A Not So Gentle Introduction to General Relativity

Benjamin Church

December 3, 2022

Contents

1 Introduction

1.1 The Equivalence Principle

In 1907, Albert Einstein had the "happiest thought of his life",

> The gravitational field has only a relative existence... Because for an observer freely falling from the roof of a house - at least in his immediate surroundings - there exists no gravitational field.

The equivalence principle refers to the idea that freely falling in a gravitational field is locally indistinguishable from being out in free space away from any sources of gravitation. Einstein's motivation came from the fact that in Newtonian mechanics the force of gravity of an object is exactly proportional to that objects inertial mass. In Newtonian physics, there seems to be no reason that gravitational mass and inertial mass would have to be exactly equal yet every experiment shows them to be identical. This proportionality has the consequence that all objects accelerate in a gravitational field at the same mass regardless of their mass. In Newtonian physics, this is the reason for weightlessness. It is not that objects in free fall do not experience any forces but rather that since all objects are being accelerated equally, the gravitational field does not cause objects to acquire relative velocities which would be measurable. However, this explanation is somewhat unsatisfying because it does not explain why such an odd scenario should occur. Einstein realized that the equivalence between free fall and no gravitational field at all is extremely fundamental. The fact that all objects behaved the same way in a gravitational field regardless of their masses suggests that gravity is not a force at all is the usual sense but an effect of geometry. The geometry of space itself is what constitutes the gravitational field and what is responsible for altering the trajectories of moving objects. Gravitational accelerations are not dependent on the mass not due to some strange cancellation but rather because it is the geometrical path through space itself which is being affected. Properties of the object in question are not relevant for all objects are simply following the natural trajectories through a curved spacetime. But what does it mean for space to be curved and how does the curvature affect trajectories of objects through it?

1.2 What is Curvature?

Fundamentally, curvature comes down to the idea that distances may be measured differently at different places. Imagine a bug living on a hotplate. Suppose that the bug has a metal ruler which it uses to measure distances and plot courses. However, unbeknownst to the bug, the ruler expands when heated and contracts when cooled. Furthermore, the hotplate is somewhat faulty so that the plate gets hotter as you move towards the center. If the bug wanted to find the shortest path between two points, it would discover that this path curved towards the hotter parts of the plate because fewer ruler lengths would be necessary to traverse through regions in which the ruler is expanded. Furthermore, if the bug measured the distance around the circumference of some circle and the radius of the circle, it might not get a ratio of π if the inner parts of the circle are at a different temperature than the rest. These are all the hallmarks of a curved space. The standard notion of curvature is to take surface and put it inside some larger ambient space such that one cannot make a plane tangent to the surface lay flat. This is certainly true but we prefer a more intrinsic notion of curvature, one which can be discussed from the viewpoint of bugs

living inside said surface. Curvature will manifest itself in strange geometrical ways such as having parallel lines diverge or converge, equal segments connected at right angles not meeting up, and circles having a ratio between their circumference and diameter different from π. Intrinsically, curvature is about having the measurement of distances be different from place to place. We need such an intrinsic definition because we ourselves are bugs constrained to live within spacetime. We cannot see the the geometry of our universe as a whole because we are stuck within it but we can measure it.

2 Tensors

2.1 The Principle of General Covariance

Once we start working with complicated curved spaces, preferred sets of coordinates such as Cartesian coordinates will no longer be possible. Most of the definitions we will make will be motivated by the principle of general covariance, that the laws of physics should look the same in all reference frames and in all coordinate systems. This principle is extremely powerful and is a guiding principle in cooking up new theories in all branches of modern physics. In essence, the principle of general covariance says that while some coordinate systems may be more convenient for a given problem, all coordinate systems are equally good from the point of view of the fundamental physical laws. The equations should take on the same form no matter what coordinates we use to compute them. There are no privileged frames of reference in physics.

2.2 Basis Vectors

For the time being we will work in a vector space equipped with a metric (or pseudo-metric). It will suffice to study R^n in this section because every real vectorspace is isomorphic to $\mathbb{R}^n$ where n is the dimension of the vectorspace. Although we will use the fact that our space is a vectorspace frequently, we may want to choose coordinates which break the additivity of the space (for example polar coordinates). With the principal of general covariance in mind, we want this discussion to be general to any set of coordinates.

Given a set of coordinates which we will denote as $(x^1, \cdots, x^n)$ where n is the dimension of the space, we can define a set of basis vectors at each point. We can think of these as vectors in the ambient $\mathbb{R}^n$ euclidean space or as vectors in the tangent space which is a copy of $\mathbb{R}^n$ glued to each point representing the fact that a tangent vector is a local notion. These subtleties will not concern us here. We define the covariant basis vectors by,

$$\vec{e}_i(x^1, \cdots, x^n) = \frac{\partial \vec{R}(x^1, \cdots, x^n)}{\partial x^i}$$

which represent the local "directions" of each coordinate. The term "covariant" will be discussed later. For the coordinate system to be nondegenerate, we require that at every point, the basis vectors from a basis of the ambient space $\mathbb{R}^n$. In general, we define the tangent space at a point to be the span (set of all linear combinations) of the covariant basis vectors.

Suppose we have two different coordinate systems, or frames, which parameterize our space, $(x^1, \cdots, x^n)$ and $(x'^1, \cdots, x'^n)$. We want to relate the covariant basis in the primed frame to that of the unprimed frame. Using the chain rule,

$$\vec{e}_i{}' = \frac{\partial \vec{R}}{\partial x'^i} = \frac{\partial \vec{R}}{\partial x^j}\frac{\partial x^j}{\partial x'^i} = \Lambda^j_{i'}\vec{e}_j$$

4

where,

$$\Lambda^j_{i'} = \frac{\partial x^j}{\partial x'^i}$$

is the lower transformation symbol and I have introduced Einstein summation conventions where repeated indices are summed over. The transformation symbol is computed by viewing the coordinates on the top as a function of the coordinates on the bottom. This is possible because both sets of coordinates can determine any point in the space. The prime on the bottom denotes the fact that the primed coordinates are on the bottom of the derivative. That is the upper transformation symbol is,

$$\Lambda^{j'}_i = \frac{\partial x'^j}{\partial x^i}$$

In general, we will call any set of objects that transforms by contracting with the lower transformation symbol covariant. For example, consider a function $F : \mathbb{R}^n \to \mathbb{R}$. We are going to consider the derivatives of this function with respect to the coordinate systems $(x_1, \cdots, x_n)$ and $(x'_1, \cdots, x'_n x)$ where we assume nothing about the relation between these coordinate systems except that each coordinate is a smooth (infinitely differentiable) function of the other. Using the chain rule,

$$\frac{\partial F}{\partial x'_i} = \frac{\partial F}{\partial x_j}\frac{\partial x_j}{\partial x'_i}$$

This sort of a transformation is know as covariant because the quantity $\frac{\partial F}{\partial x_i}$ transforms in the same way as the (covariant) basis. The opposite notion is contravariant (varying against the basis). For example, consider a curve $\gamma^i(\lambda)$ parameterized by a coordinate λ. The curve is set of points $(\gamma^1(\lambda), \cdots, \gamma^n(\lambda))$ parameterized by some variable t. Consider the tangent vector,

$$t^i = \frac{\mathrm{d}\gamma^i}{\mathrm{d}\lambda}$$

Technically, what I have defied here is just a bunch of coordinates not a vector. The real tangent vector to the curve is the vector,

$$\vec{t} = \frac{\mathrm{d}\vec{R}(\gamma^1(\lambda), \cdots, \gamma^i(\lambda))}{\mathrm{d}\lambda} = \frac{\partial \vec{R}}{\partial x^i}\frac{\mathrm{d}\gamma^i}{\mathrm{d}\lambda} = \vec{e}_i t^i$$

which lies in the tangent space. Therefore, the object t^i are the components of the true tangent vector. However, we often call t^i a vector in its own right because of how it transforms under the change of coordinates. Consider,

$$t'^i = \frac{\mathrm{d}\gamma'^i}{\mathrm{d}\lambda} = \frac{\mathrm{d}x'^i(\gamma^1(\lambda), \cdots, \gamma^n(\lambda)}{\mathrm{d}\lambda} = \frac{\partial x'^i}{\partial x^j}\frac{\mathrm{d}\gamma^j}{\mathrm{d}\lambda} = \Lambda^{i'}_j t^j$$

Therefore, t^i transforms using the upper transformation symbol so we call t^i contravariant. An object with a single contravariant (upper) index is often called a vector while one with a single covariant (lower) index is called a covector.

2.3   The Tensor Property

Definition 2.1. A type (or rank) (p, q) tensor T is an indexed collection of numbers $T^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_q}$ which transforms under change of coordinates by the rule known as the tensor property,

$$T'^{i'_1 i'_2 \cdots i'_p}_{j'_1 j'_2 \cdots j'_q} = \Lambda^{i'_1}_{i_1}\Lambda^{i'_2}_{i_2}\cdots\Lambda^{i'_p}_{i_p}\Lambda^{j_1}_{j'_1}\Lambda^{j_2}_{j'_2}\cdots\Lambda^{j_q}_{j'_q}T^{i_1 i_2 \cdots i_p}_{j_1 j_2 \cdots j_q}$$

We call p the contravariant type and q the covariant type. A type $(p, 0)$ tensor is called contravariant and a type $(0, q)$ tensor is called covariant. In particular, a type $(1, 0)$ tensor is called a vector, a type $(0, 1)$ tensor is called a covector, and a type $(0, 0)$ tensor is called a scalar or invariant because it does not transform at all. A tensor of any other type is called "mixed".

Although this definition looks fairly horrible, the idea is very simple. All it says is that T is an indexed set of numbers such that each upper index transforms with an upper transformation symbol and each lower index transforms with a lower transformation symbol. The reason that tensors are important is that if we can form our fundamental physical laws as equations of tensors then they will automatically hold in every reference frame. To see why this is true, suppose that we have the tensor equation,

$$T^{i_1 \cdots i_p}_{j_1 \cdots j_q} = S^{i_1 \cdots i_p}_{j_1 \cdots j_q}$$

what this means is that the tensors agree component-wise. For this to make sense, both sides of a tensor equation must have the same indices such that we can identify the components that have the same indices. Now suppose we evaluate the tensor T in a new coordinate system,

$$T'^{i'_1 \cdots i'_p}_{j'_1 \cdots j'_q} = \Lambda^{i'_1}_{i_1} \cdots \Lambda^{i'_p}_{i_p} \Lambda^{j_1}_{j'_1} \cdots \Lambda^{j_q}_{j'_q} T^{i_1 \cdots , i_p}_{j_1 \cdots j_q} = \Lambda^{i'_1}_{i_1} \cdots \Lambda^{i'_p}_{i_p} \Lambda^{j_1}_{j'_1} \cdots \Lambda^{j_q}_{j'_q} S^{i_1 \cdots i_p}_{j_1 \cdots j_q} = S'^{i'_1 \cdots i'_p}_{j'_1 \cdots j'_q}$$

so the tensor equation also holds in the new coordinate system. The general fact that if a tensor equation holds in one coordinate system then it must hold in all coordinate systems is, at its core, a consequence of the fact that all tensors, by definition, transform the same way. Tensors provide a perfect language to write physical laws which satisfy the principle of general covariance.

2.4 Tensor Products and Contractions

Given two tensors we can form the tensor product simply by multiplying the components. This gives us a new tensor of a higher type. Explicitly, given a type (p_1, q_1) tensor T and a type (p_2, q_2) tensor S we can form the tensor product $T \otimes S$ or just TS which is a type $(p_1 + p_2, q_1 + q_2)$ tensor with components,

$$(T \otimes S)^{i_1 \cdots i_{p_1} i_{p_1 + 1} \cdots i_{p_1 + p_2}}_{j_1 \cdots j_{q_1} j_{q_1 + 1} \cdots j_{q_1 + q_2}} = T^{i_1 \cdots i_{p_1}}_{j_1 \cdots j_{q_1}} \cdot S^{i_{p_1 + 1} \cdots i_{p_1 + p_2}}_{j_{q_1 + 1} \cdots j_{q_1 + q_2}}$$

This is clearly a $(p_1 + p_2, q_1 + q_2)$ tensor because each index gets an appropriate transformation symbol by virtue of the fact that T and S are tensors. Because it is just multiplication component-wise, the tensor product is symmetric if both the indices and the tensors are swapped.

Now we know how to make tensors of higher type but we can also make tensors with lowered type by index contraction. To contract an upper and lower index, we take the trace over those indices. For example, given the type $(2, 1)$ tensor, T^{ij}_k we can form a type $(1, 0)$ tensor S^i by taking,

$$S^i = T^{ij}_j = \sum_{j=1}^{n} T^{ij}_j$$

remembering the Einstein summation convention. I need to show that this construction satisfies the tensor property. The Kronecker delta δ^i_j is defined to be the components of the identity matrix,

$$\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

By the definition of partial derivatives,

$$\Lambda^{i'}_k \Lambda^k_{j'} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^k}{\partial x'^i} = \frac{\partial x'^i}{\partial x'^j} = \delta^i_j$$

and similarly,

$$\Lambda^i_{k'} \Lambda^{k'}_j = \frac{\partial x^i}{\partial x'^k} \frac{\partial x'^k}{\partial x^i} = \frac{\partial x^i}{\partial x^j} = \delta^i_j$$

Therefore the upper and lower transformation symbols are inverses. Now, consider the transformation of S,

$$S'^{i'} = T'^{i'j'}_{j'} = \Lambda^{i'}_i \Lambda^{j'}_{j_1} \Lambda^{j_2}_{j'} T^{ij_1}_{j_2} = \Lambda^{i'}_i \delta^{j_1}_{j_2} T^{ij_1}_{j_2} = \Lambda^{i'}_i T^{ij_1}_{j_1} = \Lambda^{i'}_i S^i$$

Therefore, S^1 is a type $(1,0)$ tensor. This construction, contracting over an upper and a lower index, works in general and transforms a type (p,q) tensor into a type $(p-1, q-1)$ tensor.

A very important special case of contraction is forming invariants form tensors. For example, given a type $(1,1)$ tensor T we can contract to form a type $(0,0)$ tensor which is by definition an invariant. Explicitly, let $S = T^i_i$ then S transforms as,

$$S' = T'^{i'}_{i'} = \Lambda^{i'}_j \Lambda^k_i T^j_k = \delta^k_j T^j_k = T^j_j = S$$

Thus, S is an invariant, it has the same value in all coordinate systems. Invariants are extremely important in physics because they have the wonderful property that you only ever need to compute them once any you can choose whatever coordinate system is most convenient in which to do this computation. There will be many applications of forming invariants out of tensors. One very important invariant is the inner product of a type $(1,0)$ tensor A and a type $(0,1)$ tensor B given by $A \cdot B = A^i B_i$ which is the contraction of the tensor product. Of course, we can also define it the other way for a type $(0,1)$ tensor A and a type $(1,0)$ tensor B in which case $A \cdot B = A_i B^i$ again the contraction of the tensor product. Once we have the metric tensor to deal with, we will be able to take the inner product of two vectors or two covectors or really any two tensors whatsoever.

