From Geometry
To
Black Holes

Otis Saint. 10584344

September 2020

Abstract

This project aims to give an insight into both the mathematics, and physics behind Einstein's famous theory of general relativity. The first section in this document tackles the subject of differential geometry, starting with an introduction to manifolds and tensors, before moving on to define important concepts such as covariant differentiation, connections, and geodesics. We conclude the first section with a discussion about curvature, and derive the Riemann curvature tensor along with the Ricci tensor and scalar, which are used to define the Einstein tensor.

In the second section, we begin by discussing the famous equivalence principle, and reformulate Newtons first law, so that it is valid in curved space-time, at which point, the importance of geodesics will become apparent. Einsteins field equations are then derived by means of minimising the Einstein-Hilbert action, and relating the result back to Newtonian gravity. Section 2.5 - 2.7 then deals with perhaps the most famous consequence of the theory, black-holes.

Acknowledgements

I would like to thank Dr. Vincent Drach for taking the time to supervise this project, whose expertise and endless ideas have been indispensable in the development of this document.

Contents

Introduction

In the year 1687, Sir Isaac Newton published his theory of gravitation in his revolutionary book, the Principia Mathematica. As the legend goes, Newton was sitting under an apple tree when an apple fell and hit him on the head, causing him to question the mysterious force that pulls everything towards the earth. Whether this legend has any truth or not, Newton was considered by his contemporaries, colleagues and rivals to be singularly driven in the pursuit of understanding the mechanics of the forces underpinning the foundations of the universe. After many experiments and hypotheses, Newton proved that the force which pulls everything to the ground is also responsible for keeping the moon in orbit around the earth, and the earth around the sun.

Newton was not the first person to think about gravity. Galileo discovered in 1589, while experimenting with dropping objects off the leaning tower of Piza, that when air resistance is ignored; all objects will fall at the same rate, regardless of the object's mass. This was further proved in 1971 when astronaut David Scott dropped a hammer and a feather on the surface of the moon. Just as as Galileo had predicted many years before, they both fell to the lunar surface at exactly the same rate.

Newton's laws of gravity remained unchallenged for over 2 centuries until in 1907 when the bright young physicist Albert Einstein decided to tackle the subject of gravity. He had already published his revolutionary 1905 paper describing special relativity, in which he introduced the idea that space and time could be better understood as a single, unified entity known as space-time. Through his famous thought experiments, and with help from many mathematicians along the way, Einstein realised that matter causes space-time to curve, and the curvature of space-time tells matter how to move. It is this curvature that we experience as gravity, thus laying the foundations for the general theory of relativity.

At first Einstein's radical new ideas about the universe were met with opposition from many in the science community. Some, such as (those) at the Lick Observatory in California, even claimed to have dis-proven the theory - although their experimental results were never published. The first piece of solid evidence supporting the theory of general relativity came when Einstein correctly calculated the perihelion precession of Mercury. Further conclusive evidence of the theory came in May of 1919 when astronomer Arthur Eddington photographed a solar eclipse and, by meticulously studying the pictures and comparing them with maps of the stars, Eddington's team was able to confirm that the light from a distant star which was located behind the sun was, as Einstein predicted, being bent by the sun's gravitational field. Einstein also correctly predicted the existence of gravitational waves 1916, and it was not until 100 years later that researches at LIGO, were able to detect these ripples in space-time caused by the collision of 2 black holes.

Due to the nature of the mathematics and counter-intuitive ideas on which the theory is based, general relativity is often seen by many as an intimidating subject to learn. In this project i have endeavoured to present the mathematics and physics of the theory in as simple terms as possible, while still paying close attention to the various subtleties that present themselves along the way.

1 Differential Geometry

> "You can get good science out of stupid questions. If someone says the world is flat, maybe in proving them wrong you can calculate the curvature of the earth more precisely" - Gavin Schmidt

1.1 Manifolds

We begin by defining the notion of a manifold; a type of topological space, which can be regarded as a generalisation of curves and surfaces to higher dimensions, such that in local neighbourhoods around points, the manifold "looks like" Euclidean space!

Definition 1 (Topological Space [18])**.** *Let $\mathcal{X} \neq \varnothing$ be some abstract set of points and let τ be a collection of open subsets of $\mathcal{X}$, which satisfy the the following conditions:*

- *The empty set $\varnothing$ and the set $\mathcal{X}$ are elements of τ.*
- *The **intersection** of a finite number of subsets $U_k \in \tau$ gives a set that is also an element of τ.*
- *The **union** of a possibly infinite amount of subsets $U_n \in \tau$ gives a set that is also an element of τ.*

*If the collection τ satisfies these axioms, then it is called a **topology** on the set $\mathcal{X}$. Thus, the tuple $(\mathcal{X}, \tau)$can be called a **topological space**.*

A topological space is said to be a **Hausdorff** space, if there exists open subsets $U_1, U_2 \in \tau$ containing the points $x_1, x_2 \in X$ respectively, and are such that $U_1 \cap U_2 = \varnothing$. In other words, any two points in the space can be seperated by neighbourhood.

Loosely put, a manifold is a Hausdorff space for which each point is contained in some open neighborhood, that "looks like" ordinary n-dimensional Euclidean space. To find a more rigorous definition, first consider a topological space $\mathbf{M} = (\mathcal{M}, \tau)$, which is such that every point, $m \in \mathcal{M}$, is contained in a finite number of open subsets $U_k \in \tau$, called the **coordinate patches** of $\mathbf{M}$. A collection of N open subsets, U_k, such that

$$\bigcup_{k=1}^{N} U_k = \mathbf{M} \,, \tag{1.1.1}$$

is called an **open covering** of $\mathbf{M}$. In order for one of these coordinate patches U_k, to "look like" $\mathbb{R}^n$, we require the existence of a continuous bijective map, $\phi_k : U_k \to \mathbb{R}^n_k \subset \mathbb{R}^n$, together with a continuous inverse map, $\phi_k^{-1} : \mathbb{R}^n_k \to U_k$. In other words, ϕ_k is a homeomorphism from a patch $U_k \subset \mathbf{M}$, to an open subset $\mathbb{R}^n_k \subset \mathbb{R}^n$, and defines a **coordinate system** or **coordinate map** on the patch U_k. This leads us to the following definition of a manifold:

Definition 2 (Topological Manifold)**.** *A topological manifold is a Hausdorff topological space $\mathbf{M}$, such that every point $m \in \mathbf{M}$ is contained in some open subset $U_k \subset \mathbf{M}$ for which there exists a homeomorphism $\phi_k : U_k \to \mathbb{R}^n_k \subset \mathbb{R}^n$. The tuple (U_k, ϕ_k), is called a **chart** on **M** and the collection of charts, whose patches together form an open-covering of $\mathbf{M}$, is called an **atlas** on $\mathbf{M}$.*

Examples of Manifolds:

- The simplest example of a manifold is just an open interval in $\mathbb{R}$.
- $\mathbb{R}^n$ itself, is clearly a manifold.
- The n-sphere, S^n, which is defined as the locus of all points which are located at a fixed distance from the origin of the space $\mathbb{R}^{n+1}$, is an n-dimensional manifold embedded in $\mathbb{R}^{n+1}$.

Consider two charts (U_α, ϕ_α) and (U_β, ϕ_β), such that $U_\alpha \cap U_\beta \neq \varnothing$. Any point lying in the overlap region, $T_{[\alpha,\beta]} = U_\alpha \cap U_\beta$, can thus be mapped to a subset of $\mathbb{R}^n$, using either coordinate map

$$\phi_\alpha : T_{[\alpha,\beta]} \to \mathbb{R}^n_\alpha \ , \tag{1.1.2}$$
$$\phi_\beta : T_{[\alpha,\beta]} \to \mathbb{R}^n_\beta \tag{1.1.3}$$

Since ϕ_α, ϕ_β are homeomorphisms, we are also assured of the existence of continuous inverse maps from the subsets of $\mathbb{R}^n$, back to the manifold, namely

$$\phi_\alpha^{-1} : \mathbb{R}^n_\alpha \to T_{[\alpha,\beta]} \ , \tag{1.1.4}$$
$$\phi_\beta^{-1} : \mathbb{R}^n_\beta \to T_{[\alpha,\beta]} \ . \tag{1.1.5}$$

Notice that we now have two different, open subsets of $\mathbb{R}^n$ being mapped to the same region on the manifold. Hence, by composing these mappings, we are now able to move between these subsets via the so called **transition functions**

$$\phi_\alpha \circ \phi_\beta^{-1} : \mathbb{R}^n_\beta \to \mathbb{R}^n_\alpha \ , \tag{1.1.6}$$
$$\phi_\beta \circ \phi_\alpha^{-1} : \mathbb{R}^n_\alpha \to \mathbb{R}^n_\beta \ . \tag{1.1.7}$$

This web of maps is summarised in the following diagram

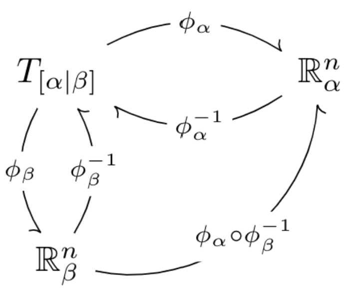

Figure 1: Transition functions allow us to move between different subsets of $\mathbb{R}^n$, making clear the notion of a coordinate transformation.

Transition functions enable us to express the position of a point in the overlap region using either coordinate system, i.e they define the notion of a coordinate transformation. If the partial derivatives of all transition functions on $\mathbf{M}$ up to order $N < \infty$ exist, and are continuous, then $\mathbf{M}$ is called a C^N manifold, and one whose transition functions are all C^∞, is called a differentiable/smooth manifold. Furthermore, if such a smooth map exists between between two spaces, then the spaces are said to be **diffeomorphic**.

1.2 Vectors and Covectors

Traditionally, a vector is a quantity with direction and magnitude that can be visualised as an arrow in space. However, if we consider a curved space, e.g the surface of a sphere, it quickly becomes apparent that this definition is insufficient, as a vector can not curve. This section introduces vectors in a slightly less intuitive way, as directional derivatives that live in the tangent space of a manifold.

1.2.1 Tangent Vectors

Consider a smooth curve $\boldsymbol{\gamma}$ defined on some patch U_k of an n-dimensional manifold, i.e $\boldsymbol{\gamma} : I \subset \mathbb{R} \to U_k$ such that $\boldsymbol{\gamma}$ is C^∞. In the previous section, we said how each point on the patch U_k can be mapped to a unique point in $\mathbb{R}^n_k$, via the mapping $\phi_k : U_k \to \mathbb{R}^n_k$, resulting in each point in U_k being endowed with a set of coordinates, for example $(x^0, ..., x^n)$ or just $(x^\mu) : \mu = 0, ..., n$. Hence, coordinates of points along the curve are given by the mapping $\phi \circ \boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^n$, thus if we parametrize $\boldsymbol{\gamma}$ with say, $t \in I$, then we can write these coordinates as $\gamma(t) = x^\mu(t)$.

We denote by $C^\infty(\mathbf{M})$, the set of smooth, linear maps from a patch $U_k \in \mathbf{M}$, to the set of real numbers, i.e $C^\infty(\mathbf{M}) = \{f : U_k \in M \to \mathbb{R}\}$. In otherwords, $C^\infty(M)$ is the set of smooth scalar fields on a patch U_k. Thus, for some $f \in C^\infty(M)$ and some smooth curve $\boldsymbol{\gamma}$ parametrised by $t \in I \subset \mathbb{R}$, the values of f along $\boldsymbol{\gamma}(t)$ are given by the mapping $f \circ \boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}$, which, in ordinary function notation reads $f(x^\mu(t))$.

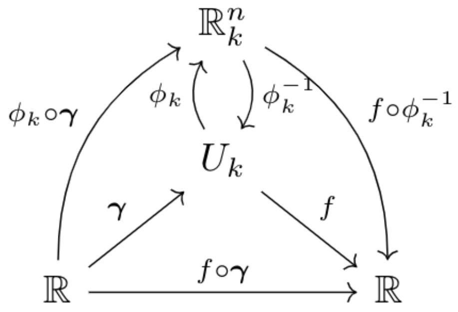

Figure 2: Constructing a curve and C^∞ scalar field on a patch U_k

Definition 3 (Tangent Vector)**.** *A **Tangent Vector** $v|_p$, is a directional derivative along a curve, $\boldsymbol{\gamma} : I \in \mathbb{R} \to U_k \in \mathbf{M}$, given by the mapping*

$$v|_p : f \in C^\infty(M) \to \mathbb{R},\ f \mapsto \frac{d}{dt}(f \circ \boldsymbol{\gamma})|_p,$$

which satisfies the following conditions:

- *Linearity :* $v|_p(f+g) = v|_p(f) + v|_p(g)$
- *Leibniz Rule:* $v|_p(fg) = fv|_p(g) + gv|_p(f)$

To see how this describes a tangent vector, we can use more conventional notation, which allows us to employ the chain rule as follows:

$$\begin{aligned} v|_p(f) &= \frac{d}{dt}(f \circ \boldsymbol{\gamma}) = \frac{d}{dt} f(x^\mu(t)) \\ &= \frac{dx^\mu}{dt}\Big|_p \partial_\mu f = v|_p(x^\mu)\partial_\mu f\ . \end{aligned} \tag{1.2.1}$$

Thus, by writing $v|_p(x^\mu) = v^\mu(p) = v^\mu$, where $p = (x^\mu)$, we obtain

$$v|_p(f) = v^\mu \partial_\mu f\ , \tag{1.2.2}$$

and since f is an arbitrary smooth scalar field on the patch U_k, we can just consider the directional derivative $v|_p$, to be an operator acting on C^∞ scalar fields at the point p,

$$v|_p = v^\mu \partial_\mu\ . \tag{1.2.3}$$

Notice, this is a linear combination of the partial derivatives ∂_μ. In addition, since $v|_p = 0$ is only true when $v^\mu = 0$ for all μ, we can say that the set of partial derivatives, $\{\partial_\mu\}$, forms a set of linearly independent basis vectors. Thus, the collection of all directional derivatives at a point p, forms a vector space!

Definition 4 (Tangent Space)**.** *The collection of directional derivatives at a point, p, on a manifold $\mathbf{M}$, forms a vector space whose dimension is equal to that of $\mathbf{M}$, called the **tangent space**, and is denoted $T_p\mathbf{M}$, or simply T_p.*

To this end, we can now express vectors in $T_p\mathbf{M}$ as

$$\vec{v} = v^\mu \partial_\mu \ , \tag{1.2.4}$$

where v^μ are the components of the vector, with respect to the basis vectors ∂_μ.

Consider two charts (U_a, ϕ_a) and (U_b, ϕ_b), and suppose we construct $T_p\mathbf{M}$ such that $p \in U_a \cap U_b$, and denote the coordinates given by the mapping $\phi_a : U_a \to \mathbb{R}^n_a$ as (x^σ), and similarly, those given by the mapping $\phi_b : U_b \to \mathbb{R}^n_b$ as $(\xi^{\lambda'})$, which are such that $x^\sigma = x^\sigma(\xi^{\lambda'})$. We can now express a vector, $\vec{v} \in T_p\mathbf{M}$ in terms of either coordinate system i.e $\vec{v} = v^\sigma \partial_\sigma$ and $\vec{v} = v^{\lambda'}\partial_{\lambda'}$. Since both of these linear combinations are describing the same vector, (note that the vector its self does not change under a change in coordinate system, only the components), namely $\vec{v}$, we can write

$$v^\sigma \partial_\sigma = v^{\lambda'}\partial_{\lambda'} \ . \tag{1.2.5}$$

Using the chain rule we can then express the right hand side in terms of the basis vector ∂_σ to obtain

$$v^\sigma \partial_\sigma = v^{\lambda'} \frac{\partial x^\sigma}{\partial \xi^{\lambda'}} \partial_\sigma \ . \tag{1.2.6}$$

Since we now have that both sides of this equation are written in terms of the same basis vector, ∂_σ, we can cancel this term from both sides, leaving us with an equation involving only the components. Then, by rearranging to make $v^{\lambda'}$ the subject, we obtain the transformation law for vectors in the tangent space

$$v^{\lambda'} = \frac{\partial \xi^{\lambda'}}{\partial x^\sigma} v^\sigma \ . \tag{1.2.7}$$

Components which transform in this way are said to be **contravariant**, and are written with the index as a **superscript**.

Remark: The union of all tangent spaces on a manifold forms the set $T\mathbf{M}$, called the tangent bundle.

1.2.2 Vector Fields

In the previous section, we introduced the idea of a vector as being an element of a tangent space at some point on a manifold. We can now extend this idea to the subject of vector fields. Intuitively, one can think of a vector field on a manifold as being a collection of tangent vectors, one from each point on the manifold. More precisely, we have the following definition:

Definition 5 (Vector field (first definition) [16])**.** *A smooth vector field X on a manifold $\mathbf{M}$, is a map $X : \mathbf{M} \to T\mathbf{M}$, such that X is C^∞. In some chart (U_k, ϕ_k) with coordinates x^α, the value of the vector field at a point $p \in U_k$ is then given by*

$$X_p = v^\mu(p)\partial_\mu|_p \ . \tag{1.2.8}$$

We can, however, write this definition in another way that is more similar to the definition of a tangent vector, given in the previous section. To do this, suppose that $X : \mathbf{M} \to T\mathbf{M}$ is smooth, and that $f \in C^\infty(\mathbf{M})$ is a scalar field defined on the manifold $\mathbf{M}$. Then, since we can view a vector as being a differential operator which acts on scalar valued functions, we can write

$$X(f) = v^\mu \partial_\mu f \ . \tag{1.2.9}$$

Thus, if we assume that the coefficients v^μ are also C^∞, then clearly the entire right hand side of (1.2.9) must also be C^∞. Furthermore, we notice that if $f, g \in C^\infty(\mathbf{M})$, then we have the following properties:

- Linearity: $X(f + g) = X(f) + X(g)$
- Leibniz rule: $X(fg) = fX(g) + gX(f)$.

With this, we arrive at the following definition:

Definition 6. *[Vector field (second definition)[16]] A smooth vector field X on a manifold $\mathbf{M}$, is a linear map $X : C^\infty(\mathbf{M}) \to C^\infty(\mathbf{M})$, such that X obeys the Leibniz rule. We denote the collection of all smooth vector fields on $\mathbf{M}$ as $\mathfrak{X}(\mathbf{M})$.*

A trivial property of the set $\mathfrak{X}(\mathbf{M})$ is that it forms a vector space which can be easily proved using the axioms for vector spaces, although, there is also an additional structure on this set.

Definition 7. *[Lie algebra [7]] A Lie algebra over $\mathbb{R}$, is a vector space ,$\mathfrak{H}$, together with the binary operator, $[.\ ,\ .] : \mathfrak{H} \times \mathfrak{H} \to \mathfrak{H}$ (called a **Lie bracket**), which obeys the axioms (where $x, y, z \in \mathfrak{H}$):*

- *Bilinearity:* $[ax + by, z] = a[x, z] + b[y, z],\ [z, ax + by] = a[z, x] + b[z, y]$ *where* $a, b \in \mathbb{R}$.
- *Alternativity:* $[x, x] = 0$ *for all* $x \in \mathfrak{H}$
- *Jacobi identity:* $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$

Notice that using the alternativity and bilinearity properties, we have

$$[x + y, x + y] = 0 \implies [x, y] + [y, x] = 0 \tag{1.2.10}$$

hence the Lie bracket is also antisymmetric, i.e

$$[x, y] = -[y, x]\ . \tag{1.2.11}$$

We will now proceed to show that the vector space, $\mathfrak{X}(\mathbf{M})$, is also a Lie algebra, over the field $\mathbb{R}$. We define the commutator of vector fields $X, Y \in \mathfrak{X}(\mathbf{M})$ to be

$$[X, Y] = X \circ Y - Y \circ X = X(Y) - Y(X)\ . \tag{1.2.12}$$

To see that the commutator is in fact a Lie bracket, one simply needs to check the axioms in definition 7.
Starting with bilinearity, we have for all $X, Y, Z \in \mathfrak{X}(\mathbf{M})$ and scalars $a, b \in \mathbb{R}$

$$\begin{aligned}[aX + bY, Z] &= (aX + bY) \circ Z - Z \circ (aX + bY) \\ &= aX \circ Z + bY \circ Z - aZ \circ X - bZ \circ Y \\ &= a(X \circ Z - Z \circ X) + b(Y \circ Z - Z \circ Y) \\ &= a[X, Z] + b[Y, Z]\ ,\end{aligned} \tag{1.2.13}$$

where we have used the linearity property of vector fields. Similarly, it can be shown that

$$[Z, aX + bY] = a[Z, X] + b[Z, Y]\ . \tag{1.2.14}$$

The alternativity property clearly holds here, since we have

$$[X, X] = X \circ X - X \circ X = 0\ . \tag{1.2.15}$$

To show that the Jacobi identity holds, we need to consider the sum of cyclic permutations of nested commutators, i.e

$$\begin{aligned}&[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = [X, Y \circ Z - Z \circ Y] + [Z, X \circ Y - Y \circ X] + [Y, Z \circ X - X \circ Z] \\ &= X \circ (Y \circ Z - Z \circ Y) - (Y \circ Z - Z \circ Y) \circ X + Z \circ (X \circ Y - Y \circ X) - (X \circ Y - Y \circ X) \circ Z \\ &+ Y \circ (Z \circ X - X \circ Z) - (Z \circ X - X \circ Z) \circ Y \\ &= X \circ Y \circ Z - X \circ Z \circ Y - Y \circ Z \circ X + Z \circ Y \circ X + Z \circ X \circ Y - Z \circ Y \circ X \\ &- X \circ Y \circ Z + Y \circ X \circ Z + Y \circ Z \circ X - Y \circ X \circ Z - Z \circ C \circ Y + X \circ Z \circ Y \\ &= 0\ .\end{aligned} \tag{1.2.16}$$

and thus, we can say that the commutator is a Lie bracket! Finally, to show that $\mathfrak{X}(\mathbf{M})$ together with the commutator forms a Lie algebra, we need to show that for all $X, Y \in \mathfrak{X}(\mathbf{M})$, we have that $[X, Y] \in \mathfrak{X}(M)$, i.e $[.\ ,\ .] : \mathfrak{X}(\mathbf{M}) \times \mathfrak{X}(\mathbf{M}) \to \mathfrak{X}(\mathbf{M})$. Consider two functions $f, g \in C^\infty(M)$, then, using the linearity property of X and Y, we see that

$$\begin{aligned}[X, Y](f + g) &= X \circ Y(f + g) - Y \circ X(f + g) \\ &= X \circ (Y(f) + Y(g)) - Y \circ (X(f) + X(g)) \\ &= X \circ Y(f) + X \circ Y(g) - Y \circ X(f) - Y \circ X(g) \\ &= [X, Y](f) + [X, Y](g)\ ,\end{aligned} \tag{1.2.17}$$

hence, the commutator of X and Y also acts linearly on scalar functions. Now, since X and Y obey the Leibniz rule, we have

$$\begin{aligned}[X,Y](fg) &= X\circ(Y(fg)) - Y\circ(X(fg)) \\ &= X\circ(Y(g)f + Y(f)g) - Y\circ(X(g)f + X(f)g) \\ &= X\circ Y(g)f + X\circ Y(f)g - Y\circ X(g)f - Y\circ X(f)g \\ &= f[X,Y](g) + g[X,Y](f)\end{aligned} \tag{1.2.18}$$

Thus, by definition 6 we can see that $[X,Y]$ does itself form a vector field in $\mathfrak{X}(\mathbf{M})$ and hence, the space of all smooth vector fields together with the commutator, forms a (real) Lie algebra. In a later section we will interpret the Lie bracket geometrically, and discuss how it relates to things such as torsion and curvature!

1.2.3 Dual Vectors

In addition to tangent vectors, there also exists quantities known as **dual vectors** or **one-forms**.

Definition 8 (Dual Vectors [3])**.** *A dual-vector is a smooth, real valued, linear function acting on vectors in the tangent space. i.e If w is a dual vector, then $w : v|_p \in T_pM \to \mathbb{R}$ and obeys the following rules:*

- $w(\vec{u}+\vec{v}) = w(\vec{u}) + w(\vec{v}) \in \mathbb{R}$
- $w(\lambda\vec{v}) = \lambda w(\vec{v}) \in \mathbb{R} \ : \lambda \in \mathbb{R}$

*The collection of dual vectors forms a vector space called the **dual space**, and is denoted $T_p^*\mathbf{M}$, or just T_p^*.*

Consider a smooth function, $f \in C^\infty(\mathbf{M})$, defined on a patch U_k endowed with a coordinate system (x^κ). At any point in U_k, we can thus write the value of the function f as $f(x^\mu)$.

Definition 9 (Exterior derivative)**.** *The exterior derivative of a smooth function, $f \in C^\infty(\mathbf{M})$, is a map*

$$d : C^\infty(\mathbf{M}) \to T_p^*\mathbf{M},\ f \mapsto df\ .$$

Taking the exterior derivative of f using the chain rule then yields

$$df = \frac{\partial f}{\partial x^\mu} dx^\mu\ , \tag{1.2.19}$$

which by definition, is an element of the dual space $T_p^*\mathbf{M}$. Thus, by letting $w = df$ and $w_\mu = \partial_\mu f$, (since our choice of function f was arbitrary), we arrive at an expression for a general dual vector

$$w = w_\mu dx^\mu\ . \tag{1.2.20}$$

Notice, this equation only vanishes iff $w_\mu = 0$ for all μ; hence, a suitable basis for the dual space $T_p^*\mathbf{M}$ is given by the set of coordinate differentials $\{dx^\mu\}$.

