Tensor Calculus

A Beginner's Introduction

Barsam Rahimi

Department of Materials Engineering and Physics
University of Toronto
August 2024

Note: Throughout this handout, I will be assuming that you are familiar with vectors, partial derivatives, sigma notation, cylindrical and spherical coordinates, and the Jacobian matrix.

1 Preliminary Definitions and Vectors

The typical way of defining a vector in high school and even in your first 2-3 years of university is an object that consists of a magnitude and direction whose components can be expressed in matrix form or bracket form. For example, some vector $\vec{A}$ with components A_x, A_y, and A_z can be expressed in either of the following forms:

$$\vec{A} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \qquad\qquad \vec{A} = \langle A_x, A_y, A_z \rangle$$

Or if you have any background in Quantum Mechanics, then you would also know that in a Hilbert space, any ket vector is represented by a column matrix and bra vectors are represented by row matrices;

$$|A\rangle = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \qquad\qquad \langle A| = \begin{pmatrix} A_x & A_y & A_z \end{pmatrix}$$

Where the inner product of the two vectors A and B is defined as:

$$\langle A|B\rangle$$

This is known as Dirac Notation and is a very handy notation. Back to defining vectors now!
So far, this is how we have all learned to define a vector; a magnitude and direction. However, what if we can define a vector in a different way ? Would there be any benefit to it ? How ?
To do that, first we need to introduce some notation.
Consider our ordinary unit vectors $\hat{\imath}, \hat{\jmath}, \hat{k}$. What is special about them ? Well, they are orthonormal. But what if we want to work in higher dimensions, or in different axis where these unit vectors are

useless ? To generalize this idea, we will introduce the following notation for unit vectors:

$$\hat{\boldsymbol{e}}_1 = \hat{i}, \quad \hat{\boldsymbol{e}}_2 = \hat{j}, \quad \hat{\boldsymbol{e}}_3 = \hat{k}$$

Now, we can express our vector A as:

$$\vec{A} = A_x \hat{\boldsymbol{e}}_1 + A_y \hat{\boldsymbol{e}}_2 + A_z \hat{\boldsymbol{e}}_3$$

To make this more compact, we can change up our notation a bit. Now, we will introduce $\hat{\boldsymbol{e}}_i$ to represent $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2$ and $\hat{\boldsymbol{e}}_3$ so:

$$\hat{\boldsymbol{e}}_m = \{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3\}$$

Another way to use this notation is defining:

$$x^i = \{x, y, z\}$$

Now what if our basis vectors are not unit vectors ? Then

$$\vec{\boldsymbol{e}}_n = \{\vec{\boldsymbol{e}}_1, \vec{\boldsymbol{e}}_2, \vec{\boldsymbol{e}}_3\}$$

So we can define our vector, $\vec{A}$ as:

$$\vec{A} = \sum_{i=1}^{3} A^i \vec{\boldsymbol{e}}_i$$

Again, where

$$A^i = \{A^1, A^2, A^3\} = \{A^x, A^y, A^z\}$$

So remember that when we say 1, we really mean the x component not being raised to the power of 1. Same thing with 2 and y, and 3 and z respectively. Now, we can define the dot product of two vectors A and B as:

$$\vec{A} \cdot \vec{B} = \left(\sum_{n=1}^{3} A^n \vec{\boldsymbol{e}}_n \right) \left(\sum_{m=1}^{3} B^m \vec{\boldsymbol{e}}_m \right) = \sum_{n=1}^{3} \sum_{m=1}^{3} (\vec{\boldsymbol{e}}_n \cdot \vec{\boldsymbol{e}}_m) A^n B^m$$

Here we come across an extremely important object which we will come back to again. The term in the parenthesis is the definition of an object called the *metric tensor*.

$$g_{mn} \equiv \vec{\boldsymbol{e}}_m \cdot \vec{\boldsymbol{e}}_n$$

We will come back and explore this object later. In Cartesian coordinates the metric is equal to something called the *Kronecker Delta Symbol*, where it is defined as

$$\delta_{mn} \equiv \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$$

We can also think of the Kronecker delta as a matrix:

$$\delta_{mn} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

But now let's think about it, if the dot product of the basis are 1, then that means they have to be orthogonal unit vectors. So it is very easy to define the Kronecker delta like the metric:

$$\delta_{mn} = \hat{\boldsymbol{e}}_m \cdot \hat{\boldsymbol{e}}_n$$

So the ordinary dot product is defined as:

$$\vec{A} \cdot \vec{B} = \sum_{n=1}^{3} \sum_{m=1}^{3} \delta_{mn} A^n B^m$$

And if we are working with non-orthonormal basis vectors we simply replace the Kronecker delta with the metric;

$$\vec{A} \cdot \vec{B} = \sum_{n=1}^{3} \sum_{m=1}^{3} g_{mn} A^n B^m$$

These objects we will come back to. Before we learn about tensors however, we must learn about how vectors transform. This will allow us to easily define tensors.

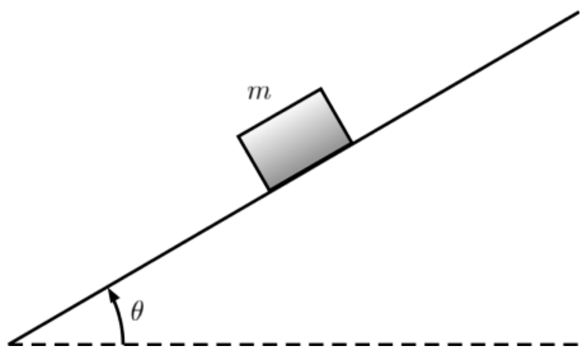
2 Transformations: A New Look At Vectors

In this section, I aim to change your understanding of the way we define an object like a vector. I will go ahead an throw the definition right at you, and then we will examine a case that will make this very easy for you to understand.

A vector is any object that transforms from some coordinate system, x^i, to another one, x^j, according to the following transformation rule:

$$(V')^j = \sum_{i=0}^{3} \frac{\partial x^j}{\partial x^i} V^i$$

There is also a version in which the indices are downstairs but we will look at that later. This is how we will also define tensors. But to get a little more confident with what this means, let us explore an example of a mass on an incline surface.