2.5 The Intrinsic Notion of a Tensor

Remark 2.5.1. I recommend skipping this section unless you have a strong background in (multi)linear algebra. The concepts expressed here are very interesting but are not necessary to the remainder of the course.

We have already seen one example of an invariant inner product, the tangent vector as an element of the tangent space. Since we saw that $\vec{e}_i$ forms a (vector-valued) rank-1 covariant tensor or covector and t^i a (real number valued) rank-1 contravariant tensor or vector. Therefore, we can take the inner product to form an invariant,

$$\vec{t} = \vec{e}_i\, t^i$$

7

which is the geometric tangent vector in the tangent space we defined earlier. This is an invariant geometric vector which means it is a geometric object which is invariant under transformations of coordinates used to define it. This geometric vector lives in the ambient vector space so it is a fundamentally different beast than the (not invariant since it is a tensor of non-trivial type) vector t^i of components. In fact, all tensors have some sort of coordinate-independent abstract geometric realization. This geometric object T is represented by the coordinate dependent indexed set of numbers $T^{i_1 \cdots i_p}_{j_1 \cdots j_q}$. The fact that T is an invariant object and has a coordinate independent meaning is captured by the specific transformation law satisfied by its component representation.

For the mathematically inclined, we can give a purely geometric definition of tensors. However, because these objects are abstract, the only way to do actual computations with them is using a coordinate representation. That said, the invariant coordinate independent notion is important conceptually for thinking as the entire tensor as a geometric "thing". Given a manifold M^1, we can define the tangent space at a point p denoted by T_pM as the span of the covariant basis. This is a finite-dimensional vectorspace. Then, the space of all type (k, ℓ) tensors at the point p is given by,

$$(T^k_\ell)_p = (T_pM)^{\otimes k} \otimes (T^*_pM)^{\otimes \ell}$$

where $T^*_pM = (T_pM)^*$ is the cotangent space which is defined to be the dual space of the tangent space. However, this is the space of all tensors at a given point. What we really care about is the bundle of all (k, l)-tensors everywhere on our space,

$$T^k_l = \coprod_{p \in M} (T_pM)^{\otimes k} \otimes (T^*_pM)^{\otimes \ell}$$

An abstract tensor can be defined by its components. Let e_i be the covariant basis at a point p and let e^i be the dual (contravariant) basis also at p. Then given a coordinate representation of a (k, ℓ) tensor we can define a type (k, ℓ) tensor at the point p by,

$$T = T^{i_1 \cdots i_k}_{j_1 \cdots j_\ell} \cdot e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes e^{j_1} \otimes \cdots \otimes e^{j_\ell} \in (T^k_\ell)_p$$

using Einstein summation convention. Because this is a total contraction of tensors, the object T is invariant. Up until now, we have only talked about tensors defined at a given point. However, what we really care about is tensor fields which are the association of a tensor to each point in space. This is where the bundle of all tensors will come in handy. We define a type (k, ℓ) tensor field σ as a section of the bundle of type (k, ℓ) tensors. That is, a map $\sigma : M \to T^k_\ell$ such that each point is sent to a tensor at that point, $\sigma(p) \in (T^k_\ell)$ for each $p \in M$.

2.6 The Metric Tensor

Because we have a dot product in our ambient space (on geometric objects) we can define the metric (rank 2 covariant) tensor given by dot products of the covariant basis,

$$g_{ij} = \vec{e}_i \cdot \vec{e}_j$$

Because the dot product is commutative, the metric tensor is symmetric, $g_{ij} = g_{ji}$. The metric tensor is the object we use to measure the lengths of vectors and the distances

---

[1]We will come back to what this means in the next chapter. For now, think of it simply as some space.

between points. To see this, consider a small geometric displacement vector,

$$\mathrm{d}\vec{R} = \frac{\partial \vec{R}}{\partial x^i}\mathrm{d}x^i = \vec{e}_i\mathrm{d}x^i$$

The squared length of this vector is given by,

$$\mathrm{d}\vec{R}^2 = (\vec{e}_i\mathrm{d}x^i)\cdot(\vec{e}_j\mathrm{d}x^j) = \vec{e}_i\cdot\vec{e}_j\mathrm{d}x^i\mathrm{d}x^j = g_{ij}\mathrm{d}x^i\mathrm{d}x^j$$

Therefore, the metric tells us how to calculate the length of small displacements given their coordinates. Now we can define the contravariant metric tensor which is a type $(2,0)$ tensor, g^{ij} defined to be the inverse of g_{ij}. That is,

$$g^{ij}g_{jk} = \delta^j_k \quad \text{and} \quad g_{ij}g^{jk} = \delta^i_k$$

We can use the covariant and contravariant versions of the metric to perform "index juggling" which consisting of moving indices up and down. Given a vector A^i we can construct a covector $A_i = g_{ij}A^j$. Similarly, given a covector B_i we can construct a vector $B^i = g^{ij}B_j$. Similarly, we can raise and lower any index of a mixed tensor using g_{ij} and g^{ij}. Because these tensors are inverses, raising and lowering indices are inverse operations.

Using this definition, we can define an invariant from any vector or covector. Given a vector A^i define the invariant,

$$A \cdot A = A^i A_i = g_{ij}A^i A^j$$

Likewise, given a covector B_i define,

$$B \cdot B = B_i B^i = g^{ij}B_i B_j$$

Furthermore, define the contravariant basis,

$$\vec{e}^{\,i} = g^{ij}\vec{e}_j$$

which transforms as a contravariant tensor of geometric vectors.

Theorem 2.1. *Properties of the Metric,*

1. *Symmetric,* $g_{ij} = g_{ji}$

2. *The contravariant basis contracts to form the contravariant metric,*

$$\vec{e}^{\,i} \cdot \vec{e}^{\,j} = g^{ik}\vec{e}_k \cdot g^{j\ell}\vec{e}_\ell = g^{ik}g^{j\ell}g_{k\ell} = g^{ik}\delta^j_k = g^{ij}$$

3. *The covariant and contravariant bases are orthonormal or dual,*

$$\vec{e}_i \cdot \vec{e}^{\,j} = g^{jk}\vec{e}_i \cdot \vec{e}_k = g^{jk}g_{ik} = \delta^j_i$$

   and similarly,

$$\vec{e}^{\,i} \cdot \vec{e}_j = g^{ik}\vec{e}_k \cdot \vec{e}_j = g^{ik}g_{kj} = \delta^i_j$$

2.7 Determinants

In general relativity, we care almost exclusively about four-dimensional spaces. Therefore, we will discuss the theory of determinants explicitly in the case $n = 4$ for rank two tensors. However, these results are all easily generalized to arbitrary dimensions. We first define the fully skew symmetric symbol, S_{ijkl} by the properties,

1. $S_{0000} = 1$

2. Swapping any two indices changes sign $S_{ijkl} = -S_{jikl} = S_{jilk}$ etc.

3. If any index is repeated then the symbol is zero $S_{iikl} = S_{ijkk} = 0$.

Definition 2.2. For the dimension $n = 4$ we can define the determinant A of a type $(0,2)$ tensor a_{ij} by,

$$S_{ijkl} a_{ai} a_{bj} a_{ck} a_{dl} = A S_{abcd}$$

where repeated indices are summed.

This is a valid definition because the left hand side is fully skew symmetric since, by,

$$S_{ijkl} a_{bi} a_{aj} a_{ck} a_{dl} = A S_{abcd} = -S_{jikl} a_{bi} a_{aj} a_{ck} a_{dl} = -S_{jikl} a_{aj} a_{bi} a_{ck} a_{dl}$$

which is equal by renaming. One can explicitly check that,

$$S_{ijkl} S_{ijkl} = 4!$$

Therefore, the determinant of a with lower indices is given by the formula,

$$\det a_{..} = A = \frac{1}{4!} S_{abcd} S_{ijkl} a_{bi} a_{aj} a_{ck} a_{dl}$$

Theorem 2.2. $\det\left(a_{.i} b^{i\cdot}\right) = \det\left(a_{..}\right) \cdot \det\left(b^{..}\right)$

Proof. (DO THIS PROOF) □

Using the same idea as the definition of the determinant, we can define a new tensor of cofactors which will give us a tensor similar to the inverse.

Definition 2.3. The cofactor tensor of $a_{..}$ is a type $(2,0)$ tensor $A^{ia} = \frac{1}{3!} S_{abcd} S_{ijkl} a_{bj} a_{ck} a_{dl}$.

The cofactor tensor is almost the inverse. It differs only by a numerical factor which happens to be exactly the determinant. Explicitly,

Theorem 2.3. $A^{im} a_{mj} = A \delta^i_j$ so if the inverse exists, $A^{ij} = A(a^{-1})^{ij}$

Proof. $A^{im} a_{mj} = \frac{1}{3!} S_{mbcd} S_{iklr} a_{mj} a_{bk} a_{cl} a_{dr} = \frac{1}{3!} S_{jklr} S_{iklr} = \delta^i_j$ □

Theorem 2.4.

$$\frac{\partial A}{\partial a_{ij}} = A^{ji} = A(a^{-1})^{ji}$$

Proof.

$$\frac{\partial}{\partial a_{ij}} \left[\frac{1}{4!} S_{abcd} S_{rskl} a_{ar} a_{bs} a_{ck} a_{dl} \right] = \frac{4}{4!} \delta^a_i \delta^r_j S_{abcd} S_{rskl} a_{bs} a_{ck} a_{dl} = \frac{1}{3!} S_{ibcd} S_{jskl} a_{bs} a_{ck} a_{dl} = A^{ji}$$

□

10

2.8 The General Transformation Symbol

Definition 2.4. $\mathfrak{T}$ is a type (p, q) relative tensor of weight W if,

$$\mathfrak{T}'^{i'_1 \cdots i'_q}_{j'_1 \cdots j'_q} = \Lambda^W \Lambda^{i'_1}_{i_1} \cdots \Lambda^{i'_p}_{i_p} \cdots \Lambda^{j_1}_{j'_1} \cdots \Lambda^{j_q}_{j'_q} \mathfrak{T}^{i_1 \cdots i_q}_{j_1 \cdots j_q}$$

where $\Lambda^W = \det (\Lambda'_\cdot)^W$. Equivalently,

$$\mathfrak{T}'^{i'_1 \cdots i'_q}_{j'_1 \cdots j'_q} \Lambda^{i_1}_{i'_1} \cdots , \Lambda^{i_p}_{i'_p} \cdots \Lambda^{j'_1}_{j_1} \cdots \Lambda^{j'_q}_{j_q} = \Lambda^W \mathfrak{T}^{i_1 \cdots i_p}_{j_1 \cdots j_q}$$

If we try to transform the transformation symbol, we find that it is a relative tensor of weight 1 because,

$$S'_{ijkl} \Lambda^i_a \Lambda^j_b \Lambda^k_c \Lambda^l_d = S_{ijkl} \Lambda^i_a \Lambda^j_b \Lambda^k_c \Lambda^l_d = \Lambda S_{abcd}$$

Let $g = \det (g_{\cdot\cdot})$. Using the fact that, $g^{ik} g_{kj} = \delta^i_j$ then,

$$\det (g^{\cdot\cdot}) \cdot \det (g_{\cdot\cdot}) = 1$$

Therefore, $\det (g^{\cdot\cdot}) = g^{-1}$.

2.9 The Levi-Civita Tensor

3 The Basic Theory of Manifolds

3.1 Definition

3.1.1 Euclidean and Minkowski Spaces

3.2 From Space to Space-Time

3.3 The Covariant Derivative

To discuss curvature, we will need to take the derivatives of tensors. However, here we reach an issue. In general, the derivative of a tensor will not be a tensor. To see the difficulty, consider,

$$\partial'_{i'} A'^{j'} = \Lambda^i_{i'} \partial_i (\Lambda^j_{j'} A^j) = \Lambda^i_{i'} \Lambda^j_{j'} \partial_i A^j + \Lambda^i_{i'} A^j \partial_i \Lambda^j_{j'}$$

which is not zero because the transformation symbol may vary from place to place. This extra term means that the derivative of A is not a tensor. We will need to fix this problem if we hope to apply our theory to differential equations of tensors.

Geometrically, the issue is one of connection. The manifold has a tangent space attached at each point but we don't have an *a priori* way to connect a tangent vector at one point to a tangent vector at another point. Since the coordinates may be quite wonky, there will not be a natural way to extend the tangent spaces to make a single vectorspace in which all vectors can be compared as they can be for Cartesian coordinates. Explicitly, the issue is that as one moves about the manifold, the basis vectors change which means that the geometrical meaning of the components of a vector changes from place to place. Without a way to compare the components of a vector at neighboring places, we cannot define a derivative that makes sense. It turns our we can solve both these problems simultaneously by making use of the ambient euclidean space. Since each tangent space is embedded in the

ambient space, we do have a natural way to compare vectors. Therefore, we should use the geometric invariant we cooked up before. I claim that the quantity,

$$\nabla_i A^k = \vec{e}^{\,k} \cdot \partial_i(\vec{e}_j A^j)$$

is a type $(1,1)$ tensor because $\vec{e}_j A^j$ is an invariant so we have no issues with differentiating a transformation symbol under changes of coordinates. This combination is known as the covariant derivative because it is a derivative which is a covariant tensor. However, it is not immediately clear that this definition reduces to the standard derivative in "nice" coordinates. We can write the covariant derivative in a more explicit form,

$$\nabla_i A^k = \vec{e}^{\,k} \cdot \partial_i(\vec{e}_j A^j) = \vec{e}^{\,k} \cdot \vec{e}_j \partial_i A^j + A^j \left(\vec{e}^{\,k} \cdot \partial_i \vec{e}_j \right) = \partial_i A^k + A^j \Gamma_{ij}^k$$

where the Christoffel symbols Γ_{ij}^k appear naturally.

Definition 3.1. The Christoffel symbol is defined the indexed set of coefficients,

$$\Gamma_{ij}^k = \vec{e}^{\,k} \cdot \frac{\partial \vec{e}_j}{\partial x^i}$$

Note that Γ_{ij}^k is called the Christoffel *symbol* **not** the Christoffel *tensor* because it does not transform as a tensor due to the derivative of an non-invariant object. However, the non-covariance of the Christoffel symbol exactly cancels the failure of $\partial_i A^k$ to be a tensor such that the combination in the covariant derivative satisfies the tensor property. The Christoffel symbol has many nice properties which we will need to exploit over and over again.