Dual-vector components also transform in a certain way. Again, consider two charts, (U_a, ϕ_a) and (U_b, ϕ_b), with coordinates x^μ and $\xi^{\lambda'}$ respectively, such that $\xi^{\lambda'} = \xi^{\lambda'}(x^\mu)$, and suppose we construct the dual-space $T_p^*\mathbf{M}$, at some point $p \in U_a \cap U_b$. Like vectors, the dual vectors themselves don't change under coordinate transformation, only the components change, allowing us to write

$$w = w_\mu dx^\mu = w_{\lambda'} d\xi^{\lambda'}\ . \tag{1.2.21}$$

Using the chain rule we can rewrite the right hand side as

$$w_\mu dx^\mu = w_{\lambda'} \frac{\partial \xi^{\lambda'}}{\partial x^\mu} dx^\mu\ . \tag{1.2.22}$$

We now have that both sides of this equation are written in terms of the same basis vector, dx^μ, and hence we can once again cancel the basis vector from both sides and write an equation involving only the dual-vector components. Rearranging for $w_{\lambda'}$ then leaves us with the transformation law for dual-vector components

$$w_{\lambda'} = \frac{\partial x^\mu}{\partial \xi^{\lambda'}} w_\mu\ . \tag{1.2.23}$$

Components which transform like this are said to be **covariant**, and are written with the index as a **subscript**.

1.3 Tensors

Tensors play an extremely important role in the theory of general relativity, since they allow to write equations that are true in any coordinate system, leading to the idea of general covariance. In this section, we discuss how tensors are formed, the types of operations which we can use to manipulate them, and how they transform. We also introduce perhaps the most important tensor, the metric tensor.

1.3.1 Tensors as Multilinear Maps

Suppose at a point ,$\mathbf{p}$, on some manifold ,M, we construct the tangent space, $T_{\mathbf{p}}$, and its dual-space, $T^*_{\mathbf{p}}$, in accordance with section 1.2 . We denote $(T_{\mathbf{p}})^n = T_{\mathbf{p}} \times ... \times T_{\mathbf{p}}$ (n times), and similarly, $(T^*_{\mathbf{p}})^m = T^*_{\mathbf{p}} \times .. \times T^*_{\mathbf{p}}$ (m times)

Definition 10. *A **tensor**, Ω, of type (m, n), over the tangent space at a point ,$\mathbf{p} \in M$, is a **multilinear** mapping*

$$\Omega : (T_p)^n \times (T^*_p)^m \to \mathbb{R} \ . \tag{1.3.1}$$

Thus, a tensor whose rank is $N = n + m$, is a function which takes as input m vectors and n dual vectors, and returns a real number. Following this line of reasoning, we can say that a vector in T_p, is in fact, a tensor of type $(1, 0)$, and similarly, a dual-vector in T^*_p is a (0,1) tensor!

Definition 11 (Tensor Product)**.** *If T is a tensor of type (k, l), and S of type (m, n), then we can define a new tensor, V, of type $(k + m, l + n)$, using the **tensor product** $T \otimes S$ i.e*

$$T^{\alpha_1,...,\alpha_k}_{\lambda_1,...,\lambda_l} \otimes S^{\nu_1,...,\nu_m}_{\beta_1,...,\beta_n} = V^{\alpha_1,...,\alpha_k,\nu_1,...,\nu_m}_{\lambda_1,...,\lambda_l,\beta_1,...,\beta_n} \ . \tag{1.3.2}$$

Recall, we defined T_p, to have a basis given by the set of partial derivatives $\{\partial_\mu\}$ and T^*_p to have a basis given by the set of coordinate differentials $\{dx^\nu\}$. Thus, using the above definition, we can construct a basis for the space of all (m, n) tensors, by taking the tensor product of basis vectors and dual vectors [3], the basis of which will consist of all tensors which have the form

$$\partial_{\mu_1} \otimes ... \otimes \partial_{\mu_m} \otimes dx^{\nu_1} \otimes ... \otimes dx^{\nu_n} \ . \tag{1.3.3}$$

Hence, we can write a tensor $\Omega \in (T_p)^m \times (T^*_p)^n$, in terms of its components and basis as

$$\Omega = \Omega^{\mu_1,...,\mu_m}_{\nu_1,...,\nu_n} \partial_{\mu_1} \otimes ... \otimes \partial_{\mu_m} \otimes dx^{\nu_1} \otimes ... \otimes dx^{\nu_n} \tag{1.3.4}$$

where $\Omega^{\mu_1,...,\mu_m}_{\nu_1,...,\nu_m}$ are the components, with respect to the basis. A defining property of tensors, just like vectors and dual vectors, is that they do not change under coordinate transformation, only their components change. Furthermore, if a tensor equation holds true in some coordinate system, then it is true in all coordinate systems!

1.3.2 Tensor Operations

Just like vectors and matrices, tensors can be added/subtracted, multiplied by a scalar, or as we have seen, multiplied by a tensor to give another tensor. In addition to these standard operations, tensors can also be contracted with each other to produce another tensor of lower rank.

Addition: Two tensors, T and S, with components $T^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q}$ and $S^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q}$ respectively, can be added or subtracted providing that they have matching indices, i.e

$$T^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q} \pm S^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q} = (T \pm S)^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q} \tag{1.3.5}$$

Scalar multiplication: Any tensor, T, can be multiplied by a scalar, ψ, by simply multiplying each component $T^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q}$ of the tensor, by ψ. i.e

$$\psi T = \psi T^{\alpha_1,...,\alpha_p}_{\beta_1,...,\beta_q} \tag{1.3.6}$$

Contraction: The contraction of two tensors is found by taking an upper index from one tensor and a lower index from the other, setting them equal to the same variable, and then summing over that variable by use of the Einstein summation convention, i.e

$$T^{\alpha_2,...,\alpha_n}_{\beta_1,...,\beta_{n-1}} = \sum_{\kappa=1}^{n} V^{\kappa,\alpha_2,...,\alpha_n} W_{\beta_1,...,\beta_{n-1},\kappa} \tag{1.3.7}$$

Remark: A single tensor can also be contracted in a similar manner. For example, given a rank 3 tensor with components $T^{\alpha}_{\beta\gamma}$, we can perform a contraction by setting $\alpha = \gamma$ and taking the sum, to obtain a rank 1 tensor, given by

$$T_\beta = \sum_\gamma T^\gamma_{\beta\gamma} \tag{1.3.8}$$

1.3.3 Transformation Laws

Consider a tensor, J, of type $(m, 0)$ i.e $J : T_p \times ... \times T_p \to \mathbb{R}$. In terms of the components and basis we have

$$J = J^{\mu_1,...,\mu_1} \partial_{\mu_1} \otimes ... \otimes \partial_{\mu_m} \ . \tag{1.3.9}$$

Under a coordinate transformation, we then obtain

$$J^{\mu_1,...,\mu_m} \partial_{\mu_1} \otimes ... \otimes \partial_{\mu_m} = J^{\lambda_1',...,\lambda_m'} \partial_{\lambda_1'} \otimes ... \otimes \partial_{\lambda_m'} \ . \tag{1.3.10}$$

Using the chain rule, we rewrite the partial derivatives in the RHS, in terms of those on the LHS.

$$\begin{aligned} J^{\mu_1,...,\mu_m} \partial_{\mu_1} \otimes ... \otimes \partial_{\mu_m} &= J^{\lambda_1',...,\lambda_m'} \frac{\partial x^{\mu_1}}{\partial \xi^{\lambda_1'}} \frac{\partial}{\partial x^{\mu_1}} \otimes ... \otimes \frac{\partial x^{\mu_m}}{\partial \xi^{\lambda_m'}} \frac{\partial}{\partial x^{\mu_m}} && (1.3.11) \\ &= J^{\lambda_1',...,\lambda_m'} \frac{\partial x^{\mu_1}}{\partial \xi^{\lambda_1'}} \frac{\partial x^{\mu_2}}{\partial \xi^{\lambda_2'}} ... \frac{\partial x^{\mu_m}}{\partial \xi^{\lambda_m'}} \partial_{\mu_1} \otimes ... \otimes \partial_{\mu_m} \ . && (1.3.12) \end{aligned}$$

Hence the components transform according to

$$J^{\lambda_1',...,\lambda_m'} = \frac{\partial \xi^{\lambda_1'}}{\partial x^{\mu_1}} \frac{\partial \xi^{\lambda_2'}}{\partial x^{\mu_2}} ... \frac{\partial \xi^{\lambda_m'}}{\partial x^{\mu_m}} J^{\mu_1,...,\mu_m} \ . \tag{1.3.13}$$

Components which transform like this, just like components of tangent vectors, are said to be contravariant, and again are written with the indices as a superscript!

A similar law holds for a tensor of type $(0, n)$, which can be written in terms of its components and basis, in either coordinate system

$$\begin{aligned} J &= J_{\sigma_1,...,\sigma_n} dx^{\sigma_1} \otimes ... \otimes dx^{\sigma_n} && (1.3.14) \\ &= J_{\rho_1',...,\rho_n'} d\xi^{\rho_1'} \otimes ... \otimes d\xi^{\rho_n'} && (1.3.15) \\ &= J_{\rho_1',...,\rho_n'} \frac{\partial \xi^{\rho_1'}}{\partial x^{\sigma_1}} \frac{\partial \xi^{\rho_2'}}{\partial x^{\sigma_2}} ... \frac{\partial \xi^{\rho_n'}}{\partial x^{\sigma_n}} dx^{\sigma_1} \otimes ... \otimes dx^{\sigma_n} && (1.3.16) \end{aligned}$$

Hence, components of this kind transform according to

$$J_{\rho_1',...,\rho_n'} = \frac{\partial x^{\sigma_1}}{\partial \xi^{\rho_1'}} \frac{\partial x^{\sigma_2}}{\partial \xi^{\rho_2'}} ... \frac{\partial x^{\sigma_n}}{\partial \xi^{\rho_n'}} J_{\sigma_1,...,\sigma_n} \tag{1.3.17}$$

Components which transform in this way, just like dual vectors, are called covariant, and are written with the index as a subscript.

Finally, for a **mixed** tensor, i.e one with both contravariant and covariant components, the components transform according to

$$J^{\alpha'_1,\ldots,\alpha'_p}_{\beta'_1,\ldots,\beta'_q} = \frac{\partial \xi^{\alpha'_1}}{\partial x^{\rho_1}} \cdots \frac{\partial \xi^{\alpha'_p}}{\partial x^{\rho_p}} \frac{\partial x^{\sigma_1}}{\partial \xi^{\beta'_1}} \cdots \frac{\partial x^{\sigma_q}}{\partial \xi^{\beta'_q}} J^{\rho_1,\ldots,\rho_p}_{\sigma_1,\ldots,\sigma_q} \tag{1.3.18}$$

Using these transformation laws, the operations defined in equations (1.3.5) - (1.3.7) are easily verified.

Example 1. *Suppose we have two tensors $V^{\rho_1,\ldots,\rho_p}$ and $W_{\sigma_1,\ldots,\sigma_q}$. In order to show that the tensor product of two tensors is in fact a tensor, we need to consider how the product transforms. i.e*

$$V^{\alpha'_1,\ldots,\alpha'_k} \otimes W_{\beta'_1,\ldots,\beta'_q} = \frac{\partial \xi^{\alpha'_1}}{\partial x^{\rho_1}} \cdots \frac{\partial \xi^{\alpha'_p}}{\partial x^{\rho_k}} \frac{\partial x^{\sigma_1}}{\partial \xi^{\beta'_1}} \cdots \frac{\partial x^{\sigma_q}}{\partial \xi^{\beta'_q}} (V^{\rho_1,\ldots,\rho_k} \otimes W_{\sigma_1,\ldots,\sigma_q}) \tag{1.3.19}$$

Notice the right hand side of this equation transforms in the same way as mixed tensor. Hence, the tensor product of a contravariant tensor and covariant tensor results in a mixed tensor!

1.3.4 The Metric Tensor

In this section we introduce a fundamental tensor, namely the metric tensor, which basically provides us with a way of measuring distances on a manifold. There exists a multitude of ways to derive this tensor; for our purposes the simplest way of defining it is by use of the line element. In later sections, we will see that many of the objects which are required to formulate the theory of general relativity (such as the connection coefficients, and the various curvature tensors) are constructed using the metric tensor!

Recall, in $\mathbb{R}^n$ we can describe a curve as a vector equation $\vec{\gamma}(t) = x^\mu(t)\hat{e}_\mu$, where $x^\mu(t)$, are the coordinates of points along the curve, with respect to the canonical Cartesian basis vectors, $\hat{e}_\mu$, satisfying $\langle \hat{e}_\mu, \hat{e}_\nu \rangle = \hat{e}_\mu \cdot \hat{e}_\nu = \delta_{\mu\nu}$. The line element can then be found by taking the inner product (dot product) of an infinitesimal displacement vector along the curve, given by $d\vec{s} = dx^\mu \hat{e}_\mu$, with itself. i.e

$$ds^2 = \langle d\vec{s}, d\vec{s} \rangle = \langle dx^\mu \hat{e}_\mu, dx^\nu \hat{e}_\nu \rangle = \langle \hat{e}_\mu, \hat{e}_\nu \rangle dx^\mu dx^\nu = \delta_{\mu\nu} dx^\mu dx^\nu \ . \tag{1.3.20}$$

By expanding out the sum over μ and ν, one then obtains a more familiar form of the line element

$$ds^2 = (dx^1)^2 + (dx^2)^2 + \ldots + (dx^n)^2 \ . \tag{1.3.21}$$

To generalise this, consider a smooth manifold $\mathbf{M}$ endowed with coordinates x^μ, and suppose that $\gamma(t) = x^\mu(t)$, is a curve passing through some point $p \in \mathbf{M}$, at which we construct the tangent space $T_p\mathbf{M}$, with basis $\{\partial_\mu\}$. We can now write an infinitesimal displacement vector at p, along the curve γ as $d\vec{s} = dx^\mu(p)\partial_\mu|_p$, and hence, in the same way as above, taking the inner product of $d\vec{s}$ with itself, produces the line element

$$\begin{aligned} ds^2 &= \langle d\vec{s}, d\vec{s} \rangle = \langle dx^\mu(p)\partial_\mu|_p, dx^\nu(p)\partial_\nu|_p \rangle \\ &= \langle \partial_\mu, \partial_\nu \rangle|_p dx^\mu(p) dx^\nu(p) \\ &= g_{\mu\nu}|_p dx^\mu(p) dx^\nu(p) \ , \end{aligned} \tag{1.3.22}$$

where $g_{\mu\nu}|_p = \langle \partial_\mu, \partial_\nu \rangle|_p$ defines the components of the **metric tensor** at the point p. Note that since our choice in γ and p were arbitrary, we can infer that $g_{\mu\nu}|_p$ exist for all $p \in \mathbf{M}$, and thus we just write

$$g_{\mu\nu} = \langle \partial_\mu, \partial_\nu \rangle \ . \tag{1.3.23}$$

Furthermore, given any two vectors, $\vec{V} = V^\mu \partial_\mu$ and $\vec{W} = W^\nu \partial_\nu$, that are elements of the same tangent space, we can write the inner of the two vectors as

$$\langle \vec{V}, \vec{W} \rangle = g_{\mu\nu} V^\mu W^\nu \tag{1.3.24}$$

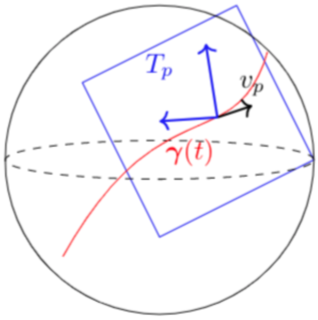

Figure 3: Tangent space of a 2-sphere

It is desirable to note that the signature of a metric is determined by the number of positive and negative eigenvalues it has: If all eigenvalues are zero, the metric is said to be degenerate, if the metric has has all positive eigenvalues, i.e it is positive definite, then the metric is said to be **Riemannian**. Similarly, a metric with a single negative eigenvalue is called **pseudo-Riemannian** or **Lorentzian** and finally, a metric which has a mix of some positive and some negative eigenvalues is said to be **indefinite**. In addition, a manifold endowed with a Riemannian metric is called a Riemannian manifold... etc. To show that $g_{\alpha\beta}$ is in fact a fully covariant rank-2 tensor, we look at how it transforms. Consider the metric in a new coordinate system $\xi^{\mu'}$, i.e

$$g_{\mu'\nu'} = \langle \partial_{\mu'}, \partial_{\nu'} \rangle \tag{1.3.25}$$

By the chain rule we can write $g_{\mu'\nu'}$ in terms of the old coordinate system x^α

$$g_{\mu'\nu'} = \left\langle \frac{\partial x^\alpha}{\partial \xi^{\mu'}} \frac{\partial}{\partial x^\alpha}, \frac{\partial x^\beta}{\partial \xi^{\nu'}} \frac{\partial}{\partial x^\alpha} \right\rangle \tag{1.3.26}$$

$$= \frac{\partial x^\alpha}{\partial \xi^{\mu'}} \frac{\partial x^\beta}{\partial \xi^{\nu'}} \langle \partial_\alpha, \partial_\beta \rangle \ , \tag{1.3.27}$$

and hence, since $g_{\alpha\beta} = \langle \partial_\alpha, \partial_\beta \rangle$, we arrive at the transformation law

$$g_{\mu'\nu'} = \frac{\partial x^\alpha}{\partial \xi^{\mu'}} \frac{\partial x^\beta}{\partial \xi^{\nu'}} g_{\alpha\beta} \ . \tag{1.3.28}$$

It is now apparent that the metric does transform, as expected, like a fully covariant rank-2 tensor! One should also note that, due to the symmetry of the inner product, the metric tensor is symmetric in its indices, i.e $g_{\mu\nu} = g_{\nu\mu}$.

Corollary 1. *In Minkowski space, the metric tensor is given by*

$$\eta_{\alpha\beta} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ . \tag{1.3.29}$$

It immediately follows that: if at some point on a pseudo-Riemannian manifold, the metric locally takes on the form of $\eta_{\alpha\beta}$, then we can write the metric for the general non-local coordinate system as

$$g_{\mu'\nu'} = \frac{\partial x^\alpha}{\partial \xi^{\mu'}} \frac{\partial x^\beta}{\partial \xi^{\nu'}} \eta_{\alpha\beta} \ . \tag{1.3.30}$$

Moreover, in the context of general relativity: at any point in space-time (also a pseudo-Riemannian manifold), one can construct a coordinate system such that the metric at that point takes on the form of $\eta_{\alpha\beta}$, in other words, space-time is locally flat. This is the idea of locally inertial coordinates and will be discussed further in a later section.

Example 2. *(Spherical Polar Coordinates) In this example, we will construct the metric tensor and line element in spherical polar coordinates. Recall, points on a 2-sphere can be described by a position vector (using the standard Cartesian basis) given by*

$$\mathbf{r}(\rho,\theta,\phi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \rho \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix} . \tag{1.3.31}$$

By differentiating, we obtain a set of linearly independent basis vectors $\{\partial_\rho, \partial_\theta, \partial_\phi\}$ which span $\mathbb{R}^3$

$$\partial_\rho = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix} \quad \partial_\theta = \rho \begin{bmatrix} \cos(\theta)\cos(\phi) \\ \cos(\theta)\sin(\phi) \\ -\sin(\theta) \end{bmatrix} \quad \partial_\phi = \rho \begin{bmatrix} -\sin(\theta)\sin(\phi) \\ \sin(\theta)\cos(\phi) \\ 0 \end{bmatrix} , \tag{1.3.32}$$

and hence, the non-zero components of the metric tensor are then easily found to be

$$g_{\rho\rho} = \langle \partial_\rho, \partial_\rho \rangle = 1 , \tag{1.3.33}$$

$$g_{\theta\theta} = \langle \partial_\theta, \partial_\theta \rangle = \rho^2 , \tag{1.3.34}$$

$$g_{\phi\phi} = \langle \partial_\phi, \partial_\phi \rangle = \rho^2 \sin^2(\theta) . \tag{1.3.35}$$

Notice that all off-diagonal components are zero! Hence, the metric tensor can be written as the matrix

$$g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho^2\sin^2(\theta) \end{bmatrix} . \tag{1.3.36}$$

Therefore, using equation (1.3.22), we obtain the line element in spherical polar coordinates

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2(\theta) d\phi^2 . \tag{1.3.37}$$

Remark: If we consider a sphere of constant radius, i.e $\rho = const$, then clearly the component $g_{\rho\rho}$ will vanish and so the line element on the surface of the sphere will be given by

$$ds^2 = \rho^2 (d\theta)^2 + \rho^2 \sin^2(d\phi)^2 . \tag{1.3.38}$$

Suppose that in some coordinate system x^γ, we construct the metric $g_{\gamma\delta}$ and define its inverse to be $g^{\delta\rho}$ (we assume it is nondegenerate) so that in this coordinate system we have

$$g_{\gamma\delta} g^{\delta\rho} = \delta^\rho_\gamma . \tag{1.3.39}$$

Since $g^{\delta\rho}$ transforms as a fully contravariant rank-2 tensor, we have in another coordinate system, ξ^κ,

$$\begin{aligned} g_{\kappa'\varsigma'} g^{\varsigma'\nu'} &= \frac{\partial x^\gamma}{\partial \xi^{\kappa'}} \frac{\partial x^\delta}{\partial \xi^{\varsigma'}} g_{\gamma\delta} \frac{\partial \xi^{\varsigma'}}{\partial x^\delta} \frac{\partial \xi^{\nu'}}{\partial x^\rho} g^{\delta\rho} = \frac{\partial x^\gamma}{\partial \xi^{\kappa'}} \frac{\partial x^\delta}{\partial \xi^{\varsigma'}} \frac{\partial \xi^{\varsigma'}}{\partial x^\delta} \frac{\partial \xi^{\nu'}}{\partial x^\rho} g_{\gamma\delta} g^{\delta\rho} \\ &= \frac{\partial x^\gamma}{\partial \xi^{\kappa'}} \delta^\delta_\delta \frac{\partial \xi^{\nu'}}{\partial x^\rho} g_{\gamma\delta} g^{\delta\rho} = \frac{\partial x^\gamma}{\partial \xi^{\kappa'}} \frac{\partial \xi^{\nu'}}{\partial x^\rho} g_{\gamma\delta} g^{\delta\rho} . \end{aligned} \tag{1.3.40}$$

We can now apply equation (1.3.39) to obtain

$$g_{\kappa'\varsigma'} g^{\varsigma'\nu'} = \frac{\partial x^\gamma}{\partial \xi^{\kappa'}} \frac{\partial \xi^{\nu'}}{\partial x^\rho} \delta^\rho_\gamma = \frac{\partial x^\rho}{\partial \xi^{\kappa'}} \frac{\partial \xi^{\nu'}}{\partial x^\rho} = \frac{\partial \xi^{\nu'}}{\partial \xi^{\kappa'}} = \delta^{\nu'}_{\kappa'} . \tag{1.3.41}$$

Hence, we have shown that equation (1.3.39), as we would expect from a tensor equation, holds true in all coordinate systems.