For an object at rest with friction, the components of the net force are equivalent to the following expression:

$$F^x = mg\sin\theta - f_{fr} \qquad F^y = mg\cos\theta - N$$

However, this is only if we consider this system with our x axis on the inclined surface and our y axis perpendicular to the incline surface. What if we consider an alternate set of axis ? Let us solve this problem again but with a different set of axis. This time, we will set our axis such that the horizontal ground is the x axis and the y axis is a vertical line, perpendicular to the ground. This time, the components of the force will have different forms:

$$F^{x'} = mg\sin\theta - f_{fr} \qquad F^{y'} = mg - N\cos\theta$$

Both of these solutions work and that's a problem; how can we have two different solutions that give the same answer to the same physical phenomena ? There must be an explanation! The answer is being invariant under transformations. Let us map out the transformation rule: $(x, y) \to (x', y')$:

$$
\begin{aligned}
x' &= x\cos\theta + y\sin\theta \\
y' &= -x\sin\theta + y\cos\theta
\end{aligned}
$$

Of course, these are rotational matrices and this makes physical sense since we are rotating around the (invisible) z axis. Now what good does this do ?

These two sets of coordinates will only give us the same results if the space-time interval between two infinitesimally small points is equal. In other words,

$$(dS)^2 = (dS')^2$$

From the rotation rule we get:

$$
\begin{aligned}
dx' &= dx\cos\theta + dy\sin\theta \\
dy' &= -dx\sin\theta + dy\cos\theta
\end{aligned}
$$

and of course we also have $dz' = dz$ but we do not need to worry about that. Now, let us see if the *invariance of spacetime interval* holds true:

$$
\begin{aligned}
(dS')^2 &= (dx')^2 + (dy')^2 + (dz')^2 \\
&= (dx\cos\theta + dy\sin\theta)^2 + (-dx\sin\theta + dy\cos\theta)^2 + dz^2 \\
&= dx^2\cos^2\theta + dy^2\sin^2\theta + dx^2\sin\theta + dy^2\cos^2\theta + dz^2 \\
&= dx^2 + dy^2 + dz^2 = dS^2
\end{aligned}
$$

Since we have proven that the spacetime interval is invariant under this transformation rule, we have proven that these two sets of solutions actually do represent the same thing! In other words, we have proven that since F transforms according to the transformation rule for vectors, it is a vector! That is the same we will define tensors.

3 Tensors: Definitions and The Metric

3.1 Defining Tensors

Conceptually, a tensor is a generalization of scalars and vectors. A scalar is also called a rank 0 tensor, vectors are rank 1 tensors but now we are interested in tensors of higher ranks. Let us take some rank 2 tensor, T^{mn}. This object is only a tensor if it follows the transformation rule for all tensors of 2nd rank. If the following holds true about T, then it is a tensor;

$$(T^{pq})' = \frac{\partial x^p}{\partial x^m}\frac{\partial x^q}{\partial x^n}T^{mn} \qquad\qquad (T_{pq})' = \frac{\partial x^m}{\partial x^p}\frac{\partial x^n}{\partial x^q}T_{mn}$$

We will discuss the case where there are indices downstairs later. Tensors are not very easy to understand this way. Geometrically, a rank 2 tensor can be thought of as an object that assigns a set of basis to a point and a tensor field assigns a different set of basis to each point in the space. So you can think of rank 2 tensors geometrically as this field of basis. They are particularly good in physics because they have the property that they are invariant under a coordinate transformation, meaning they will always obey their transformation rule and do not change between reference frames. Some examples of tensors in physics are the Inertia tensor, Electromagnetic Field Tensor, Stress-strain Tensor, etc. but we will focus on the metric tensor which we defined earlier. We defined the metric tensor to be the dot product of our basis vectors. But how can we use it ?

3.2 The Generalized Pythagorean Theorem

Consider Pythagoras' Theorem but in differential form:

$$dS^2 = dx^2 + dy^2 + dz^2$$

Now let us consider the form of this formula when we write it out in different coordinates, in this case using spherical and cylindrical coordinates:

$$\begin{aligned} dS^2 &= dx^2 + dy^2 + dz^2 \\ &= dr^2 + rd\theta^2 + dz^2 \\ &= dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \end{aligned}$$

If we consider the units, on the left hand side, we need to have some distance squared and therefore, so should we on the right. However, differentials like $d\theta$ and $d\phi$ are not distances and so we need scaling factors to help fix their dimension and to make the theorem work correctly. But here is the problem, how can we have an object that fixes the differentials that need fixing but leave the other ones as they are ? The answer is that this information is encoded within the metric tensor! And we can see that from our derivation of the generalized dot product. So, the generalized Pythagorean theorem, using the notation we introduced in section 1, is:

$$dS^2 = \sum_{ij} g_{ij} dx^i dx^j = g_{ij} dx^i dx^j$$

You may have noticed that the second expression is without a summation sign. This is where we will bring the Einstein Summation Convention. This convention tells us that whenever we have an index repeating upstairs and downstairs, we assume that it is being summed over and we no longer need to write the summation sign for it. This will make things a lot easier in the subject of Geodesics.

Back to our discussion, let us write out a couple of metrics. For the metric of Cartesian coordinates, we can see the coefficients are all 1, so the metric is equivalent to the Kronecker Delta. What about cylindrical coordinates ?