Lemma 3.1. *Properties of the Christoffel Symbol,*

  1. $\Gamma_{ij}^k = \Gamma_{ji}^k$

  2. $\vec{e}_k \cdot \frac{\partial \vec{e}^{\,j}}{\partial x^i} = -\Gamma_{ik}^j$

Proof. (DO PROOF) $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can similarly define a covariant derivative for covectors by considering the invariant geometric covector which has a tonsorial derivative,

$$\nabla_i A_k = \vec{e}_k \cdot \partial_i(\vec{e}^{\,j} A_j) = \vec{e}_k \cdot \vec{e}^{\,j} \partial_i A_j + A_j \left(\vec{e}_k \cdot \partial_i \vec{e}^{\,j} \right) = \partial_i A_k - A_j \Gamma_{ik}^j$$

This suggests the following definition,

Definition 3.2. The covariant derivative of a type (p,q) tensor T is the type $(p, q+1)$ tensor,

$$\nabla_i T_{j_1 \cdots j_q}^{i_1 \cdots i_p} = \partial_i T_{j_1 \cdots j_q}^{i_1 \cdots i_p} + T_{j_1 \cdots j_q}^{k \cdots i_p} \Gamma_{ki}^{i_1} + \cdots + T_{j_1 \cdots j_q}^{i_1 \cdots k} \Gamma_{ki}^{i_p} - T_{k \cdots j_q}^{i_1 \cdots i_p} \Gamma_{j_1 i}^k + \cdots - T_{j_1 \cdots k}^{i_1 \cdots i_p} \Gamma_{j_q i}^k$$

Using the invariant geometric notion of the tensor which has a coordinate independent derivative because it is an invariant, we can prove that the above definition always produces a tensor. Alternatively, this can be checked explicitly by evaluating it in a new coordinate system. I absolutely do not recommend doing this because it is beyond horrible.

Lemma 3.2 (The Product Rule). *For any two tensors T and S, the covariant derivative satisfies the product rule,*

$$\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$$

Proof. This follows simply by the product rule for normal derivatives and factoring. We will show this explicitly for the special case of type $(0,1)$ tensors,

$$
\begin{aligned}
\nabla_k(A_i B_j) &= \partial_k(A_i B_j) - A_\ell B_j \Gamma^\ell_{ik} - A_i B_\ell \Gamma^\ell_{kj} \\
&= B_j \partial_k A_i + A_i \partial_k B_j - A_\ell B_j \Gamma^\ell_{ik} - A_i B_\ell \Gamma^\ell_{kj} \\
&= B_j \left(\partial_k A_i - A_\ell \Gamma^\ell_{ik} \right) + A_i \left(\partial_k B_j - B_\ell \Gamma^\ell_{kj} \right) \\
&= (\nabla_k A_i) B_j + A_i (\nabla_k B_j)
\end{aligned}
$$

$\square$

3.4 Covariant Derivative of the Metric

The covariant derivative of the metric can be explicitly computed,

$$
\begin{aligned}
\nabla_k g_{ij} &= \frac{\partial(\vec{e}_i \cdot \vec{e}_j)}{\partial x^k} - g_{mj} \Gamma^m_{ik} - g_{im} \Gamma^m_{jk} \\
&= \frac{\partial \vec{e}_i}{\partial x^k} \cdot \vec{e}_j + \vec{e}_i \cdot \frac{\partial \vec{e}_j}{\partial x^k} - g_{mj} \vec{e}^m \cdot \frac{\partial \vec{e}_i}{\partial x^k} - g_{im} \vec{e}^m \cdot \frac{\partial \vec{e}_j}{\partial x^k} \\
&= \frac{\partial \vec{e}_i}{\partial x^k} \cdot \vec{e}_j + \vec{e}_i \cdot \frac{\partial \vec{e}_j}{\partial x^k} - \vec{e}_j \cdot \frac{\partial \vec{e}_i}{\partial x^k} - \vec{e}_i \cdot \frac{\partial \vec{e}_j}{\partial x^k} = 0
\end{aligned}
$$

Therefore, the covariant derivative of the metric is always zero. This implies that if we can find a coordinate system in which the Christoffel symbols vanish at a point, then the coordinate derivatives of the metric also vanish. Such a reference frame is known as a locally inertial frame in general relativity and will be very important for simplifying calculations. This important property makes our definition of the covariant derivative well-defined under index juggling. That is, the covariant derivative commute with index juggling. Consider,

$$\nabla_k A_i = \nabla_k g_{ij} A^j = A^j \nabla_k g_{ij} + g_{ij} \nabla_k A^j = g_{ij} \left[\frac{\partial A^j}{\partial x^k} + A^m \Gamma^j_{mk} \right]$$

Furthermore,

$$
\begin{aligned}
g_{ij} \left[\frac{\partial A^j}{\partial x^k} + A^m \Gamma^j_{mk} \right] &= \frac{\partial A_i}{\partial x_k} - A^j \frac{\partial g_{ij}}{\partial x^k} + g_{ij} A^m \Gamma^j_{mk} \\
&= \frac{\partial A_i}{\partial x^k} - A^j_{mj} \Gamma^m_{lk} - A^j g_{im} \Gamma^m_{jk} + A^m g_{ij} \Gamma^j_{mk} \\
&= \frac{\partial A_i}{\partial x^k} - A_m \Gamma^m_{ik}
\end{aligned}
$$

which is consistent with our definition that,

$$\nabla_k A_i = \frac{\partial A_i}{\partial x^k} - A_m \Gamma^m_{ik}$$

Index juggling in the other direction also works because $\nabla_k g^{ij} = 0$. Therefore, we can raise and lower indices inside a covariant derivative without issue since raising with the metric inside the derivative is the same as raising with the metric outside.

13

There is one further important set of quantities we would like to know the covariant derivative of, the basis vectors. Consider,

$$\nabla_i \vec{e}_j = \frac{\partial \vec{e}_j}{\partial x^i} - \vec{e}_k \Gamma_{ij}^k = \left(1 - \vec{e}_k \otimes \vec{e}^k\right) \cdot \frac{\partial \vec{e}_j}{\partial x^i}$$

The operator $\vec{e}_k \otimes \vec{e}^k$ is a map from the ambient space of geometric vectors to the tangent space acting by the dot product with a vector,

$$(\vec{e}_k \otimes \vec{e}^k) \cdot \vec{v} = \vec{e}_k \left(\vec{e}^k \cdot \vec{v}\right)$$

Because $\vec{e}_k$ is a basis of the tangent space and $\vec{e}_i \cdot \vec{e}^j = \delta_i^j$ we know that if $\vec{v} = v^i \vec{e}_i$ is in the tangent space then,

$$(\vec{e}_k \otimes \vec{e}^k) \cdot \vec{v} = (\vec{e}_k \otimes \vec{e}^k) \cdot v^i \vec{e}_i = \vec{e}_k v^i \left(\vec{e}^k \cdot \vec{e}_i\right) = \vec{e}_k v^i \delta_i^k = \vec{e}_k v^k = \vec{v}$$

Thus, $\vec{e}_k \otimes \vec{e}^k$ is the projection operator from the ambient space onto the tangent space since it taken any vector into the ambient space and preserves any vector already in the tangent space. Therefore,

$$1 - \vec{e}_k \otimes \vec{e}^k$$

is the orthogonal projector to the tangent space. This operator takes any vector and subtracts out the component in the tangent space leaving only the component perpendicular to the tangent space. Therefore, the covariant derivative of the basis,

$$\nabla_i \vec{e}_j = \frac{\partial \vec{e}_j}{\partial x^i} - \vec{e}_k \Gamma_{ij}^k = \left(1 - \vec{e}_k \otimes \vec{e}^k\right) \cdot \frac{\partial \vec{e}_j}{\partial x^i}$$

gives the component of the change in $\vec{e}_j$ in the direction x^i which is pointing out of the tangent space. If the tangent space has the same dimension as the ambient space we are embedded in then the basis $\vec{e}_i$ spans the entire space so the covariant derivative of the basis vectors is identically zero. Furthermore,

$$\nabla_i \vec{e}^j = \frac{\partial \vec{e}^j}{\partial x^i} + \vec{e}^k \Gamma_{ik}^j = \frac{\partial \vec{e}^j}{\partial x^i} - \vec{e}^k \cdot \left(\vec{e}_k \cdot \frac{\partial \vec{e}^j}{\partial x^i}\right) = \left(1 - \vec{e}^k \otimes \vec{e}_k\right) \cdot \frac{\partial \vec{e}^j}{\partial x^i}$$

which has the exact same properties but is a map from the ambient space of covectors to the cotangent space. Similarly, if the dimension of the cotangent space (which is equal to the dimension of the tangent space because they are finite dimensional) is equal to the dimension of the ambient space then the covariant derivative of the contravariant basis is also zero,

$$\nabla_i \vec{e}^j = 0$$

3.5   Calculating Christoffel symbols

From the above derivation we find,

1. $\nabla_k g_{ij} = 0 \implies \frac{\partial g_{ij}}{\partial x^k} = g_{mj} \Gamma_{ik}^m + g_{im} \Gamma_{jk}^m$

2. $\nabla_i g_{kj} = 0 \implies \frac{\partial g_{kj}}{\partial x^k} = g_{mj} \Gamma_{ki}^m + g_{km} \Gamma_{ji}^m$

3. $\nabla_j g_{ki} = 0 \implies \frac{\partial g_{ki}}{\partial x^j} = g_{mi} \Gamma_{kj}^m + g_{km} \Gamma_{ij}^m$

14

Then, noting that $\Gamma_{ij}^m = \Gamma_{ij}^m$ taking $(2) + (3) - (1)$ we get,

$$\frac{\partial g_{kj}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} = 2g_{km}\Gamma_{ij}^m$$

We multiply both sides of this equation by $\frac{1}{2}g^{rk}$ and thus,

$$\Gamma_{ij}^r = \frac{1}{2}g^{rk}\left(\frac{\partial g_{kj}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right)$$

This can be used to calculate the Christoffel symbols explicitly from the metric.

3.6 The Covariant Derivative Along a Curve

Given a curve $\vec{\gamma}(\lambda) = \vec{R}(\gamma_1(\lambda), \cdots, \gamma_n(\lambda))$ parameterized by λ, we have the tangent vector:

$$\vec{t} = \frac{\mathrm{d}\vec{\gamma}}{\mathrm{d}\lambda} = \frac{\mathrm{d}x^k}{\mathrm{d}\lambda}\frac{\partial\vec{R}}{\partial x^k} = t^k\vec{e}_k$$

If we have a vector $\vec{v} = v^i\vec{e}_i$ in the tangent space defined at each point on the curve,

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}\lambda} = \frac{\mathrm{d}v^k}{\mathrm{d}\lambda}\vec{e}_i + v^i\frac{\partial\vec{e}_i}{\partial x^k}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda} = \frac{\mathrm{d}v^i}{\mathrm{d}\lambda}\vec{e}_i + v^i\vec{e}_m\Gamma_{ik}^m t^k = \left(\frac{\mathrm{d}v^i}{\mathrm{d}\lambda} + v^m\Gamma_{mk}^i t^k\right)\vec{e}_i$$

where we changed the indices of the second term. We now define, the derivative along a curve by,

$$D_\lambda v^i = \nabla_{\gamma(\lambda)}v^i = \frac{\mathrm{d}v^i}{\mathrm{d}\lambda} + v^m\Gamma_{mk}^i t^k$$

which produces tensors. Similarly, we may write this in terms of lowered indices as,

$$\nabla_{\gamma(\lambda)}v_i = \frac{\mathrm{d}v_i}{\mathrm{d}\lambda} - v_m\Gamma_{ik}^m t^k$$

This derivative along a curve is related to the covariant derivative by,

$$\nabla_{\gamma(\lambda)}v^i = t^k\nabla_k v^i = \frac{\mathrm{d}x^k}{\mathrm{d}\lambda}\frac{\partial v^i}{\partial x^k} + v^m\Gamma_{mk}^i t^k = \frac{\mathrm{d}v^i}{\mathrm{d}\lambda} + v^m\Gamma_{mk}^i t^k$$

This allows us to find the acceleration in any coordinates that transforms. So if $\lambda = $ time, and $t^k = v^k$, the velocity, then we get the acceleration:

$$D_t v^k = \frac{\mathrm{d}^2 x^k}{\mathrm{d}t^2} + v^m\Gamma_{mn}^k v^n = a^k$$

Remark 3.6.1 (Notation). In the next section I will begin to use the Greek convention for spacetime indices. The only reason is to get familiar with the convention. I will use Greek indices for any index for a spacetime coordinate (of a Minkowski-like manifold) in the range $0, 1, 2, 3$ for $3 + 1$ dimensions. I will use Latin indices for spatial coordinates (of a Euclidean-like manifold) in the range $1, 2, 3$ for 3 dimensions. If a tensor in spacetime appears with Latin indices it means the index runs only over spatial coordinates and not the time component(s). The results of the following section are generally true for any space or space-time whether it be Euclidean or Minkowski. As such, I will illustrate with an example in 3-dimensional Euclidean space.

In the coming sections there will be many derivatives which will quickly get very tiresome to write out in full. Here I introduce the following shorthand notation,

$$\partial_\mu = \frac{\partial}{\partial x^\mu}$$

for a derivative with respect to the coordinates. Furthermore, given quantities which depend on a parameter λ, I introduce the shorthand notation,

$$\partial_\lambda = \frac{d}{d\lambda}$$

for the total derivative with respect to λ. While these notations are very standard, they have the unfortunate ambiguity that ∂_λ looks like a coordinate derivative with respect to the spacetime index λ. Unfortunately, we have run out of letters for parameters since any Greek letter can represent a spacetime index and any Latin letter can represent a spatial index. To ameliorate this failure of notation, I will never use λ as an index, only as a parameter.

3.7 The Voss-Weyl Formula

We need to further investigate the properties of determinants under differentiation.