1.4 The Covariant Derivative and Christoffel Symbols

Suppose we define a vector, $\vec{V} = V^\alpha \partial_\alpha$, where $V^\alpha = V^\alpha(x^\beta)$, are the components of the vector with respect to the covariant basis $\{\partial_\alpha\}$. Differentiating $\vec{V}$ with respect to x^β using the product rule then yields

$$\frac{\partial \vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\partial_\alpha + V^\alpha \frac{\partial}{\partial x^\beta}(\partial_\alpha) \tag{1.4.1}$$

Notice, the first term is a vector whose components are $\partial_\beta(V^\alpha)$ with respect to the basis vector ∂_α and hence transforms as a rank-1 tensor. The second term however, does not transform as a true tensor unless the basis from which we construct our vectors happens to be constant! We therefore require some new machinery to allow us to differentiate a tensor to obtain a new tensor. In the case where the basis vectors are dependent on position, we define a set of tensor-like objects (not true tensors) called the **connection coefficients** $\Gamma^\gamma_{\beta\alpha}$, which are given to represent the magnitude of change in the direction ∂_γ, of the basis vector ∂_α due to a change in the coordinate x^β. We define

$$\frac{\partial}{\partial x^\beta}(\partial_\alpha) = \Gamma^\gamma_{\beta\alpha}\partial_\gamma \ . \tag{1.4.2}$$

Hence, upon a second inspection of equation (1.4.1) we see that the second term can be written using the connection coefficients. Thus, we now have

$$\frac{\partial \vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\partial_\alpha + V^\alpha \Gamma^\gamma_{\beta\alpha}\partial_\gamma \ . \tag{1.4.3}$$

Relabeling the dummy indices in the second term $\alpha \to \gamma$, $\gamma \to \alpha$ then allows us to factor out the basis vector ∂_α as follows

$$\begin{aligned} \frac{\partial \vec{V}}{\partial x^\beta} &= \frac{\partial V^\alpha}{\partial x^\beta}\partial_\alpha + V^\gamma \Gamma^\alpha_{\beta\gamma}\partial_\alpha \\ &= \left[\frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\gamma}V^\gamma\right]\partial_\alpha \ . \end{aligned} \tag{1.4.4}$$

The term in the square brackets, as we will show, is now in fact a true tensor and thus we are left with a tidy expression for the **covariant derivative** of contravariant vector components

$$\nabla_\beta V^\alpha = \frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\gamma}V^\gamma \ . \tag{1.4.5}$$

This expression tells us how the vector field $\vec{V} = V^\alpha \partial_\alpha$ changes due to change in the x^β coordinate. It is worth noting that for some scalar field, say f, the covariant derivative of f is the same as just taking the partial derivative, since a scalar will always transform as a scalar (rank 0 tensor) i.e

$$\nabla_\beta f = \frac{\partial f}{\partial x^\beta} \ . \tag{1.4.6}$$

Another property of the covariant derivative is that, like ordinary derivatives, it obeys the Leibniz rule (or product rule) [3], i.e

$$\nabla(U \otimes V) = U \otimes \nabla V + V \otimes \nabla U \ . \tag{1.4.7}$$

Using these results, we can now derive a similar expression for the covariant derivative of covariant components by considering the covariant derivative of the scalar quantity produced by contracting a vector V^α, with a dual vector W_α. By equation (1.4.7) we can write this as

$$\nabla_\beta(V^\alpha W_\alpha) = V^\alpha \nabla_\beta W_\alpha + W_\alpha \nabla_\beta V^\alpha \ , \tag{1.4.8}$$

and hence,

$$V^\alpha \nabla_\beta W_\alpha = \nabla_\beta(V^\alpha W_\alpha) - W_\alpha \nabla_\beta V^\alpha \ . \tag{1.4.9}$$

Now, since the first term on the right hand side is the covariant derivative of a scalar quantity, we can replace the covariant derivative with ordinary partial derivatives here. Notice also that the second term contains the covariant derivative given in equation (1.4.5). Thus, we now have

$$\begin{aligned} V^\alpha \nabla_\beta W_\alpha &= \frac{\partial}{\partial x^\beta}(V^\alpha W_\alpha) - W_\alpha \Big(\frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\gamma} V^\gamma \Big) \\ &= V^\alpha \frac{\partial W_\alpha}{\partial x^\beta} + W_\alpha \frac{\partial V^\alpha}{\partial x^\beta} - W_\alpha \frac{\partial V^\alpha}{\partial x^\beta} - W_\alpha \Gamma^\alpha_{\beta\gamma} V^\gamma \\ &= V^\alpha \frac{\partial W_\alpha}{\partial x^\beta} - W_\alpha \Gamma^\alpha_{\beta\gamma} V^\gamma \ . \end{aligned} \tag{1.4.10}$$

Relabelling $\alpha \to \gamma$ and $\gamma \to \alpha$ in the last term now gives

$$V^\alpha \nabla_\beta W_\alpha = V^\alpha \frac{\partial W_\alpha}{\partial x^\beta} - W_\gamma \Gamma^\gamma_{\beta\alpha} V^\alpha \tag{1.4.11}$$

$$\implies \qquad (\nabla_\beta W_\alpha - \frac{\partial W_\alpha}{\partial x^\beta} + W_\gamma \Gamma^\gamma_{\beta\alpha}) V^\alpha = 0 \ , \tag{1.4.12}$$

and since V^α in general is non-zero, it follows that

$$\nabla_\beta W_\alpha - \frac{\partial W_\alpha}{\partial x^\beta} + W_\gamma \Gamma^\gamma_{\beta\alpha} = 0 \ . \tag{1.4.13}$$

Hence, upon rearranging this equation, we arrive at the covariant derivative of covariant components

$$\nabla_\beta W_\alpha = \frac{\partial W_\alpha}{\partial x^\beta} - \Gamma^\gamma_{\beta\alpha} W_\gamma \ . \tag{1.4.14}$$

To this end, we can now generalise the covariant derivative to an arbitrary (m, n) tensor by simply writing an extra connection coefficient for each tensor index: $+\Gamma$ for contravariant indices and $-\Gamma$ for covariant indices [3], and so

$$\begin{aligned} \nabla_\beta T^{\mu_1,\dots,\mu_m}_{\nu_1,\dots,\nu_n} = \frac{\partial T^{\mu_1,\dots,\mu_m}_{\nu_1,\dots,\nu_n}}{\partial x^\beta} &+ \Gamma^{\mu_1}_{\beta\alpha} T^{\alpha,\dots,\mu_m}_{\nu_1,\dots,\nu_n} + \Gamma^{\mu_2}_{\beta\alpha} T^{\mu_1,\alpha,\dots,\mu_m}_{\nu_1,\dots,\nu_n} + \dots \\ &- \Gamma^\alpha_{\beta\nu_1} T^{\mu_1,\dots,\mu_m}_{\alpha,\dots,\nu_n} - \Gamma^\alpha_{\beta\nu_2} T^{\mu_1,\dots,\mu_m}_{\nu_1,\alpha,\dots,\nu_n} - \dots \end{aligned} \tag{1.4.15}$$

1.4.1 Transformation of the Connection Coefficients

In order to show that the covariant derivative of a vector is actually a tensor, we need to consider how each term in (1.4.5) transforms. Using equation (1.4.2), we can define the connection coefficients in a new coordinate system $\xi^{\eta'}$, as

$$\frac{\partial}{\partial \xi^{\alpha'}}(\partial_{\mu'}) = \Gamma^{\rho'}_{\alpha'\mu'} \partial_{\rho'} \ . \tag{1.4.16}$$

Thus, using the transformation laws from section 1.2.2, we can write $\partial_{\mu'}$ in terms of the coordinate system x^ν to give

$$\frac{\partial}{\partial \xi^{\alpha'}}\Big(\frac{\partial x^\beta}{\partial \xi^{\mu'}} \partial_\beta \Big) = \Gamma^{\rho'}_{\alpha'\mu'} \frac{\partial x^\beta}{\partial \xi^{\rho'}} \partial_\beta \ . \tag{1.4.17}$$

Carrying out the differentiation on the left hand side of this equation by use of the product rule now yields

$$\frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} \partial_\beta + \frac{\partial x^\beta}{\partial \xi^{\mu'}} \frac{\partial}{\partial \xi^{\alpha'}}(\partial_\beta) = \Gamma^{\rho'}_{\alpha'\mu'} \frac{\partial x^\beta}{\partial \xi^{\rho'}} \partial_\beta \ . \tag{1.4.18}$$

Notice, using the chain rule we can express the partial derivative acting on the basis vector ∂_β as

$$\frac{\partial}{\partial \xi^{\alpha'}} = \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial}{\partial x^\tau} \ , \tag{1.4.19}$$

and hence, our equations becomes

$$\frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} \partial_\beta + \frac{\partial x^\beta}{\partial \xi^{\mu'}} \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial}{\partial x^\tau}(\partial_\beta) = \Gamma^{\rho'}_{\alpha'\mu'} \frac{\partial x^\beta}{\partial \xi^{\rho'}} \partial_\beta \ . \tag{1.4.20}$$

By spotting the connection coefficient in the second term on the left hand side, we can now use (1.4.2) to obtain

$$\frac{\partial^2 x^\beta}{\partial\xi^{\alpha'}\partial\xi^{\mu'}}\partial_\beta + \frac{\partial x^\beta}{\partial\xi^{\mu'}}\frac{\partial x^\tau}{\partial\xi^{\alpha'}}\Gamma^\rho_{\tau\beta}\partial_\rho = \Gamma^{\rho'}_{\alpha'\mu'}\frac{\partial x^\beta}{\partial\xi^{\rho'}}\partial_\beta \ . \tag{1.4.21}$$

Observe that in this second term we are summing over the dummy variables β, ρ and can thus re-label $\beta \to \rho$ and $\rho \to \beta$ to give a common factor of ∂_β on both sides of the equation and hence, we now have

$$\frac{\partial^2 x^\beta}{\partial\xi^{\alpha'}\partial\xi^{\mu'}}\partial_\beta + \frac{\partial x^\rho}{\partial\xi^{\mu'}}\frac{\partial x^\tau}{\partial\xi^{\alpha'}}\Gamma^\beta_{\tau\rho}\partial_\beta = \Gamma^{\rho'}_{\alpha'\mu'}\frac{\partial x^\beta}{\partial\xi^{\rho'}}\partial_\beta$$

$$\implies \quad \left(\Gamma^{\rho'}_{\alpha'\mu'}\frac{\partial x^\beta}{\partial\xi^{\rho'}} - \frac{\partial^2 x^\beta}{\partial\xi^{\alpha'}\partial\xi^{\mu'}} - \frac{\partial x^\rho}{\partial\xi^{\mu'}}\frac{\partial x^\tau}{\partial\xi^{\alpha'}}\Gamma^\beta_{\tau\rho}\right)\partial_\beta = 0 \ . \tag{1.4.22}$$

Since the partial derivatives ∂_β form a set of linearly independent basis vectors then we must have that the quantity inside the brackets vanishes and so we can write

$$\Gamma^{\rho'}_{\alpha'\mu'}\frac{\partial x^\beta}{\partial\xi^{\rho'}} = \frac{\partial^2 x^\beta}{\partial\xi^{\alpha'}\partial\xi^{\mu'}} + \frac{\partial x^\rho}{\partial\xi^{\mu'}}\frac{\partial x^\tau}{\partial\xi^{\alpha'}}\Gamma^\beta_{\tau\rho} \ . \tag{1.4.23}$$

Finally, by multiplying both sides of this equation by $\partial_\beta\xi^{\sigma'}$ and noticing that

$$\frac{\partial x^\beta}{\partial\xi^{\rho'}}\frac{\partial\xi^{\sigma'}}{\partial x^\beta} = \delta^{\sigma'}_{\rho'} \ , \tag{1.4.24}$$

we obtain the transformation law for the connection coefficients [9]

$$\Gamma^{\sigma'}_{\alpha'\mu'} = \frac{\partial\xi^{\sigma'}}{\partial x^\beta}\frac{\partial x^\rho}{\partial\xi^{\mu'}}\frac{\partial x^\tau}{\partial\xi^{\alpha'}}\Gamma^\beta_{\tau\rho} + \frac{\partial\xi^{\sigma'}}{\partial x^\beta}\frac{\partial^2 x^\beta}{\partial\xi^{\alpha'}\partial\xi^{\mu'}} \ . \tag{1.4.25}$$

It is important to note that **the connection coefficients alone are not tensors**, since a true tensor of type $(1, 2)$ would transform according to the rule

$$T^{\sigma'}_{\alpha'\mu'} = \frac{\partial\xi^{\sigma'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial\xi^{\alpha'}}\frac{\partial x^\rho}{\partial\xi^{\mu'}}T^\mu_{\nu\rho} \ . \tag{1.4.26}$$

1.4.2 Transformation of the Covariant Derivative

We now consider the covariant derivative expressed in a new coordinate system $\xi^{\mu'}$, which can be written as

$$\nabla_{\alpha'}V^{\sigma'} = \frac{\partial V^{\sigma'}}{\partial\xi^{\alpha'}} + \Gamma^{\sigma'}_{\alpha'\mu'}V^{\mu'} \ . \tag{1.4.27}$$

Using the transformation rule for contravariant vector components derived in section (1.2.1), we can write the first term on the right hand side as

$$\frac{\partial V^{\sigma'}}{\partial\xi^{\alpha'}} = \frac{\partial}{\partial\xi^{\alpha'}}\left(\frac{\partial\xi^{\sigma'}}{\partial x^\rho}V^\rho\right) , \tag{1.4.28}$$

Then, by invoking the chain rule, we can also re-write the partial derivative acting on the term inside the brackets in terms of the un-primed coordinates to give

$$\frac{\partial V^{\sigma'}}{\partial\xi^{\alpha'}} = \frac{\partial x^\beta}{\partial\xi^{\alpha'}}\frac{\partial}{\partial x^\beta}\left(\frac{\partial\xi^{\sigma'}}{\partial x^\rho}V^\rho\right) . \tag{1.4.29}$$

Thus, carrying out the differentiation by use of the product rule gives the result

$$\frac{\partial V^{\sigma'}}{\partial\xi^{\alpha'}} = \frac{\partial x^\beta}{\partial\xi^{\alpha'}}\frac{\partial\xi^{\sigma'}}{\partial x^\rho}\frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial x^\beta}{\partial\xi^{\alpha'}}\frac{\partial^2\xi^{\sigma'}}{\partial x^\beta\partial x^\rho}V^\rho \ . \tag{1.4.30}$$

We can now substitute this expression along with equation (1.4.25) in to the covariant derivative given by equation (1.4.27) to obtain

$$\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\rho} \frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\beta \partial x^\rho} V^\rho + \Big(\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\rho}{\partial \xi^{\mu'}} \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \Gamma^\beta_{\tau\rho} + \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} \Big) V^{\mu'} \; . \tag{1.4.31}$$

Writing $V^{\mu'}$ in terms of the un-primed coordinate system now gives

$$\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\rho} \frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\beta \partial x^\rho} V^\rho + \Big(\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\rho}{\partial \xi^{\mu'}} \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \Gamma^\beta_{\tau\rho} + \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} \Big) \Big(\frac{\partial \xi^{\mu'}}{\partial x^\rho} V^\rho \Big) \; . \tag{1.4.32}$$

In order to simplify this equation, consider the identity

$$\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\beta}{\partial \xi^{\mu'}} = \delta^{\sigma'}_{\mu'} \; . \tag{1.4.33}$$

If we differentiate (1.4.33) with respect to $\xi^{\alpha'}$ we then obtain

$$\frac{\partial}{\partial \xi^{\alpha'}} \Big(\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\beta}{\partial \xi^{\mu'}} \Big) = \frac{\partial}{\partial \xi^{\alpha'}} \Big(\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \Big) \frac{\partial x^\beta}{\partial \xi^{\mu'}} + \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} = 0 \; . \tag{1.4.34}$$

Hence, by the chain rule we get

$$\frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial}{\partial x^\tau} \Big(\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \Big) \frac{\partial x^\beta}{\partial \xi^{\mu'}} + \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} = 0 \; , \tag{1.4.35}$$

and we can now evaluate the remaining derivative to obtain

$$\frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\tau \partial x^\beta} \frac{\partial x^\beta}{\partial \xi^{\mu'}} + \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} = 0 \; . \tag{1.4.36}$$

Notice that the second term in this expression is present in equation (1.4.32). Rearranging this equation then gives

$$\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial^2 x^\beta}{\partial \xi^{\alpha'} \partial \xi^{\mu'}} = - \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\tau \partial x^\beta} \frac{\partial x^\beta}{\partial \xi^{\mu'}} \; , \tag{1.4.37}$$

and hence, by substituting this expression in to (1.4.32) we obtain

$$\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\rho} \frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\beta \partial x^\rho} V^\rho + \Big(\frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\rho}{\partial \xi^{\mu'}} \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \Gamma^\beta_{\tau\rho} - \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\tau \partial x^\beta} \frac{\partial x^\beta}{\partial \xi^{\mu'}} \Big) \Big(\frac{\partial \xi^{\mu'}}{\partial x^\rho} V^\rho \Big) \; . \tag{1.4.38}$$

By expanding the brackets and noticing that we have

$$\frac{\partial x^\rho}{\partial \xi^{\mu'}} \frac{\partial \xi^{\mu'}}{\partial x^\rho} = 1 \quad \text{and} \quad \frac{\partial x^\beta}{\partial \xi^{\mu'}} \frac{\partial \xi^{\mu'}}{\partial x^\rho} = \delta^\beta_\rho \; , \tag{1.4.39}$$

we then see that

$$\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\rho} \frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\beta \partial x^\rho} V^\rho + \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \Gamma^\beta_{\tau\rho} V^\rho - \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\tau \partial x^\beta} \delta^\beta_\rho V^\rho \; . \tag{1.4.40}$$

Using δ^β_ρ to sum over ρ in the last term, and also relabeling $\tau \to \beta$ and $\beta \to \tau$ in the third and last terms now gives

$$\begin{aligned}\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\rho} \frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\beta \partial x^\rho} V^\rho \\ + \frac{\partial \xi^{\sigma'}}{\partial x^\tau} \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \Gamma^\tau_{\beta\rho} V^\rho - \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\tau \partial x^\rho} V^\rho \ .\end{aligned} \tag{1.4.41}$$

Hence, by cancelling the second and last terms we obtain

$$\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\rho} \frac{\partial V^\rho}{\partial x^\beta} + \frac{\partial \xi^{\sigma'}}{\partial x^\tau} \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \Gamma^\tau_{\beta\rho} V^\rho \ , \tag{1.4.42}$$

and finally, by relabeling $\rho \to \tau$ in the first term, we are able to factor out the partial derivatives in front of the connection coefficient, i.e

$$\begin{aligned}\nabla_{\alpha'} V^{\sigma'} &= \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\tau} \frac{\partial V^\tau}{\partial x^\beta} + \frac{\partial \xi^{\sigma'}}{\partial x^\tau} \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \Gamma^\tau_{\beta\rho} V^\rho \\ &= \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\tau} \Big(\frac{\partial V^\tau}{\partial x^\beta} + \Gamma^\tau_{\beta\rho} V^\rho \Big) \ .\end{aligned} \tag{1.4.43}$$

With this, we arrive at the transformation law for the covariant derivative

$$\nabla_{\alpha'} V^{\sigma'} = \frac{\partial x^\beta}{\partial \xi^{\alpha'}} \frac{\partial \xi^{\sigma'}}{\partial x^\tau} \nabla_\beta V^\tau \ , \tag{1.4.44}$$

and upon comparison with the transformation law for a tensor of type $(1,1)$,

$$T^{\lambda'}_{\rho'} = \frac{\partial x^\nu}{\partial \xi^{\rho'}} \frac{\partial \xi^{\lambda'}}{\partial x^\mu} T^\mu_\nu \ , \tag{1.4.45}$$

we can conclude that the covariant derivative of a vector (with contravariant components) is in fact a true tensor of type $(1,1)$, since it obeys the correct transformation law! Of course, we have only shown this for the single case of differentiating contravariant components, however similar transformation laws for covariant components, and higher rank tensors can be derived in a similar manner.

Corollary 2. *From equation (1.4.38) it is apparent that an alternative way of writing the transformation law for the connection coefficients is*

$$\Gamma^{\sigma'}_{\alpha'\mu'} = \frac{\partial \xi^{\sigma'}}{\partial x^\beta} \frac{\partial x^\rho}{\partial \xi^{\mu'}} \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \Gamma^\beta_{\tau\rho} - \frac{\partial x^\tau}{\partial \xi^{\alpha'}} \frac{\partial^2 \xi^{\sigma'}}{\partial x^\tau \partial x^\beta} \frac{\partial x^\beta}{\partial \xi^{\mu'}} \tag{1.4.46}$$

This alternative transformation law further supports the notion that the connection coefficients are not tensors.

1.4.3 Parallel Transport

In section 1.2.1 we defined vectors as being elements in a tangent space at some point on a manifold, however by the very nature of such spaces, it is evident that on a curved manifold such as a 2-sphere, there can exist an infinite amount of tangent spaces, the collection of which is called the tangent bundle. The question to be asked now is: how do we compare vectors that live in different spaces of the tangent bundle?

Recall, the directional derivative of a scalar field f along a curve $x^\mu(t)$ is defined to be

$$\frac{df}{dt} = \frac{dx^\mu}{dt} \partial_\mu f \ . \tag{1.4.47}$$

This equation intuitively tells us how much the scalar field changes from point to point along the curve. However, we have now developed the tools needed, namely the covariant derivative, to allow us to extend this definition to vectors and more generally, tensors.

We define **directional covariant derivative** of a vector V^σ, along a curve $x^\mu(t)$, to be

$$\frac{DV^\sigma}{Dt} = \frac{dx^\mu}{dt}\nabla_\mu V^\sigma \ . \tag{1.4.48}$$

In a similar way to ordinary directional derivatives, equation (1.4.48) tells us how much the vector changes from point to point along the curve $x^\mu(t)$. Moreover, by demanding that the directional covariant derivative vanishes, we arrive at the notion of **parallel transport**, i.e

$$\frac{dx^\mu}{dt}\nabla_\mu V^\sigma = 0 \ . \tag{1.4.49}$$

Using the expression for the covariant derivative in (1.4.5) we can now write this equation as

$$\frac{dx^\mu}{dt}\Big(\frac{\partial V^\sigma}{\partial x^\mu} + \Gamma^\sigma_{\mu\lambda}V^\lambda\Big) = 0 \ , \tag{1.4.50}$$

and hence, we obtain an alternative form of the parallel transport equation

$$\frac{dV^\sigma}{dt} + \frac{dx^\mu}{dt}\Gamma^\sigma_{\mu\lambda}V^\lambda = 0 \ . \tag{1.4.51}$$

This equation furnishes us with a way of taking a vector from one tangent space of a manifold and transporting it to another, while essentially keeping the vector unchanged as much as possible and thus allows us to compare vectors from different tangent spaces in the usual way (addition/subtraction, dot product...etc).

Remark: Consider two vector fields $V, W \in \mathfrak{X}(\mathbf{M})$. If we define the vector field W such that the integral/flow curves are given by $x^\mu(t)$, then equation (1.4.48) gives the covariant derivative of the vector field V in the direction of W. We write

$$(\nabla_W V)^\alpha = W^\beta \nabla_\beta V^\alpha \tag{1.4.52}$$

Lemma 1. *Parallel transport of a vector in $\mathbb{R}^n$ does not change the vector components, and is therefore path independent.*

Proof. Since, in $\mathbb{R}^n$ the basis vectors $\hat{e}_\alpha$ are constant, equation (1.4.2) would yield

$$\Gamma^\sigma_{\mu\lambda} = 0 \ , \tag{1.4.53}$$

and hence, by (1.4.51), the equation for parallel transporting a vector in $\mathbb{R}^n$ would reduce to simply

$$\frac{dV^\sigma}{dt} = 0 \ . \tag{1.4.54}$$

Suppose that initially, the vector components are given by $V^\sigma(t_0) = V^\sigma_0 = const$. Then, upon solving (1.4.54) we find that for all values of t, the vector components are given by

$$V^\sigma(t) = V^\sigma_0 \ . \tag{1.4.55}$$

Therefore, parallel transport of a vector in $\mathbb{R}^n$ does not change the vector components, and hence, is path independent! □

In section 1.6, we will see that on a curved manifold, such as a sphere, lemma 1 no longer holds true. In fact, curvature directly influences how a vector changes as it undergoes parallel transport.

1.4.4 The Torsion Tensor

In the theory of surfaces, torsion is thought of as a measure of how much a surface tends to twist about a curve. Perhaps the simplest example of a surface with non-zero torsion is the Möbius strip: a strip that has been twisted 180^o and glued together at the ends. If one were to ride a bicycle around the surface while holding one arm perpendicular to the direction of travel, then, after moving around one full circle, you would arrive back at the same point you started at, however you would be on the opposite side of the strip, and your arm would be pointing in the opposite direction to which it started. This change in orientation is a direct consequence of the torsion of the surface.

In the context of manifolds, torsion is given as the twisting of a tangent space as it is parallel transported along a curve. To begin casting this idea into a mathematical statement, consider two vector fields $V, W \in \mathfrak{X}(\mathbf{M})$ and recall, the commutator (Lie bracket) of V and W is given by

$$[V, W] = V(W) - W(V) \;, \tag{1.4.56}$$

which, in component notation can be expressed as

$$[V, W]^\alpha = V^\beta \partial_\beta W^\alpha - W^\beta \partial_\beta V^\alpha \tag{1.4.57}$$

To see what is happening here, consider the tangent vectors, which are elements of the vector fields V and W, given by $v_p = V^\alpha(p)\partial_\alpha$ and $w_p = W^\beta(p)\partial_\beta$, located at some point $p \in \mathbf{M}$.

From the point p, suppose we move an infinitesimal distance ε along v_p, w_p to the points $q_v = p + \varepsilon v_p$ and $q_w = p + \varepsilon w_p$ respectively. At the point q_v, the value of the vector field W is given by $w_q = W(q_v)$, and similarly, at q_w, the vector field V has the value $v_q = V(q_w)$. Once again, we move a distance ε along each of these vectors. We denote the point at the end of the vector εw_q as r_{vw}, and the point at the end of εv_q as r_{wv}.

$$r^\mu_{vw} = p^\mu + \varepsilon V^\mu(p) + \varepsilon W^\mu(p + \varepsilon v_p), \tag{1.4.58}$$
$$r^\mu_{wv} = p^\mu + \varepsilon W^\mu(p) + \varepsilon V^\mu(p + \varepsilon w_p) \;. \tag{1.4.59}$$

In general, these points may not be equivalent and in which case, the vector vector difference between them is given by

$$(r_{vw} - r_{wv})^\mu = \varepsilon V^\mu(p) + \varepsilon W^\mu(p + \varepsilon v_p) - \varepsilon W^\mu(p) - \varepsilon V^\mu(p + \varepsilon w_p) \;. \tag{1.4.60}$$

Since we took ε to be small, we can now Taylor the second and fourth terms around the point p, which up to first order yields

$$(r_{vw} - r_{wv})^\mu = \varepsilon V^\mu(p) + \varepsilon W^\mu(p) + \varepsilon^2 V^\nu(p)\partial_\nu W^\mu|_p - \varepsilon W^\mu(p) - \varepsilon V^\mu(p) - \varepsilon^2 W^\nu(p)\partial_\nu V^\mu|_p + ... \tag{1.4.61}$$
$$= \varepsilon^2 (V^\nu(p)\partial_\nu W^\mu|_p - W^\nu(p)\partial_\nu V^\mu|_p) + ... \tag{1.4.62}$$

Hence, on comparison with equation (1.4.57), we see that the vector between r_{wv} and r_{vw} is in fact given by the Lie bracket!

$$(r_{vw} - r_{wv})^\mu = \varepsilon^2 [V, W]^\mu|_p \tag{1.4.63}$$

Furthermore, if $[V, W] = 0$ then this implies that $r_{vw} = r_{wv}$ and it follows that the infinitesimal parallelogram formed in the way described above, is closed.