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

But what if we are not given the Pythagorean theorem ? How can we calculate the components of the metric ? Let us try this with spherical coordinates. If we write some vector $\vec{r}$ in Cartesian coordinates and then apply the transformation rule for spherical coordinates:

$$\begin{aligned} \vec{r} &= x\hat{x} + y\hat{y} + z\hat{z} \\ &= r\sin\theta\cos\phi\hat{x} + r\sin\theta\sin\phi\hat{y} + r\cos\theta\hat{z} \end{aligned}$$

And now constructing our basis vectors, we know that $\vec{\boldsymbol{e}}_i = \partial\vec{r}/\partial x^i$. Performing the calculations we get:

$$\vec{\boldsymbol{e}}_r = \frac{\partial\vec{r}}{\partial r} = \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + \cos\theta\hat{z}$$

$$\implies g_{rr} = \sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta = 1$$

Now, before we move on, it is important to remember since our basis are orthonormal, the components of the metric that are not on the diagonal will be 0. So there is no need to calculate those. However, in more curvy and complicated coordinate systems some of these components may not be zero and so we will be able to make use of the non-diagonal components of the metric. Now, let us calculate the other 2 diagonal components:

$$\vec{\boldsymbol{e}}_\theta = \frac{\partial\vec{r}}{\partial\theta} = r\cos\theta\cos\phi\hat{x} + r\cos\theta\sin\phi\hat{y} - r\sin\theta\hat{z}$$

$$\implies g_{\theta\theta} = r^2\cos^2\theta\cos^2\phi + r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta = r^2$$

$$\vec{\boldsymbol{e}}_\phi = \frac{\partial\vec{r}}{\partial\phi} = -r\sin\theta\sin\phi\hat{x} + r\sin\theta\cos\phi\hat{y} + 0\hat{z}$$

$$\implies g_{\phi\phi} = r^2\sin^2\theta\sin^2\phi + r^2\sin^2\theta\cos^2\phi = r^2\sin^2\theta$$

And for $i \neq j$, we have

$$g_{ij} = 0$$

And there we have our spherical metric:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

So we know that one way we can use the metric is defining a generalized Pythagorean theorem which allows us to tell whether the spacetime interval will be invariant under a coordinate transformation;

$$dS^2 = g_{ij}dx^i dx^j$$

3.3 A Note On The Kronecker Delta; Is the Metric a Tensor ?

We have said, previously, that the metric is a tensor. However, we have not proved it yet. How can we do that ? We will go back to the definition; if it transforms as a tensor, then it is a tensor. If it is a tensor, then:

$$g_{\mu\nu} = \frac{\partial x^\alpha}{\partial x^\mu}\frac{\partial x^\beta}{\partial x^\nu} g_{\alpha\beta}$$

You may have noticed that the transformation rule we used is slightly different, we will come back to this later to find out the difference between having indices downstairs vs upstairs but for now let us stick to this definition. Now, we will substitute in this into the Pythagorean theorem to see if it will hold:

$$\begin{aligned}
(dS')^2 &= dS^2 \\
\implies \frac{\partial x^\alpha}{\partial x^\mu}\frac{\partial x^\beta}{\partial x^\nu} g_{\alpha\beta} dx^\alpha dx^\beta &= g_{\mu\nu} dx^\mu dx^\nu \\
\implies \left(g_{\mu\nu} - \frac{\partial x^\alpha}{\partial x^\mu}\frac{\partial x^\beta}{\partial x^\nu} g_{\alpha\beta} \right) dx^\mu dx^\nu &= 0 \\
\therefore g_{\mu\nu} &= \frac{\partial x^\alpha}{\partial x^\mu}\frac{\partial x^\beta}{\partial x^\nu} g_{\alpha\beta}
\end{aligned}$$

Since this statement holds up the invariance of spacetime intervals, the metric must be a tensor. Now, let us consider the case where we are not really transforming between two different systems but rather the same coordinate system. In that case:

$$g'_{\alpha\beta} = \frac{\partial x^\alpha}{\partial x^\beta}\frac{\partial x^\beta}{\partial x^\alpha} g_{\alpha\beta}$$

Where both alpha and beta are being summed over. But we also know that the metric should not change in the same coordinate system! Therefore the transformation coefficient must be 1 when alpha and beta are equal, and 0 when they are not. This is equivalent to the Kronecker Delta!

$$\delta_{\alpha\beta} = \frac{\partial x^\alpha}{\partial x^\beta}\frac{\partial x^\beta}{\partial x^\alpha}$$

This is also a great way to see why the Kronecker delta is not a tensor.

4 Contravariant vs. Covariant

So far, we have been using the Upstairs and Downstairs terminology for indices. But what actually is the difference ? What makes an upstairs index different from downstairs index ? The question we must ask ourselves is what happens to a vector's components as we change the magnitude of the basis vectors ? For example, would our vector seem twice as long if we make the basis half as long, would it be the same size as before or would it also shrink with the basis ?

If the component compresses along with our basis it is a covariant component denoted by a downstairs index. On the other hand, if the component expands, contrary to the basis, then it is a contravariant component denoted by an upstairs index.

Okay, now that we have tackled the geometric concept, let us look at the mathematical definition. This will be quite helpful. Let us look back at the Pythagorean theorem:

$$dS^2 = g_{\mu\nu} dx^\mu dx^\nu$$

If we isolate one part of it;

$$g_{\mu\nu} dx^\mu$$

We know from the summation convention that the index μ is being summed over. Meaning that in the end, our object will have only 1 downstairs index. Hence, we have found the definition of the covariant differential with respect to the contravariant differential:

$$dx_\nu = g_{\mu\nu} dx^\mu$$

Now what about defining the contravariant component in terms of the covariant component ? We cannot just divide by the metric. Instead we must apply the inverse metric to both sides in order to get the correct index form:

$$g_{\mu\nu} \otimes g^{\lambda\nu} = \delta^\lambda_\mu$$

So we get that:

$$dx^\mu = dx_\nu g^{\mu\nu}$$

Now how can we describe this using "ordinary" vector components ? To do that, let us construct some vector A:

$$\vec{\tilde{A}} = \tilde{A}_r\,\hat{r} + \tilde{A}_\theta\,\hat{\theta} + \tilde{A}_z\,\hat{z}$$

Recalling that metric is diagonal and its elements are either 0 or some positive value (in this case). Hence we can re-write the metric as a product of some coefficients h_μ squared multiplied by the Kronecker Delta;

$$g_{\mu\nu} = (h_\mu)^2\delta_{\mu\nu}$$

It is important to note that we are not summing over any index. Using this, we get that the contravariant component is related to the ordinary component through the following relation:

$$A^\mu = \frac{\tilde{A}_\mu}{h_\mu}$$

This is easy to prove by simply expanding the dot product of two vectors and re-writing the metric in this way. What about the covariant components ?

$$A_\mu = \sum_\nu g_{\mu\nu}A^\nu = \sum_\nu (h_\mu)^2\delta_{\mu\nu}\frac{\tilde{A}_\nu}{h_\nu}$$

$$= (h_\mu)^2\left(\delta_{\mu 1}\frac{\tilde{A}_1}{h_1} + \delta_{\mu 2}\frac{\tilde{A}_2}{h_2} + \delta_{\mu 3}\frac{\tilde{A}_3}{h_3}\right)$$

And now considering each component one by one, we see that:

$$A_\mu = h_\mu\tilde{A}_\mu$$

This question now becomes, how can we make money out of this ? What is the point ?