Lemma 3.3.
$$\frac{\partial \sqrt{g}}{\partial g_{\mu\nu}} = \tfrac{1}{2}\sqrt{g}\, g^{\mu\nu} \quad and \quad \frac{\partial \sqrt{g}}{\partial g^{\mu\nu}} = -\tfrac{1}{2}\sqrt{g}\, g_{\mu\nu}$$

Proof.
$$\frac{\partial \sqrt{g}}{\partial g_{\mu\nu}} = \frac{1}{2\sqrt{g}}\frac{\partial g}{\partial g_{\mu\nu}} = \frac{1}{2\sqrt{g}} g\, g^{\nu\mu} = \tfrac{1}{2}\sqrt{g}\, g^{\mu\nu}$$

Similarly, (FINISH PROOF) □

Lemma 3.4.
$$\frac{\partial \sqrt{g}}{\partial x^\mu} = \sqrt{g}\, \Gamma^\nu_{\nu\mu}$$

Proof.
$$\frac{\partial \sqrt{g}}{\partial x^\mu} = \frac{1}{2\sqrt{g}}\frac{\partial g}{\partial g_{\alpha\beta}}\frac{\partial g_{\alpha\beta}}{\partial x^\mu}$$

However, we know that,

$$\frac{\partial g}{\partial g_{\alpha\beta}} = g g^{\alpha\beta} \quad \text{and} \quad \frac{\partial g_{\alpha\beta}}{\partial x^\mu} = \vec{e}_\alpha \cdot \frac{\partial \vec{e}_\beta}{\partial x^\mu} + \frac{\partial \vec{e}_\alpha}{\partial x^\mu} \cdot \vec{e}_\beta$$

Therefore,

$$\frac{\partial \sqrt{g}}{\partial x^\mu} = \tfrac{1}{2}\sqrt{g}\, g^{\alpha\beta}\left[\vec{e}_\alpha \cdot \frac{\partial \vec{e}_\beta}{\partial x^\mu} + \frac{\partial \vec{e}_\alpha}{\partial x^\mu} \cdot \vec{e}_\beta \right] = \tfrac{1}{2}\sqrt{g}\left[\vec{e}^\beta \cdot \frac{\partial \vec{e}_\beta}{\partial x^\mu} + \vec{e}^\alpha \cdot \frac{\partial \vec{e}_\alpha}{\partial x^\mu} \right]$$

$$= \tfrac{1}{2}\sqrt{g}\left[\Gamma^\beta_{\beta\mu} + \Gamma^\alpha_{\alpha\mu} \right] = \sqrt{g}\, \Gamma^\alpha_{\alpha\mu}$$

□

Theorem 3.5 (Voss-Weyl)**.**

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g}\, V^\mu \right)$$

Theorem 3.6.

$$\frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g}\, V^\mu \right) = \partial_\mu V^\mu + \frac{1}{\sqrt{g}} V^\mu \partial_\mu(\sqrt{g}) = \partial_\mu V^\mu + V^\mu \Gamma^\nu_{\nu\mu}$$

$$= \partial_\mu V^\mu + V^\nu \Gamma^\mu_{\nu\mu} = \nabla_\mu V^\mu$$

A very important corollary is where we take $V^\mu = \nabla^\mu F = \partial^\mu F = g^{\mu\nu} \partial_\nu F$ where F is a scalar. This is a type $(1,0)$ tensor so we can apply the Voss-Weyl formula,

$$\nabla_\mu \nabla^\mu F = \frac{1}{\sqrt{g}} \partial_\mu(\sqrt{g}\, g^{\mu\nu} \partial_\nu F)$$

This gives an explicit formula for the covariant Laplacian, $\nabla^2 = \nabla_\mu \nabla^\mu$ which does not require evaluating the Christoffel symbols. When doing computations, the Voss-Weyl formula is far more convenient to evaluate than proceeding from the definition of the covariant derivative. For example,
     (SPHERICAL EXAMPLE)

4 Intrinsic Curvature

4.1 Parallel Transport

4.2 Geodesics

A Geodesic is the notion of the "straightest possible" path on a manifold which may be curved in some nontrivial way. There are two different ways to capture the idea of a path being straight. We could call a path between two points "straight" if it minimizes the distance between the points. There is a slight subtlety here which is that we don't just want paths that minimize the distance but actually any path for which any *nearby* path (such that the difference between the two is always small) is longer. We call such a path "extremal" because it is a local extremum of the distance function rather than minimal. In flat space, minimizing and extremalizing the distance is the same. However, consider two points on a sphere. In general, there are two "straight" great circle segments which join the points going opposite ways around the sphere. Only one of these two paths will be the global minimum distance. However, we would like to call both paths geodesics. I will call this the global definition of a geodesic because knowing that a path extremalizes the distance between two points requires global knowledge of the path as a whole. In symbols, a path $\gamma^\mu(\lambda)$ parameterized by t is a geodesic if,

$$\delta \int \sqrt{g_{\alpha\beta} \mathrm{d}x^\alpha \mathrm{d}x^\beta} = \delta \int \sqrt{g_{\alpha\beta} \frac{\mathrm{d}\gamma^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}\gamma^\beta}{\mathrm{d}\lambda}} \mathrm{d}\lambda = 0$$

which says that if you make small variations away from the actual geodesic the variation in the path length[2].

---

[2] Remember the squared distance of a small displacement is given by $g_{\alpha\beta} \mathrm{d}x^\alpha \mathrm{d}x^\beta$ so the entire distance is the integral of the square root of this quantity along the path.

There is also a local criterion for what it means for a path to be straight. A geodesic is a path that parallel transports its own tangent vector. This is the statement that a geodesic is a path which accelerates as little as possible in the sense that the tangent vector stays in the same direction as much as possible. The best way to visualize this local criterion for being a geodesic is to imagine a car driving along the surface. The path the car takes if we lock the steering wheel dead ahead so the car turns as little as possible as it goes over the curved terrain is a geodesic. The local criterion for a path $\gamma^\mu(\lambda)$ to be a geodesic is that the tangent vector $t^i = \frac{d\gamma^\mu}{d\lambda}$ is parallel transported along the curve. This requires that,

$$\nabla_{\gamma(\lambda)} t^\mu = \frac{dt^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta} t^\alpha t^\beta = 0$$

This is known as the geodesic equation. A beautiful and important fact is that the global and local definitions of a geodesic are equivalent.[3] Writing out the tangent vector explicitly,

$$\frac{d^2\gamma^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{d\gamma^\alpha}{d\lambda} \frac{d\gamma^\beta}{d\lambda} = 0$$

In general relativity, particles not under the influence of external forces follow geodesics in spacetime. This can be taken as an axiom or can be argued from the equivalence principle. The idea is that a particle without external forces (excluding gravity of course) should accelerate as little as possible. This is how we should think of a geodesic. In fact, we cannot force the particle's trajectory to have zero acceleration,

$$\frac{d^2\gamma^\mu}{d\lambda^2} = 0$$

in all reference frames because the quantity $\frac{dt^i}{d\lambda}$ is not a tensor. However, we can get close. We will now use a common trick in general relativity which is to evaluate some quantity we care about in a *local inertial reference frame* about some point of interest. A local inertial reference frame or LIF is the "flattest possible" reference frame at a certain point. Imagine taking the surface and gluing a flat sheet of paper with a Cartesian coordinate system on it to some point x_0. This sheet of paper represents the tangent space at x_0. Now we glue the entire sheet to the surface while stretching it as little as possible. Anyone who has tried to make a map from a globe knows that this is impossible without stretching the coordinate system due to the intrinsic curvature of the surface. However, we do get something from an LIF because we fix $g_{\mu\nu}(x_0) = \eta_{\mu\nu}$ the metric at the point of interest is exactly that of the flat tangent space and the Christoffel symbols at x_0 all vanish because to first order the metric is constant at x_0. However, as we will see shortly, we cannot make the second derivatives of the metric, or equivalently the first derivatives of the Christoffel symbols, vanish as well because of the nonzero intrinsic curvature which is invariant under changes of coordinate system. Consider a point x_0 on the trajectory of a particle in spacetime. In an LIF at x_0, we know that $\Gamma^\mu_{\alpha\beta} = 0$ therefore,

$$\frac{d^2\gamma^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{d\gamma^\alpha}{d\lambda} \frac{d\gamma^\beta}{d\lambda} = \frac{d^2\gamma^\mu}{d\lambda^2} = 0$$

so the geodesic equation reduces to the statement that the coordinate acceleration is zero in a LIF. In fact, we know this must be true from the equivalence principle which tells us that

---

[3] The proof of this fact uses calculus of variations. See Appendix A. for the derivation of the geodesic equation from extremalizing the total distance along the curve.

a particle in free-fall "feels" no acceleration. A frame in free-fall is an LIF because by the equivalence principle it looks exactly like flat space close to the point of interest. Therefore, the equivalence principle tells us that,

$$\frac{\mathrm{d}^2\gamma^\mu}{\mathrm{d}\lambda^2} = 0 \implies \frac{\mathrm{d}^2\gamma^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}\gamma^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}\gamma^\beta}{\mathrm{d}\lambda} = 0$$

because the Christoffel symbols vanish in an LIF. Thus, the geodesic equation holds in these specific coordinates. However, the geodesic equation is a equivalent to $\nabla_{\gamma(\lambda)}t^\mu = 0$ which is a tensor equation. Therefore, if the geodesic equation holds in any coordinate system, it holds in all coordinate systems. The equivalence principle is enough to give us this equation in a specific reference frame and then the power of the tensor property means it must be true in all reference frames.

The geodesic equation is the fundamental law governing the motion of a particle in curved spacetime. We can write it in a somewhat more suggestive manner,

$$m\frac{\mathrm{d}^2\gamma^\mu}{\mathrm{d}\lambda^2} = ma^\mu = -m\Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}\gamma^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}\gamma^\beta}{\mathrm{d}\lambda}$$

where m is the mass of the particle and,

$$a^\mu = \frac{\mathrm{d}^2\gamma^\mu}{\mathrm{d}\lambda^2}$$

is the *coordinate acceleration*. This looks a lot like Newton's second law with a force,

$$F^\mu_{\mathrm{grav}} = -m\Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}\gamma^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}\gamma^\beta}{\mathrm{d}\lambda}$$

Thus the Christoffel symbols take on an interpretation as the gravitational force. We see that the gravitational force is proportional to the mass or equivalently that the gravitational acceleration is independent of the mass a fundamental observation which lead Einstein to the equivalence principle. In the formalism of GR, adding this factor of m is actually extraneous and only done to make analogies to Newtonian physics.

4.3 Geodesic Deviation

An second notion of what it means for a space to be curved is to consider nearby geodesics. In flat space, geodesics are straight lines and two parallel lines remain parallel forever and do not intersect. However, on a sphere, two great circles always intersect even if they look parallel at a given point. The fact that parallel geodesics converge or diverge is a manifestation of the intrinsic curvature of the manifold. Furthermore, we will be able to extract an invariant measure of this curvature, the Riemann tensor, from the behavior of nearby geodesics.

Consider two geodesic paths $x^\mu(\lambda)$ and $\bar{x}^\mu(\lambda)$ each satisfying the geodesic equation,

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}(x^\mu)t^\alpha t^\beta = 0 \quad \text{and} \quad \frac{\mathrm{d}^2\bar{x}^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}(\bar{x}^\mu)\bar{t}^\alpha\bar{t}^\beta = 0$$

Let us define a new tensor $\xi^m u$ as:

$$\xi^\mu(\lambda) = \bar{x}^\mu(\lambda) - x^\mu(\lambda)$$

In this case, $\bar{x}$ and x will only transform the same way if their difference is infinitesimal. Thus, ξ^μ is a tensor. We expand in a Taylor series, $\Gamma(\bar{x}^\mu) = \Gamma(x^\mu) + \xi^\nu \partial_\nu \Gamma(x^\mu) + O(\xi^2)$. The higher order terms are negligible for small ξ. Therefore, via substituting into the geodesic, equation for $\bar{x}$ we get,

$$\frac{\mathrm{d}^2 \bar{x}^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}(\bar{x}^\mu) \bar{t}^\alpha \bar{t}^\beta = \partial_\lambda^2 (x^\mu + \xi^\mu) + \left(\Gamma^\mu_{\alpha\beta} + \xi^\nu \partial_\nu \Gamma^\mu_{\alpha\beta} \right) \partial_\lambda (x^\alpha + \xi^\alpha) \partial_\lambda \left(x^\beta + \xi^\beta \right)$$

$$= \partial_\lambda^2 \xi^\mu + \Gamma^\mu_{\alpha\beta} \left(\partial_\lambda x^\alpha \partial_\lambda \xi^\beta + \partial_\lambda x^\beta \partial_\lambda \xi^\alpha \right) + \xi^\nu \partial_\nu \Gamma^\mu_{\alpha\beta} \partial_\lambda x^\alpha \partial_\lambda x^\beta$$

$$+ \partial_\lambda^2 x^\mu + \Gamma^\mu_{\alpha\beta} \partial_\lambda x^\alpha \partial_\lambda x^\beta + O(\xi^2)$$

$$= 0$$

However, because x^μ also satisfies the geodesic equation,

$$\partial_\lambda^2 x^\mu + \Gamma^\mu_{\alpha\beta} \partial_\lambda x^\alpha \partial_\lambda x^\beta = 0$$

the last line of the above equation is zero to first order in ξ. Therefore,

$$\partial_\lambda^2 \xi^\mu + \Gamma^\mu_{\alpha\beta} \left(\partial_\lambda x^\alpha \partial_\lambda \xi^\beta + \partial_\lambda x^\beta \partial_\lambda \xi^\alpha \right) + \xi^\nu \partial_\nu \Gamma^\mu_{\alpha\beta} \partial_\lambda x^\alpha \partial_\lambda x^\beta$$

$$= \partial_\lambda^2 \xi^\mu + 2\Gamma^\mu_{\alpha\beta} \partial_\lambda x^\alpha \partial_\lambda \xi^\beta + \xi^\nu \partial_\nu \Gamma^\mu_{\alpha\beta} \partial_\lambda x^\alpha \partial_\lambda x^\beta = 0$$

where I have used the fact that $\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{\beta\alpha}$. Thus, we have the differential equation for ξ,

$$\partial_\lambda^2 \xi^\mu + 2\Gamma^\mu_{\alpha\beta} t^\alpha \partial_\lambda \xi^\beta + t^\alpha t^\beta \xi^\nu \partial_\nu \Gamma^\mu_{\alpha\beta} = 0$$

To write this in tensor form, we calculate the double derivative along the curve $x^\mu(\lambda)$,

$$D_\lambda^2 \xi^\mu = \nabla_{x^\mu(\lambda)} \left(\nabla_{x^\mu(\lambda)} \xi^\mu \right) = \partial_\lambda^2 \xi^\mu + \partial_\lambda \left[\xi^\alpha t^\beta \Gamma^\mu_{\alpha\beta} \right] + \left(\partial_\lambda \xi^\delta + \xi^\alpha t^\beta \Gamma^\delta_{\alpha\beta} \right) t^\gamma \Gamma^\mu_{\gamma\delta}$$

$$= \partial_\lambda^2 \xi^\mu + 2t^\alpha \partial_\lambda \xi^\beta \Gamma^\mu_{\alpha\beta} + \xi^\alpha \partial_\lambda t^\beta \Gamma^\mu_{\alpha\beta} + \xi^\alpha t^\beta \partial_\lambda \Gamma^\mu_{\alpha\beta} + \xi^\alpha t^\beta t^\gamma \Gamma^\delta_{\alpha\beta} \Gamma^\mu_{\gamma\delta}$$

Now we use the facts that, $\partial_\lambda t^\beta + t^\rho t^\sigma \Gamma^\beta_{\rho\sigma} = 0$ and $\xi^\alpha t^\beta \partial_\lambda \Gamma^\mu_{\alpha\beta} = \xi^\alpha t^\beta t^\gamma \partial_\gamma \Gamma^\mu_{\alpha\beta}$ we get,

$$D_\lambda^2 \xi^\mu = \partial_\lambda^2 \xi^\mu + 2t^\alpha \partial_\lambda \xi^\beta \Gamma^\mu_{\alpha\beta} + \xi^\alpha \left(-t^\rho t^\sigma \Gamma^\beta_{\rho\sigma} \right) \Gamma^\mu_{\alpha\beta} + \xi^\alpha t^\beta t^\gamma \partial_\gamma \Gamma^\mu_{\alpha\beta} + \xi^\alpha t^\beta t^\gamma \Gamma^\delta_{\alpha\beta} \Gamma^\mu_{\gamma\delta}$$

Subtracting the first equation, we find that,

$$D_\lambda^2 \xi^\mu = -\xi^\nu t^\alpha t^\beta \partial_\nu \Gamma^\mu_{\alpha\beta} + \xi^\alpha t^\beta t^\gamma \partial_\gamma \Gamma^\mu_{\alpha\beta} - \xi^\alpha t^\rho t^\sigma \Gamma^\beta_{\rho\sigma} \Gamma^\mu_{\alpha\beta} + \xi^\alpha t^\beta t^\gamma \Gamma^\delta_{\alpha\beta} \Gamma^\mu_{\gamma\delta}$$

$$= -\xi^\nu t^\alpha t^\beta \left(\partial_\nu \Gamma^\mu_{\alpha\beta} - \partial_\beta \Gamma^\mu_{\nu\alpha} + \Gamma^\delta_{\alpha\beta} \Gamma^\mu_{\nu\delta} - \Gamma^\delta_{\nu\alpha} \Gamma^\mu_{\beta\delta} \right)$$

Because the left hand side is a tensor since the covariant derivative produces tensors and we have shown equality in any coordinate systems, the right hand side must also be a tensor. This motivates Riemann's definition of the curvature tensor.