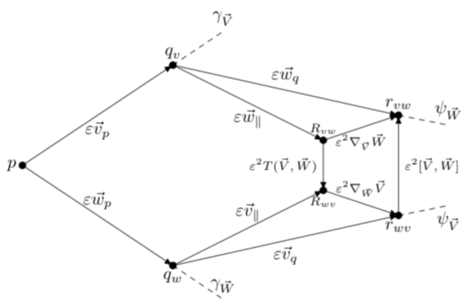

Figure 4: Geometric interpretation of the Lie bracket and Torsion tensor

Suppose now that from the point p we parallel transport the vector w_p along the integral curve γ_V, generated by the vector field V, by a distance ε, i.e to the point q_v. Denote the vector after it has been parallel transported as $w_\|$. If we move a distance ε along $w_{\parallel}$, we arrive at the point R_{vw}. Similarly, parallel transporting v_p a distance ε along the integral curve γ_W, generated by the vector field W, produces the vector $v_{\parallel}$ and at a distance ε along $v_{\parallel}$ lies the point R_{wv}.

To proceed, it it useful to consider the covariant derivative of a vector field Y in the direction of another vector field X as the following limit:

$$\nabla_X Y = \lim_{\varepsilon \to 0} \left[\frac{Y_{p+\varepsilon X} - Y_{\parallel}}{\varepsilon} \right] \tag{1.4.64}$$

Now, since we have already taken ε to be infinitesimal, we thus have that the vector which connects the point R_{vw} to the point r_{vw} is given by

$$\varepsilon w_q - \varepsilon w_{\parallel} = \varepsilon^2 \left(\frac{w_{p+\varepsilon v_p} - w_{\parallel}}{\varepsilon} \right) = \varepsilon^2 \nabla_V W \ . \tag{1.4.65}$$

In the same way, the vector connecting the points R_{wv} and r_{wv} is given by $\varepsilon^2 \nabla_W V$. To this end, we define the torsion tensor $T(V, W)$, which can be viewed as the vector connecting the points R_{vw} to R_{wv}, given by

$$T(V, W) = \nabla_V W - \nabla_W V - [V, W] \ . \tag{1.4.66}$$

From this, we can immediately see that in order for the torsion to vanish, we must have

$$\nabla_V W - \nabla_W V = [V, W] \ . \tag{1.4.67}$$

In terms of components we can write this as

$$V^\beta \partial_\beta W^\alpha + V^\beta \Gamma^\alpha_{\beta\mu} W^\mu - W^\beta \partial_\beta V^\alpha - W^\beta \Gamma^\alpha_{\beta\mu} V^\mu = V^\beta \partial_\beta W^\alpha - W^\beta \partial_\beta V^\alpha \ . \tag{1.4.68}$$

Notice, the terms on the right hand side cancel with the ordinary partial derivatives on the left hand side, leaving us with

$$V^\beta W^\mu \Gamma^\alpha_{\beta\mu} - V^\mu W^\beta \Gamma^\alpha_{\beta\mu} = 0 \ . \tag{1.4.69}$$

Relabelling $\mu \to \beta, \beta \to \mu$ in the second term and moving this term over to the right hand side then gives

$$V^\beta W^\mu \Gamma^\alpha_{\beta\mu} = V^\beta W^\mu \Gamma^\alpha_{\mu\beta} \ . \tag{1.4.70}$$

Hence, by cancelling the common factors of $V^\beta W^\mu$, we obtain a condition on the connection coefficients

$$\Gamma^\alpha_{\beta\mu} = \Gamma^\alpha_{\mu\beta} \ . \tag{1.4.71}$$

Hence, we see that for a torsion free connection, the connection coefficients are symmetric in the lower indices! Furthermore, we can write the torsion tensor acting on the vector fields $\vec{V}$ and $\vec{W}$ as

$$[T(\vec{V}, \vec{W})]^\alpha = T^\alpha_{\beta\mu} V^\beta W^\mu \ , \tag{1.4.72}$$

where

$$T^\alpha_{\beta\mu} = \Gamma^\alpha_{\beta\mu} - \Gamma^\alpha_{\mu\beta} \ , \tag{1.4.73}$$

and we can now see that the torsion tensor is entirely dependent on the connection.

1.4.5 The Levi-Civita Connection

We can now use the idea of torsion to define a unique connection. As the name suggests, a connection provides us with a way to compare vectors that live in different tangent spaces, by connecting different elements of the tangent bundle. Indeed, there can exist a huge number of connections on a given manifold, however, the **fundamental theorem of Riemannian geometry** states that on any Riemannian (or pseudo-Riemannian) manifold, there exists a unique connection that is **torsion free** and also **metric compatible** [15], i.e the metric tensor is covariantly constant in the sense that $\nabla_\beta g_{\mu\nu} = 0$. The compatibility condition thus leads to the following equation:

$$\nabla_\beta g_{\mu\nu} = \partial_\beta g_{\mu\nu} - \Gamma^\alpha_{\beta\mu} g_{\alpha\nu} - \Gamma^\alpha_{\beta\nu} g_{\mu\alpha} = 0 \ . \tag{1.4.74}$$

Notice, we can also write two more equations by simply permuting the indices, namely

$$\nabla_\mu g_{\nu\beta} = \partial_\mu g_{\nu\beta} - \Gamma^\alpha_{\mu\nu} g_{\alpha\beta} - \Gamma^\alpha_{\mu\beta} g_{\nu\alpha} = 0 \,, \tag{1.4.75}$$

$$\nabla_\nu g_{\beta\mu} = \partial_\nu g_{\beta\mu} - \Gamma^\alpha_{\nu\beta} g_{\alpha\mu} - \Gamma^\alpha_{\nu\mu} g_{\beta\alpha} = 0 \,. \tag{1.4.76}$$

Hence, by subtracting (1.4.75) and (1.4.76) from (1.4.74) we obtain

$$\begin{aligned} \partial_\beta g_{\mu\nu} - \Gamma^\alpha_{\beta\mu} g_{\alpha\nu} - \Gamma^\alpha_{\beta\nu} g_{\mu\alpha} - \partial_\mu g_{\nu\beta} + \Gamma^\alpha_{\mu\nu} g_{\alpha\beta} + \Gamma^\alpha_{\mu\beta} g_{\nu\alpha} \\ - \partial_\nu g_{\beta\mu} + \Gamma^\alpha_{\nu\beta} g_{\alpha\mu} + \Gamma^\alpha_{\nu\mu} g_{\beta\alpha} = 0 \,, \end{aligned} \tag{1.4.77}$$

and then using the symmetry condition stated in equation (1.4.71), we can simplify this to

$$\partial_\beta g_{\mu\nu} - \partial_\mu g_{\nu\beta} - \partial_\nu g_{\beta\mu} + 2\Gamma^\alpha_{\mu\nu} g_{\alpha\beta} = 0 \,. \tag{1.4.78}$$

Rearranging this equation for $\Gamma^\alpha_{\mu\nu} g_{\alpha\beta}$ then gives

$$\Gamma^\alpha_{\mu\nu} g_{\alpha\beta} = \frac{1}{2}(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\beta\mu} - \partial_\beta g_{\mu\nu}) \,, \tag{1.4.79}$$

and finally, by multiplying both sides by $g^{\alpha\beta}$, we can isolate the connection coefficient $\Gamma^\alpha_{\mu\nu}$, thus obtaining

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2} g^{\alpha\beta}(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\beta\mu} - \partial_\beta g_{\mu\nu}) \,. \tag{1.4.80}$$

This connection, called the **Levi-Civita connection** or **metric connection** is fundamental to the theory of classical general relativity, and also to the study of Riemannian geometry. The corresponding connection coefficients $\Gamma^\alpha_{\mu\nu}$ are called the **Christoffel symbols**, which in the following section will be used to describe geodesics on manifolds.

Lemma 2. *The Christoffel symbols vanish in $\mathbb{R}^n$, and thus the covariant derivative reduces to ordinary partial derivatives*

Proof. Since the metric tensor in $\mathbb{R}^n$ is given by the Kronecker delta $\delta_{\alpha\beta}$, we would have that all the partial derivatives, $\partial_\eta \delta_{\alpha\beta}$, and hence the Christoffel symbols vanish, and so the covariant derivative reduces to

$$\nabla_\beta V^\alpha = \frac{\partial V^\alpha}{\partial x^\beta} \,, \tag{1.4.81}$$

which are simply the partial derivatives of the vector components V^α. □

Example 3. *(Christoffel symbols in spherical coordinates) Recall, in example 2 we found the metric tensor in spherical polar coordinates to be*

$$g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho^2 \sin^2(\theta) \end{bmatrix} \,. \tag{1.4.82}$$

Since g is diagonal, its inverse is then easily found to be

$$g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho^2} & 0 \\ 0 & 0 & \frac{1}{\rho^2 \sin^2(\theta)} \end{bmatrix} \,, \tag{1.4.83}$$

or, in terms of components

$$g_{\rho\rho} = 1, \quad g_{\theta\theta} = \rho^2, \quad g_{\phi\phi} = \rho^2 \sin^2(\theta) \tag{1.4.84}$$

$$g^{\rho\rho} = 1, \quad g^{\theta\theta} = \frac{1}{\rho^2}, \quad g^{\phi\phi} = \frac{1}{\rho^2 \sin^2(\theta)} \,. \tag{1.4.85}$$

By equation (1.4.80) we can write

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\beta}(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\beta\mu} - \partial_\beta g_{\mu\nu}) \,, \tag{1.4.86}$$

and since $g^{\rho\beta} = 0$ except for when $\beta = \rho$, we thus have

$$\Gamma^{\rho}_{\mu\nu} = \frac{1}{2} g^{\rho\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) = \frac{1}{2}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) \ , \tag{1.4.87}$$

where $\mu, \nu \in \{\rho, \theta, \phi\}$. Using the same reasoning, we can also define

$$\Gamma^{\theta}_{\mu\nu} = \frac{1}{2} g^{\theta\theta}(\partial_\mu g_{\nu\theta} + \partial_\nu g_{\theta\mu} - \partial_\theta g_{\mu\nu}) = \frac{1}{2\rho^2}(\partial_\mu g_{\nu\theta} + \partial_\nu g_{\theta\mu} - \partial_\theta g_{\mu\nu}) \ , \tag{1.4.88}$$

$$\Gamma^{\phi}_{\mu\nu} = \frac{1}{2} g^{\phi\phi}(\partial_\mu g_{\nu\phi} + \partial_\nu g_{\phi\mu} - \partial_\phi g_{\mu\nu}) = \frac{1}{2\rho^2 \sin^2(\theta)}(\partial_\mu g_{\nu\phi} + \partial_\nu g_{\phi\mu} - \partial_\phi g_{\mu\nu}) \ . \tag{1.4.89}$$

In order to cut down the amount of algebra, we notice that the only non-vanishing partial derivatives of the metric tensor are

$$\partial_\rho g_{\theta\theta} = 2\rho, \quad \partial_\rho g_{\phi\phi} = 2\rho \sin(\theta), \quad \partial_\theta g_{\phi\phi} = 2\rho^2 \sin(\theta)\cos(\theta) \tag{1.4.90}$$

Hence, the only non-vanishing Christoffel symbols are given by

$$\Gamma^{\rho}_{\theta\theta} = \frac{1}{2}(\partial_\theta g_{\theta\rho} + \partial_\theta g_{\rho\theta} - \partial_\rho g_{\theta\theta}) = -\rho \ , \tag{1.4.91}$$

$$\Gamma^{\rho}_{\phi\phi} = \frac{1}{2}(\partial_\phi g_{\phi\rho} + \partial_\phi g_{\rho\phi} - \partial_\rho g_{\phi\phi}) = -\rho \sin^2(\theta) \ , \tag{1.4.92}$$

$$\Gamma^{\theta}_{\rho\theta} = \Gamma^{\theta}_{\theta\rho} = \frac{1}{2\rho^2}(\partial_\rho g_{\theta\theta} + \partial_\theta g_{\theta\rho} - \partial_\theta g_{\rho\theta}) = \frac{1}{\rho} \ , \tag{1.4.93}$$

$$\Gamma^{\theta}_{\phi\phi} = \frac{1}{2\rho^2}(\partial_\phi g_{\theta\phi} + \partial_\phi g_{\theta\phi} - \partial_\theta g_{\phi\phi}) = -\sin(\theta)\cos(\theta) \ , \tag{1.4.94}$$

$$\Gamma^{\phi}_{\rho\phi} = \Gamma^{\phi}_{\phi\rho} = \frac{1}{2\rho^2 \sin^2(\theta)}(\partial_\rho g_{\phi\phi} + \partial_\phi g_{\phi\rho} - \partial_\phi g_{\rho\phi}) = \frac{1}{\rho} \ , \tag{1.4.95}$$

$$\Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} \frac{1}{2\rho^2 \sin^2(\theta)}(\partial_\theta g_{\phi\phi} + \partial_\phi g_{\phi\theta} - \partial_\phi g_{\theta\phi}) = \cot(\theta) \ . \tag{1.4.96}$$

We can then summarise all 27 Christoffel symbols for a sphere (including those that are zero) by writing them as the following matrices

$$\Gamma^{\rho} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\rho & 0 \\ 0 & 0 & -\rho\sin^2(\theta) \end{bmatrix} \ , \tag{1.4.97}$$

$$\Gamma^{\theta} = \begin{bmatrix} 0 & \frac{1}{\rho} & 0 \\ \frac{1}{\rho} & 0 & 0 \\ 0 & 0 & -\sin(\theta)\cos(\theta) \end{bmatrix} \ , \tag{1.4.98}$$

$$\Gamma^{\phi} = \begin{bmatrix} 0 & 0 & \frac{1}{\rho} \\ 0 & 0 & \cot(\theta) \\ \frac{1}{\rho} & \cot(\theta) & 0 \end{bmatrix} \ . \tag{1.4.99}$$

1.5 Geodesics

In ordinary Cartesian space, the shortest distance between any two points is undoubtedly a straight line. On a curved manifold, the shortest path is not necessarily as simple, unless however we were to zoom in enough for the manifold to appear flat in which case the shortest distance would again be a straight line. With this we arrive at the notion of a geodesic: a curve which gives the shortest distance between two points on a manifold.

Consider a manifold endowed with a positive-definite metric $g_{\mu\nu}$ and suppose that on some coordinate patch of the manifold, the line element is given by $d\tau^2 = g_{\mu\nu}dx^\mu dx^\nu$, where due to the metric being positive-definite we have $d\tau^2 > 0$. we can now construct a functional for the arc-length of a curve $x^\mu(t)$ between $x^\mu(\alpha)$ and $x^\mu(\beta)$ given by

$$S[x^\mu(t)] = \int d\tau = \int_\alpha^\beta \frac{d\tau}{dt} dt = \int_\alpha^\beta \sqrt{g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}} \ dt \ . \tag{1.5.1}$$

To calculate the geodesics on the patch, we need to determine the function $x^\mu(t)$, such that this functional is minimised! Suppose we vary S with respect to the curve x^μ, i.e $x^\mu \to x^\mu + \delta x^\mu$, such that

$$S[x^\mu(t) + \delta x^\mu(t)] = S[x^\mu(t)] + \delta S[x^\mu(t)] \ , \tag{1.5.2}$$

and $\delta x^\mu(\alpha) = \delta x^\mu(\beta) = 0$, where

$$\delta S = \int_\alpha^\beta \delta \sqrt{g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}\, dt = \frac{1}{2}\int_\alpha^\beta \frac{\delta\left(g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\right)}{\sqrt{g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}} dt \ . \tag{1.5.3}$$

From equation (1.5.1), we see that

$$\frac{d\tau}{dt} = \sqrt{g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}} \qquad \Longrightarrow \qquad \frac{dt}{d\tau} = \left(\sqrt{g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}\right)^{-1} \tag{1.5.4}$$

and thus, our functional variation can be written as

$$\delta S = \frac{1}{2}\int_\alpha^\beta \delta\left(g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\right)\frac{dt}{d\tau}dt \tag{1.5.5}$$

Carrying out the variation of the integrand using the product rule and re-writing dt, using the chain rule, then gives

$$\delta S = \frac{1}{2}\int_\alpha^\beta \left[\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\delta g_{\mu\nu} + \frac{d\delta x^\mu}{dt}\frac{dx^\nu}{dt}g_{\mu\nu} + \frac{dx^\mu}{dt}\frac{d\delta x^\nu}{dt}g_{\mu\nu}\right]\left(\frac{dt}{d\tau}\right)^2 d\tau \ , \tag{1.5.6}$$

and hence, we can now rewrite the entire integral in terms of the parameter τ, as

$$\delta S = \frac{1}{2}\int_\alpha^\beta \left[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\delta g_{\mu\nu} + \frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}g_{\mu\nu} + \frac{dx^\mu}{d\tau}\frac{d\delta x^\nu}{d\tau}g_{\mu\nu}\right] d\tau \ . \tag{1.5.7}$$

Then, by employing the principle of least action ($\delta S = 0$), we obtain

$$\frac{1}{2}\int_\alpha^\beta \left[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\delta g_{\mu\nu} + \frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}g_{\mu\nu} + \frac{dx^\mu}{d\tau}\frac{d\delta x^\nu}{d\tau}g_{\mu\nu}\right] d\tau = 0 \ . \tag{1.5.8}$$

Note that, since $g_{\mu\nu}$ is symmetric in its indices, the second and third terms can be combined into single term, and we now obtain

$$\frac{1}{2}\int_\alpha^\beta \left[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\delta g_{\mu\nu} + 2\frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}g_{\mu\nu}\right] d\tau = 0 \ . \tag{1.5.9}$$

By the chain rule, we can now write $\delta g_{\mu\nu} = \partial_\sigma g_{\mu\nu}\delta x^\sigma$, and hence our integral becomes

$$\frac{1}{2}\int_\alpha^\beta \left[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\partial_\sigma g_{\mu\nu}\delta x^\sigma + 2\frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}g_{\mu\nu}\right] d\tau = 0 \ . \tag{1.5.10}$$

Performing integration by parts on the second term yields

$$\int_\alpha^\beta \left[\frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}g_{\mu\nu}\right] d\tau = \delta x^\mu \frac{dx^\nu}{d\tau}g_{\mu\nu}\Big|_\alpha^\beta - \int_\alpha^\beta \delta x^\mu \frac{d}{d\tau}\left[\frac{dx^\nu}{d\tau}g_{\mu\nu}\right] d\tau \ . \tag{1.5.11}$$

Then, since $\delta x^\mu(\alpha) = \delta x^\mu(\beta) = 0$, we have that the first term here must vanish. Hence, by carrying out the differentiation inside the remaining integral, we obtain

$$\begin{aligned}\int_\alpha^\beta \left[\frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}g_{\mu\nu}\right] d\tau &= -\int_\alpha^\beta \left[\frac{dx^\nu}{d\tau}\frac{dg_{\mu\nu}}{d\tau} + \frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}\right]\delta x^\mu d\tau \\ &= -\int_\alpha^\beta \left[\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau}\partial_\sigma g_{\mu\nu} + \frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}\right]\delta x^\mu d\tau \ ,\end{aligned} \tag{1.5.12}$$

which can now be substituted back into equation (1.5.10), to give

$$\begin{aligned}&\frac{1}{2}\int_\alpha^\beta\Big[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\partial_\sigma g_{\mu\nu}\delta x^\sigma-2\Big(\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau}\partial_\sigma g_{\mu\nu}+\frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}\Big)\delta x^\mu\Big]d\tau\\&=\frac{1}{2}\int_\alpha^\beta\Big[-2\frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}\delta x^\mu+\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\partial_\sigma g_{\mu\nu}\delta x^\sigma-2\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau}\partial_\sigma g_{\mu\nu}\delta x^\mu\Big]d\tau=0\end{aligned}\tag{1.5.13}$$

Splitting up the last term into a sum of its symmetrical parts, and relabeling the dummy indices $\mu\to\sigma,\sigma\to\mu$ in the second term, then gives

$$\begin{aligned}&\frac{1}{2}\int_\alpha^\beta\Big[-2\frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}\delta x^\mu+\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}\partial_\mu g_{\sigma\nu}\delta x^\mu-\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau}\partial_\sigma g_{\mu\nu}\delta x^\mu-\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}\partial_\nu g_{\mu\sigma}\delta x^\mu\Big]d\tau\\&=\int_\alpha^\beta\Big[-\frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}+\frac{1}{2}\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}\Big(\partial_\mu g_{\sigma\nu}-\partial_\sigma g_{\mu\nu}-\partial_\nu g_{\mu\sigma}\Big)\Big]\delta x^\mu d\tau=0\;.\end{aligned}\tag{1.5.14}$$

Since we require this integral to vanish, then clearly we must have that the integrand itself is zero, and hence

$$\begin{aligned}&-\frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}+\frac{1}{2}\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}\Big(\partial_\mu g_{\sigma\nu}-\partial_\sigma g_{\mu\nu}-\partial_\nu g_{\mu\sigma}\Big)=0\\\Longrightarrow\quad&\frac{d^2x^\nu}{d\tau^2}g_{\mu\nu}+\frac{1}{2}\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}\Big(\partial_\sigma g_{\mu\nu}+\partial_\nu g_{\mu\sigma}-\partial_\mu g_{\sigma\nu}\Big)=0\;.\end{aligned}\tag{1.5.15}$$

Finally, by multiplying by the inverse metric tensor $g^{\rho\mu}$, we obtain the **geodesic equation**

$$\frac{d^2x^\rho}{d\tau^2}+\Gamma^\rho_{\sigma\nu}\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}=0\;,\tag{1.5.16}$$

where we recognise

$$\Gamma^\rho_{\sigma\nu}=\frac{1}{2}g^{\rho\mu}\Big(\partial_\sigma g_{\mu\nu}+\partial_\nu g_{\mu\sigma}-\partial_\mu g_{\sigma\nu}\Big)\;,\tag{1.5.17}$$

is the Levi-Civita connection, derived in the previous section.

Corollary 3. *A curve which has the property of parallel transporting its own tangent vector, is a geodesic.*

Proof. Recall, the equation for parallel transporting a vector along a curve $x^\mu(t)$ is given by

$$\frac{dV^\sigma}{dt}+\frac{dx^\mu}{dt}\Gamma^\sigma_{\mu\lambda}V^\lambda=0\;.\tag{1.5.18}$$

If we consider the vector V^σ, to be tangent to the curve $x^\mu(t)$ for all t, i.e

$$V^\sigma=\frac{dx^\sigma}{dt}\;,\tag{1.5.19}$$

then, upon substituting this in to the parallel transport equation, we obtain

$$\frac{d}{dt}\Big(\frac{dx^\sigma}{dt}\Big)+\frac{dx^\mu}{dt}\Gamma^\sigma_{\mu\lambda}\frac{dx^\lambda}{dt}=0\;.\tag{1.5.20}$$

Hence, carrying out the differentiation then gives, once again, the geodesic equation

$$\frac{d^2x^\sigma}{dt^2}+\Gamma^\sigma_{\mu\lambda}\frac{dx^\mu}{dt}\frac{dx^\lambda}{dt}=0\;.\tag{1.5.21}$$

□

Example 4. *(Geodesics on a sphere) In this example it is assumed, without loss of generality, that the sphere on which we wish to find geodesics, is a unit sphere, i.e $\rho = 1$. In this case, the non-zero components of the metric tensor are given by*

$$g_{\theta\theta} = 1, \qquad g_{\phi\phi} = \sin^2(\theta) \ , \tag{1.5.22}$$

and hence, the only non-zero derivative is

$$\partial_\theta g_{\phi\phi} = 2\sin(\theta)\cos(\theta) \ . \tag{1.5.23}$$

Equations (1.4.91)-(1.4.96) then give the corresponding Christoffel symbols to be

$$\Gamma^\theta_{\phi\phi} = -\sin(\theta)\cos(\theta), \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot(\theta) \ . \tag{1.5.24}$$

Using the geodesic equation (1.5.16) and by assuming $x^\rho = \rho$, $x^\theta = \theta$ and $x^\phi = \phi$, we now obtain

$$\frac{d^2\theta}{dt^2} + \Gamma^\theta_{\phi\phi}\Big(\frac{d\phi}{dt}\Big)^2 = 0 \ , \tag{1.5.25}$$

$$\frac{d^2\phi}{dt^2} + (\Gamma^\phi_{\theta\phi} + \Gamma^\phi_{\phi\theta})\frac{d\theta}{dt}\frac{d\phi}{dt} = 0 \ . \tag{1.5.26}$$

Hence, plugging in the Christoffel symbols from (1.5.24) gives a system of coupled, nonlinear, ordinary differential equations

$$\frac{d^2\theta}{dt^2} - \sin(\theta)\cos(\theta)\Big(\frac{d\phi}{dt}\Big)^2 = 0 \ , \tag{1.5.27}$$

$$\frac{d^2\phi}{dt^2} + 2\cot(\theta)\frac{d\theta}{dt}\frac{d\phi}{dt} = 0 \ . \tag{1.5.28}$$

These equations are in general, difficult to solve. However, by assuming $\phi(t) = \phi_0 = constant$, for all t, so that $\dot\phi = \ddot\phi = 0$, and by also imposing the initial conditions

$$\theta(0) = \theta_0, \quad \dot\theta(0) = \dot\theta_0 = 1 \ , \tag{1.5.29}$$

it is clear that equation (1.5.28) vanishes along with the second term in (1.5.27), thus leaving us with the simple equation in θ

$$\frac{d^2\theta}{dt^2} = 0 \tag{1.5.30}$$

. The general solution subject to the conditions in (1.5.29) , is then easily found to be

$$\theta(t) = \dot\theta_0 t + \theta_0 \implies \theta(t) = t + \theta_0 \tag{1.5.31}$$

Hence, together with our equation for $\phi(t)$ we have the solution

$$\theta(t) = t + \theta_2 \tag{1.5.32}$$

$$\phi(t) = \phi_0 \tag{1.5.33}$$

These equations describe great circles, i.e circles whose radius is equal to that of the sphere, which intersect at the poles (circles of longitude) and are such that for any two points which lie on these paths, the great circle minimises the distance between them and is thus a geodesic. Hence, by rotational symmetry we can deduce that all geodesics on a sphere are in fact great circles.