5 Derivatives of Tensors and Tensor Algebra

5.1 Directional Derivatives of Tensors

It is very easy to define the derivative of a tensor. Say we want to differentiate some tensor $T_{\mu\nu}$ with respect to some coordinate, x^μ. Then the derivative is very easy to write:

$$\frac{\partial T_{\mu\nu}}{\partial x^\mu}$$

But how do we actually compute this ? As we know, different coordinates have different coefficients that have to be taken into account. Hence, we must find some definition that will fix this problem. To do that we define:

$$\partial_\mu = h_\mu \nabla_\mu \quad \Longrightarrow \quad \nabla_\mu = \frac{1}{h_\mu}\partial_\mu$$

Where ∂_μ is a shorthand notation for $\partial/\partial x^\mu$ that we will use more and more from now. So if we consider the metric of spherical coordinates as an example, then:

$$\begin{aligned}
\nabla_r &= \partial_r \\
\nabla_\theta &= \frac{1}{r}\partial_\theta \\
\nabla_\phi &= \frac{1}{r\sin\theta}\partial_\phi
\end{aligned}$$

That is how we compute directional derivatives of tensors. Now, let us look at some basic rules of tensor algebra which we will make use of later. It is also worthwhile noting that

$$\partial^\mu = g^{\mu\nu}\partial_\nu \qquad\qquad \nabla^\mu = g^{\mu\nu}\nabla_\nu$$

Even though they are operators, they are related just like contravariant and covariant components.

5.2 Tensor Algebra

Note that for some mixed tensor T^α_β, the quantity $\alpha + \beta$ is known as the tensor's rank. The rule of scalar multiplication of tensors states that for some tensor T and scalar α:

$$\alpha T^\alpha_\beta = S^\alpha_\beta$$

Where the elements of S are all elements of T scaled by a factor of α. The addition property states that for two tensors T and S:

$$T^{\alpha}_{\beta} + S^{\alpha}_{\beta} = A^{\alpha}_{\beta}$$

Where the elements of A are just the sum of the elements of T and S. This next one is an interesting one. Tensor contraction is the property where for two tensors with 1 or more repeating index where one is up and one is down;

$$T^{\alpha\gamma} S_{\gamma\beta} = R^{\alpha}_{\beta}$$

The logic here is that the index γ is just summed over so it will be incorporated as a scalar within R's elements. Hence, it will be eliminated from the indices. This is tensor contraction and it is in a way a generalization of the dot product. With the dot product, we had:

$$V_i W^i = S$$

where S is some scalar. Here we have contracted the index i.
Finally, the tensor (outer) product states that for tow tensors T and S:

$$T^{\alpha\beta} S^{\gamma\lambda} = R^{\alpha\beta\gamma\lambda}$$

The same thing can also be done if there are tensors of mixed covariant and contravariant indices or solely covariant indices. Here:

$$(\alpha + \beta) + (\gamma + \lambda) = \text{rank of } R$$

The tensor product or outer product can also be denoted by:

$$T^{\alpha\beta} \otimes S^{\gamma\lambda} = R^{\alpha\beta\gamma\lambda}$$

The last property of tensors is symmetry and anti-symmetry. If a tensor is such that:

$$S_{mn} = S_{nm}$$

Then it is symmetric. Similarly, if it has the property that:

$$S_{mn} = -S_{nm}$$

Then it is anti-symmetric. The switching of the indices is another way of saying the transpose of the original tensor. So this could also be re-written as:

$$\begin{aligned} \text{if} \quad S = S^T \implies & \quad S \text{ is symmetric} \\ \text{if} \quad S = -S^T \implies & \quad S \text{ is anti-symmetric} \end{aligned}$$

An interesting note for anti-symmetric tensors is that if $\mu = \nu$ then;

$$T^{\mu\mu} = -T^{\mu\mu} \implies T^{\mu\mu} = 0$$

So for all anti-symmetric cases, the diagonal components are 0. A general rule to keep in mind is that if an index is ever being summer over, you can change it to whatever other index you want to simplify the expression. Here is an example: Prove that R^{μ}_{μ} is a scalar.
We know that if it is a scalar, then it must obey its transformation rule and the transformation coefficients must be 1 or simplify to the Kronecker delta.

$$(R')^{\mu}_{\mu} = \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} R^{\rho}_{\sigma} = \delta^{\sigma}_{\rho} R^{\rho}_{\sigma} = R^{\rho}_{\rho}$$

which proves that R is indeed a scalar. Now that we have all these tools to work with, we must address a quite big problem with the directional derivative of tensors. The issue with it is that while it is "a" derivative of the tensor, we cannot expect it to also be a tensor.

6 The Covariant Derivative

So far, we have constructed our first important tool which will come up frequently in all forms of tensor calculus. Now comes the second tool which guarantees us that the derivative of our tensor will indeed transform as a tensor. But how can we achieve that ?

6.1 The Affine Connection and The Geodesic Equation

To construct the tool that we need to define our tensor derivative, first we have to look at the motion of objects. From now on, when we use Latin indices, we are summing over spatial components only. On the other side, when we use Greek indices, it means we are also summing over the time component so from 0 to 3. From now on, we will no longer use Latin indices as we wish to work in all 4 dimensions. Also, the Pythagorean theorem from now on will take the form:

$$c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu \qquad dS^2 = -d\tau^2$$

And sometimes we will use the Minkowski metric, $\eta_{\mu\nu}$, which is the metric in special relativity. Let us construct our equations of motion from the very first thought that started the entire theory of relativity, that gravity and acceleration are the same thing, that there is no experiment that can be done for some one in a spaceship in deep space to tell if they are in the middle of deep space accelerating at 9.8m/s or if they are on earth, stationary with respect to the ground. What this implies is that our second derivative of position is equal to zero in all reference frames, since if it was not, then we would not be in an inertial reference frame. Let our position vector be x^α. Then the equivalence principle states that

$$\frac{d^2 x^\alpha}{d\tau^2} = 0$$

Now we will use the chain rule to expand the expression into:

$$\frac{d^2 x^\alpha}{d\tau^2} = \frac{d}{d\tau}\left(\frac{dx^\alpha}{d\tau}\right) = \frac{d}{d\tau}\left(\frac{\partial x^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\right)$$