4.4 The Riemann Tensor

Definition 4.1. The Riemann Tensor defined as,

$$R^\mu_{\alpha\nu\beta} = \partial_\nu \Gamma^\mu_{\alpha\beta} - \partial_\beta \Gamma^\mu_{\nu\alpha} + \Gamma^\delta_{\alpha\beta} \Gamma^\mu_{\nu\delta} - \Gamma^\delta_{\nu\alpha} \Gamma^\mu_{\beta\delta}$$

Using this notation, our equation for geometric deviation becomes,

$$D_\lambda^2 \xi^\mu + \xi^\nu t^\alpha t^\beta R_{\alpha\nu\beta}^\mu = 0$$

Therefore, parallel geodesics generically remain parallel only when $R_{\alpha\nu\beta}^\mu = 0$. This justifies calling any space with $R_{\alpha\nu\beta}^\mu = 0$ flat.

In locally inertial frames, The Christoffel symbols vanish at x_0^μ, and using the fact that, $\Gamma_{\alpha\beta}^\mu = \frac{1}{2} g^{\mu\nu} (\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\alpha\nu} - \partial_\nu g_{\alpha\beta})$, we find,

$$R_{\alpha\beta\gamma}^\mu (x_0^\mu) = \partial_\beta \Gamma_{\alpha\gamma}^\mu - \partial_\gamma \Gamma_{\alpha\beta}^\mu$$

$$\implies R_{\mu\alpha\beta\gamma}(x_0^\mu) = \frac{1}{2} (\partial_\beta \partial_\alpha g_{\gamma\mu} + \partial_\beta \partial_\gamma g_{\alpha\mu} - \partial_\beta \partial_\mu g_{\alpha\gamma} - \partial_\gamma \partial_\alpha g_{\beta\mu} - \partial_\gamma \partial_\beta g_{\alpha\mu} + \partial_\gamma \partial_\mu g_{\alpha\beta})$$

$$\implies R_{\mu\alpha\beta\gamma}(x_0^\mu) = \frac{1}{2} (\partial_\beta \partial_\alpha g_{\gamma\mu} + \partial_\gamma \partial_\mu g_{\alpha\beta} - \partial_\beta \partial_\mu g_{\alpha\gamma} - \partial_\gamma \partial_\alpha g_{\beta\mu})$$

Remember! This result only holds in an LIF.

4.4.1 Properties of the Riemann Tensor

In a LIF, we found above, that,

$$R_{\alpha\beta\gamma\delta}(x_0^\mu) = \frac{1}{2} \left(\underbrace{\partial_\gamma \partial_\beta g_{\alpha\delta}}_{(1)} + \underbrace{\partial_\alpha \partial_\delta g_{\beta\gamma}}_{(2)} - \underbrace{\partial_\gamma \partial_\alpha g_{\beta\gamma}}_{(3)} - \underbrace{\partial_\delta \partial_\beta g_{\alpha\gamma}}_{(4)} \right)$$

We note that there are specific symmetries in exchanging indices. In other words, in swapping the Greek indices we find that terms switch roles:

$$\alpha \leftrightarrow \beta \implies \text{ Term (1)} \leftrightarrow \text{ Term (3) and } \text{Term (2)} \leftrightarrow \text{ Term (4)}$$

$$\gamma \leftrightarrow \delta \implies \text{ Term (1)} \leftrightarrow \text{ Term (4) and } \text{Term (2)} \leftrightarrow \text{ Term (3)}$$

Therefore, we get the property that,

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma} = R_{\beta\alpha\delta\gamma}$$

Similarly,

$$\alpha \leftrightarrow \gamma \quad \text{and} \quad \beta \leftrightarrow \delta \implies \text{ Term (1)} \leftrightarrow \text{ Term (2) and } \text{Term (3)} \leftrightarrow \text{ Term (4)}$$

which implies that,

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} = R_{\delta\gamma\beta\alpha} = -R_{\delta\gamma\alpha\beta} = -R_{\gamma\delta\beta\alpha}$$

Next, consider, at x_0,

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2} (\partial_\gamma \partial_\beta g_{\alpha\delta} + \partial_\alpha \partial_\delta g_{\beta\gamma} - \partial_\gamma \partial_\alpha g_{\beta\gamma} - \partial_\delta \partial_\beta g_{\alpha\gamma})$$
$$R_{\alpha\gamma\delta\beta} = \frac{1}{2} (\partial_\delta \partial_\gamma g_{\alpha\beta} + \partial_\alpha \partial_\beta g_{\gamma\delta} - \partial_\delta \partial_\alpha g_{\gamma\delta} - \partial_\beta \partial_\gamma g_{\alpha\delta})$$
$$R_{\alpha\delta\beta\gamma} = \frac{1}{2} (\partial_\beta \partial_\delta g_{\alpha\gamma} + \partial_\alpha \partial_\gamma g_{\delta\beta} - \partial_\beta \partial_\alpha g_{\delta\beta} - \partial_\gamma \partial_\delta g_{\alpha\beta})$$

Let (a, b) refer to the bth term of the ath line. Now,

1. (1,1) and (2,4) cancel

2. (1,2) and (2,3) cancel

3. (1,3) and (3,2) cancel

4. (1,4) and (3,1) cancel

5. (2,1) and (3,4) cancel

6. (2,2) and (3,3) cancel

Therefore, all terms cancel to zero leaving,

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} = 0$$

which is known as the **First or Arithmetic Bianchi Identity**.

4.4.2   Bianchi Identity

In LIF, we recall that $\nabla_\mu = \partial_\mu$ when evaluated at x_0^μ. Now, let us introduce the notation, $\partial_{\alpha\beta} = \partial_\alpha \partial_\beta$. From the definition of the Riemann Tensor along and the gradient in LIF, we have,

$$\nabla_\epsilon R_{\alpha\beta\gamma\delta} = \tfrac{1}{2} \left(\partial_{\epsilon\gamma\beta} g_{\alpha\delta} + \partial_{\epsilon\alpha\delta} g_{\beta\gamma} - \partial_{\epsilon\gamma\alpha} g_{\beta\delta} - \partial_{\epsilon\delta\beta} g_{\alpha\gamma} \right)$$
$$\nabla_\gamma R_{\alpha\beta\delta\epsilon} = \tfrac{1}{2} \left(\partial_{\gamma\delta\beta} g_{\alpha\epsilon} + \partial_{\gamma\alpha\epsilon} g_{\beta\delta} - \partial_{\gamma\delta\alpha} g_{\beta\epsilon} - \partial_{\gamma\epsilon\beta} g_{\alpha\delta} \right)$$
$$\nabla_\delta R_{\alpha\beta\epsilon\gamma} = \tfrac{1}{2} \left(\partial_{\delta\epsilon\beta} g_{\alpha\gamma} + \partial_{\delta\alpha\gamma} g_{\beta\epsilon} - \partial_{\delta\epsilon\alpha} g_{\beta\gamma} - \partial_{\delta\gamma\beta} g_{\alpha\epsilon} \right)$$

Now,

1. (1,1) and (2,4) cancel

2. (1,2) and (3,3) cancel

3. (1,3) and (2,2) cancel

4. (1,4) and (3,1) cancel

5. (2,1) and (3,4) cancel

6. (3,2) and (2,3) cancel

These terms all cancel out just right so that the sum vanishes. This gives the **Second or Differential Bianchi Identity**,

$$\nabla_\epsilon R_{\alpha\beta\gamma\delta} + \nabla_\gamma R_{\alpha\beta\delta\epsilon} + \nabla_\delta R_{\alpha\beta\epsilon\gamma} = 0 \tag{1}$$

We note that these derivations assumed that we were working in a LIF, but since they are tensor identities, they are true in any coordinates.

4.5 Ricci Curvature

We now define the Ricci Curvature tensor

Definition 4.2. The Ricci Tensor, $R_{\mu\nu}$ is defined as,

$$R_{\mu\nu} = R^{\alpha}_{\mu\alpha\nu} = g^{\alpha\gamma} R_{\alpha\mu\gamma\nu} = g^{\gamma\alpha} R_{\gamma\nu\alpha\mu} = R_{\nu\mu}$$

We therefore have from above that the Ricci Tensor is symmetric.

Definition 4.3. From the Ricci Tensor, we define the Ricci Curvature scalar as,

$$R = R^{\mu}_{\mu} = g^{\mu\nu} R_{\mu\nu}$$

From these definitions, we get the following lemma:

Lemma 4.1.
$$\nabla_\mu R^\mu_\nu = \frac{1}{2} \nabla_\nu R$$

Proof. We begin with the Bianchi identity as in 1 and contract it to get,

$$g^{\beta\delta} g^{\alpha\gamma} \left(\nabla_\epsilon R_{\alpha\beta\gamma\delta} + \nabla_\gamma R_{\alpha\beta\delta\epsilon} + \nabla_\delta R_{\alpha\beta\epsilon\gamma} \right) = 0$$

We contract and get,

$$
\begin{aligned}
0 g^{\beta\delta} g^{\alpha\gamma} & \left(\nabla_\epsilon R_{\alpha\beta\gamma\delta} + \nabla_\gamma R_{\alpha\beta\delta\epsilon} + \nabla_\delta R_{\alpha\beta\epsilon\gamma} \right) \\
& = g^{\beta\gamma} \left(\nabla_\mu R_{\beta\delta} + \nabla^\gamma R_{\gamma\beta\delta\mu} + \nabla_\delta R^{\gamma}_{\beta\mu\gamma} \right) \\
& = \nabla_\mu R^{\delta}_{\delta} + \nabla_\gamma R^{\delta\gamma}_{\delta\mu} - \nabla_\delta R^{\delta}_{\mu}
\end{aligned}
$$

Thus, solving for $\nabla_\mu R$

$$\nabla_\mu R = \nabla_\gamma R^{\gamma}_{\mu} + \nabla_\delta R^{\delta}_{\mu} \implies \nabla_\mu R^\mu_\nu = \frac{1}{2} \nabla_\nu R$$

where we took advantage of substituting dummy variables. $\qquad\square$

As a corollary, we find,

$$\nabla_\mu R^{\mu\nu} = \frac{1}{2} g^{\mu\nu} \nabla_\mu R$$

5 The Einstein Field Equations

5.1 Motivation

5.2 Einstein-Hilbert Action

David Hilbert, one of the greatest mathematicians of the 20th century showed that the field equations of general relativity could be obtained by minimizing a beautifully simple action. We consider all possible metrics and compute some number called the action for each of possible metric. We want to find a metric which extremalizes this action. It turns out that the condition that a metric extremalizes some action will be that it satisfies some field equation. If we start with an invariant action, then we must derive conditions which satisfy the principle of general covariance because they must have the same form in any reference

frame. A truly remarkable fact about general relativity is that the action which gives rise the Einstein field equations is the simplest possible curvature invariant we can cook up. It turns out that satisfying the Einstein field equations is exactly the condition for a metric to minimize (really extremalize) the total Ricci curvature scalar. If there is matter than we want the metric which extremalizes the difference between the Lagrangian density of the matter and the Ricci scalar.

We begin to construct the Einstein-Hilbert action by first considering an invariant volume. In normal Cartesian coordinates, the infinitesimal volume of a parallelogram with sides $dx_1^\mu, dx_2^\mu, dx_3^\mu, dx_2^\mu$, is $S_{ijkl}dx_1^i dx_2^j dx_3^k dx_4^l$ where S is the symmetry symbol. We may makes this quantity invariant by defining the volume element of the same parallelogram as,

$$dV = \epsilon_{ijkl}dx_1^i dx_2^j dx_3^k dx_4^l$$

Therefore, to integrate over such a volume element, we use a volume element spanned by the four infinitesimal increments in each coordinate: $dV = \sqrt{q}\, dx^1 dx^2 dx^3 dx^4$ which we shall write as $dV = \sqrt{g}d^4 x$.

The action S defined as,

$$S = \int \mathcal{L}\sqrt{g}\, d^4 x$$

is invariant of $\mathcal{L}$ is invariant. Thus, all laws derived from S will hold in all reference frames.

Definition 5.1 (Einstein-Hilbert Action)**.** This is the simplest action connecting curvature to Matter/Energy is the corresponding action from the

$$\mathcal{L}_{GR} = -\frac{c^4}{16\pi G}R + \mathcal{L}_M$$

where R is the Ricci scalar, the simplest invariant of $R_{\alpha\beta\gamma\delta}$.

