we were to trace back a light signal from any one of the Absolute Zero's we would find that the signal goes to the causal past of that Absolute Zero and it's future infinity, i.e., $\mathcal{J}^-(\mathcal{I}^+)$. Let's now position a black hole at the boundary so that all the Absolute Zero's reach it and travel into it's interior. Now if we try and trace back a light signal from any one of the clones we would find out that they have no future since the events lying withing the black hole interior cannot be in the causal past of $\mathcal{I}^+$. In other words, once all the clones reach the boundary $\dot{\mathcal{J}}^-(\mathcal{I}^+)$, there is no going back for both them and the light signals. With this thought experiment we have just proven Penrose's theorem on *the structure of future horizons*. Formally put, the theorem states the following

The future horizon $\dot{\mathcal{J}}^-(\mathcal{I}^+)$ is generated by null geodesics that have no future end points.

The null geodesics which lie in $\dot{\mathcal{J}}^-(\mathcal{I}^+)$ are called the *generators* of $\dot{\mathcal{J}}^-(\mathcal{I}^+)$ and the point (or event) at which the generator leaves is called a *caustic* of $\dot{\mathcal{J}}^-(\mathcal{I}^+)$. It is obvious that when we follow all the light signals from an instantaneous point a priori to them entering the black hole, we would observe that none of them intersect each other. This gives rise to the condition that generators can never intersect each other. After all the clones enter the black hole, they are spaghettified towards the singularity, where we have positioned the boundary $\dot{\mathcal{J}}^-(\mathcal{I}^+)$, and all the light signals or generators intersect at the caustic. Form this we can conclude that generators can only intersect at a caustic. Thus, the entry points into the event horizon are caustics of congruence of null generators. By following the generators of each of the clones locally we have obtained a global picture.

Penrose's theorem implies that the event horizon $\mathcal{H}$ of a black hole is a *null hypersurface* that is generated by null geodesics which have no future end points. For a Schwarzschild black hole, the generators of $\dot{\mathcal{J}}^-(\mathcal{I}^+)$ are the world lines of radially outgoing photons at the Schwarzschild radius, $r_s = 2M$. Penrose's theorem is not myopic, it is valid for any black hole- dynamic or static; rotating or stationary; coalescing with another black hole or isolated- in any asymptotically flat spacetime.

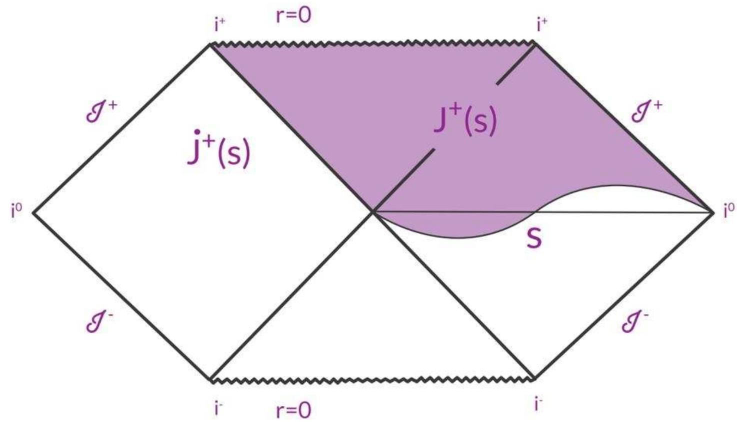


Fig. 6.3. S is a spacelike slice which extends from i^0 in $r = r_s$ but does not include $r = r_s$. All the Absolute Zeros traverse in this region. $J^+(S)$ does not include the leftmost horizon, but $\dot{\mathcal{J}}^+(S)$ is the leftmost horizon.