Remark: In corollary 1, we mentioned the idea of locally inertial coordinates. Further to this discussion, we can now see that due to the constancy of the metric components, $\eta_{\alpha\beta}$, in the local coordinate system, all partial derivatives $\partial_\sigma\eta_{\alpha\beta}$ will vanish, leading to the Christoffel symbols being null, and hence, the geodesic equation (1.5.16) reduces to

$$\frac{d^2x^\rho}{dt^2} = 0 \ . \tag{1.5.34}$$

Thus, geodesics in this (local) coordinate system will become straight lines!

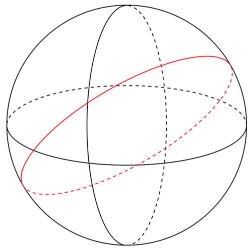

Figure 5: Geodesics on a sphere are great circles

1.6 Curvature

Now that we have developed important tools such as covariant derivatives, connections, and parallel transport, we are finally in a position to tackle the subject of curvature! In Lemma 2, we showed that for any vector in $\mathbb{R}^n$, parallel transport leaves the vector unchanged, and thus parallel transport in $\mathbb{R}^n$ is path independent. On a curved manifold however, the result of parallel transporting a vector from one point to another will depend on the path taken between the points [3].

One can intuitively notice the effects of curvature on parallel transport of vectors by considering the transport of a vector around a closed path on a sphere. Suppose we choose our initial vector to lie at the north pole, and then parallel transport it down to the equator along a great circle of constant longitude, we then transport it along the equator by an angle θ, before finally transporting it back to the initial position along another great circle(see figure 6).

It can now be seen that, although the vector has been returned to its initial position, it is no longer pointing in the same direction. In fact, it has been rotated by an angle θ. This leads us to the idea that the parallel transport of a vector around a closed loop on a curved manifold will result in a transformation (namely a rotation) of the vector [3], and as a consequence of this, by measuring how much a vector changes as it undergoes parallel transport around a closed loop, we can in fact calculate the curvature of a manifold!

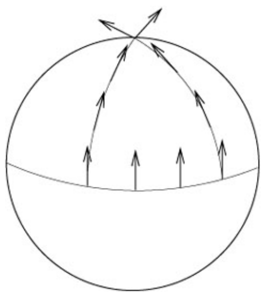

Figure 6: Parallel transport of vectors on a sphere [4]

1.6.1 The Riemann Curvature tensor

Consider the parallel transport of a vector, V^γ, along a curve $x^\eta(t)$, defined on some arbitrary manifold endowed with a Levi-Civita connection. The corresponding equation is thus

$$\frac{dV^\beta}{dt} + \Gamma^\beta_{\mu\nu} V^\mu \frac{dx^\nu}{dt} = 0 \,. \tag{1.6.1}$$

Furthermore, suppose that the initial vector starts at a point P, on the curve given by $x^\eta(t_0) = x^\eta_P$, so that the vector has initial components $V^\gamma(t_0) = V^\gamma_P$. Thus, by directly integrating (1.6.1) we find that the components at time t, of the vector undergoing parallel transport become

$$V^\beta(t) = V^\beta_P - \int_{t_0}^{t} \Gamma^\beta_{\mu\nu} V^\mu \frac{dx^\nu}{dt'} dt' \,. \tag{1.6.2}$$

It should be noted here that $\Gamma^{\beta}_{\mu\nu}$ and V^{μ}, are both dependent on t'. If we now assume the curve between t_0 and t to be infinitesimal in nature, we can then Taylor expand $\Gamma^{\beta}_{\mu\nu}(t')$ and $V^{\mu}(t')$, up to first order, about the point P, to obtain

$$V^{\mu}(t') \approx V^{\mu}_P + \frac{\partial V^{\mu}}{\partial x^{\alpha}}\Big|_P (x^{\alpha}(t') - x^{\alpha}_P) \ , \tag{1.6.3}$$

$$\Gamma^{\beta}_{\mu\nu}(t') \approx (\Gamma^{\beta}_{\mu\nu})_P + \Big(\frac{\partial}{\partial x^{\alpha}}\Gamma^{\beta}_{\mu\nu}\Big|_P\Big)(x^{\alpha}(t') - x^{\alpha}_P) \ . \tag{1.6.4}$$

Notice, by using the chain rule we can write (1.6.1) in the form of (1.4.50) i.e

$$\frac{dx^{\alpha}}{dt}\Big(\frac{\partial V^{\beta}}{\partial x^{\alpha}} + \Gamma^{\beta}_{\mu\alpha}V^{\mu}\Big) = 0 \ . \tag{1.6.5}$$

If we now assume that as t increases, the value of $x^{\nu}(t)$ continuously changes, then clearly the derivative outside of the brackets will be non-zero and the term in the bracket must therefore vanish. Hence, at t_0 (i.e the point P) we can write

$$\frac{\partial V^{\mu}}{\partial x^{\alpha}}\Big|_P = -(\Gamma^{\mu}_{\beta\alpha})_P V^{\beta}_P \ , \tag{1.6.6}$$

where we have relabeled $\beta \to \mu$, $\mu \to \beta$. We can now plug this equation in to (1.6.3) to obtain

$$V^{\mu}(t') = V^{\mu}_P - (\Gamma^{\mu}_{\beta\alpha})_P V^{\beta}_P (x^{\alpha}(t') - x^{\alpha}_P) \ . \tag{1.6.7}$$

Substituting (1.6.7) along with (1.6.4) in to the integral in (1.6.2), then yields (note that in the following equations it is assumed that $x^{\alpha} = x^{\alpha}(t')$)

$$V^{\beta}(t) \approx V^{\beta}_P - \int_{t_0}^{t} \Big((\Gamma^{\beta}_{\mu\nu})_P + \Big(\frac{\partial \Gamma^{\beta}_{\mu\nu}}{\partial x^{\alpha}}\Big|_P\Big)(x^{\alpha} - x^{\alpha}_P)\Big)\Big(V^{\mu}_P - (\Gamma^{\mu}_{\beta\alpha})_P V^{\beta}_P (x^{\alpha} - x^{\alpha}_P)\Big)\frac{dx^{\nu}}{dt'}dt' \tag{1.6.8}$$

Thus, expanding out the brackets and discarding quadratic terms we then gives

$$V^{\beta}(t) \approx V^{\beta}_P - \int_{t_0}^{t} \Big[V^{\mu}_P(\Gamma^{\beta}_{\mu\nu})_P + \Big(V^{\mu}_P\Big(\frac{\partial \Gamma^{\beta}_{\mu\nu}}{\partial x^{\alpha}}\Big|_P\Big) - V^{\beta}_P(\Gamma^{\beta}_{\mu\nu})_P(\Gamma^{\mu}_{\beta\alpha})_P\Big)(x^{\alpha} - x^{\alpha}_P)\Big]\frac{dx^{\nu}}{dt'}dt' \ . \tag{1.6.9}$$

Inside the integral, we now relabel $\mu \to \lambda$ in the first and second term, and $\beta \to \lambda$ in the third term, being careful to notice that V^{β}_P is contracted with only **one** of the connection coefficients, giving

$$V^{\beta}(t) \approx V^{\beta}_P - \int_{t_0}^{t} \Big[V^{\lambda}_P(\Gamma^{\beta}_{\lambda\nu})_P + V^{\lambda}_P\Big(\frac{\partial \Gamma^{\beta}_{\lambda\nu}}{\partial x^{\alpha}} - \Gamma^{\beta}_{\mu\nu}\Gamma^{\mu}_{\lambda\alpha}\Big)_P(x^{\alpha} - x^{\alpha}_P)\Big]\frac{dx^{\nu}}{dt'}dt' \ . \tag{1.6.10}$$

Finally, by relabeling $\mu \to \tau$ and rewriting the equation as a sum of two separate integrals, we obtain

$$V^{\beta}(t) \approx V^{\beta}_P - V^{\lambda}_P(\Gamma^{\beta}_{\lambda\nu})_P \int_{t_0}^{t}\frac{dx^{\nu}}{dt'}dt' - \Big(\frac{\partial \Gamma^{\beta}_{\lambda\nu}}{\partial x^{\alpha}} - \Gamma^{\beta}_{\tau\nu}\Gamma^{\tau}_{\lambda\alpha}\Big)_P V^{\lambda}_P \int_{t_0}^{t}(x^{\alpha} - x^{\alpha}_P)\frac{dx^{\nu}}{dt'}dt' \ . \tag{1.6.11}$$

This equation gives an approximation of the components $V^{\beta}(t)$, as the vector undergoes parallel transport along a curve. Suppose now that the curve on which the vector is being transported, is an infinitesimal closed path and denote the vector components after they have been transported and returned to their original position as $V^{\beta}_{\parallel P}$.

The change in the vector components after parallel transport can now be approximates as

$$V^{\beta}_{\parallel P} - V^{\beta}_P \approx -V^{\lambda}_P(\Gamma^{\beta}_{\lambda\nu})_P \oint \frac{dx^{\nu}}{dt'}dt' - \Big(\frac{\partial \Gamma^{\beta}_{\lambda\nu}}{\partial x^{\alpha}} - \Gamma^{\beta}_{\tau\nu}\Gamma^{\tau}_{\lambda\alpha}\Big)_P V^{\lambda}_P \oint (x^{\alpha} - x^{\alpha}_P)\frac{dx^{\nu}}{dt'}dt' \ . \tag{1.6.12}$$

Recall, for an integral around a closed path we have

$$\oint \frac{dx^{\nu}}{dt'}dt' = \oint dx^{\nu} = 0 \ . \tag{1.6.13}$$

Hence, since x^α_P is constant, our approximation becomes

$$V^\beta_{\parallel P} - V^\beta_P \approx -\Big(\frac{\partial \Gamma^\beta_{\lambda\nu}}{\partial x^\alpha} - \Gamma^\beta_{\tau\nu}\Gamma^\tau_{\lambda\alpha}\Big)_P V^\lambda_P \oint x^\alpha \frac{dx^\nu}{dt'} dt' \ , \tag{1.6.14}$$

By relabeling the dummy indices $\nu \to \alpha$ and $\alpha \to \nu$ we obtain another equation

$$V^\beta_{\parallel P} - V^\beta_P \approx -\Big(\frac{\partial \Gamma^\beta_{\lambda\alpha}}{\partial x^\nu} - \Gamma^\beta_{\tau\alpha}\Gamma^\tau_{\lambda\nu}\Big)_P V^\lambda_P \oint x^\nu \frac{dx^\alpha}{dt'} dt' \ . \tag{1.6.15}$$

We now have that the integrals in (1.6.14) and (1.6.15) differ by only the switched indices, but we will find another way of relating the two equations. Observe that by the product rule we have

$$0 = \oint d(x^\nu x^\alpha) = \oint (x^\nu dx^\alpha + x^\alpha dx^\nu) = \oint x^\nu \frac{dx^\alpha}{dt'} dt' + \oint x^\alpha \frac{dx^\nu}{dt'} dt' \ , \tag{1.6.16}$$

and hence,

$$\oint x^\alpha \frac{dx^\nu}{dt'} dt' = -\oint x^\nu \frac{dx^\alpha}{dt'} dt' \ . \tag{1.6.17}$$

Plugging this equation into (1.6.15) now gives

$$V^\beta_{\parallel P} - V^\beta_P \approx \Big(\frac{\partial \Gamma^\beta_{\lambda\alpha}}{\partial x^\nu} - \Gamma^\beta_{\tau\alpha}\Gamma^\tau_{\lambda\nu}\Big)_P V^\lambda_P \oint x^\alpha \frac{dx^\nu}{dt'} dt' \ . \tag{1.6.18}$$

Notice that the integrals in (1.6.14) and (1.6.18) are now identical and we therefore can add both equations together and divide by 2 to obtain

$$V^\beta_{\parallel P} - V^\beta_P \approx \frac{1}{2}\Big(\frac{\partial \Gamma^\beta_{\lambda\alpha}}{\partial x^\nu} - \frac{\partial \Gamma^\beta_{\lambda\nu}}{\partial x^\alpha} + \Gamma^\beta_{\tau\nu}\Gamma^\tau_{\lambda\alpha} - \Gamma^\beta_{\tau\alpha}\Gamma^\tau_{\lambda\nu}\Big)_P V^\lambda_P \oint x^\alpha \frac{dx^\nu}{dt'} dt' \ , \tag{1.6.19}$$

which we write as

$$V^\beta_{\parallel P} - V^\beta_P \approx \frac{1}{2}(R^\beta_{\lambda\nu\alpha})_P V^\lambda_P \oint x^\alpha \frac{dx^\nu}{dt'} dt' \ . \tag{1.6.20}$$

where the quantity

$$R^\beta_{\lambda\nu\alpha} = \frac{\partial \Gamma^\beta_{\lambda\alpha}}{\partial x^\nu} - \frac{\partial \Gamma^\beta_{\lambda\nu}}{\partial x^\alpha} + \Gamma^\beta_{\tau\nu}\Gamma^\tau_{\lambda\alpha} - \Gamma^\beta_{\tau\alpha}\Gamma^\tau_{\lambda\nu} \ , \tag{1.6.21}$$

defines the **Riemann curvature tensor**, expressed in terms of the Christoffel symbols.
Clearly, due to the Christoffel symbols being dependent on the metric tensor, we can see that the curvature tensor is also entirely dependent on the metric as well. It straightforward to check using the transformation laws that $R^\beta_{\lambda\nu\alpha}$ is a true tensor, although we admit the proof here, see [1] for details. This tensor tells us how much a vector changes as it is parallel transported around a closed path, which recall, is a direct consequence of curvature. An immediate and also intuitive result is that since vectors in $\mathbb{R}^n$ do not change under parallel transport, we have

$$R^\beta_{\lambda\nu\alpha} = 0 \ , \tag{1.6.22}$$

and hence, we can finally justify our saying that $\mathbb{R}^n$ is a **flat space**! Moreover, in any coordinate system for which the metric tensor components are constant, the curvature tensor will vanish.

Remark: This statement directly leads to the important result that Minkowski space-time with metric $\eta_{\alpha\beta}$, is indeed flat!

Another immediate property which can bee seen from (1.6.21) is that $R^\beta_{\lambda\nu\alpha}$ is antisymmetric in the last two lower indices, that is

$$R^\beta_{\lambda\nu\alpha} = -R^\beta_{\lambda\alpha\nu} \ . \tag{1.6.23}$$

To investigate further symmetries of the curvature tensor, it is desirable to first write it in a fully covariant form, which we can do by means of the metric tensor;

$$R_{\sigma\lambda\nu\alpha} = g_{\sigma\beta} R^{\beta}_{\lambda\nu\alpha} \ , \tag{1.6.24}$$

which, by (1.6.21), can be written as

$$R_{\sigma\lambda\nu\alpha} = g_{\sigma\beta} \frac{\partial \Gamma^{\beta}_{\lambda\alpha}}{\partial x^{\nu}} - g_{\sigma\beta} \frac{\partial \Gamma^{\beta}_{\lambda\nu}}{\partial x^{\alpha}} + g_{\sigma\beta}(\Gamma^{\beta}_{\tau\nu}\Gamma^{\tau}_{\lambda\alpha} - \Gamma^{\beta}_{\tau\alpha}\Gamma^{\tau}_{\lambda\nu}) \tag{1.6.25}$$

Now, since at any point on a pseudo-Riemannian manifold we can construct a set of locally flat coordinates such that at that point the Christoffel symbols vanish (but not their derivatives)[3], we can simplify this expression to

$$R_{\sigma\lambda\nu\alpha} = g_{\sigma\beta} \frac{\partial \Gamma^{\beta}_{\lambda\alpha}}{\partial x^{\nu}} - g_{\sigma\beta} \frac{\partial \Gamma^{\beta}_{\lambda\nu}}{\partial x^{\alpha}} = \frac{\partial \Gamma_{\sigma\lambda\alpha}}{\partial x^{\nu}} - \frac{\partial \Gamma_{\sigma\lambda\nu}}{\partial x^{\alpha}} \ , \tag{1.6.26}$$

where

$$\Gamma_{\sigma\lambda\alpha} = \frac{1}{2}(\partial_{\lambda} g_{\sigma\alpha} + \partial_{\alpha} g_{\sigma\lambda} - \partial_{\sigma g_{\lambda\alpha}}) \ , \tag{1.6.27}$$

$$\Gamma_{\sigma\lambda\nu} = \frac{1}{2}(\partial_{\lambda} g_{\sigma\nu} + \partial_{\nu} g_{\sigma\lambda} - \partial_{\sigma} g_{\lambda\nu}) \tag{1.6.28}$$

define the Christoffel symbols of the first kind. To this end, the Curvature tensor can now be expressed as

$$R_{\sigma\lambda\nu\alpha} = \frac{1}{2}(\partial_{\nu}\partial_{\lambda} g_{\sigma\alpha} + \cancel{\partial_{\nu}\partial_{\alpha} g_{\sigma\lambda}} - \partial_{\nu}\partial_{\sigma} g_{\lambda\alpha} - \partial_{\alpha}\partial_{\lambda} g_{\sigma\nu} - \cancel{\partial_{\alpha}\partial_{\nu} g_{\sigma\lambda}} + \partial_{\alpha}\partial_{\sigma} g_{\lambda\nu}) \ , \tag{1.6.29}$$

and hence, we obtain the fully covariant form of the curvature tensor

$$R_{\sigma\lambda\nu\alpha} = \frac{1}{2}(\partial_{\nu}\partial_{\lambda} g_{\sigma\alpha} - \partial_{\nu}\partial_{\sigma} g_{\lambda\alpha} - \partial_{\alpha}\partial_{\lambda} g_{\sigma\nu} + \partial_{\alpha}\partial_{\sigma} g_{\lambda\nu}) \ . \tag{1.6.30}$$

In this form, it is now clear that the following additional symmetries hold

$$R_{\sigma\lambda\nu\alpha} = -R_{\sigma\lambda\alpha\nu} = -R_{\lambda\sigma\alpha\nu} \ , \tag{1.6.31}$$

$$R_{\sigma\lambda\nu\alpha} = R_{\nu\alpha\sigma\lambda} \ . \tag{1.6.32}$$

It is also easy to show that by performing the sum of cyclic permutations of the last three indices we obtain

$$R_{\sigma\lambda\nu\alpha} + R_{\sigma\nu\alpha\lambda} + R_{\sigma\alpha\lambda\nu} = 0 \ . \tag{1.6.33}$$

This is sometimes referred to as the **first Bianchi identity** and can be written more succinctly as

$$R_{\sigma[\lambda\nu\alpha]} = 0 \ . \tag{1.6.34}$$

Remark: One should note that although these properties have been derived in locally flat coordinates, due to the nature of tensor equations they must also hold in any other coordinate system!

Remark: In four dimensional space-time, the curvature tensor has $4^4 = 256$ components! However, due to the symmetry properties, this number reduces to just 20 independent components.

From the Riemann tensor in (1.6.21) we can also obtain another important quantity known as the **Ricci curvature tensor**, which is found by contracting the upper index with one of the lower indices, i.e

$$R_{\lambda\alpha} = R^{\beta}_{\lambda\beta\alpha} \ . \tag{1.6.35}$$

This tensor is also symmetric in its two lower indices. We can then perform a further contraction using the metric to produce the **Ricci scalar**

$$R = g^{\lambda\alpha} R_{\lambda\alpha} \ . \tag{1.6.36}$$

This quantity is perhaps the simplest measure of curvature as it assigns a single value to represent the curvature at each point. Both the Ricci tensor and Ricci scalar, as we will see, play an important part in the formulation of the Einstein tensor and therefore, the field equations of general relativity!

1.6.2 The second Bianchi identity and Einstein Tensor

In addition to the algebraic symmetries, the Riemann tensor also obeys a differential identity, known as the **second Bianchi identity**. Consider the covariant derivative of the Riemann tensor in locally flat coordinates (recall that in this case, the covariant derivative reduces to ordinary partial derivatives), that is

$$\nabla_\rho R_{\sigma\lambda\nu\alpha} = \frac{1}{2}(\partial_\rho\partial_\nu\partial_\lambda g_{\sigma\alpha} - \partial_\rho\partial_\nu\partial_\sigma g_{\lambda\alpha} - \partial_\rho\partial_\alpha\partial_\lambda g_{\sigma\nu} + \partial_\rho\partial_\alpha\partial_\sigma g_{\lambda\nu}) \ . \tag{1.6.37}$$

From here, it is easily shown that by permuting the first three indices and taking the sum, we obtain the following Bianchi identity

$$\nabla_\rho R_{\sigma\lambda\nu\alpha} + \nabla_\sigma R_{\lambda\rho\nu\alpha} + \nabla_\lambda R_{\rho\sigma\nu\alpha} = 0 \ . \tag{1.6.38}$$

Suppose now that we twice contract (1.6.38) using the metric tensor, i.e

$$g^{\lambda\alpha} g^{\nu\rho}(\nabla_\rho R_{\sigma\lambda\nu\alpha} + \nabla_\sigma R_{\lambda\rho\nu\alpha} + \nabla_\lambda R_{\rho\sigma\nu\alpha}) = 0 \ . \tag{1.6.39}$$

By using the symmetry property $R_{\lambda\rho\nu\alpha} = -R_{\rho\lambda\nu\alpha}$, on the middle term and then carrying out the first contraction, we obtain

$$g^{\lambda\alpha}(\nabla^\nu R_{\sigma\lambda\nu\alpha} - \nabla_\sigma R_{\lambda\alpha} + \nabla_\lambda R_{\sigma\alpha}) = 0 \ . \tag{1.6.40}$$

The second contraction then gives

$$\nabla^\nu R_{\sigma\nu} - \nabla_\sigma R + \nabla^\alpha R_{\sigma\alpha} = 0 \ , \tag{1.6.41}$$

and we can now relabel $\alpha \to \nu$ and $\sigma \to \mu$ to give

$$2\nabla^\nu R_{\mu\nu} - \nabla_\mu R = 0 \ . \tag{1.6.42}$$

Thus, dividing both sides of this equation by 2 yields

$$\nabla^\nu R_{\mu\nu} - \frac{1}{2}\nabla_\mu R = 0 \ , \tag{1.6.43}$$

and since $\nabla_\mu = g_{\mu\nu}\nabla^\nu$, we can therefore write

$$\nabla^\nu R_{\mu\nu} - \frac{1}{2} g_{\mu\nu}\nabla^\nu R = 0 \ . \tag{1.6.44}$$

Finally, due to the linearity of ∇^ν we can write this expression as

$$\nabla^\nu (R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}) = 0 \ . \tag{1.6.45}$$

With this, we define the **Einstein tensor**

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \ , \tag{1.6.46}$$

which gives us an expression for the curvature of a pseudo-Riemannian manifold and is such that

$$\nabla^\nu G_{\mu\nu} = 0 \ . \tag{1.6.47}$$

In other words, the Einstein tensor is divergence free! Due to the symmetry of the Ricci tensor and metric tensor, it is clear that $G_{\mu\nu}$ is also symmetric in its indices and is also, just like the Riemann tensor, entirely dependent on the metric.

2 General Relativity

2.1 The Equivalence Principle

After the publication of his 1905 paper ," *On The Electrodynamics of Moving Bodies*" [6], which covered the theory of special relativity, Albert Einstein quickly realised that his current theory only applied to objects moving at constant speeds, i.e it only covered inertial motion. To begin generalising his theory, Einstein turned to his old faithful: his famous thought experiments. One day in 1907, from his third floor office in Bern, Einstein looked out of his window and began to imagine some people working on the roof of a building across the street and wondered what one might experience if they were to fall from the roof. This lead to what he later called the happiest thought of his life: he noticed that someone in free fall would no longer be experiencing their own weight! This lead him to the realisation that if you were in a box with no windows, such as an elevator, an observer would not be able to tell the difference between being in free fall and being in deep space away from any gravitational influence of large bodies.

He later developed this idea further. Instead of the elevator being in free fall, he pictured the elevator being uniformly accelerated through deep space. He realised that the observer in the elevator would not be able to tell the difference between being stationary on the surface of the earth, experiencing the force of gravity, and being accelerated through space at a constant rate of $9.81ms^{-2}$. Due to these statements, it becomes clear that an observer in an elevator with no connection to the outside, will be unable to detect whether or not they are being influenced by a gravitational field. As a consequence of this, Einstein realised that in small enough (local) regions of space-time, the laws of physics reduce to those of special relativity. In other words, space-time can be considered flat in such a local region. This is the famed **Einstein equivalence principle**.