We now proceed to find the functional variation of this Einstein-Hilbert action. For notational convenience[4] set $\kappa = \frac{8\pi G}{c^4}$. Now we extremalize S by setting its variation equal to zero,

$$\delta S = 0$$

$$= \delta\left(\int\left[-\frac{1}{2\kappa}R + \mathcal{L}_M\right]\sqrt{g}\, d^4 x\right) = \int\left[-\frac{1}{2\kappa}\frac{\delta\sqrt{g}R}{\delta g^{\mu\nu}} + \frac{\delta\sqrt{g}L_M}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu}\, d^4 x$$

$$= \int\left[-\frac{\delta R}{\delta g^{\mu\nu}} + \frac{R}{\sqrt{g}}\frac{\delta\sqrt{g}}{\delta g^{\mu\nu}} + \frac{2\kappa}{\sqrt{g}}\frac{\delta\sqrt{g}L_M}{\delta g^{\mu\nu}}\right]\delta g^{\mu\nu}\, d^4 x$$

Now, we solve and find, δR.

$$\delta R = \delta R_{\mu\nu}g^{\mu\nu} + R_{\mu\nu}\delta g^{\mu\nu}$$

where,

$$\delta R_{\mu\nu} = \delta R_{\mu\sigma\nu}^\sigma = \delta\left[\partial_\mu\Gamma_{v\sigma}^\sigma - \partial_\nu\Gamma_{\mu\sigma}^\sigma + \Gamma_{\sigma\nu}^\rho\Gamma_{\mu\rho}^\sigma - \Gamma_{\mu\sigma}^\delta\Gamma_{\nu\delta}^\sigma\right]$$

This may be simplified as, (FIX THIS DERIVATION)

$$\delta R_{\mu\nu} = \partial_\sigma\delta\Gamma_{\mu\nu}^\alpha - \partial_\nu\delta\Gamma_{\mu\sigma}^\alpha + \delta\Gamma_{\mu\nu}^\rho\Gamma_{\mu\nu}^\rho - \delta\Gamma_{\mu\sigma}^\delta\Gamma_{\nu\delta}^\alpha - \Gamma_{\mu\sigma}^\delta\Gamma_{\nu\delta}^\alpha$$

$$= \nabla_\sigma\delta\Gamma_{\mu\nu}^\alpha + \delta\Gamma_{\mu\rho}^\alpha\Gamma_{\nu\sigma}^\rho - \nabla_\nu\delta\Gamma_{\mu\sigma}^\alpha + \delta\Gamma_{\mu\rho}^\alpha\Gamma_\nu^\rho\sigma$$

---

[4]The strange combination of symbols is a historical accident. If general relativity were discovered first then κ rather than G would be called the universal gravitational constant

Therefore, we plug this in to find,

$$\delta R = R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\left[\nabla_\sigma \delta\Gamma^\sigma_{\mu\nu} + \nabla_\nu \delta\Gamma^\sigma_{\mu\sigma}\right]$$
$$= R_{\mu\nu}\delta g^{\mu\nu} + \nabla_\sigma\left[g^{\mu\nu}\delta\Gamma^\sigma_{\mu\nu} - g^{\mu\sigma}\delta\Gamma^\sigma_{\mu\sigma}\right]$$

We note however that since, $(\nabla_\sigma A^\sigma)\sqrt{g} = \partial_\sigma\sqrt{g}A^\sigma$, we get, that,

$$\frac{\delta R}{\delta g^{\mu\nu}} = R_{\mu\nu}$$

Similarly, we find therefore that,

$$\delta S = \int\left[-R_{\mu\nu} + \frac{1}{2}Rg_{\mu\nu} + \kappa\frac{2}{\sqrt{g}}\frac{\delta\sqrt{g}\mathcal{L}_{\mathcal{M}}}{\delta^{\mu\nu}}\right]\delta g^{\mu\nu}\sqrt{g}\mathrm{d}^4x$$

this holds for any $\delta g^{\mu\nu}$ and therefore, the integrand is zero.
We now define the Stress-Energy-Momentum Tensor.

Definition 5.2. The Stress-Energy-Momentum Tensor is defined as,

$$T_{\mu\nu} = \frac{2}{\sqrt{g}}\frac{\delta\sqrt{g}\mathcal{L}_{\mathcal{M}}}{\delta g^{\mu\nu}}$$

We note that this tensor is symmetric such that, $T_{\mu\nu} = T_{\nu\mu}$

Substituting into the above condition we get by setting the variation of the Einstein-Hilbert action to 0, we get the equation,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}$$

Therefore, taking the gradient with respect to μ of each side, we find that the equations of motion imply,

$$\nabla_\mu T^{\mu\nu} = 0$$

5.3 Properties of the Field Equations

Now we have the fundamental law of general relativity which describes the shape of spacetime,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}$$

The energy-momentum tensor acts as a source term for this set of coupled differential equations. The left hand size of the field equations is a function of $R_{\mu\nu}$ which is built from the Riemann tensor which is built from the Christoffel symbols which themselves are built from the metric and its derivatives. Therefore, the Einstein field equations give us a set of differential equations which can be used to, in principle, compute the metric everywhere. Once we know the metric, we can find the geodesics and therefore know everything about the spacetime geometry we could think to ask. However, this is an extremely difficult calculation to do even in simple cases and effectively impossible in general. Therefore, we will put quite a bit of effort into finding approximate solutions to the field equations and to extracting information from them without writing down an explicit solution for the metric.

First, from the properties of the geometric objects on the left, we can deduce properties of the energy-momentum tensor which must be satisfied if Einstein's equations are to hold. For example, we know that,

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu}$$

is a symmetric tensor because both $R_{\mu\nu}$ and $g_{\mu\nu}$ are. Thus, any valid stress-energy tensor $T_{\mu\nu}$ must be symmetric as well if Einstein's field equations are satisfied. Furthermore, we want the energy-momentum tensor to satisfy a local conservation equation,

$$\nabla_\mu T^{\mu\nu} = 0$$

which is immediate for true solutions of the field equations for the metric (metrics that extremalize the Einstein-Hilbert action) because,

$$\nabla_\mu R^{\mu\nu} = \tfrac{1}{2} g^{\mu\nu} \nabla_\mu R = \tfrac{1}{2} \nabla_\mu (R g^{\mu\nu}) \implies \nabla_\mu G^{\mu\nu} = 0$$

This property of the geometric objects was part of Einstein's original motivation for the form of the field equations.

Taking the trace of the field equations,

$$R - \tfrac{1}{2} R g^{\mu\nu} g_{\mu\nu} = \frac{8\pi G}{c^4} T^\mu_\mu$$

Denoting the energy-momentum trace, $T = T^\mu_\mu$ and seeing that $g^{\mu\nu} g_{\mu\nu} = \delta^\mu_\mu = n = 4$ we get that,

$$R - 2R = \frac{8\pi G}{c^4} T \implies R = -\frac{8\pi G}{c^4} T$$

Therefore, we can swap R with T in the Einstein field equations to get,

$$R_{\mu\nu} = \frac{8\pi G}{c^4} (T_{\mu\nu} - \tfrac{1}{2} T g_{\mu\nu})$$

Suppose we consider the gravitational field in a vacuum, no matter or energy anywhere such that $T_{\mu\nu} = 0$. Using the above relations, $T = R = 0$ so Einstein's field equations reduce to $R_{\mu\nu} = 0$. Therefore, "vacuum solutions" to Einstein's field equations are exactly those metrics which give rise to vanishing Ricci curvature $R_{\mu\nu}$. However, this does not mean that a vacuum solution is necessarily flat. It turns out that the Ricci tensor vanishing, $R_{\mu\nu} = 0$, does not imply that the Riemann curvature tensor $R_{\alpha\beta\gamma\delta}$ also vanishes, only that its partial traces are all zero. Therefore, we can have vacuum solutions with nontrivial Riemannian curvature and thus interesting geometry. This fact is very important because it allows for gravitational waves to propagate though a vacuum. We will spend quite a lot of time working with vacuum solutions in the coming sections.

6 The Newtonian Limit

We want to consider gravity in the weak field limit meaning when gravity is very weak and therefore the metric is approximately flat. We can choose almost Minkowski coordinates such that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where we thing of $h_{\mu\nu}$ as being a small perturbation away from being flat.

7 Black Holes: The Schwarzschild Metric

7.1 Deriving the Metric

We want to find a spherically symmetric vacuum solution to Einstein's field equations in the empty space surrounding a compact massive object. We can always choose an orthogonal coordinate system in which the basis vectors are orthogonal and therefore the metric is diagonal everywhere. Since we require the solution to have spherical symmetry, we will proceed by generalizing the metric for spherical coordinates (t, r, θ, ϕ),

$$g_{\mu\nu}^{\text{sph.}} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$

By spherical symmetry and time translation invariance, that is we want the solution to be static in time, we want the metric to only be a function of r. Therefore, we can write,

$$g_{\mu\nu} = \begin{pmatrix} c^2 A(r) & 0 & 0 & 0 \\ 0 & -B(r) & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$

where A and B are yet undetermined functions of r. One might wonder why we only let the first two terms depend on arbitrary functions of r. By spherical symmetry, the last two terms representing the interval along a sphere in terms of the angles θ and ϕ must be of the form,

$$C(r)\mathrm{d}\Omega^2 = C(r)r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)$$

since otherwise there would be preferred directions on the sphere. However, in this case we can simply redefine the coordinate r such that $r'^2 = C(r)r^2$ and therefore the metric can be put in the above form. We can next scale the time component to fix the determinant of the metric. It is convenient to choose,

$$\det g_{\mu\nu} = -c^2 r^4 \sin^2\theta$$

which is the determinant of the metric for flat spherical coordinates. However,

$$\det g_{\mu\nu} = -c^2 A(r) B(r) r^4 \sin^2\theta$$

so our determinant condition forces $B(r) = 1/A(r)$.

7.2 The Christoffel Symbols and the Ricci Tensor

Now we need to calculate the Christoffel symbols. This is a tedious and horrible exercise which Mathematica is perfectly suited for. Therefore, I will simply quote the results,

$$\Gamma^t_{\alpha\beta} = \begin{pmatrix} 0 & \frac{A'(r)}{2A(r)} & 0 & 0 \\ \frac{A'(r)}{2A(r)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \Gamma^r_{\alpha\beta} = \begin{pmatrix} \frac{1}{2}c^2 A'(r)A(r) & 0 & 0 & 0 \\ 0 & -\frac{A'(r)}{2A(r)} & 0 & 0 \\ 0 & 0 & -rA(r) & 0 \\ 0 & 0 & 0 & -rA(r)\sin^2\theta \end{pmatrix}$$

$$\Gamma^{\theta}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\cos\theta\sin\theta \end{pmatrix} \qquad \Gamma^{r}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \frac{\cos\theta}{\sin\theta} \\ 0 & \frac{1}{r} & \frac{\cos\theta}{\sin\theta} & 0 \end{pmatrix}$$

Next we need to calculate the Ricci tensor,

$$R_{\mu\nu} = \partial_{\alpha}\Gamma^{\alpha}_{\mu\nu} - \partial_{\nu}\Gamma^{\alpha}_{\alpha\mu} + \Gamma^{\beta}_{\mu\nu}\Gamma^{\alpha}_{\alpha\beta} - \Gamma^{\beta}_{\alpha\mu}\Gamma^{\alpha}_{\nu\beta}$$

again, this is a terrible and boring calculation which I will let Mathematica chew on. The result is,

$$R_{\mu\nu} = \begin{pmatrix} \frac{1}{2}c^2 A(r)\left(\frac{2A'(r)}{r} + A''(r)\right) & 0 & 0 & 0 \\ 0 & -\frac{2A'(r)+rA''(r)}{2rA(r)} & 0 & 0 \\ 0 & 0 & 1 - A(r) - rA'(r) & 0 \\ 0 & 0 & 0 & (1 - A(r) - rA'(r))\sin^2\theta \end{pmatrix}$$

because we are looking for a vacuum solution, we set,

$$R_{\mu\nu} = 0$$

Looking at the equation $R_{\theta\theta} = 0$ we get,

$$1 - A(r) - rA'(r) = 0 \implies \frac{\mathrm{d}}{\mathrm{d}r}(rA(r)) = 1$$

Therefore, the generic solution for A is,

$$A(r) = \frac{r - r_s}{r} = 1 - \frac{r_s}{r}$$

where r_s is an integration constant. Thus, the metric becomes,

$$g_{\mu\nu} = \begin{pmatrix} c^2\left(1 - \frac{r_s}{r}\right) & 0 & 0 & 0 \\ 0 & -\left(1 - \frac{r_s}{r}\right)^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}$$

so the proper time interval can be written as,

$$c^2\mathrm{d}\tau^2 = \left(1 - \frac{r_s}{r}\right)c^2\mathrm{d}t^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2$$

And finally, the Christoffel symbols become,

$$\Gamma^{t}_{\alpha\beta} = \begin{pmatrix} 0 & \frac{r_s}{2r^2}\left(1 - \frac{r_s}{r}\right)^{-1} & 0 & 0 \\ \frac{r_s}{2r^2}\left(1 - \frac{r_s}{r}\right)^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Gamma^{r}_{\alpha\beta} = \begin{pmatrix} \frac{c^2\left(1 - \frac{r_s}{r}\right)}{2r^2} & 0 & 0 & 0 \\ 0 & -\frac{r_s}{2r^2}\left(1 - \frac{r_s}{r}\right)^{-1} & 0 & 0 \\ 0 & 0 & -(r - r_s) & 0 \\ 0 & 0 & 0 & -(r - r_s)\sin^2\theta \end{pmatrix}$$

$$
\Gamma^{\theta}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\cos\theta\sin\theta \end{pmatrix} \qquad \Gamma^{\phi}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \frac{\cos\theta}{\sin\theta} \\ 0 & \frac{1}{r} & \frac{\cos\theta}{\sin\theta} & 0 \end{pmatrix}
$$

We can determine the constant r_s from the mass of the compact object. Using the geodesic equation parameterized by proper time,

$$
\frac{\mathrm{d}^2 x^{\mu}}{\mathrm{d}\tau^2} = -\Gamma^{\mu}_{\alpha\beta} \frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau} \frac{\mathrm{d}x^{\beta}}{\mathrm{d}\tau}
$$

In particular suppose that θ and ϕ are constant. Then, in the r direction,

$$
a^r = \frac{\mathrm{d}^2 x^r}{\mathrm{d}\tau^2} = -\Gamma^{r}_{\alpha\beta} \frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau} \frac{\mathrm{d}x^{\beta}}{\mathrm{d}\tau} = -\frac{r_s}{2r^2} \left[c^2 \left(1 - \frac{r_s}{r}\right) \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(1 - \frac{r_s}{r}\right)^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 \right]
$$

However, for $\mathrm{d}\theta = \mathrm{d}\phi = 0$ we see from the expression for an increment of proper time,

$$
c^2 = c^2 \left(\frac{\mathrm{d}\tau}{\mathrm{d}\tau}\right)^2 = c^2 \left(1 - \frac{r_s}{r}\right) \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(1 - \frac{r_s}{r}\right)^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2
$$

Therefore, the radial acceleration becomes,

$$
a^r = -\frac{r_s c^2}{2r^2}
$$

which has the form of Newton's universal law of gravitation,

$$
a = -\frac{MG}{r^2}
$$

Since we want general relativity to reproduce Newton's laws in the low energy limit and we only have one parameter to play with, we are forced to take,

$$
r_s = \frac{2MG}{c^2}
$$

This is known as the Schwarzschild radius. Matching the low energy limit to Newtonian theory may be an unsatisfying argument and if so one can fix the constant r_s by solving Einstein's field equations inside the compact body where $T_{\mu\nu} \neq 0$. However, this is even more difficult and does not give us nearly enough to be worth the trouble. From above, it may seem that general relativity exactly reproduces Newton's law of gravitation but this is not true. Firstly, we derived the acceleration in the special case that θ and ϕ were constant, that is the particle is moving along a radial trajectory. For more general trajectories the acceleration will not have this simple form and the angular momentum will enter the expression. Furthermore, a^r is the acceleration with respect to *proper* time not the *coordinate* acceleration one gets by simply differentiating the coordinate r with respect to the coordinate t twice. That said, in the limit of velocities much less than the speed of light and radii much larger than r_s, only the the term Γ^r_{00} contributes in the geodesic equation because it has a relative factor of c^2 so the acceleration does not depend on angular velocity and the proper and coordinate accelerations agree with each other and with the Newtonian value.

7.3 A Black Hole

7.3.1 Killing Vectors

Solving the geodesic equation is difficult and annoying. We want to make use of the symmetry of the Schwarzschild metric to find conserved quantities of the motion. One method of doing this is to find so called Killing vectors named for Wilhelm Killing.

Definition 7.1. A vector field K_μ is a Killing field if,

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0$$

Such a vector field is a manifestation of some symmetry of the spacetime. The essential fact about Killing vectors is that they give rise to conserved quantities along geodesics.