And then we go ahead and use our handy product rule:

$$\frac{d}{d\tau}\left(\frac{\partial x^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\right) = \frac{\partial x^\alpha}{\partial x^\mu}\frac{d^2 x^\mu}{d\tau^2} + \frac{d}{d\tau}\left(\frac{\partial x^\alpha}{\partial x^\mu}\right)\frac{dx^\mu}{d\tau}$$

Expanding the expression in the parenthesis with the chain rule again, we get

$$\frac{\partial x^\alpha}{\partial x^\mu}\frac{d^2 x^\mu}{d\tau^2} + \frac{d}{d\tau}\left(\frac{\partial x^\alpha}{\partial x^\mu}\right)\frac{dx^\mu}{d\tau} = \frac{\partial x^\alpha}{\partial x^\mu}\frac{d^2 x^\mu}{d\tau^2} + \frac{\partial^2 x^\alpha}{\partial x^\mu \partial x^\nu}\frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau} = 0$$

Now, we will multiply the whole thing by a 1 or it's analog which is a Kronecker delta to get:

$$\delta_\mu^\lambda \frac{d^2 x^\mu}{d\tau^2} + \frac{\partial x^\lambda}{\partial x^\alpha} \frac{\partial^2 x^\alpha}{\partial x^\mu \partial x^\nu} \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} = 0$$

$$\frac{d^2 x^\lambda}{d\tau^2} + \frac{\partial x^\lambda}{\partial x^\alpha} \frac{\partial^2 x^\alpha}{\partial x^\mu \partial x^\nu} \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} = 0$$

Now let us look at the first 2 derivatives in the second term. We are summing over α so its corresponding symbol should not have an α in it. There is also no summation for μ, ν, λ. So if we were to introduce a symbol to represent this part, it must have a λ index in the numerator, and a μ and ν in the denominator. We let the letter capital gamma represent this term:

$$\Gamma_{\mu\nu}^\lambda \equiv \frac{\partial x^\lambda}{\partial x^\alpha} \frac{\partial^2 x^\alpha}{\partial x^\mu \partial x^\nu}$$

This gives us what is called the geodesic equation;

$$\frac{d^2 x^\lambda}{d\tau^2} = -\Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}$$

This equation tells us the path that particles take in all sorts of curvature in spacetime. The symbol Γ is known as an "affine connection". When we take derivatives of tensors, problems arise, we cannot expect the result to be a tensor. There are problems with curl and divergence too. It turns out that if the affine connection is not 0, then the derivative does not obey the tensor transformation rule. This brings up the question, is the affine connection itself a tensor ? Well, if it is a tensor, then it must transform according to the tensor transformation rule. Hence, let us consider a primed affine connection and see how it is related to the unprimed frame;

$$\Gamma_{\mu\nu}^{\prime\lambda} = \frac{\partial x^{\prime\lambda}}{\partial x^\alpha} \frac{\partial^2 x^\alpha}{\partial x^{\prime\mu} \partial x^\nu}$$

Now we use chain rule to transform to unprimed frame

$$\begin{aligned} x^{\alpha} &= x^{\alpha}(x^{\sigma}) \\ x'^{\mu} &= x'^{\mu}(x^{\rho}) \\ \implies \Gamma'^{\lambda}_{\mu\nu} &= \left(\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x^{\alpha}}\right)\frac{\partial}{\partial x'^{\mu}}\left(\frac{\partial x^{\alpha}}{\partial x^{\sigma}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\right) \\ &= \left(\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x^{\alpha}}\right)\left(\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\frac{\partial}{\partial x^{\sigma}}\frac{\partial x^{\alpha}}{\partial x^{\eta}}\frac{\partial x^{\eta}}{\partial x'^{\mu}} + \frac{\partial x^{\alpha}}{\partial x^{\sigma}}\frac{\partial^2 x^{\sigma}}{\partial x'^{\mu}\partial x'^{\nu}}\right) \\ &= \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x^{\alpha}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\frac{\partial^2 x^{\alpha}}{\partial x^{\sigma}\partial x^{\eta}}\frac{\partial x^{\eta}}{\partial x'^{\mu}} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x^{\alpha}}\frac{\partial x^{\alpha}}{\partial x^{\sigma}}\frac{\partial^2 x^{\sigma}}{\partial x'^{\mu}\partial x'^{\nu}} \end{aligned}$$

Now let us examine each term individually. Let us look at term 1 first and see if we can manipulate it to look more familiar. We begin by moving the second fraction beside the 4th fraction to get:

$$\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\eta}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\frac{\partial x^{\rho}}{\partial x^{\alpha}}\frac{\partial^2 x^{\alpha}}{\partial x^{\sigma}\partial x^{\eta}}$$

And now we see that the last two fractions are the definition of the affine connection with indices of alpha, eta and sigma. Hence we get that the first term is equal to:

$$\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\eta}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\Gamma^{\rho}_{\sigma\eta}$$

So far this seems consistent with a tensor transformation. But what about the second term ? The second term's middle 2 fractions are simply a Kronecker delta. Hence:

$$\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x^{\alpha}}\frac{\partial x^{\alpha}}{\partial x^{\sigma}}\frac{\partial^2 x^{\sigma}}{\partial x'^{\mu}\partial x'^{\nu}} = \delta^{\rho}_{\sigma}\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial^2 x^{\sigma}}{\partial x'^{\mu}\partial x'^{\nu}} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial^2 x^{\sigma}}{\partial x'^{\mu}\partial x'^{\nu}}$$

And so we get that the general transformation of an affine connection obeys the following rule:

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\eta}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\Gamma^{\rho}_{\sigma\eta} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial^2 x^{\sigma}}{\partial x'^{\mu}\partial x'^{\nu}}$$