2.2 Newtons First Law Revisited

Consider a freely falling particle in an inertial frame. According to Newtons first law: unless the particle is acted on by an external force, its motion will remain uniform, i.e it will move along a straight line. This law however, is only true for particles living in flat space! In this short section, we aim to generalise this law so that it is valid in curved space-time. Suppose the position of a particle is given by the locally flat coordinates x^μ(t), then, since the particle is in free fall (not accelerating) we can write

$$\frac{d^2x^\mu}{dt^2} = 0 \; . \tag{2.2.1}$$

In order to generalise this equation to curved space-time, we need to write it in a covariant form, i.e as a tensor equation. Using the chain rule we can write

$$\frac{d^2x^\mu}{dt^2} = \frac{dx^\nu}{dt}\frac{\partial}{\partial x^\nu}\frac{dx^\mu}{dt} \; . \tag{2.2.2}$$

and thus, to make this a tensor equation, we simply need to switch out the ordinary partial derivative for a covariant one to obtain

$$\frac{dx^\nu}{dt}\nabla_\nu\frac{dx^\mu}{dt} = 0 \; . \tag{2.2.3}$$

Evaluating the covariant derivative then gives

$$\frac{dx^\nu}{dt}\nabla_\nu\frac{dx^\mu}{dt} = \frac{dx^\nu}{dt}\left(\frac{\partial}{\partial x^\nu}\left[\frac{dx^\mu}{dt}\right] + \Gamma^\mu_{\nu\alpha}\frac{dx^\alpha}{dt}\right) = 0 \; , \tag{2.2.4}$$

and hence, we once again arrive at the geodesic equation!

$$\frac{d^2x^\mu}{dt^2} + \Gamma^\mu_{\nu\alpha}\frac{dx^\nu}{dt}\frac{dx^\alpha}{dt} = 0 \; . \tag{2.2.5}$$

With this realisation, we can therefore write an equivalent statement of newtons first law, but for curved space-time: Freely falling particles move along geodesics in space-time.

2.3 The Newtonian Limit

Before moving on to formulating the field equations for general relativity, we first need to investigate some aspects of classical Newtonian gravity, with the aim of showing that space-time curvature is in fact a viable way of describing gravity. Recall, the differential form of Gauss' law of gravity, which can be derived from Newtons law of universal gravitation [20], is given by

$$\vec{\nabla} \cdot \vec{g} = -4\pi G\rho \ , \tag{2.3.1}$$

where $\vec{g}$ is the acceleration due to gravity, G is Newtons gravitational constant, and ρ is the mass density. Also, since the gravitational field is such that $\vec{\nabla} \times \vec{g} = 0$, we can write $\vec{g}$ in terms of a potential, ϕ, as

$$\vec{g} = -\vec{\nabla}\phi \ . \tag{2.3.2}$$

Hence, by substituting this equation in to Gauss' law, we obtain Poisson's equation of gravity

$$\vec{\nabla} \cdot (-\vec{\nabla}\phi) = -4\pi G\rho \implies \nabla^2\phi = 4\pi G\rho \ , \tag{2.3.3}$$

which can alternatively be written using index notation as

$$\delta^{ij}\partial_i\partial_j\phi = 4\pi G\rho \ . \tag{2.3.4}$$

This is an important result, since in order for the field equations of general relativity to be consistent with Newtonian gravity, we must have that in the non-relativistic case, the field equations reduce to Poisson's equation.

We define the **Newtonian limit** by the following requirements [3]:

- The gravitational field is weak.
- Particles are moving slowly with respect to the speed of light (we will thus include the speed of light, c, in our equations).
- The field is static.

The assumption of a weak gravitational field allows us to approximate space-time as being inertial in a local region around a point, and thus we can write the metric tensor as $\eta_{\alpha\beta}$, plus a small perturbation $h_{\alpha\beta}$ which represents the gravitational effects, i.e

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \quad \text{such that} \quad |h_{\alpha\beta}| \ll 1 \ . \tag{2.3.5}$$

Similarly, we can express the inverse metric up to first order in h as

$$g^{\alpha\beta} = \eta^{\alpha\beta} - h^{\alpha\beta} \ , \tag{2.3.6}$$

since,

$$\begin{aligned} g^{\alpha\beta}g_{\beta\gamma} &= (\eta^{\alpha\beta} - h^{\alpha\beta})(\eta_{\beta\gamma} + h_{\beta\gamma}) \\ &= \eta^{\alpha\beta}\eta_{\beta\gamma} + \eta^{\alpha\beta}h_{\beta\gamma} - \eta_{\beta\gamma}h^{\alpha\beta} + \mathcal{O}(h^2) \\ &= \delta^\alpha_\gamma + \eta^{\alpha\beta}h_{\beta\gamma} - \eta_{\beta\gamma}\eta^{\alpha\sigma}\eta^{\beta\rho}h_{\rho\sigma} + \mathcal{O}(h^2) \\ &= \delta^\alpha_\gamma + \eta^{\alpha\beta}h_{\beta\gamma} - \delta^\rho_\gamma\eta^{\alpha\sigma}h_{\sigma\rho} + \mathcal{O}(h^2) \\ &= \delta^\alpha_\gamma + \eta^{\alpha\beta}h_{\beta\gamma} - \eta^{\alpha\sigma}h_{\sigma\gamma} + \mathcal{O}(h^2) \ . \end{aligned} \tag{2.3.7}$$

Relabelling $\sigma \to \beta$ then gives

$$g^{\alpha\beta}g_{\beta\gamma} = \delta^\alpha_\gamma + \eta^{\alpha\beta}h_{\beta\gamma} - \eta^{\alpha\beta}h_{\beta\gamma} + \mathcal{O}(h^2) = \delta^\alpha_\gamma + \mathcal{O}(h^2) \ . \tag{2.3.8}$$

Thus, up to first order in h, the usual identity for the inverse metric holds

$$g^{\alpha\beta} g_{\beta\gamma} = \delta^{\alpha}_{\gamma} \qquad \square \, . \tag{2.3.9}$$

Suppose now that the position of a particle is given by the coordinates $x^{\mu}(\tau)$, where τ is the proper time. The condition that the particle is moving slowly then means that we can essentially neglect the spatial velocities, since

$$\frac{dx^i}{d\tau} \ll \frac{dt}{d\tau} \quad i = 1, 2, 3 \tag{2.3.10}$$

and hence by (2.2.5), the particle will move along the space-time geodesic described by

$$\frac{d^2 x^{\mu}}{d\tau^2} + \Gamma^{\mu}_{00} c^2 \left(\frac{dt}{d\tau}\right)^2 = 0 \, . \tag{2.3.11}$$

Now, since the field is static this means that the field does not change with time, which corresponds to

$$\partial_0 g_{\mu\nu} = 0 \, , \tag{2.3.12}$$

and hence, the Christoffel symbols reduce to

$$\Gamma^{\mu}_{00} = \frac{1}{2} g^{\mu\sigma} (\partial_0 g_{\sigma 0} + \partial_0 g_{0\sigma} - \partial_{\sigma} g_{00}) = -\frac{1}{2} g^{\mu\sigma} \partial_{\sigma} g_{00} \, . \tag{2.3.13}$$

Up to first order in h we thus have

$$\Gamma^{\mu}_{00} = -\frac{1}{2} \eta^{\mu\sigma} \partial_{\sigma} h_{00} \, , \tag{2.3.14}$$

which, upon plugging into (2.3.11) now gives

$$\frac{d^2 x^{\mu}}{d\tau^2} = \frac{1}{2} c^2 \left(\frac{dt}{d\tau}\right)^2 \eta^{\mu\sigma} \partial_{\sigma} h_{00} \, . \tag{2.3.15}$$

By considering only the time component of this equation we see that

$$\frac{d^2 x^0}{d\tau^2} = \frac{1}{2} c^2 \eta^{0\sigma} \partial_{\sigma} h_{00} \, . \tag{2.3.16}$$

Notice, the only non-zero component of $\eta^{0\sigma}$, occurs when $\sigma = 0$, however this value of σ will lead to the partial derivative $\partial_0 h_{00}$, which by our assumption, must vanish and therefore

$$\frac{d^2 x^0}{d\tau^2} = \frac{d^2 t}{d\tau^2} = 0 \qquad \Longrightarrow \qquad t = \tau \, . \tag{2.3.17}$$

Additionally, since the space-like components of $\eta^{\mu\sigma}$ are given by δ^{ij}, equation (2.3.15) now simplifies to

$$\frac{d^2 x^i}{dt^2} = \frac{1}{2} c^2 \partial_i h_{00} \, . \tag{2.3.18}$$

Thus, by noticing the left hand side of this equation gives the components of acceleration, we can write express the left hand side terms of a potential, ϕ, to obtain a Poisson equation

$$[\vec{\nabla}\phi]_i = \partial_i \phi = -\frac{1}{2} c^2 \partial_i h_{00} \, . \tag{2.3.19}$$

Integrating both sides of this equation then yields an expression for the perturbation, h_{00}, in terms of the potential

$$h_{00} = -\frac{2\phi}{c^2} \, . \tag{2.3.20}$$

Hence, by substituting this back in to (2.3.5), we find that the time component of the metric tensor is given by

$$g_{00} = \eta_{00} + h_{00} = -(1 + \frac{2\phi}{c^2}) \, , \tag{2.3.21}$$

and thus, we have shown that in the Newtonian limit, the geodesic equation reduces to the form of (2.3.2), providing the metric takes on the form (2.3.21). Hence, we can conclude that space-time curvature is indeed sufficient to describe gravity in the Newtonian limit!

2.4 Einstein-Hilbert Action

Of course, the theory of general relativity is usually accredited purely to Albert Einstein. However, while Einstein was still working on his theory, the great mathematician David Hilbert, also began working on a theory of gravity, and although Hilbert supposedly was the first to formulate the correct field equations, he agreed that the theory truly belonged to Einstein. This race to relativity has been the source of many disputes over the years as to whether the field equations should really be known as the Einstein-Hilbert field equations. In this section, we explore the method set out by Hilbert to find the equations of gravity.

Consider the action, S, defined over the unbounded space-time volume V

$$S[g_{\mu\nu}] = \int_V \mathcal{L}\, d^4V;, \tag{2.4.1}$$

where $\mathcal{L}$, is a scalar Lagrange density, and d^4V is a four volume element.

In order to find an expression for the general 4-volume element, we first consider d^4V in Minkowski space, given by

$$d^4V = d^4x^\mu = dx^0dx^1dx^2dx^3\ . \tag{2.4.2}$$

Since we can consider a local region of space-time be flat with metric $\eta_{\alpha\beta}$, we can thus write the global metric $g_{\mu\nu}$, with coordinates ξ^ρ using the transformation law for the metric tensor

$$g_{\mu\nu} = \frac{\partial x^\alpha}{\partial \xi^\mu}\frac{\partial x^\beta}{\partial \xi^\nu}\eta_{\alpha\beta} = J^\alpha_\mu J^\beta_\nu \eta_{\alpha\beta}\ , \tag{2.4.3}$$

where $J^\alpha_\mu, J^\beta_\nu$ are components of the jacobian matrix

$$J = \frac{\partial(x^0, x^1, x^2, x^3)}{\partial(\xi^0, \xi^1, \xi^2\xi^3)}\ . \tag{2.4.4}$$

Thus, if we let g be the determinant of the metric tensor, we see that

$$g = \det(J^\alpha_\mu J^\beta_\nu \eta_{\alpha\beta}) = \det(J^\alpha_\mu J^\beta_\nu)\det(\eta_{\alpha\beta}) = -(\det(J^\alpha_\mu))^2 \tag{2.4.5}$$

and hence,

$$\det(J^\alpha_\mu) = \sqrt{-g}\ . \tag{2.4.6}$$

With this, we can now write the four volume element in terms of the coordinates ξ^ρ, as

$$d^4V = \det(J^\alpha_\mu)\, d^4\xi = \sqrt{-g}\, d^4\xi \tag{2.4.7}$$

Recall, in section 1.6.1 we introduced the Ricci tensor, and more importantly the Ricci scalar, which turns out to be the simplest scalar valued function that depends wholly on the metric tensor and its derivatives. Due to this, we can make a guess that the Ricci scalar is a suitable Lagrangian for this system. Thus, by inserting $\mathcal{L} = R$, along with (2.4.7), in to the action given in equation (2.4.1), we obtain the **Einstein-Hilbert action** [3]

$$S[g_{\mu\nu}] = \int_V R\sqrt{-g}\, d^4\xi\ . \tag{2.4.8}$$

By varying this action, and writing the Ricci scalar as $g^{\beta\sigma}R_{\beta\sigma}$, we obtain

$$\delta S = \int_V \delta(g^{\beta\sigma}R_{\beta\sigma}\sqrt{-g})\, d^4\xi \tag{2.4.9}$$

$$= \int_V \delta g^{\beta\sigma}R_{\beta\sigma}\sqrt{-g} + g^{\beta\sigma}R_{\beta\sigma}\delta\sqrt{-g}\ + g^{\beta\sigma}\sqrt{-g}\,\delta R_{\beta\sigma}\, d^4\xi \tag{2.4.10}$$

$$= \int_V \delta g^{\beta\sigma}R_{\beta\sigma}\sqrt{-g} + R\,\delta\sqrt{-g}\, d^4\xi + \int_V g^{\beta\sigma}\sqrt{-g}\,\delta R_{\beta\sigma}\, d^4\xi\ . \tag{2.4.11}$$

In order to progress further, we need to find out how the determinant of the metric, and also the Ricci tensor vary. This will be the focus of the next two sections.

2.4.1 Variation of the Metric Determinant

Consider, for simplicity, a 2×2 diagonal matrix

$$A = \begin{bmatrix} a_0 & 0 \\ 0 & a_1 \end{bmatrix} , \tag{2.4.12}$$

and recall, to exponentiate or take the logarithm of such a matrix, we can simply apply the desired operation to the non-zero components, i.e

$$e^A = \begin{bmatrix} e^{a_0} & 0 \\ 0 & e^{a_1} \end{bmatrix} \quad \text{and} \quad \ln(A) = \begin{bmatrix} \ln(a_0) & 0 \\ 0 & \ln(a_1) \end{bmatrix} . \tag{2.4.13}$$

If we now define a new matrix $B = e^A$, so that $A = \ln(B)$, then since B is also a diagonal matrix, its determinant is easily found to be

$$\det(B) = e^{\mathrm{tr}(A)} = e^{\mathrm{tr}(\ln(B))} . \tag{2.4.14}$$

Thus, taking the logarithm of both sides then yields

$$\ln(\det(B)) = \mathrm{tr}(\ln(B)) . \tag{2.4.15}$$

We can now vary both sides of this equation to give

$$\delta \ln(\det(B)) = \mathrm{tr}(\delta \ln(B)) \quad \Longrightarrow \quad \frac{\delta \det(B)}{\det(B)} = \mathrm{tr}(B^{-1}\delta B) . \tag{2.4.16}$$

Hence, by substituting $B = g_{\beta\sigma}$, we obtain

$$\frac{\delta g}{g} = g^{\beta\sigma}\delta g_{\beta\sigma} \quad \Longrightarrow \quad \delta g = g g^{\beta\sigma}\delta g_{\beta\sigma} . \tag{2.4.17}$$

We can now apply this to find an expression for the variation $\delta\sqrt{-g}$,

$$\delta\sqrt{-g} = -\frac{1}{2\sqrt{-g}}\delta g = -\frac{1}{2\sqrt{-g}} g g^{\beta\sigma}\delta g_{\beta\sigma} . \tag{2.4.18}$$

Thus,

$$\delta\sqrt{-g} = \frac{1}{2\sqrt{-g}}(\sqrt{-g})^2 g^{\beta\sigma}\delta g_{\beta\sigma} \quad \Longrightarrow \quad \delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}\, g^{\beta\sigma}\delta g_{\beta\sigma} . \tag{2.4.19}$$

Notice that since $g_{\beta\sigma}g^{\beta\sigma} = \delta^\mu_\mu = 1$, then by the product rule we get

$$\delta(g_{\beta\sigma}g^{\beta\sigma}) = g_{\beta\sigma}\delta g^{\beta\sigma} + g^{\beta\sigma}\delta g_{\beta\sigma} = 0 \tag{2.4.20}$$

$$\Longrightarrow \quad g_{\beta\sigma}\delta g^{\beta\sigma} = -g^{\beta\sigma}\delta g_{\beta\sigma} . \tag{2.4.21}$$

By plugging this equation in to (2.4.19) we thus obtain the result

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\, g_{\beta\sigma}\delta g^{\beta\sigma} . \tag{2.4.22}$$

2.4.2 Variation of the Ricci tensor: The Palatini Identity

Recall, the Riemann curvature tensor is defined to be

$$R^\alpha_{\beta\gamma\sigma} = \partial_\gamma\Gamma^\alpha_{\beta\sigma} - \partial_\sigma\Gamma^\alpha_{\beta\gamma} + \Gamma^\alpha_{\gamma\rho}\Gamma^\rho_{\sigma\beta} - \Gamma^\alpha_{\sigma\rho}\Gamma^\rho_{\beta\gamma} , \tag{2.4.23}$$

and hence, we can write the Ricci tensor by contracting the upper index with the second lower index

$$R_{\beta\sigma} = R^\alpha_{\beta\alpha\sigma} = \partial_\alpha\Gamma^\alpha_{\beta\sigma} - \partial_\sigma\Gamma^\alpha_{\beta\alpha} + \Gamma^\alpha_{\alpha\rho}\Gamma^\rho_{\beta\sigma} - \Gamma^\alpha_{\sigma\rho}\Gamma^\rho_{\beta\alpha} . \tag{2.4.24}$$

Applying a variation to the Ricci tensor, and by using the product rule where appropriate, we obtain

$$\delta R_{\beta\sigma} = \partial_\alpha \delta\Gamma^\alpha_{\beta\sigma} - \partial_\sigma \delta\Gamma^\alpha_{\beta\alpha} + \Gamma^\rho_{\beta\sigma}\delta\Gamma^\alpha_{\alpha\rho} + \Gamma^\alpha_{\alpha\rho}\delta\Gamma^\rho_{\beta\sigma} - \Gamma^\rho_{\beta\alpha}\delta\Gamma^\alpha_{\sigma\rho} - \Gamma^\alpha_{\sigma\rho}\delta\Gamma^\rho_{\beta\alpha} \ . \tag{2.4.25}$$

Consider now the following covariant derivatives

$$\nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) = \partial_\alpha\delta\Gamma^\alpha_{\beta\sigma} + \Gamma^\alpha_{\alpha\rho}\delta\Gamma^\rho_{\beta\sigma} - \Gamma^\rho_{\beta\alpha}\delta\Gamma^\alpha_{\sigma\rho} - \Gamma^\rho_{\alpha\sigma}\delta\Gamma^\alpha_{\beta\rho} \ , \tag{2.4.26}$$

$$\nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}) = \partial_\sigma\delta\Gamma^\alpha_{\beta\alpha} + \Gamma^\alpha_{\sigma\rho}\delta\Gamma^\rho_{\beta\alpha} - \Gamma^\rho_{\beta\sigma}\delta\Gamma^\alpha_{\alpha\rho} - \Gamma^\rho_{\sigma\alpha}\delta\Gamma^\alpha_{\beta\rho} \ . \tag{2.4.27}$$

Taking the difference of these two equations gives

$$\begin{aligned}\nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}) = \ & \partial_\alpha\delta\Gamma^\alpha_{\beta\sigma} + \Gamma^\alpha_{\alpha\rho}\delta\Gamma^\rho_{\beta\sigma} - \Gamma^\rho_{\beta\alpha}\delta\Gamma^\alpha_{\sigma\rho} - \cancel{\Gamma^\rho_{\alpha\sigma}\delta\Gamma^\alpha_{\beta\rho}} \\ & -\partial_\sigma\delta\Gamma^\alpha_{\beta\alpha} - \Gamma^\alpha_{\sigma\rho}\delta\Gamma^\rho_{\beta\alpha} + \Gamma^\rho_{\beta\sigma}\delta\Gamma^\alpha_{\alpha\rho} + \cancel{\Gamma^\rho_{\sigma\alpha}\delta\Gamma^\alpha_{\beta\rho}} \ ,\end{aligned} \tag{2.4.28}$$

where we have used the symmetry of the Christoffel symbols to cancel the fourth and last terms. Thus, we are left with

$$\nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}) = \partial_\alpha\delta\Gamma^\alpha_{\beta\sigma} - \partial_\sigma\delta\Gamma^\alpha_{\beta\alpha} + \Gamma^\rho_{\beta\sigma}\delta\Gamma^\alpha_{\alpha\rho} + \Gamma^\alpha_{\alpha\rho}\delta\Gamma^\rho_{\beta\sigma} - \Gamma^\rho_{\beta\alpha}\delta\Gamma^\alpha_{\sigma\rho} - \Gamma^\alpha_{\sigma\rho}\delta\Gamma^\rho_{\beta\alpha} \ . \tag{2.4.29}$$

Hence, upon comparison with (2.4.25), we obtain the **Palatini identity** [2], which gives the variation of the Ricci tensor, in terms of covariant derivatives

$$\delta R_{\beta\sigma} = \nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}) \ . \tag{2.4.30}$$

2.4.3 Einstein's Field Equations in Vacuum

Now that we have developed the necessary variations, we can continue with varying the Einstein-Hilbert action. By substituting (2.4.22) and (2.4.30) in to equation (2.4.11) we now obtain

$$\begin{aligned}\delta S &= \int_V \delta g^{\beta\sigma} R_{\beta\sigma}\sqrt{-g} - \frac{1}{2}R\sqrt{-g}\, g_{\beta\sigma}\delta g^{\beta\sigma}\, d^4\xi + \int_V g^{\beta\sigma}\sqrt{-g}\,(\nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}))\, d^4\xi \\ &= \int_V \delta g^{\beta\sigma}\sqrt{-g}(R_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}R)\, d^4\xi + \int_V g^{\beta\sigma}\sqrt{-g}\,(\nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}))\, d^4\xi \\ &= \delta S_1 + \delta S_2 \ .\end{aligned} \tag{2.4.31}$$

Looking at the second integral, δS_2, we see that

$$\delta S_2 = \int_V g^{\beta\sigma}\sqrt{-g}\,(\nabla_\alpha(\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^\alpha_{\beta\alpha}))\, d^4\xi = \int_V \sqrt{-g}\,(\nabla_\alpha(g^{\beta\sigma}\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\sigma(g^{\beta\sigma}\delta\Gamma^\alpha_{\beta\alpha}))\, d^4\xi \tag{2.4.32}$$

and hence, relabeling $\sigma \to \alpha$ and $\alpha \to \sigma$ in the second covariant derivative gives

$$\delta S_2 = \int_V \sqrt{-g}\,(\nabla_\alpha(g^{\beta\sigma}\delta\Gamma^\alpha_{\beta\sigma}) - \nabla_\alpha(g^{\beta\alpha}\delta\Gamma^\sigma_{\beta\sigma}))\, d^4\xi = \int_V \sqrt{-g}\,\nabla_\alpha(\underbrace{g^{\beta\sigma}\delta\Gamma^\alpha_{\beta\sigma} - g^{\beta\alpha}\delta\Gamma^\sigma_{\beta\sigma}}_{W^\alpha})\, d^4\xi \ . \tag{2.4.33}$$

Notice that by contracting the indices σ, β the term inside the covariant derivative becomes a tensor of type (1,0), call it W^α. Using a generalised version of Stokes' theorem [14], we then obtain

$$\delta S_2 = \int_V \sqrt{-g}\,\nabla_\alpha W^\alpha\, d^4\xi = \int_{\partial V} \sqrt{-h}\, W^\alpha n_\alpha\, d^3\xi \ . \tag{2.4.34}$$

Here, h is the induced metric, and n_α is a unit vector that is orthogonal to the boundary manifold ∂V. This integral is now over the boundary of space-time, and thus, by demanding that there be no variation at infinity (on the boundary), we can conclude that this integral must vanish and the total action is therefore

$$\delta S = \int_V \delta g^{\beta\sigma}\sqrt{-g}(R_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}R)\, d^4\xi \ . \tag{2.4.35}$$

Thus, by the principle of least action $\delta S = 0$ we have

$$\int_V \delta g^{\beta\sigma} \sqrt{-g}(R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R)\, d^4\xi = 0 \tag{2.4.36}$$

and hence, in order for this integral to vanish for all possible variations, the integrand its self must vanish. Therefore, we arrive at the **Einstein field equations in vacuum**

$$R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R = 0\ . \tag{2.4.37}$$

This equation describes space-time in the absence of mass and energy.

Alternatively, we can write this equation in terms of the functional variation of the action since,

$$\delta S = \int_V \delta g^{\beta\sigma} \sqrt{-g}(R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R)\, d^4\xi = \int_V \frac{\delta S}{\delta g^{\beta\sigma}} \delta g^{\beta\sigma} d^4\xi \tag{2.4.38}$$

where

$$\frac{\delta S}{\delta g^{\beta\sigma}} = \sqrt{-g}(R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R) \tag{2.4.39}$$

Hence,

$$\frac{\delta S}{\sqrt{-g}\delta g^{\beta\sigma}} = R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R \tag{2.4.40}$$

This form of the equations will be used to derive the full field equations in the next section.