Theorem 7.1. Let K_μ be a killing vector and $x^\mu(\lambda)$ a geodesic. Then the quantity,

$$Q = K_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$$

is conserved along the geodesic. Explicitly,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(K_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right) = 0$$

Proof. Using the geodesic equation,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(K_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right) = \frac{\mathrm{d}K_\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} + K_\mu \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2}$$

$$= \frac{\partial K_\mu}{\partial x^\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} - K_\mu \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}$$

However,

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0$$

and therefore,

$$\partial_\mu K_\nu + \partial_\nu K_\mu - K_\alpha \Gamma^\alpha_{\mu\nu} - K_\alpha \Gamma^\alpha_{\nu\mu} = 0$$

However, the Christoffel symbol is symmetric so,

$$\partial_\mu K_\nu + \partial_\nu K_\mu = 2K_\alpha \Gamma^\alpha_{\mu\nu}$$

Now we doubly contract both sides of this equation with the tangent vector of the geodesic $t^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ which gives,

$$\partial_\mu K_\nu t^\mu t^\nu + \partial_\nu K_\mu t^\mu t^\nu = 2K_\alpha \Gamma^\alpha_{\mu\nu} t^\mu t^\nu$$

The two terms on the right hand side are equal by renaming indices so,

$$2\partial_\mu K_\nu t^\mu t^\nu = 2K_\alpha \Gamma^\alpha_{\mu\nu} t^\mu t^\nu$$

which implies that.

$$\partial_\mu K_\nu t^\mu t^\nu = K_\alpha \Gamma^\alpha_{\mu\nu} t^\mu t^\nu$$

which are, up to renaming summed over indices, exactly the two terms in the derivative of Q. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(K_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right) = \frac{\partial K_\mu}{\partial x^\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} - K_\mu \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = \partial_\alpha K_\mu t^\alpha t^\mu = K_\mu \Gamma^\mu_{\alpha\beta} t^\alpha t^\beta = 0$$

$\square$

The real power of the Killing vector method comes from our ability to identify Killing vectors from the coordinate form of the metric. In particular, if the metric does not depend on some coordinate, then there will be a Killing vector corresponding to that coordinate.

Theorem 7.2. *Then, the vector $K^\mu = \delta_\kappa^\mu$ for some fixed index κ is a Killing vector if and only if $\partial_\kappa g_{\alpha\beta} = 0$, that is $g_{\alpha\beta}$ goes not depend on x^κ.*

Proof. With lowered indices, $K_\mu = g_{\mu\nu}K^\mu = g_{\mu\nu}\delta_\kappa^\nu = g_{\mu\kappa}$. Now consider the Killing term,

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = \partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - 2g_{\alpha\kappa}\Gamma_{\mu\nu}^\alpha$$

Now we use the explicit form of the Christoffel symbols in terms of the metric,

$$\Gamma_{\mu\nu}^\alpha = \frac{1}{2}g^{\alpha\beta}\left(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}\right)$$

Plugging in,

$$\begin{aligned}
\nabla_\mu K_\nu + \nabla_\nu K_\mu &= \partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - g_{\alpha\kappa}g^{\alpha\beta}\left(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}\right) \\
&= \partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - \delta_\kappa^\beta\left(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}\right) \\
&= \partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - \left(\partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - \partial_\kappa g_{\mu\nu}\right) = \partial_\kappa g_{\mu\nu}
\end{aligned}$$

Therefore, the Killing criterion,

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \iff \partial_\kappa g_{\mu\nu} = 0$$

is satisfied if and only if $\partial_\kappa g_{\mu\nu} = 0$. $\qquad\square$

Corollary 7.3. *If $g_{\alpha\beta}$ does not depend on the coordinate κ, that is $\partial_\kappa g_{\alpha\beta} = 0$ then for any geodesic $x^\mu(\lambda)$ the quantity,*

$$P_\kappa = g_{\kappa\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$$

is conserved along the geodesic. This quantity is called the canonical momentum in the direction κ.

Proof. If $\partial_\kappa g_{\alpha\beta} = 0$ then by the above theorem, $K^\mu = \delta_\kappa^\mu$ is a Killing vector. Therefore, by Killing's theorem, the quantity,

$$P_\kappa = K_\mu\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = g_{\mu\nu}\delta_\kappa^\nu\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = g_{\kappa\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$$

is conserved along the geodesic. $\qquad\square$

We will now apply Killing's theorems to the Schwarzschild geometry.

7.4 Schwarzschild Geodesics

The Schwarzschild metric,

$$g_{\mu\nu} = \begin{pmatrix} c^2\left(1 - \frac{r_s}{r}\right) & 0 & 0 & 0 \\ 0 & -\left(1 - \frac{r_s}{r}\right)^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}$$

is independent of t and ϕ. Therefore, we immediately get two killing vectors $\xi^\alpha = (1, 0, 0, 0)$ and $\zeta^\alpha = (0, 0, 0, 1)$. This give the conserved quantities,

$$h = P_t = g_{\alpha\beta}\xi^\alpha \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = g_{00} \frac{\mathrm{d}x^0}{\mathrm{d}\lambda} = c^2 \left(1 - \frac{r_s}{r}\right) \frac{\mathrm{d}t}{\mathrm{d}\lambda}$$

and

$$\ell = -P_\theta = -g_{\alpha\beta}\zeta^\alpha \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = -g_{33} \frac{\mathrm{d}x^3}{\mathrm{d}\lambda} = r^2 \frac{\mathrm{d}\theta}{\mathrm{d}\lambda}$$

7.4.1 Time-Like Geodesics

We will first consider the conserved quantities along the special geodesics parameterized by proper time,

$$H = mP_t = mg_{\alpha\beta}\xi^\alpha \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = g_{00} \frac{\mathrm{d}x^0}{\mathrm{d}\tau} = mc^2 \left(1 - \frac{r_s}{r}\right) \frac{\mathrm{d}t}{\mathrm{d}\tau}$$
$$= \frac{mc^2 \left(1 - \frac{r_s}{r}\right)}{\sqrt{\left(1 - \frac{r_s}{r}\right) - \frac{\dot{r}^2}{c^2}\left(1 - \frac{r_s}{r}\right)^{-1} - r^2 \left(\frac{\dot{\theta}^2}{c^2} + \sin^2\theta \frac{\dot{\phi}^2}{c^2}\right)}}$$

which turns out be be exactly the energy of the particle and

$$L = -mP_\phi = -mg_{\alpha\beta}\zeta^\alpha \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = -mg_{33} \frac{\mathrm{d}x^3}{\mathrm{d}\tau} = mr^2 \sin^2\theta \frac{\mathrm{d}\phi}{\mathrm{d}\tau}$$
$$= \frac{m(r^2 \sin^2\theta)\,\dot{\phi}}{\sqrt{\left(1 - \frac{r_s}{r}\right) - \frac{\dot{r}^2}{c^2}\left(1 - \frac{r_s}{r}\right)^{-1} - r^2 \left(\frac{\dot{\theta}^2}{c^2} + \sin^2\theta \frac{\dot{\phi}^2}{c^2}\right)}}$$

where dots represent derivatives with respect to coordinate time. From the proper time interval,

$$c^2 = c^2 \left(1 - \frac{r_s}{r}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - r^2 \left(\frac{\mathrm{d}\theta}{\mathrm{d}\tau}\right)^2 - r^2 \sin^2\theta \left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right)^2$$

calling $\frac{\partial r}{\partial \tau} = u^r$ and plugging in our invariants,

$$c^2 \left(1 - \frac{r_s}{r}\right) = c^2 \left(\frac{H}{mc^2}\right)^2 - (u^r)^2 - r^2 \left(1 - \frac{r_s}{r}\right)(u^\theta)^2 - \frac{L^2}{m^2 r^2 \sin^2\theta}\left(1 - \frac{r_s}{r}\right)$$

Since the orbits must be planar due to spherical symmetry, we can require that $\theta = \frac{\pi}{2}$ is constant. Thus, $u^\theta = 0$ and $\sin\theta = 1$. Therefore, we get the so called radial equation,

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + \frac{L^2}{m^2 r^2}\left(1 - \frac{r_s}{r}\right) - \frac{r_s c^2}{r} = c^2 \left[\left(\frac{H}{mc^2}\right)^2 - 1\right]$$

Rewriting this equation in terms of physical quantities and using $H = E + mc^2$ where E is the mechanical energy,

$$\frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + \frac{L^2}{2m^2 r^2}\left(1 - \frac{2MG}{c^2 r}\right) - \frac{MG}{r} = \frac{E}{m} + \frac{E^2}{m^2 c^2}$$

32

If we compare this to the Newtonian equation for the energy,

$$\frac{1}{2}\dot{r}^2 + \frac{L^2}{2m^2r^2} - \frac{MG}{r} = \frac{E}{m}$$

we see the two are nearly identical except the general relativity equation has derivatives with respect to proper time rather than coordinate time and there is an extra term,

$$-\frac{L^2}{2m^2r^2}\frac{2MG}{c^2r} = -\frac{MGL^2}{m^2c^2r^3}$$

So there is an extra r^{-3} potential showing up in GR. We call the combination,

$$V_{\text{eff}}(r) = \frac{L^2}{2m^2r^2}\left(1 - \frac{2MG}{c^2r}\right) - \frac{MG}{r}$$

the Schwarzschild effective potential. The extrema of this function correspond to circular orbits.

7.4.2   Light-Like Geodesics

For a light-like geodesic, the proper time is exactly zero so it is not an available parameter. Here, we will take an arbitrary parameter λ. From the condition for a null-geodesic,

$$c^2\left(1 - \frac{r_s}{r}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 - r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}\lambda}\right)^2 - r^2\sin^2\theta\left(\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\right)^2 = 0$$

As before, for simplicity we restrict to planar orbits with $\theta = \frac{\pi}{2}$ is constant. Plugging in the invariants,

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 + \frac{\ell^2}{r^2}\left(1 - \frac{r_s}{r}\right) - \frac{h^2}{c^2} = 0$$

We want to find a circular orbit where,

$$\frac{\mathrm{d}r}{\mathrm{d}\lambda} = \frac{\mathrm{d}^2r}{\mathrm{d}\lambda^2} = 0$$

This tells us that,

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{\ell^2}{r^2}\left(1 - \frac{r_s}{r}\right) - \frac{h^2}{c^2}\right] = -2\frac{\ell^2}{r^3}\left(1 - \frac{r_s}{r}\right) + \frac{\ell^2}{r^2}\frac{r_s}{r^2} = 0$$

Therefore,

$$r_s = 2(r - r_s) \implies r = \tfrac{3}{2}r_s$$

This is known as the photon sphere, the only place photons can make circular orbits.

7.5   The Action Approach to Geodesics of the Schwarzschild Metric

8   Gravitational Waves

Mention LIGO

8.1 Linearized Gravity

As before, we want to consider gravity in the weak field limit meaning when gravity is very weak and therefore the metric is approximately flat. We can choose coordinates such that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where we thing of $h_{\mu\nu}$ as being a small perturbation away from being flat.

It turns out that when we drop a

8.2 Gravitational Wave Solutions

9 Experimental Tests of General Relativity

10 Dust and Perfect Fluids

11 Electromagnetism in General Relativity

11.1 Maxwell's Equations in Curved Spacetime

11.2 The Reissner-Nordström Metric of a Charged Black Hole

12 Cosmological Models

12.1 Early Models

12.2 Dark Energy and the Cosmological Constant

12.3 The Friedmann–Lemaître–Robertson–Walker Metric

12.4 The ΛCDM Model

13 Gravitoelectromagnetism

14 Kaluza-Klein Theory

In a 1919 letter to Einstein, Kaluza proposed that a 5-dimensional theory of general relativity could unify gravity and electromagnetism through a purely geometric theory. Although Kaluza-Klein theory makes incorrect predictions about the charges and masses of elementary particles, the framework has become one the foundational pillars of modern physics leading to Yang-Mills theories.

14.1 A Five-Dimensional Space-Time

Remark 14.1.1. It will be convenient to use the $(-, +, +, +)$ convention for this chapter.

Consider a metric,

$$\tilde{g}_{\mu\nu} = \begin{pmatrix} g_{\mu\nu} & g_{\mu 5} \\ g_{\nu 5} & g_{55} \end{pmatrix}$$

where tildes will indicate 5-dimensional quantities. Adding a whole new space-time dimension provides far too much dynamical freedom. We require the fifth dimension to be compactified, rolled-up in a small circle, explaining the fact that an extra spatial dimension cannot be observed. The 5-dimensional space-time we construct has the topology of a fiber bundle of the circle S^1 over a usual four dimensional Minkowski-like manifold M^4. This

means the space looks locally like $M^4 \times S^1$. Often, we will take the space to globally have the topology $M^4 \times S^1$ which we refer to a "trivial" bundle.

In general, the 5th basis vector will not be orthogonal to the basis $\vec{e}_\mu$ so $g_{\mu 5} = \vec{e}_\mu \cdot \vec{e}_5 \neq 0$. We can decompose $\vec{e}_\mu = \vec{e}_\mu^\parallel + \vec{e}_\mu^\perp$ where $\vec{e}_\mu^\parallel \parallel \vec{e}_5$ and $\vec{e}_\mu^\perp \perp \vec{e}_5$. We then write, $\vec{e}_\mu^\parallel = B_\mu \vec{e}_5$ and thus,

$$\tilde{g}_{\mu\nu} = (\vec{e}_\mu^\perp + \vec{e}_\mu^\parallel) \cdot (\vec{e}_\nu^\perp + \vec{e}_\nu^\parallel) = \vec{e}_\mu^\perp \cdot \vec{e}_\nu^\perp + \vec{e}_\mu^\parallel \cdot \vec{e}_\nu^\parallel = g_{\mu\nu} + \phi^2 B_\mu B_\nu$$

Where,

$$g = \vec{e}_\mu^\perp \cdot \vec{e}_\nu^\perp \quad \text{and} \quad \phi^2 = \tilde{g}_{55} = \vec{e}_5 \cdot \vec{e}_5 \quad \text{and} \quad B_\mu = \frac{\vec{e}_\mu^\parallel \cdot \vec{e}_5}{\vec{e}_5 \cdot \vec{e}_5} = \frac{\vec{e}_\mu \cdot \vec{e}_5}{\vec{e}_5 \cdot \vec{e}_5} = \frac{\tilde{g}_{\mu 5}}{\tilde{g}_{55}}$$

Therefore, the entire 5D metric becomes,

$$\tilde{g}_{\mu\nu} = \begin{pmatrix} g_{\mu\nu} + \phi^2 B_\mu B_\nu & \phi^2 B_\mu \\ \phi^2 B_\nu & \phi^2 \end{pmatrix}$$

A simple calculation shows that,

$$\tilde{g}^{\mu\nu} = \begin{pmatrix} g^{\mu\nu} & -B^\mu \\ -B^\nu & \phi^{-2} + B_\alpha B^\alpha \end{pmatrix}$$

Furthermore,

$$\tilde{g} = \det \tilde{g}_{\mu\nu} = \phi^2 g$$

where index juggling is performed with the 4D metric $g_{\mu\nu}$. To be able to project physics down into the base 4D space-time, we impose the cylinder condition,

$$\partial_5 \tilde{g}_{\mu\nu} = 0$$

so physical quantities should not change while moving along the circular fifth dimension.