As we can see, this proves that the affine connection does not transform as a tensor, and hence is not a tensor. An alternate approach to the transformation rule is following the Kronecker delta's definition:

$$\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x'^{\nu}} = \delta^{\lambda}_{\nu} \qquad\qquad \frac{\partial}{\partial x'^{\mu}}\left(\frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x'^{\nu}}\right) = 0$$

$$\implies \frac{\partial x'^\lambda}{\partial x^\rho}\frac{\partial^2 x^\rho}{\partial x'^\mu \partial x'^\nu} + \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\nu} = 0$$

$$\implies \Gamma'^\lambda_{\mu\nu} = \frac{\partial x'^\lambda}{\partial x^\rho}\frac{\partial x^\eta}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}\Gamma^\rho_{\sigma\eta} - \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\nu}$$

This is an alternate transformation rule for affine connections. You may also see the affine connection being called a Christoffel symbols. We will see a different definition later and we will call that the Christoffel symbol. However, we must first learn about the covariant derivative. This will be a quite important tool. But first, a note on Jacobians.

6.2 Relating the Jacobian to the Metric

As we know, Jacobians are used in order to transform between coordinate systems when solving integrals. For example:

$$\iint f(x,y)\,dxdy = \iint f(r,\theta)r\,drd\theta$$

Where the r term comes from the Jacobian matrix such that:

$$dx'dy' = \mathcal{J}\,dxdy$$

Where

$$\mathcal{J} = \left|\frac{\partial x'}{\partial x}\right|$$

For example, in polar coordinates, the Jacobian will be:

$$\mathcal{J}(r,\theta) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r$$

This idea seems very familiar to the concept of the metric. So naturally, we ask, is there a connection between the two ? Let us see!

We know that the metric's transformation rule is:

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

Now, let us take the determinant of both sides

$$|g'| = \left|\frac{\partial x}{\partial x'}\right| \left|\frac{\partial x}{\partial x'}\right| |g| = \mathcal{J}^{-2}|g|$$

and therefore we get that the Jacobian is related to the metric through the relation

$$\mathcal{J} = \sqrt{\frac{|g|}{|g'|}}$$

This shows that the Jacobian can be described as a quotient of the metric and its primed version, and if you substitute in the transformation rule, then you get back the definition of the Jacobian.

6.3 The Covariant Derivative

Earlier in this section, we discussed taking directional derivatives of tensors. We also discussed the issue with them; you cannot expect the directional derivative of a tensor to be a tensor as well. This is going to cause lots of problems if we do not fix it. The issue is the main motivation behind the covariant derivative. How can we make a notion of the derivative which still includes all the information about the directional derivative, but also guarantees that the derivative of the tensor also transforms as a tensor ?

To answer this question, we must examine the notion of differentiation. What does it mean to differentiate ? We are looking at an object and comparing its value at 2 different points. For vectors in Euclidean space, we translate one vector to the origin of the second point and then compare the two. In curved space, however, the basis are not constant. So how should we go about this ? Well, if we extend the direction of the vector, we get a tangent vector. To use this, we must be aware of some problems. To compare the two vectors, we have to move one to the origin of the other but the basis are not constant so the differentiation must also involve differentiating the basis. On the other hand, the path we take from one point to another matters in curved space. Different paths will change the vector differently. Hence, we must rely on parallel transporting. Lastly, the tangent spaces of the two vectors are different. How do we compare vectors in different spaces ?

When there is curvature, we need to incorporate affine connections into the derivative. Let us begin by taking the derivative of some vector, A'^{λ}, with respect to coordinate x'^{μ} in terms of unprimed coordinates, given that $A'\mu$ is a four vector;

$$\begin{aligned}\frac{\partial A'^{\lambda}}{\partial x'^{\mu}} &= \frac{\partial}{\partial x'^{\mu}}\left(\frac{\partial x'^{\lambda}}{\partial x^{\nu}}A^{\nu}\right)\\ &= \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial}{\partial x^{\sigma}}\left(\frac{\partial x'^{\lambda}}{\partial x^{\nu}}A^{\nu}\right)\\ &= \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma}\partial x^{\nu}}A^{\nu} + \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial x'^{\lambda}}{\partial x^{\nu}}\frac{\partial A^{\nu}}{\partial x^{\sigma}}\end{aligned}$$

Here, the second term is the transformation coefficient of some second rank tensor and the first term looks similar to term 2 of the alternate transformation of our affine connection. The question is, can we somehow add it to our definition of the derivative ?

Since $\partial'_{\mu}A'^{\lambda}$ has 2 indices, we need to contract one of the indices of the right hand side's affine connection's bottom indices. Hence, we define the covariant derivative to be:

$$\nabla'_{\mu}A'^{\lambda} \equiv \partial'_{\mu}A'^{\lambda} + \Gamma'^{\lambda}_{\mu\nu}A'^{\nu}$$

There are 2 main notations for the covariant derivative; $\nabla_{\mu}A^{\nu}$ and $D_{\mu}A^{\nu}$ but for the sake of consistency, we will use the nabla symbol to denote the covariant derivative. Our initial goal was to construct a derivative of the tensor that transforms as a tensor. Let us observe how this quantity transforms to see if it is consistent with our goal;

$$\begin{aligned}\nabla'_{\mu}A'^{\lambda} &= \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma}\partial x^{\nu}}A^{\nu} + \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial x'^{\lambda}}{\partial x^{\nu}}\partial_{\sigma}A^{\nu} + \Gamma'^{\lambda}_{\mu\nu}A'^{\nu}\\ &= \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma}\partial x^{\nu}}A^{\nu} + \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial x'^{\lambda}}{\partial x^{\nu}}\partial_{\sigma}A^{\nu} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\eta}}{\partial x'^{\mu}}\frac{\partial x^{\alpha}}{\partial x'^{\nu}}\Gamma^{\rho}_{\alpha\eta}\frac{\partial x'^{\nu}}{\partial x^{\beta}}A^{\beta} - \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho}\partial x^{\alpha}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\rho}}{\partial x'^{\nu}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}A^{\beta}\end{aligned}$$