2.4.4 The Full Einstein Field Equations

In order to compute the field equations for non-empty space time, we need to introduce a new term in the action to account for matter and energy, call it $S_{\mathcal{M}}$. The full action is now [12]

$$S = \frac{1}{2\kappa} S_{EH} + S_{\mathcal{M}}\ , \tag{2.4.41}$$

where S_{EH}, denotes the Einstein-Hilbert action developed in the previous section, and κ is some constant to be determined. Varying this action with respect to $g^{\beta\sigma}$, and dividing each term by $\sqrt{-g}$, then gives

$$\frac{\delta S}{\delta g^{\beta\sigma}\sqrt{-g}} = \frac{1}{2\kappa\sqrt{-g}} \frac{\delta S_{EH}}{\delta g^{\beta\sigma}} + \frac{1}{\sqrt{-g}} \frac{\delta S_{\mathcal{M}}}{\delta g^{\beta\sigma}}\ . \tag{2.4.42}$$

Thus, by equation (2.4.40)

$$\frac{\delta S}{\delta g^{\beta\sigma}\sqrt{-g}} = \frac{1}{2\kappa}(R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R) + \frac{1}{\sqrt{-g}} \frac{\delta S_{\mathcal{M}}}{\delta g^{\beta\sigma}}\ . \tag{2.4.43}$$

Hence, by the principle of least action, $\delta S = 0$, the right hand side of this equation will vanish and we now obtain

$$\frac{1}{2\kappa}(R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R) = -\frac{1}{\sqrt{-g}} \frac{\delta S_{\mathcal{M}}}{\delta g^{\beta\sigma}} = \frac{1}{2} T_{\beta\sigma}\ , \tag{2.4.44}$$

where we have defined the **energy-momentum tensor** to be given by [3]

$$T_{\beta\sigma} = -\frac{2}{\sqrt{-g}} \frac{\delta S_{\mathcal{M}}}{\delta g^{\beta\sigma}}\ . \tag{2.4.45}$$

and hence, by substituting this in to (2.4.44) and multiplying through by κ we get

$$R_{\beta\sigma} - \frac{1}{2} g_{\beta\sigma} R = \kappa T_{\beta\sigma}\ . \tag{2.4.46}$$

This equation is now in the correct form, with all terms involving curvature of space-time on the left and the source term, namely the energy tensor, on the right hand side.

2.4.5 Determining The Value of κ Using The Newtonian Limit

In order to determine the value of the constant κ we need to again look at the Newtonian limit.

By contracting (2.4.46) using the inverse metric tensor, it can be shown that

$$R = -\kappa T, \tag{2.4.47}$$

We can then substitute this expression for the Ricci scalar back in to (2.4.46), allowing us to write the Ricci tensor completely in terms of the $T_{\beta\sigma}$ and $T = g^{\beta\sigma}T_{\beta\sigma}$

$$R_{\beta\sigma} = \kappa\Big(T_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}T\Big) . \tag{2.4.48}$$

Thus, the time component of this equation is given by

$$R_{00} = \kappa\Big(T_{00} - \frac{1}{2}g_{00}T\Big) , \tag{2.4.49}$$

where as we found in section 2.3,

$$g_{00} = \eta_{00} + h_{00} = -1 + h_{00} , \tag{2.4.50}$$
$$g^{00} = \eta^{00} - h^{00} = -1 - h^{00} \tag{2.4.51}$$

and hence, up to lowest order we have

$$T = g^{00}T_{00} = -T_{00} . \tag{2.4.52}$$

To find the value of T_{00}, we consider the energy-momentum tensor for a perfect fluid, written in a more conventional form

$$T_{\beta\sigma} = (\rho + p)U_\beta U_\sigma + pg_{\beta\sigma}; \quad U^\mu = (c, 0, 0, 0) . \tag{2.4.53}$$

In the Newtonian limit however, we can neglect the pressure terms, and since we are considering each particle to have no spatial velocity, the only non-zero component is given by the **rest energy** [3]

$$T_{00} = \rho c^2 . \tag{2.4.54}$$

Hence,

$$T = g^{00}T_{00} \approx \eta^{00}T_{00} = -\rho c^2 . \tag{2.4.55}$$

Plugging this result along with (2.4.52), back in to (2.4.49) thus gives

$$R_{00} = \kappa\Big(T_{00} - \frac{1}{2}T_{00}\Big) = \frac{1}{2}\kappa\rho c^2 . \tag{2.4.56}$$

We now need to compute the time-like component of the Ricci tensor, which by definition can be written in terms of the Riemann tensor as

$$R_{00} = R^j_{0j0} = \partial_j\Gamma^j_{00} - \partial_0\Gamma^j_{j0} + \Gamma^i_{ik}\Gamma^k_{00} - \Gamma^i_{0k}\Gamma^k_{i0} . \tag{2.4.57}$$

Since we are considering a static field, then due to the second term containing a time derivative, we can see that this term vanishes. We can also neglect the last two terms since they are dependent on the metric and would thus contribute terms of second order. This leaves us with, up to first order in h

$$R_{00} \approx \partial_j\Gamma^j_{00} = \partial_j\Big(-\frac{1}{2}\delta^{ji}\partial_i h_{00}\Big) . \tag{2.4.58}$$

Using equation (2.3.14) and by noticing that in three dimensions, η^{ji} becomes δ^{ji}, we obtain

$$R_{00} = \partial_j\Big(-\frac{1}{2}\delta^{ji}\partial_i h_{00}\Big) = -\frac{1}{2}\partial_i\partial_i h_{00} = -\frac{1}{2}\nabla^2 h_{00} . \tag{2.4.59}$$

Hence, by equation (2.3.20) we have

$$R_{00} = \frac{1}{c^2}\nabla^2\phi \; , \tag{2.4.60}$$

which upon inserting into equation (2.4.56) yields

$$\frac{1}{c^2}\nabla^2\phi = \frac{1}{2}\kappa\rho c^2 \; . \tag{2.4.61}$$

Then, using Poisson's equation, we can rewrite the left hand side to obtain

$$\frac{4\pi G\rho}{c^2} = \frac{1}{2}\kappa\rho c^2 \; . \tag{2.4.62}$$

Finally, by solving this equation for κ we find that

$$\kappa = \frac{8\pi G}{c^4} \; . \tag{2.4.63}$$

Substituting this value in to equation (2.4.46) gives, at last, the full **Einstein field equations**

$$R_{\beta\sigma} - \frac{1}{2}Rg_{\beta\sigma} = \frac{8\pi G}{c^4}T_{\beta\sigma} \tag{2.4.64}$$

The left hand side of this equation relates purely to the curvature of space-time, whereas the right hand side purely relates to matter and energy. Thus, matter and energy tells space-time how to curve. It is this curvature that we experience as gravity!

Remark: Inserting the value for κ in to equation (2.4.48) gives the **trace reversed** form of Einstein's equations

$$R_{\beta\sigma} = \frac{8\pi G}{c^4}\left(T_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}T\right) \; . \tag{2.4.65}$$

Remark: When Einstein first published his theory of general relativity, it was widely believed that the universe was static. However, Einstein quickly realised that his equations did not agree with this belief, and thus led him to introduce a new term in the equations, known as the **cosmological constant**, Λ, and the field equations took on the form [3]

$$R_{\beta\sigma} - \frac{1}{2}Rg_{\beta\sigma} + \Lambda g_{\beta\sigma} = \frac{8\pi G}{c^4}T_{\beta\sigma} \; . \tag{2.4.66}$$

It was only when, in 1931, Edwin Hubble discovered that the universe is not static, rather, it is expanding, that the cosmological constant was again discarded, and the field equations reverted to their original form.

2.5 The Schwarzschild solution

In this section we consider perhaps the most obvious type of gravitational field: one that is spherically symmetric and is also static. Immediate examples of such gravitational fields are that of stars and planets. Moreover, to keep things as simple as possible, we will be concerned here with exterior solutions i.e solutions in the **empty** space outside of the gravitating object. In addition, such a solution will lead directly to possibly the most famous consequence of general relativity: Black holes.

2.5.1 Setting up the metric

Since we are considering only exterior solutions to Einstein's equations, one can immediately infer; since the space around the object is free of matter and energy, the energy-momentum tensor will vanish at all points outside of the object. We therefore need only be concerned with the field equations in vacuum

$$R_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}R = 0 \; . \tag{2.5.1}$$

Furthermore, equation (2.4.65) leads us to the the simple requirement that only the Ricci tensor needs to vanish in order for the field equations in vacuum to be satisfied, i.e

$$R_{\beta\sigma} = 0 \ . \tag{2.5.2}$$

To begin solving this equation we consider with the condition that the space-time surrounding the object is static, that is to say, the components of the metric must be time independent and also the line element must be invariant under time reversal, i.e $t \to -t$. Due to these assumptions, the line element must necessarily take the form (with metric signature $(-,+,+,+)$)

$$ds^2 = -g_{tt}dt^2 + g_{ij}dx^i dx^j \tag{2.5.3}$$

where $g_{tt}, g_{ij} > 0$ are functions of only the spatial coordinates. In order to impose spherical symmetry, we assume a uniform S^2-foliation of space-time, that is, we can decompose space-time into uniformly stacked 2-spheres. In general one could then endow each sphere with its own set of angular coordinates $\{\theta, \phi\}$, however, we will assume that they are aligned such that we can use the same coordinates on every sphere. Thus, by our previous discussion on the metric tensor in spherical polar coordinates, we can write the line element for one such sphere of constant radius ρ

$$ds^2_{sphere} = \rho^2 d\Omega^2 \tag{2.5.4}$$

where $d\Omega^2 = d\theta^2 + \sin^2(\theta)d\phi^2$ is the line element on the surface of the unit sphere. With this, we can write down a general form for the full space-time line element

$$ds^2 = -A^2dt^2 + B^2d\rho^2 + \rho^2 d\Omega^2 \tag{2.5.5}$$

where due to the isotropic nature of this situation, the arbitrary functions A and B should depend only on the radial coordinate. Also, since the radial direction is orthogonal to the spheres of constant radius ρ, there cannot be any $d\rho d\theta$ or $d\rho d\phi$ terms [11]. In order to simplify things later on, we write $A(\rho)$ and $B(\rho)$ in terms of an exponential, i.e the line element becomes

$$ds^2 = -e^{2\alpha(\rho)}dt^2 + e^{2\beta(\rho)}d\rho^2 + \rho^2 d\Omega^2 \ , \tag{2.5.6}$$

and the corresponding components of the metric tensor are

$$g_{tt} = -e^{2\alpha(\rho)}, \ g_{\rho\rho} = e^{2\beta(\rho)}, \ g_{\theta\theta} = \rho^2, \ g_{\phi\phi} = \rho^2\sin^2(\theta) \tag{2.5.7}$$

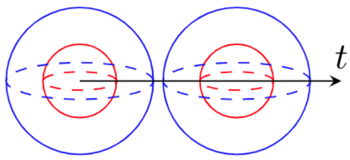

Figure 7: S^2-foliation of space-time

2.5.2 The Ricci tensor

Now that we have found the general form of the line element along with its corresponding metric components, we can begin to calculate the components of the Ricci tensor. Using (1.4.80) together with the metric components from the previous section, we find that the only non-zero Christoffel symbols for this system are given by [3]

$$\begin{aligned}
&\Gamma^{\rho}_{tt} = e^{2(\alpha-\beta)}\partial_\rho\alpha, \quad \Gamma^{\rho}_{\rho\rho} = \partial_\rho\beta, \quad \Gamma^{\rho}_{\theta\theta} = -\rho e^{-2\beta}, \quad \Gamma^{\rho}_{\phi\phi} = -\rho e^{-2\beta}\sin^2(\theta),\\
&\Gamma^{\theta}_{\rho\theta} = \frac{1}{\rho}, \quad \Gamma^{\theta}_{\phi\phi} = -\sin(\theta)\cos(\theta),\\
&\Gamma^{\phi}_{\rho\phi} = \frac{1}{\rho}, \quad \Gamma^{\phi}_{\theta\phi} = \frac{\cos(\theta)}{\sin(\theta)},\\
&\Gamma^{t}_{t\rho} = \partial_\rho\alpha \ .
\end{aligned}$$

Recall, the Ricci tensor is formed by contracting the upper index of the Riemann curvature tensor, with one one of its lower indices

$$R_{\lambda\alpha} = R^{\beta}_{\lambda\beta\alpha} = \partial_\beta \Gamma^{\beta}_{\lambda\alpha} - \partial_\alpha \Gamma^{\beta}_{\lambda\beta} + \Gamma^{\beta}_{\tau\beta}\Gamma^{\tau}_{\lambda\alpha} - \Gamma^{\beta}_{\tau\alpha}\Gamma^{\tau}_{\lambda\beta} \tag{2.5.8}$$

where $\beta, \lambda, \alpha \in \{t, \rho, \theta, \phi\}$.Observe, the case where $\lambda = \alpha = t$ gives the component

$$R_{tt} = \partial_\beta \Gamma^{\beta}_{tt} - \partial_t \Gamma^{\beta}_{t\beta} + \Gamma^{\beta}_{\tau\beta}\Gamma^{\tau}_{tt} - \Gamma^{\beta}_{\tau t}\Gamma^{\tau}_{t\beta} \, . \tag{2.5.9}$$

Hence, by explicitly writing the sums over β, we obtain

$$R_{tt} = \partial_t \Gamma^{t}_{tt} + \partial_\rho \Gamma^{\rho}_{tt} + \partial_\theta \Gamma^{\theta}_{tt} + \partial_\phi \Gamma^{\phi}_{tt} + \partial_t(\Gamma^{t}_{tt} + \Gamma^{\rho}_{t\rho} + \Gamma^{\theta}_{t\theta} + \Gamma^{\phi}_{t\phi}) \tag{2.5.10}$$

$$+ (\Gamma^{t}_{\tau t}\Gamma^{\tau}_{tt} + \Gamma^{\rho}_{\tau\rho}\Gamma^{\tau}_{tt} + \Gamma^{\theta}_{\tau\theta}\Gamma^{\tau}_{tt} + \Gamma^{\phi}_{\tau\phi}\Gamma^{\tau}_{tt}) - (\Gamma^{t}_{\tau t}\Gamma^{\tau}_{tt} + \Gamma^{\rho}_{\tau t}\Gamma^{\tau}_{t\rho} + \Gamma^{\theta}_{\tau t}\Gamma^{\tau}_{t\theta} + \Gamma^{\phi}_{\tau t}\Gamma^{\tau}_{t\phi}) \, . \tag{2.5.11}$$

It is now immediately clear that all terms being acted on by ∂_t are zero and that setting $\tau \to \rho$ gives the only combination of indices such that we get non vanishing terms in the last two sets of brackets (notice also that the terms first terms in these brackets will cancel each other out), thus we have

$$R_{tt} = \partial_t \Gamma^{t}_{tt} + \Gamma^{\theta}_{\rho\theta}\Gamma^{\rho}_{tt} + \Gamma^{\rho}_{\rho\rho}\Gamma^{\rho}_{tt} + \Gamma^{\phi}_{\rho\phi}\Gamma^{\rho}_{tt} - \Gamma^{\rho}_{tt}\Gamma^{t}_{\rho t} \, . \tag{2.5.12}$$

We can now substitute in for the Christoffel symbols to give

$$R_{tt} = \partial_t (e^{2(\alpha-\beta)}\partial_\rho \alpha) + e^{2(\alpha-\beta)} \Big[\frac{2}{\rho}\partial_\rho \alpha + \partial_\rho \alpha \partial_\rho \beta - (\partial_\rho \alpha)^2 \Big] \, . \tag{2.5.13}$$

By carrying out the differentiation in the first term we now obtain

$$R_{tt} = e^{2(\alpha-\beta)}\partial^2_\rho \alpha + [2(\partial_\rho \alpha)^2 - 2\partial_\rho \alpha \partial_\rho \beta] + e^{2(\alpha-\beta)} \Big[\frac{2}{\rho}\partial_\rho \alpha + \partial_\rho \alpha \partial_\rho \beta - (\partial_\rho \alpha)^2 \Big] \tag{2.5.14}$$

and hence, we find the tt component of the Ricci tensor is given by

$$R_{tt} = e^{2(\alpha-\beta)} \Big(\partial^2_\rho \alpha + (\partial_\rho \alpha)^2 - \partial_\rho \alpha \partial_\rho \beta + \frac{2}{\rho}\partial_\rho \alpha \Big) \, . \tag{2.5.15}$$

Similarly, one finds that the only other non-zero components are

$$R_{\rho\rho} = -\partial^2_\rho \alpha - (\partial_\rho \alpha)^2 + \partial_\rho \alpha \partial_\rho \beta + \frac{2}{\rho}\partial_\rho \beta \tag{2.5.16}$$

$$R_{\theta\theta} = e^{-2\beta}[\rho(\partial_\rho \beta - \partial_\rho \alpha) - 1] + 1 \tag{2.5.17}$$

$$R_{\phi\phi} = \sin^2(\theta) R_{\theta\theta} \, . \tag{2.5.18}$$

Furthermore, although not relevant for this calculation, one can compute the curvature scalar to be [3]

$$R = -2e^{-2\beta} \Big[\partial^2_\rho \alpha + (\partial_\rho \alpha)^2 - \partial_\rho \alpha \partial_\rho \beta + \frac{2}{\rho}(\partial_\rho \alpha - \partial_\rho \beta) + \frac{1}{\rho^2}(1 - e^{2\beta}) \Big] \, . \tag{2.5.19}$$

We will however use this quantity when looking at interior solutions.

2.5.3 Solving the Einstein Equations

With the Ricci tensor components calculated, we can now apply equation (2.5.2) to obtain a set of four, coupled differential equations.

$$R_{tt} = e^{2(\alpha-\beta)} \Big(\partial^2_\rho \alpha + (\partial_\rho \alpha)^2 - \partial_\rho \alpha \partial_\rho \beta + \frac{2}{\rho}\partial_\rho \alpha \Big) = 0 \tag{2.5.20}$$

$$R_{\rho\rho} = -\partial^2_\rho \alpha - (\partial_\rho \alpha)^2 + \partial_\rho \alpha \partial_\rho \beta + \frac{2}{\rho}\partial_\rho \beta = 0 \tag{2.5.21}$$

$$R_{\theta\theta} = e^{-2\beta}[\rho(\partial_\rho \beta - \partial_\rho \alpha) - 1] + 1 = 0 \tag{2.5.22}$$

$$R_{\phi\phi} = e^{-2\beta}\sin^2(\theta)[\rho(\partial_\rho \beta - \partial_\rho \alpha) - 1] + 1 = 0 \tag{2.5.23}$$

Notice, since (2.5.20) and (2.5.21) vanish independently, we can add both equations to give

$$0 = R_{tt} + R_{\rho\rho} = \frac{2}{\rho} e^{\alpha-\beta}(\partial_\rho \alpha + \partial_\rho \beta) \tag{2.5.24}$$

and hence,

$$\partial_\rho \alpha + \partial_\rho \beta = 0 \implies \alpha = -\beta + c \tag{2.5.25}$$

where c is an arbitrary constant which we can eliminate by re-scaling $dt \to e^{-c} dt$ [3], leaving us with

$$\alpha = -\beta \, . \tag{2.5.26}$$

Plugging this result into equation (2.5.22) now gives us

$$2\rho e^{2\alpha} \partial_\rho \alpha + e^{2\alpha} = 1 \, . \tag{2.5.27}$$

Observe that the left hand side of this equation is a total derivative and we therefore can write

$$\partial_\rho (\rho e^{2\alpha}) = 1 \, . \tag{2.5.28}$$

Hence, by solving for $e^{2\alpha}$ we obtain

$$e^{2\alpha} = 1 - \frac{R_s}{\rho} \tag{2.5.29}$$

where R_s is some constant to be determined. With this equation, together with the result in (2.5.26), our line element can now be expressed as

$$ds^2 = -\Big(1 - \frac{R_s}{\rho}\Big) dt^2 + \Big(1 - \frac{R_s}{\rho}\Big)^{-1} d\rho^2 + \rho^2 d\Omega^2 \, . \tag{2.5.30}$$

To determine R_s, we recall that in Newtonian limit

$$g_{00} = g_{tt} = -(1 + \frac{2\phi}{c^2}) \tag{2.5.31}$$

where the gravitational potential ϕ, is given by [11]

$$\phi = -\frac{GM}{\rho} \, . \tag{2.5.32}$$

Hence, by comparing these results with the g_{tt} component in (2.5.30), we see that $R_s = 2GM/c^2$, and we thus arrive at the **Schwarzschild solution** written in terms of **Schwarschild coordinates** [11]

$$ds^2 = -\Big(1 - \frac{2GM}{c^2 \rho}\Big) dt^2 + \Big(1 - \frac{2GM}{c^2 \rho}\Big)^{-1} d\rho^2 + \rho^2 d\Omega^2 \, . \tag{2.5.33}$$

The quantity R_s, denotes the **Schwarzschild radius**, that is, the radius of the event horizon of a non-rotating black hole. Furthermore, any object whose radius is smaller than its Schwarzschild radius will be a black hole (we will discuss this in greater detail in the next section).

An important feature of this solution is that it is asymptotically flat, i.e

$$\rho \to \infty \quad \implies \quad ds^2 \to \eta_{\mu\nu} dx^\mu dx^\nu \, . \tag{2.5.34}$$

2.6 Interior Solutions

In order to further motivate the existence of blackholes, we first take a dive into interior solutions. Whereas in the exterior solution we found that the energy-momentum tensor vanishes, we now have to reckon with the fact that that for interior solutions, $T_{\beta\sigma} \neq 0$. Thus, we now have to consider the Einstein equations in their full form

$$G_{\beta\sigma} \equiv R_{\beta\sigma} - \frac{1}{2}Rg_{\beta\sigma} = 8\pi G\, T_{\beta\sigma} \ . \tag{2.6.1}$$

Note that here we have set $c = 1$. In the previous section we discussed the idea of a static, spherically symmetric space-time, for which we defined the general line element to be

$$ds^2 = g_{tt}dt^2 + g_{\rho\rho}d\rho^2 + g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2 = -e^{2\alpha(\rho)}dt^2 + e^{2\beta(\rho)}d\rho^2 + \rho^2 d\Omega^2 \ . \tag{2.6.2}$$

Here, we begin by assuming a time-independent solution and at the end we will find a condition for which this assumption fails.

Following from our calculation of the Ricci tensor and curvature scalar: after a little algebra, the non-zero components of the Einstein tensor are found to be [3]

$$G_{tt} = \frac{1}{\rho^2}e^{2(\alpha-\beta)}(2\rho\partial_\rho\beta - 1 + e^{2\beta}) \ , \tag{2.6.3}$$

$$G_{\rho\rho} = \frac{1}{\rho^2}(2\rho\partial_\rho\beta + 1 - e^{2\beta}) \ , \tag{2.6.4}$$

$$G_{\theta\theta} = \rho^2 e^{-2\beta}\left[\partial_\rho^2\alpha + (\partial_\rho\alpha)^2 - \partial_\rho\alpha\partial_\rho\beta + \frac{1}{\rho}(\partial_\rho\alpha - \partial_\rho\beta)\right] \ , \tag{2.6.5}$$

$$G_{\phi\phi} = \sin^2(\theta)G_{\theta\theta} \ . \tag{2.6.6}$$

With the left hand side of Einstein's equation computed, we now turn our attention to the energy-momentum tensor. Suppose we model a star (of radius R) as a perfect fluid so that the energy momentum tensor takes the form

$$T_{\beta\sigma} = (\varrho + p)U_\beta U_\sigma + pg_{\beta\sigma} \ , \tag{2.6.7}$$

where the energy density ϱ and pressure p are dependent on the radial coordinate ρ. Assuming the fluid particles are at rest with respect to each other, the four velocity will move only in the timelike direction. Thus, if the four velocity is normalised to $U_\mu U^\mu = -1$, then we have

$$-1 = U_\mu U^\mu = g^{\mu\nu}U_\mu U_\nu \implies g^{tt}U_t^2 = -e^{2\alpha}U_t^2 = -1 \ , \tag{2.6.8}$$

and hence, $U_t = e^\alpha$. Using this result together with the components of the metric tensor, we can express the energy momentum tensor as the diagonal matrix

$$T_{\beta\sigma} = \begin{bmatrix} \varrho e^{2\alpha} & 0 & 0 & 0 \\ 0 & pe^{2\beta} & 0 & 0 \\ 0 & 0 & p\varrho^2 & 0 \\ 0 & 0 & 0 & \varrho^2 p\sin^2(\theta) \end{bmatrix} \tag{2.6.9}$$

Consider now the tt component of Einstein's equation

$$\frac{1}{\rho^2}e^{-2\beta}(2\rho\partial_\rho\beta - 1 + e^{2\beta}) = 8\pi G\varrho \ . \tag{2.6.10}$$

Although this may not be the most difficult equation to solve, it is convenient to introduce a new function to replace β, given by [11]

$$m(\rho) = \frac{\rho}{2G}(1 - e^{-2\beta}) \ , \tag{2.6.11}$$

such that $m(\rho = R) = M$. Upon rearranging we get

$$e^{2\beta} = \left[1 - \frac{2Gm(\rho)}{\rho}\right]^{-1} \ , \tag{2.6.12}$$

and we now see that when $\rho = R$, this function agrees with the exterior Schwarzschild solution. By differentiating the reciprocal of both sides of this equation with respect to ρ, we obtain

$$2e^{-2\beta}\partial_\rho\beta = 2G\,\frac{\rho\partial_\rho m - m(\rho)}{\rho^2}\,. \tag{2.6.13}$$

Plugging this result in to (2.6.10) now gives

$$\frac{1}{\rho^2}\Big[2G\,\frac{\rho\partial_\rho m - m(\rho)}{\rho^2} - \Big(1 - \frac{2Gm(\rho)}{\rho}\Big) + 1\Big] = 8\pi G\varrho\,, \tag{2.6.14}$$

and hence, simplifying leaves us with

$$\partial_\rho m = \frac{dm}{d\rho} = 4\pi\rho^2\varrho\,, \tag{2.6.15}$$

which can be integrated to give the result

$$m(\rho) = 4\pi\int_0^\rho \varrho(\rho')\rho'^2 d\rho'\,. \tag{2.6.16}$$

This can be interpreted as saying; the mass of the star within a radius ρ, is just the integral of the energy density over that region.