References

1. Hawking, Stephen W., and George Francis Rayner Ellis. The large scale structure of space-time. Vol. 1. Cambridge university press, 1973.
2. Choquet-Bruhat, Yvonne. General relativity and the Einstein equations. Oxford University Press, 2009.
3. Misner, C. W., Thorne, K. S., & Wheeler, J. A. (2017). Gravitation. Princeton University Press.
4. Spivak, M. D. (1970). A comprehensive introduction to differential geometry. Publish or perish.
5. Schutz, B. F. (1980). Geometrical methods of mathematical physics. Cambridge university press.
6. Warner, F. W. (2013). Foundations of differentiable manifolds and Lie groups (Vol. 94). Springer Science & Business Media.
7. Choquet-Bruhat, Y., de Witt, C., DeWitt, C. M., DeWitt-Morette, C., Dillard-Bleick, M., & Dillard-Bleick, M. (1982). Analysis, man- ifolds, and physics. Gulf Professional Publishing.
8. Nakahara, M. (2003). Geometry, topology and physics. CRC Press.
9. Dray, T. (2014). Differential forms and the geometry of general relativity. AK Peters/CRC Press.
10. Landau, L. D. (Ed.). (2013). The classical theory of fields (Vol. 2). Elsevier.
11. Griffiths, J. B., & Podolský, J. (2009). Exact space-times in Einstein's general relativity. Cambridge University Press.
12. Kramer, D., Stephani, H., MacCallum, M., & Herlt, E. (1980). Exact solutions of Einstein's field equations. Berlin.
13. Plebanski, J., & Krasinski, A. (2006). An introduction to general relativity and cosmology. Cambridge University Press.
14. De Sabbata, V., & Gasperini, M. (1985). Introduction to gravitation. World scientific.
15. Christodoulou, D. (2008). Mathematical problems of general relativity I (Vol. 1). European Mathematical Society.
16. Poisson, E. (2004). A relativist's toolkit: the mathematics of black-hole mechanics. Cambridge university press.
17. Sachs, R. K., & Wu, H. H. (2012). General relativity for mathematicians (Vol. 48). Springer Science & Business Media.
18. Wald, R. M. (1984). General Relativity, Chicago, Usa: Univ. Pr. 491p.

19. Paston, S. A., & Sheykin, A. A. (2012). Embeddings for the Schwarzschild metric: classification and new results. Classical and Quantum Gravity, 29(9), 095022.
20. Kruskal, M. D. (1960). Maximal extension of Schwarzschild metric. Physical review, 119(5), 1743.
21. Fronsdal, C. (1959). Completion and embedding of the Schwarzschild solution. Physical Review, 116(3), 778.
22. Visser, M. (2005). Heuristic approach to the Schwarzschild geometry. International Journal of Modern Physics D, 14(12), 2051-2067.
23. Benzoni-Gavage, S. (2012). The Noether Theorems, Invariance and Conservation Laws in the Twentieth Century.
24. Petrov, A. Z. (2016). Einstein spaces. Elsevier.
25. Winberg, S. (1972). Gravitation and cosmology. ed. John Wiley and Sons, New York.
26. Goldstein, H. (2011). Classical mechanics. Pearson Education India.
27. Taylor, J. R. (2005). Classical mechanics. University Science Books.
28. Susskind, L., & Hrabovsky, G. (2014). Classical mechanics: the theoretical minimum. Penguin Books.
29. Kibble, T. W., & Berkshire, F. H. (2004). Classical mechanics. World Scientific Publishing Company.
30. Blandford, R. D., & Thorne, K. S. (2008). Applications of classical physics. lecture notes, California Institute of Technology, 12.
31. Tangherlini, F. R. (1963). Schwarzschild field inn dimensions and the dimensionality of space problem. Il Nuovo Cimento (1955-1965), 27(3), 636-651.
32. Price, R. H. (1972). Nonspherical perturbations of relativistic gravitational collapse. I. Scalar and gravitational perturbations. Physical Review D, 5(10), 2419.
33. Banados, M., Henneaux, M., Teitelboim, C., & Zanelli, J. (1993). Geometry of the 2+ 1 black hole. Physical Review D, 48(4), 1506.

Index