Now, I know this may look like it is very tedious and impossible to make something out of. However, let us examine the 3rd and 4th terms carefully; Recall that when we get a Kronecker delta we are allowed to perform an index manipulation. With that in mind:

$$\begin{aligned}\frac{\partial x'^{\nu}}{\partial x^{\beta}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}} &= \delta^{\alpha}_{\beta} \qquad \implies \qquad \beta \to \alpha\\ \frac{\partial x^{\rho}}{\partial x'^{\nu}}\frac{\partial x'^{\nu}}{\partial x^{\beta}} &= \delta^{\rho}_{\beta} \qquad \implies \qquad \beta \to \rho\end{aligned}$$

$$\implies \nabla'_{\mu}A'^{\lambda} = \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\lambda}}{\partial x^{\sigma}\partial x^{\nu}}A^{\nu} + \frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial x'^{\lambda}}{\partial x^{\nu}}\partial_{\sigma}A^{\nu} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\eta}}{\partial x'^{\mu}}\Gamma^{\rho}_{\eta\alpha}A^{\alpha} - \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho}\partial x^{\alpha}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}A^{\rho}$$

Let us now use our second rule of thumb, when we are summing and contracting indices, it does not matter what we name them So we can do even more renaming!

$$\alpha \to \sigma \qquad \rho \to \nu \qquad \eta \to \sigma$$

Doing this shows us that the first and last term simply cancel. Leaving us with the expression:

$$\nabla_{'\mu} A'^{\lambda} = \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \frac{\partial x'^{\lambda}}{\partial x^{\nu}} \partial_{\sigma} A^{\nu} + \frac{\partial x'^{\lambda}}{\partial x^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \Gamma^{\nu}_{\sigma\alpha} A^{\alpha} = \frac{\partial x'^{\lambda}}{\partial x^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \left(\partial_{\sigma} A^{\nu} + \Gamma^{\nu}_{\sigma\alpha} A^{\alpha}\right) = \frac{\partial x'^{\lambda}}{\partial x^{\nu}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} (\nabla_{\sigma} A^{\nu})$$

And this shows us that the covariant derivative does indeed transform as a tensor! Hence we have succeeded in constructing a tensor derivative that transforms as a tensor. This will be of great advantage in the future. For covariant indices:

$$\nabla_{\mu} A_{\lambda} = \partial_{\mu} A_{\lambda} - \Gamma^{\nu}_{\mu\lambda} A_{\nu}$$

What if we differentiate the vector with respect to a scalar ? Then,

$$\frac{DA^{\lambda}}{d\tau} = \partial_{\tau} A^{\lambda} + \Gamma^{\lambda}_{\mu\nu} \frac{dx^{\mu}}{d\tau} A^{\nu}$$

For tensors of rank 2:

$$\nabla_{\alpha} T_{\mu\nu} = \partial_{\alpha} T_{\mu\nu} - \Gamma^{\rho}_{\alpha\mu} T_{\rho\nu} - \Gamma^{\rho}_{\alpha\nu} T_{\rho\mu}$$

The covariant derivatives guarantees that the derivative of the vector or tensor will increase the rank of the object by 1 just like gradient does with vectors. But is there a connection between this and the metric tensor ?

6.4 Christoffel Symbols

So far, we have not really delved deep into the affine connections. We only know how they transform. On this object, we impose that there is not torsion, so:

$$\Gamma^{\lambda}_{\alpha\beta} = \Gamma^{\lambda}_{\beta\alpha}$$

Our goal now, is to set another condition on Γ to narrow it to just 1 set. This leads to Christoffel symbols. Before we continue, we must know about metric compatibility. This is a condition imposed on the metric tensor. The importance of this condition follows from the second postulate of Special Relativity; laws of physics must have the same form in all reference frames. This means that the covariant derivative of the metric must be 0. This is metric compatibility;

$$\nabla_\alpha g_{\mu\nu} = 0$$

On the other hand, mathematically, when we parallel transport the basis, their dot product should not change. Hence the metric compatibility. Now let us write down the covariant derivative of the metric in 3 different permutations of the indices and see what we can get out of it;

$$\begin{aligned}
\nabla_\alpha g_{\mu\nu} &= \partial_\alpha g_{\mu\nu} - \Gamma^\rho_{\alpha\mu} g_{\rho\nu} - \Gamma^\rho_{\alpha\nu} g_{\rho\mu} = 0 \\
\nabla_\mu g_{\nu\alpha} &= \partial_\mu g_{\nu\alpha} - \Gamma^\rho_{\mu\nu} g_{\rho\alpha} - \Gamma^\rho_{\mu\alpha} g_{\rho\nu} = 0 \\
\nabla_\nu g_{\alpha\mu} &= \partial_\nu g_{\alpha\mu} - \Gamma^\rho_{\nu\alpha} g_{\rho\mu} - \Gamma^\rho_{\nu\mu} g_{\rho\alpha} = 0
\end{aligned}$$

If we subtract the second and third equations from the first, we get that:

$$\begin{aligned}
\partial_\alpha g_{\mu\nu} - \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu} + 2\Gamma^\rho_{\mu\nu} g_{\rho\alpha} &= 0 \\
\implies \Gamma^\rho_{\mu\nu} g_{\rho\alpha} &= \frac{1}{2}\left[\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}\right] \\
\implies \Gamma^\lambda_{\mu\nu} &= \frac{1}{2} g^{\lambda\alpha}\left[\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}\right]
\end{aligned}$$

This is the definition of the christoffel symbols of the second kind. Geometrically, the variations of the basis vectors along some path are captured in the christoffel symbols as it involves derivatives of the metric. So far, we have built all the necessary, basic tools that we can build off of in order to get to our final goal; deriving the Einstein Field Equations. We will spend the next section to build some more tools off of what we have so far. Then, we will use them to derive the equations of Einstein's general relativity. I hope that you can see by now, that while the mathematics looks quite intimidating, once you get past all the indexing and all the partial derivatives it is all the easy, simple ideas that we have all learned in our introductory calculus and linear algebra courses but more generalized. So do not hesitate to take your time in order to internalize each topic.

7 Other Differential Operators In Covariant Form

So far we have established the most fundamental operation in tensor calculus which is the derivative. Now we will use the covariant derivative to define other differential operators; divergence, curl and laplacian.