We can now move on to the $\rho\rho$ component of the Einstein equation, which in terms of α and β is given by

$$\rho^{-2}[1 - e^{2\beta} + 2\rho\partial_\rho\alpha] = 8\pi Ge^{2\beta}\,, \tag{2.6.17}$$

and in a similar way to the tt component, we can use the function $m(\rho)$, to write

$$\partial_\rho\alpha = \frac{Gm(\rho) - 4\pi G\rho^3 p}{\rho(\rho - 2Gm(\rho))}\,. \tag{2.6.18}$$

Instead of now analysing the $\theta\theta$ and $\phi\phi$ components, it is convenient instead to use the following property of the energy momentum tensor (see [3] for proof)

$$\nabla_\beta T^{\beta\sigma} = 0\,. \tag{2.6.19}$$

For our metric, the only non-trivial component of this equation occurs when $\nu = \rho$, and is given by

$$(\varrho + p)\frac{d\alpha}{d\rho} = -\frac{dp}{d\rho}\,. \tag{2.6.20}$$

Substituting in our equation for $\partial_\rho\alpha$, leaves us with the **Tolman-Oppenheimer-Volkoff** equation of hydrostatic equilibrium [3]

$$\frac{dp}{d\rho} = -\frac{(\varrho + p)(Gm(\rho) + 4\pi G\rho^3 p)}{\rho(\rho - 2Gm(\rho))} \tag{2.6.21}$$

We now have all the equations we need to determine $m(\rho), p(\rho)$ and $\alpha(\rho)$ given an energy density distribution $\varrho(\rho)$. Suppose the star in question has constant energy density

$$\varrho(\rho) = \begin{cases} \tilde{\varrho}, & \rho < R \\ 0, & r > R \end{cases} \tag{2.6.22}$$

It is now straightforward to integrate equation (2.6.16) to obtain

$$m(\rho) = \begin{cases} \frac{4}{3}\pi\rho^3\tilde{\varrho}, & r < R \\ M, & r > R \end{cases} \tag{2.6.23}$$

since we defined $m(\rho = R) = M$. Plugging this result into the Tolman-Oppenheimer-Volkoff equation thus gives, after a "little bit" of work [11]

$$p(\rho) = \tilde{\varrho}\left[\frac{\sqrt{1-\frac{2GM}{R}} - \sqrt{1-\frac{2GM\rho^2}{R^3}}}{\sqrt{1-\frac{2GM\rho^2}{R^3}} - 3\sqrt{1-\frac{2GM}{R}}}\right], \tag{2.6.24}$$

and finally, using equation (2.6.18) we can find the metric component $g_{tt} = -e^{2\alpha(\rho)}$, and we see that [3]

$$e^{\alpha(\rho)} = \frac{3}{2}\sqrt{1-\frac{2GM}{R}} - \frac{1}{2}\sqrt{1-\frac{2GM\rho^2}{R^3}}, \quad \rho < R \tag{2.6.25}$$

and as we would expect, when $\rho > R$ the solution agrees with the exterior solution from the previous section

$$e^{2\alpha(\rho)} = 1 - \frac{2GM}{\rho}, \quad \rho > R\,. \tag{2.6.26}$$

Observations: Of course, when thinking about a star, we would expect that the point of highest pressure to lie in the centre, i.e at $\rho = 0$. It is clear from (2.6.24) that the pressure is indeed constant at $\rho = 0$, however it is not immediately obvious that it is in fact the maximum! To check this, consider a star which is such that $R \gg 2GM$ so that we can let

$$\Delta = \frac{2GM}{R} \ll 1\,. \tag{2.6.27}$$

With this, equation (2.6.24) becomes

$$p(\rho) = \tilde{\varrho}\left[\frac{\sqrt{1-\Delta} - \sqrt{1-\frac{\Delta\rho^2}{R^2}}}{\sqrt{1-\frac{\Delta\rho^2}{R^2}} - 3\sqrt{1-\Delta}}\right]. \tag{2.6.28}$$

By Taylor expanding each term in Δ, up to first order we now obtain, after simplifying

$$p(\rho) \approx \tilde{\varrho}\left[\frac{\frac{\Delta\rho^2}{R^2} - \Delta}{3\Delta - \frac{\Delta\rho^2}{R^2} - 4}\right] + \mathcal{O}(\Delta^2)\,. \tag{2.6.29}$$

The second derivative of the pressure can then be found to be

$$\frac{d^2p}{d\rho^2} = \frac{4\Delta\tilde{\varrho}}{R^2}(\Delta - 2)\left[\frac{3\Delta - 4 + \frac{3\Delta\rho^2}{R^2}}{(3\Delta - 4 - \frac{\Delta\rho^2}{4})}\right] + \mathcal{O}(\Delta^2)\,, \tag{2.6.30}$$

and thus, at the centre of the star we have

$$\left.\frac{d^2p}{d\rho^2}\right|_{\rho=0} = \frac{4\Delta\tilde{\varrho}(\Delta - 2)}{(3\Delta - 4)^2} + \mathcal{O}(\Delta^2)\,. \tag{2.6.31}$$

Since we assumed $\Delta \ll 1$, then clearly $\Delta - 2 < 0$, leading to the right hand side being negative. Hence, we can conclude that $\rho = 0$, gives the maximum pressure, and is given by

$$p(0) = \tilde{\varrho}\left[\frac{\sqrt{1-\frac{2GM}{R}} - 1}{1 - 3\sqrt{1-\frac{2GM}{R}}}\right]. \tag{2.6.32}$$

Another interesting property to notice is that

$$M \to M_{max} = \frac{4}{9G}R \quad \implies p(0) \to \infty\,. \tag{2.6.33}$$

In other words; if the mass of a star is bigger than or equal to M_{max}, the pressure in the centre of the star would become infinite, in which case, our assumption of a time independent solution will no longer hold: a star which shrinks to such a size will inevitably keep shrinking until a black hole is formed [3]. Notice also, that in the given form M_{max} appears to have dimensions $ML^{-2}T^2$, however recall that we set the speed of light $c = 1$. If we instead were to carry through the factor of c in our calculations we would then find

$$M_{\max} = c^2 \frac{4}{9G} R \, , \tag{2.6.34}$$

and hence, M_{max} does indeed have the dimensions of mass. Of course, in reality this is not always the case, for example our sun has radius of approximately 6.96×10^8m, meaning that according to our model

$$M_{max} \approx 1.38 \times 10^{27} kg \, , \tag{2.6.35}$$

however, the sun has mass $M_\odot \approx 2 \times 10^{30} kg$, which is clearly greater than M_{max}. So the natural question to ask now is: why does the sun not collapse? The answer being that stars are really just giant fusion engines; the energy created from fusing elements together in the core pushes outwards to balance the inward force of gravity. Undoubtedly, the nuclear fuel in a star will eventually run out and as a consequence, gravity will win over and the star will begin to cave. Depending on the size of a given star the collapse may be halted at a certain point, resulting in different outcomes.

For a star like our sun, eventually the electrons will be so tightly packed together that they will resist any further compression (Electron degeneracy pressure) resulting in a **white dwarf**.

A star with enough mass however, will reach what is known as the **Chandrasekhar limit** (about 1.4 $M_\odot$) [3] and in this case, even electron degeneracy pressure will not withstand the force of gravity. This will eventually lead to protons and electrons being smashed together to form neutrons. The result of such a collapse is known as a **neutron star** and are the most dense object currently known, they are so dense that just a single teaspoon of neutron star material would weigh 1 billion tons!

Finally, for a star which is so massive $(3 - 4M_\odot)$, the collapse will continue until the entire mass of the star is compacted to a single point, creating a space-time **singularity**. At this point, the gravitational field becomes so strong that not even light can escape its pull, resulting in the famous **black hole**!

2.7 Schwarzschild Black Holes

2.7.1 Eddington-Finkelstein coordinates

At the end of section 2.5.3, we mentioned the idea that any object whose radius is smaller than its corresponding Schwarzschild radius is automatically a black hole, i.e the Schwarzschild radius defines the radius of the **event horizon** of a non-rotating black hole! Now that we have explored how these objects are formed under gravitational collapse, we can begin to investigate the characteristics of the space-time generated by them.

Recall, the Schwarzchild solution is given by (note that we have again set $c = 1$)

$$ds^2 = -\Big(1 - \frac{2GM}{\rho}\Big) dt^2 + \Big(1 - \frac{2GM}{\rho}\Big)^{-1} d\rho^2 + \rho^2 d\Omega^2 \, . \tag{2.7.1}$$

Notice, at $\rho = R_s$ and $\rho = 0$, we appear to run into a problem: the $\rho\rho$ component of the metric becomes infinite, thus suggesting that these points define a singularity. To find out what is really happening here we can consider what happens to the curvature at each of these points. Of course, deciding whether or not the Riemann curvature tensor is infinite may in general be difficult, however as we have seen, it is possible to contract the curvature tensor to instead produce scalar quantities to measure the curvature; the simplest being the Ricci scalar (although other scalar quantities can be formed by different contractions). For the Schwarschild metric, it is possible to show that one such scalar quantity is given by [3]

$$R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} = \frac{48 G^2 M^2}{\rho^6} \, . \tag{2.7.2}$$

With this, we can now see that $\rho = 0$ is in fact a curvature singularity, however at $\rho = 2GM$ the curvature appears to be well defined! This leads us to say that $\rho = 2GM$ is really just a coordinate singularity, meaning that by implementing a suitable change in coordinates, we can eliminate this problem while maintaining the same overall geometry. To this end, we introduce a coordinate system known as **Eddington-Finkelstein** coordinates, and are defined by the transformation $(t, \rho, \theta, \phi) \to (v, \rho, \theta, \phi)$ where

$$v = t + \rho^* = t + \rho + 2GM \ln\left|\frac{\rho}{2GM} - 1\right|, \tag{2.7.3}$$

and $\rho^* = \rho + 2GM \ln\left|\frac{\rho}{2GM} - 1\right|$ is a **tortoise coordinate**. By rearranging this equation for t, we find dt in the new coordinates to be

$$dt = dv - \left(1 - \frac{2GM}{\rho}\right)^{-1} d\rho, \tag{2.7.4}$$

and hence, the Schwarzchild solution in the new coordinates is given by

$$ds^2 = -\left(1 - \frac{2GM}{\rho}\right)dv^2 + 2dvd\rho + \rho^2 d\Omega^2 \ . \tag{2.7.5}$$

Notice that $\rho = 2GM$ no longer presents an issue, yet we have maintained the curvature singularity at the origin!

2.7.2 Light Cones: Schwarzschild vs Eddington-Finkelstein

To further understand the geometry of what is happening here, we can look at what happens to future light cones as we approach the Schwarzschild radius and beyond! We will begin by analysing what happens in Schwarzschild coordinates. To this end, we consider **radial light-like geodesics** [3] (i.e curves for which θ and ϕ are constant and $ds^2 = 0$) so that we can write

$$ds^2 = 0 = -\left(1 - \frac{2GM}{\rho}\right)dt^2 + \left(1 - \frac{2GM}{\rho}\right)^{-1} d\rho^2 \ . \tag{2.7.6}$$

and hence, the slope of the light-like geodesics is given by

$$\frac{dt}{d\rho} = \pm\left(1 - \frac{2GM}{\rho}\right)^{-1} \ . \tag{2.7.7}$$

A light beam which is moving away from $\rho = 2GM$ will then follow the geodesic described by this equation with a positive right hand side and those moving towards the black hole will follow the geodesics described by the converse. These geodesics allow us to construct future light cones for an observer approaching a black hole from the point of view of another observer located far away.

Another property which can be seen from equation (2.7.6) is that, as $\rho \to 2GM$, the time component tends to zero. In other words, close to the schwarzchild radius, time appears to slow down, an effect known as **gravitational time-dilation**. It can be shown that the amount by which the flow of time changes is given by

$$\Delta t_0 = \Delta t_f \sqrt{1 - \frac{2GM}{\rho}} \tag{2.7.8}$$

where Δt_0 is the time interval between two events, as seen by an observer close to the spherical body, and Δt_f is the time interval between events, measured by an observer that is far away from the body.

From equation (2.7.7), it is clear that far away from $\rho = 2GM$, i.e as $\rho \to \infty$, the slope of the geodesics becomes ± 1 and the light cones will be just like those in Minkowski space. If we now consider approaching $\rho = 2GM$ from the positive side, it becomes apparent that the gradients approach $\pm\infty$ and the light beams appear to be asymptotic at $\rho = 2GM$. With this, we see that the light cones "close up" as they approach the Schwarzschild radius (see figure 8). To put this in to some context, if we observed someone falling in to a black hole, we would never actually see them cross $\rho = 2GM$: they would appear to us to just move slower and slower as they get closer to that point. In other words, if the observer falling into the black hole was sending signals back to us at evenly spaced intervals with respect to his proper time, we would see the time between these signals increase as he approaches $\rho = 2GM$. As a result of this, even the frequency of light waves will decrease and consequently cause an increase in the wave length and our friend would appear more and more red as they approach the black hole.

This phenomenon is know as **gravitational red-shift**, the mathematical details of which are beyond the scope of this project.

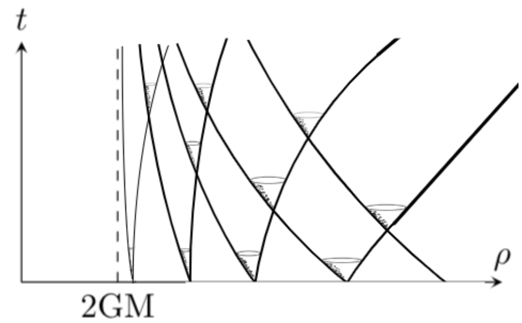

Figure 8: Future light cones approaching $\rho = R_s$ in Schwarzschild coordinates

In order to determine what is happening from the point of view of the observer falling into the black hole, we turn our attention to Eddington-Finkelstein coordinates. In this coordinate system, the radial light-like geodesics satisfy the condition

$$ds^2 = 0 = -\Big(1 - \frac{2GM}{\rho}\Big)dv^2 + 2dvd\rho = dv\Big[2d\rho - \Big(1 - \frac{2GM}{\rho}\Big)dv\Big] \ . \tag{2.7.9}$$

which is solved by solutions to the following equation [3]

$$\frac{dv}{d\rho} = \begin{cases} 0, & \text{(infalling)} \\ 2\Big(1 - \frac{2GM}{\rho}\Big)^{-1}, & \text{(outgoing)} \ . \end{cases} \tag{2.7.10}$$

Introducing a new time-like coordinate $\tilde{t} = v - \rho$ then gives the slope of the light-like geodesics in terms of coordinates $(\tilde{t}, \rho)$

$$\frac{d\tilde{t}}{d\rho} = \begin{cases} -1, & \text{(infalling)} \\ 2\Big(1 - \frac{2GM}{\rho}\Big)^{-1} - 1, & \text{(outgoing)} \ . \end{cases} \tag{2.7.11}$$

Similar analysis to that carried out on the Schwarzchild coordinates now yields the space-time diagram in Figure 9. From this picture we can deduce that from the point of view of an observer falling into a black hole, the ingoing light rays always appear to move at a constant velocity just like in special relativity, whereas the outgoing rays seem to move slower and slower in the radial direction as they approach $\rho = 2GM$ until eventually, outgoing light rays emitted from $\rho = 2GM$ appear to orbit the black hole! i.e They appear stationary in the radial direction.

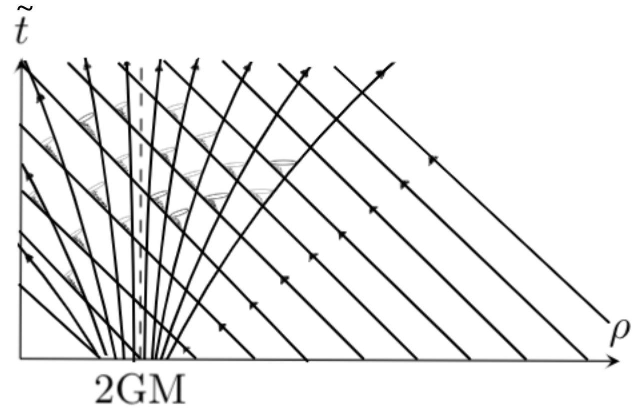

Figure 9: Future light cones approaching a black hole in Eddington-Finkelstein coordinates

Once $\rho < 2GM$, even the outgoing light rays fall towards the singularity and the future light cones of an observer point purely towards the centre of the black hole. In other words, events which happen past $\rho = 2GM$ cannot observed from outside of this point. Finally, we can say that the surface described by $\rho = 2GM$ defines the **event horizon** of a black hole! Evidently, in reality someone falling into a black hole would be torn apart by the ever growing tidal forces long before they reach the event horizon, a process known as **spaghettification**!

Figure 10: Image of a black hole in the centre of the galaxy M87 [21]

In 2019, astronomers used the Event Horizon radio telescope (EHT) to take a picture of a black hole located at the centre of the galaxy M87, some 55 million lightyears from earth! Of course, by the very nature of a black hole, we cannot actually observe the object its self, rather, as matter falls towards the event horizon, the combination of extreme gravitational and frictional forces cause the matter to heat up and emit electromagnetic radiation which, when detected by the correct type of telescope, clearly shows the silhouette of the event horizon!

Of course, the type of black hole outlined in this project describes only one special case of one that is both stationary and not charged. Other solutions include the Kerr-Newmann solution which describes a gravitating body that both rotating and electrically charged, however as one would expect, the introduction of these extra conditions leads to a lot more complications than that of the Schwarschild solution.

3 Final Remarks

The theory of general relativity has remained the best theory of gravity since it was first conceived 105 years ago, and although we have only discussed the most famous consequence of theory, black-holes, there are many other profound discoveries that have been made, thanks to Einstein's theory, such as the existence of gravitational waves (see [3] for further insight into this subject). General relativity is also to thank for many things we take for granted today, such as the global positioning system (GPS). Since GPS satellites orbit high above the earth, time as they experience it, moves faster than it does here on the surface of the earth and therefore needs to be regularly corrected. Therefore, without the knowledge of gravitational time-dilation, GPS would simply fail to work.

One of the biggest remaining problems in modern physics is the disagreement between general relativity, and quantum mechanics. For example, as we have seen, a black hole is formed when the gravity of a massive star wins over the nuclear forces, causing the star to shrink in to a singularity, something which according to quantum mechanics cannot exist.

To conclude this project, i have attempted to generalise the Einstein field equations to include the effects of spin and torsion, however, this has proved to be a difficult undertaking and is something i would like to further investigate in the future. What follows is a summary of my attempt.

3.1 Einstein-Cartan Theory

General relativity as we have formulated in this project, relies heavily on the assumptions that the connection be both metric compatible and torsion free. However, it is possible to formulate the field equations in such a way that they include the effects of torsion. We begin by defining a new connection, which follows from the metric compatibility condition, given by

[10]

$$\Gamma^{\alpha}_{\mu\nu} = \{^{\alpha}_{\mu\nu}\} - g^{\alpha\beta}K_{\mu\nu\beta} \ , \tag{3.1.1}$$

where the quantity $\{^{\alpha}_{\mu\nu}\}$, denotes the symmetric Levi-Civita connection, and

$$K_{\mu\nu\beta} = T_{\nu\beta\mu} + T_{\mu\beta\nu} - T_{\mu\nu\beta} \ , \tag{3.1.2}$$

defines the **contorsion tensor** [13], written in terms of the fully covariant torsion tensor. Notice, when the the torsion vanishes, then so does the contorsion tensor, and we would once again obtain the Levi-Civita connection. Furthermore, a pseudo-Riemannian manifold, **M**, endowed with a connection of the form (3.1.1) is called an **Riemann-Cartan manifold** [8].

Using this connection, one can then set up an action, akin to the Einstein-Hilbert action used to derive the field equations of general relativity. Let

$$\mathcal{L}_G = \frac{1}{\kappa}R\sqrt{-g} = \frac{1}{\kappa}g^{\beta\sigma}R_{\beta\sigma}\sqrt{-g} \tag{3.1.3}$$

be a Lagrangian density of the gravitational field, and let $\mathcal{L}_m$ be the Lagrangian density of matter. We then define an action S, given by [10]

$$S = \int_{\mathbf{M}} \frac{1}{2\kappa}g^{\beta\sigma}R_{\beta\sigma}\sqrt{-g} + \mathcal{L}_m\sqrt{-g}\ d^4x \ . \tag{3.1.4}$$

It should be noted that due to the non-vanishing torsion, the Ricci tensor is no longer symmetric and the variation of $R_{\beta\sigma}$ will now also include torsion term. Following equation (2.4.28), it can be seen that in this case

$$\nabla_\alpha(\delta\Gamma^{\alpha}_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^{\alpha}_{\beta\alpha}) = \delta R_{\beta\sigma} - T^{\rho}_{\alpha\sigma}\delta\Gamma^{\alpha}_{\beta\rho} \ , \tag{3.1.5}$$

and hence,

$$\delta R_{\beta\sigma} = \nabla_\alpha(\delta\Gamma^{\alpha}_{\beta\sigma}) - \nabla_\sigma(\delta\Gamma^{\alpha}_{\beta\alpha}) + T^{\rho}_{\alpha\sigma}\delta\Gamma^{\alpha}_{\beta\rho} \ . \tag{3.1.6}$$

Thus, after a little bit of work, it can be shown that varying the action, S, gives

$$\delta S = \int_{\mathbf{M}} \sqrt{-g}\Big[\frac{1}{\sqrt{-g}}\delta(\sqrt{-g}\mathcal{L}_m) + \frac{1}{2\kappa}\Big(R_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}R\Big)\delta g^{\beta\sigma} + \frac{1}{2\kappa}g^{\beta\sigma}T^{\rho}_{\alpha\sigma}\delta\Gamma^{\alpha}_{\beta\rho}\Big]\ d^4x \ . \tag{3.1.7}$$

From this point, i have been unsuccessful in completing the calculation. However, it can be shown that by extremizing this equation with respect to the inverse metric tensor, and torsion tensor, one obtains the **Einstein-Cartan** equations (sometimes called the Einstein-Cartan-Sciama-Kibble equations) [13]

$$R_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}R = \kappa\Sigma_{\beta\sigma} \ , \tag{3.1.8}$$

$$T^{\rho}_{\beta\sigma} + \delta^{\rho}_{\beta}T^{\lambda}_{\sigma\lambda} - \delta^{\rho}_{\sigma}T^{\lambda}_{\beta\lambda} = \kappa S^{\rho}_{\beta\sigma} \ , \tag{3.1.9}$$

where $\Sigma_{\beta\sigma}$, denotes the **canonical energy-momentum tensor** and $S^{\rho}_{\beta\sigma}$, denotes the **spin-angular momentum tensor**. These tensors are related to each other through the following equation:

$$\Sigma^{\beta\sigma} = T^{\beta\sigma} + (\nabla_\rho + 2T^{\lambda}_{\rho\lambda})[T^{\beta\sigma\rho} - T^{\sigma\rho\beta} - T^{\rho\beta\sigma}] \ , \tag{3.1.10}$$

where $T^{\beta\sigma}$, is the symmetric energy momentum tensor given by equation (2.4.45). Finally, it should be noted that at the macroscopic level, i.e if we let $S^{\rho}_{\beta\sigma} \to 0$, then $\Sigma_{\beta\sigma} \to T_{\beta\sigma}$, we once again recover Einstein's equation

$$R_{\beta\sigma} - \frac{1}{2}g_{\beta\sigma}R = \kappa T_{\beta\sigma} \ . \tag{3.1.11}$$

References

[1] S. Balbus. (2017). Oxford Physics Department Notes on General Relativity.

[2] M.Blau. Lecture Notes on General Relativity

[3] S.Carroll. (2019). Spacetime and Geometry: An Introduction to General Relativity.

[4] S.Carroll. Lecture Notes on General Relativity

[5] S.Chakraborty. Boundary Terms of the Einstein-Hilbert Action

[6] A. Einstein 1905. On the Electrodynamics of Moving Bodies

[7] F. Fitzmaurice. Differential Geometry

[8] I.Gordeeva. V.Pan'zhenskii. Stepanov, Sergey. (2010). Riemann–Cartan manifolds. Journal of Mathematical Sciences.

[9] Howard E. Haber. PHYSICS 171 Available at: http://scipp.ucsc.edu/ haber/ph171/connection15v2.pdf

[10] P.J.A.Harrison. Einstein-Cartan Theory and Its Formulation As A Quantaum Field Theory

[11] L.P.Hughston and K.P.Tod. (1992). An introduction to general relativity.

[12] C.Kaeonikhom. Variational Approach to General Relativity

[13] L. N. KATKAR. Einstein-Cartan Theory of Gravitation: Kinematical Parameters and Maxwell Equations

[14] D.Martila. Do Stokes' and Gauss's theorems hold in curved spacetime?

[15] P.Peterson. Riemannian Geometry: Springer.

[16] L.Tu, 2010. An Introduction To Manifolds. New York: Springer.

[17] R.M.Wald. General Relativity

[18] UCDAVIS 2021. Topological Spaces. Available at: https://www.math.ucdavis.edu/ hunter/book/ch4.pdf

[19] University Of Toronto. Differential Geometry Notes

[20] https://en.wikipedia.org/wiki/Gauss

[21] https://www.nasa.gov/mission_pages/chandra/news/black-hole-image-makes-history