14.2   Gauge Transformations

Suppose we make a coordinate transformation by rotating about the local copy of S^1 by an amount dependent on 4D space-time position. As we will see, such a transformation is exactly a local gauge transformation which is why Kaluza-Klein is considered a $U(1)$ gauge theory. Explicitly, consider the transformation, $x'^\mu = x^\mu$ and $x'^5 = x^5 - \chi(x^\mu)$. Then, consider the change in components of the metric,

$$\tilde{g}'_{\mu\nu} = \frac{\partial \tilde{x}^\alpha}{\partial \tilde{x}'^\mu} \frac{\partial \tilde{x}^\beta}{\partial \tilde{x}'^\nu} \tilde{g}_{\alpha\beta}$$

In particular,

$$\phi'^2 = \tilde{g}'_{55} = \frac{\partial \tilde{x}^\alpha}{\partial \tilde{x}'^5} \frac{\partial \tilde{x}^\beta}{\partial \tilde{x}'^5} \tilde{g}_{\alpha\beta} = \tilde{g}_{55} = \phi^2$$

since the coordinate transformation does not depend on x^5. However,

$$\phi'^2 B'_\mu = \tilde{g}'_{\mu 5} = \frac{\partial \tilde{x}^\alpha}{\partial \tilde{x}'^\mu} \frac{\partial \tilde{x}^\beta}{\partial \tilde{x}'^5} \tilde{g}_{\alpha\beta} = \tilde{g}_{\mu 5} + \frac{\partial x^5}{\partial x'^\mu} \tilde{g}_{55} = \tilde{g}_{\mu 5} + \phi^2 \, \partial_\mu \chi = \phi^2 (B_\mu + \partial_\mu \chi)$$

Therefore, since $\phi = \phi'$, we see that B_μ transforms as a gauge field,

$$B'_\mu = B_\mu + \partial_\mu \chi$$

Finally, the 4D part of the metric transforms as,

$$\tilde{g}'_{\mu\nu} = \tilde{g}_{\mu\nu} + \frac{\partial \tilde{x}^5}{\partial x'^\mu} \tilde{g}_{5\nu} + \frac{\partial \tilde{x}^5}{\partial x'^\nu} \tilde{g}_{\mu5} + \frac{\partial \tilde{x}^5}{\partial x'^\mu} \frac{\partial \tilde{x}^5}{\partial x'^\nu} \tilde{g}_{55} = \tilde{g}_{\mu\nu} + \partial_\mu \chi \tilde{g}_{5\nu} + \tilde{g}_{\mu5} \partial_\nu \chi + \tilde{g}_{55} \partial_\mu \chi \partial_\nu \chi$$

Using our introduced fields,

$$g'_{\mu\nu} + \phi'^2 B'_\mu B'_\nu = g_{\mu\nu} + \phi^2 \left[B_\mu B_\nu + \partial_\mu \chi B_\nu + B_\mu \partial_\nu \chi \right] = g_{\mu\nu} + \phi^2 B'_\mu B'_\nu$$

and thus the 4D space-time metric is left invariant under the gauge transformation.

$$g'_{\mu\nu} = g_{\mu\nu}$$

Therefore, rotating locally around the S^1 components only changes the gauge field B_μ,

$$\phi \mapsto \phi \qquad B_\mu \mapsto B_\mu + \partial_\mu \chi \qquad g_{\mu\nu} \mapsto g_{\mu\nu}$$

14.3 The Five-Dimensional Christoffel Symbols

Starting directly from the 5D metric, the Christoffel symbols are easily calculated in terms of their 4D counterparts. For the 4D components,

$$\tilde{\Gamma}^\alpha_{\mu\nu} = \frac{1}{2} \tilde{g}^{\alpha\beta} \left(\partial_\mu \tilde{g}_{\beta\nu} + \partial_\nu \tilde{g}_{\mu\beta} - \partial_\beta \tilde{g}_{\mu\nu} \right)$$

$$= \Gamma^\alpha_{\mu\nu} + \frac{1}{2} g^{\alpha\beta} \left(\partial_\mu \phi^2 B_\beta B_\nu + \partial_\nu \phi^2 B_\mu B_\beta - \partial_\beta \phi^2 B_\mu B_\nu \right) + \frac{1}{2} \tilde{g}^{\alpha5} \left(\partial_\mu \tilde{g}_{5\nu} + \partial_\nu \tilde{g}_{\mu5} \right)$$

$$= \Gamma^\alpha_{\mu\nu} + \phi^2 g^{\alpha\beta} \left(W_{\mu\beta} B_\nu + B_\mu W_{\nu\beta} - B_\mu B_\nu \partial_\beta \ln \phi^2 \right)$$

where I have used the cylindrical condition $\partial_5 g_{\mu\nu} = 0$ and defined the quantity,

$$W_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

Furthermore,

$$\tilde{\Gamma}^\alpha_{\mu5} = \frac{1}{2} \tilde{g}^{\alpha\beta} \left(\partial_\mu \tilde{g}_{\beta5} - \partial_\beta \tilde{g}_{\mu5} \right) = \frac{1}{2} g^{\alpha\beta} \phi^2 \left(\partial_\mu B_\beta - \partial_\beta B_\mu + B_\beta \partial_\mu \ln \phi^2 - B_\mu \partial_\beta \ln \phi^2 \right) + \frac{1}{2} \tilde{g}^{\alpha5} \partial_\mu \tilde{g}_{55}$$

$$= \frac{1}{2} g^{\alpha\beta} \phi^2 \left(W_{\mu\beta} - B_\mu \partial_\beta \ln \phi^2 \right)$$

and finally,

$$\tilde{\Gamma}^\alpha_{55} = -\frac{1}{2} \tilde{g}^{\alpha\beta} \partial_\beta \tilde{g}_{55} = -\frac{1}{2} g^{\alpha\beta} \partial_\beta \phi^2$$

where I can restrict to $\beta \neq 5$ by the cylindrical condition.

14.4 The Einstein-Hilbert Action for Kaluza-Klein Theory

To study the dynamics of the Kaluza-Klein metric, we need to calculate the Christoffel symbols and then the curvature tensor. This is a straightforward yet exceedingly tedious calculation so I will simply write down the results. We postulate a 5D Einstein-Hilbert action proportional to the simplest curvature invariant, R the Ricci curvature scalar. This is the exact same form as the action for 4D general relativity. Consider the action,

$$\tilde{S}_{KK} = \int \frac{1}{2\tilde{\kappa}} \tilde{R} \sqrt{\tilde{g}} \, \mathrm{d}^5 x$$

Therefore, we need to examine the form of the 5-dimensional quantity $\tilde{R}$. The expression for $\tilde{R}$ follows formally from the form of the metric and the Christoffel symbols,

$$\tilde{R} = R - \frac{1}{4} \phi^2 W_{\mu\nu} W^{\mu\nu} - \frac{2}{\phi} \partial_\mu \partial^\mu \phi$$

This is our first hint of the so called "Kaluza Miracle" since term,

$$\mathcal{L}_{\text{"EM"}} = -\frac{1}{4} W_{\mu\nu} W^{\mu\nu}$$

has exactly the form of the electromagnetic Lagrangian. Therefore, the action becomes,

$$\tilde{S}_{KK} = \int \left[\frac{1}{2\tilde{\kappa}} R - \frac{1}{8\tilde{\kappa}} W_{\mu\nu} W^{\mu\nu} \right] \phi \sqrt{g} \, \mathrm{d}^5 x - \int \partial_\mu \partial^\mu \phi \, \sqrt{g} \, \mathrm{d}^5 x$$

By the cylinder condition, none of these quantities depends on x^5 so we can integrate out by x^5. Suppose that C is the period of x^5 or equivalently the circumference of the compactified dimension. Then the Kaluza-Klein action becomes,

$$\tilde{S}_{KK} = \int \left[\frac{C}{2\tilde{\kappa}} R - \frac{C}{8\tilde{\kappa}} \phi^2 W_{\mu\nu} W^{\mu\nu} \right] \phi \sqrt{g} \, \mathrm{d}^4 x - C \int \partial_\mu \partial^\mu \phi \, \sqrt{g} \, \mathrm{d}^4 x$$

Now, one last bit of alchemy. Since the constant $\tilde{\kappa}$ is arbitrary, set,

$$\tilde{\kappa} = \frac{8\pi G C}{c^4}$$

Then we define the vector potential $A_\mu = \sqrt{\frac{C}{2\tilde{\kappa}}} B_\mu$ and the force tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Then,

$$F_{\mu\nu} F^{\mu\nu} = \frac{C}{2\tilde{\kappa}} \phi^2 W_{\mu\nu} W^{\mu\nu}$$

And at last, the Kaluza Miracle,

$$\tilde{S}_{KK} = \int \frac{c^4}{16\pi G} R \phi \sqrt{g} \mathrm{d}^4 x - \int \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \phi^3 \sqrt{g} \mathrm{d}^4 x - C \int \partial_\mu \partial^\mu \phi \sqrt{g} \mathrm{d}^4 n = S_{GR} + S_{EM} + S_{SF}$$

the Kaluza-Klein action becomes the action of general relativity plus the action of electromagnetism plus the action of a scalar field. However, there is one subtly, the GR and EM Lagrangians are multiplied by the scalar field ϕ. However, if ϕ is slowly varying then we can approximate it as a constant and absorb it as a constant multiple of the entire Lagrangian. When $\phi \approx \phi_0$ a constant, the action literally reduces to,

$$\tilde{S}_{KK} = S_{GR} + S_{EM}$$

14.5 The Field Equations

We have seen that the Einstein-Hilbert action for the 5D Kaluza-Klein theory gives rise to a 4D Einstein-Hilbert action and the Lagrangian of an electromagnetic-like theory. Varying this action with respect to the fields, we recover the Einstein field equations with an electromagnetic source term from varying the 4D metric, Maxwell's equations from varying the field A_μ and an equation of motion for the scalar field by varying ϕ. These results can also be obtained with a slightly different flavor by ruining the derivation with the opposite order. Since the Kaluza-Klein action is simple the 5D Einstein Hilbert action,

$$\tilde{S}_{KK} = \int \left[\frac{1}{2\tilde{\kappa}} \tilde{R} + \tilde{\mathcal{L}}_M \right] \sqrt{\tilde{g}}\, \mathrm{d}^5 x$$

varying with respect to the 5D metric $\tilde{g}_{\mu\nu}$ will recover a 5D version of the Einstein field equations with exactly the same form as its 4D general relativity counterpart,

$$\tilde{R}_{\mu\nu} - \frac{1}{2} \tilde{R} \tilde{g}_{\mu\nu} = \tilde{\kappa} \tilde{T}_{\mu\nu}$$

As before, I will restrict to 5D "vacuua" in which $T_{\mu\nu} = 0$. We shall consider the different components of the Einstein field equations. First, the 4D space-time part,

$$\tilde{R}_{\mu\nu} - \frac{1}{2} \tilde{R} \tilde{g}_{\mu\nu} = 0$$

gives when expanded in term of 4D quantities,

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} \phi^2 \left(F_{\mu\alpha} F_\nu{}^\alpha - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right) + \frac{1}{\phi} \left(\nabla_\mu \nabla_\nu \phi - g_{\mu\nu} \phi \right)$$

which is the 4D Einstein field equations with an electromagnetic energy-momentum tensor and a strange scalar field energy-momentum. Using the fact that,

$$\tilde{R}_{\mu\nu} - \frac{1}{2} \tilde{R} \tilde{g}_{\mu\nu} = 0 \implies \tilde{R}_{\mu\nu} = 0$$

by the standard trace trick then we get,

$$\tilde{R}^{5\mu} = 0 \implies \frac{1}{2} \nabla_\nu \left(\phi^3 F^{\mu\nu} \right) = 0$$

For a slowly varying ϕ field this tells us that,

$$\nabla_\nu F^{\mu\nu} = 0$$

which is the Maxwell equations. Finally,

$$\tilde{R}_{55} = 0 \implies \nabla_\mu \nabla^\mu \phi = \frac{1}{4} \phi^4 F_{\alpha\beta} F^{\alpha\beta}$$

which tells us that the electromagnetic field is the source of this strange new scalar field.

14.6 The Lorentz Force Law

We now want to consider the geodesics in this 5D Kaluza-Klein spacetime. In particular, we care about the 4D projections of these trajectories. Unfortunately, the full geodesic equation has many many terms. Specifically, we want to look at the terms contracted with $\tilde{\Gamma}^\mu_{\alpha\beta}$ and $\tilde{\Gamma}^\mu_{\alpha 5}$. We get,

$$\frac{\mathrm{d}\tilde{U}^\mu}{\mathrm{d}s} + \Gamma^\mu_{\alpha\beta}\tilde{U}^\alpha\tilde{U}^\beta + \sqrt{\frac{8\pi G}{c^4}}g^{\mu\nu}\phi^2\left(F_{\alpha\nu} - A_\alpha\partial_\nu\ln\phi^2\right)\tilde{U}^\alpha\tilde{U}^5 + O(A^3) = 0$$

Furthermore, the cylinder condition tells us that $\partial_5 g_{\alpha\beta} = 0$ therefore, the constant vector $\xi^\mu = (0,0,0,0,1)$ is a Killing vector so the quantity,

$$\xi^\mu\tilde{U}_\mu = \tilde{U}_5$$

is conserved along geodesics. Therefore, we can write,

$$\frac{\mathrm{d}\tilde{U}^\mu}{\mathrm{d}s} + \Gamma^\mu_{\alpha\beta}\tilde{U}^\alpha\tilde{U}^\beta = \sqrt{\frac{8\pi G}{c^4}}\,\phi^2\,\tilde{U}^5\,F^\mu{}_\alpha\tilde{U}^\alpha + O(A^2)$$

This is the Lorentz force law if we set,

$$\frac{q}{mc} = \sqrt{\frac{8\pi G}{c^4}}\tilde{U}^5$$

The velocity in the compactified dimension is the charge to mass ratio of the particle!

14.7 Charge Quantization

Now we add a tad bit of quantum mechanics! We have particles constrained to a closed periodic circle in the x^5 direction. Therefore, a quantum particle cannot have an arbitrary momentum in the x^5 direction. If we want to have definite momentum (corresponding to a state with definite charge) then we need to have a standing wave in the x^5 direction. Therefore we need to have C be an integer number of wavelengths $n\lambda$. Thus,

$$\lambda = \frac{h}{p} = \frac{h}{m\tilde{U}^5} = \frac{C}{n}$$

Therefore, solving for $\tilde{U}^5$ we arrive at a condition on the charge,

$$q = \frac{n}{C}\frac{h\sqrt{8\pi G}}{c}$$

and therefore charge is quantized in units of,

$$q_Q = \frac{1}{C}\frac{h\sqrt{8\pi G}}{c}$$

15 Appendix A. The Calculus of Variations