7.1 Divergence

For a scalar function, f, then $\nabla_\mu f = \vec{\nabla} f$. The divergence is denoted by $\nabla_\mu A^\mu$. Notice that the indices are the same. Now, this is quite interesting. If we write out the expression for the divergence;

$$\nabla_\mu A^\mu = \partial_\mu A^\mu + \Gamma^\mu_{\mu\lambda} A^\lambda$$

We see that two of the indices in the christoffel symbol are the same, meaning:

$$\Gamma^\mu_{\mu\lambda} = \frac{1}{2} g^{\mu\kappa} \left[\partial_\mu g_{\lambda\kappa} + \partial_\lambda g_{\mu\kappa} - \partial_\kappa g_{\mu\lambda}\right]$$

You can see that we can perform a change of indices for the first and last derivatives of the metric: $\kappa \to \mu$. This gives us;

$$\begin{aligned}
\Gamma^\mu_{\mu\lambda} &= \frac{1}{2} g^{\mu\kappa} \left[\partial_\mu g_{\lambda\mu} + \partial_\lambda g_{\mu\kappa} - \partial_\mu g_{\mu\lambda}\right] \\
&= \frac{1}{2} g^{\mu\kappa} \partial_\lambda g_{\mu\kappa}
\end{aligned}$$

And then using a formula we will prove later;

$$\frac{1}{2} g^{\mu\kappa} \partial_\lambda g_{\mu\kappa} = \frac{1}{\sqrt{g}} \partial_\lambda \sqrt{g}$$

We can write the covariant divergence as:

$$\nabla_\mu A^\mu = \partial_\mu A^\mu + \frac{1}{\sqrt{g}} \partial_\lambda \sqrt{g} A^\lambda$$

Since λ is being summed over, we can change it to μ;

$$\lambda \to \mu \quad \implies \quad \nabla_\mu A^\mu = \partial_\mu A^\mu + \frac{1}{\sqrt{g}}\partial_\mu \sqrt{g} A^\mu$$

Now, factoring out the inverse root g, and using the product rule gives us the ultimate form of the covariant divergence:

$$\nabla_\mu A^\mu = \frac{1}{\sqrt{g}}[\sqrt{g}\partial_\mu A^\mu + \partial_\mu \sqrt{g} A^\mu]$$

$$\text{Recall: } \frac{d}{dx} fg = f'g + fg'$$

$$\therefore \nabla_\mu A^\mu = \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g} A^\mu)$$

Note that $\partial_\mu A^\mu$ does not necessarily have unit of $1/L$. It is a derivative with respect to coordinates not spatial position.

7.2 Laplacian

The laplacian is quite simple. We know that we are taking the divergence of a scalar field's gradient or in other words;

$$\nabla^2 \phi = \nabla \cdot \vec{\nabla} \phi$$

Since gradient of phi is just a vector, we are taking the divergence of a vector;

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

And since ϕ is just a scalar field, $\nabla^\mu \phi = \partial^\mu \phi$ Before we jump into learning about the covariant curl - which will take a bit more work than divergence and Laplacian - we will discuss an example of computing the divergence of a tensor just to get an idea of how the calculation is carried out. Then we will learn the covariant curl, followed by the Bianchi identities and finally, the Einstein Field Equations. Note that when learning tensor calculus, it is important to sometimes work with actual numbers and functions just to get an idea of how to actually compute a result from tensor equations. Sometimes, too many symbols causes physicists and mathematicians to forget the initial goal of the whole shebang, computing a result or prediction.

7.3 Curl

To understand curl, we must first redefine the dot product for tensors. To do this, we will have to use the Levi-Civita symbol defined as:

$$\varepsilon^{ijk} = \begin{cases} 1 & \text{for cyclic permutations of i,j,k} \\ -1 & \text{for anti-cyclic permutations of i,j,k} \\ 0 & i = j,\ j = k,\ i = k \text{ or } i = j = k \end{cases}$$

Cyclic permutations means the indices are in any of the orders (1,2,3), (2,3,1) or (3,1,2). On the other hand, Anti-cyclic means they are in any of the orders (3,2,1), (2,1,3) or (1,3,2). And when any 2 or more indices repeat, then the Levi-Civita symbol (or connection) is 0. On the other hand, where the indices of the Levi-Civita go do not matter. They are purely for using the summation convention. Using this, we define the cross product of two vectors $\vec{A}$ and $\vec{B}$'s ith component as;

$$(\vec{A} \times \vec{B})^i = \sum_{k=1}^{3} \sum_{j=1}^{3} \varepsilon^{ijk} A^j B^k$$

Now, going back to curl, we know that the curl is just the gradient operator's cross product with a field. Using our notion of the gradient we define the ith component of the curl as;

$$(\nabla \times \tilde{A})_i = \varepsilon^{ijk} \partial_j \tilde{A}_k$$

We normally promote everything to its tensor and covariant form;

$$\partial_j \to D \qquad \tilde{A}_k \to A_k$$

But what about the Levi-Civita connection ? This symbol is known as *tensor density*. Let us recall;

$$\mathcal{J} = \sqrt{\frac{g}{g'}} \implies \sqrt{g'} = \frac{1}{\mathcal{J}} \sqrt{g}$$

This means that g or $\sqrt{g}$ is not a scalar, but a scalar or tensor density. It is weighted by some power of the Jacobian. Suppose that some tensor T^{ij} is defined as

$$T^{ij} = \int \tau^{ij} d^n x$$

Then if T is a tensor, what kind of object is τ ? It is the tensor density of T. Since,

$$d' x^n = \mathcal{J} dx^n$$

Then τ should also not be a tensor. For example, if we are going from some system to Cartesian, then $\sqrt{g'} = 1$ which implies that $\sqrt{g}\tau^{ij}$ must be a tensor. The transformation of the Levi-Civita follows the following rule;

$$\varepsilon_{ijk} = \mathcal{J} \frac{\partial x^l}{\partial x^i} \frac{\partial x^m}{\partial x^j} \frac{\partial x^n}{\partial x^k} \varepsilon_{lmn}$$

Now, if the Jacobian of the symbol is 1, how can we make a tensor ? If $\sqrt{g'} = \frac{1}{\mathcal{J}}\sqrt{g}$ then we know that $\sqrt{g}\,\varepsilon_{ijk}$ is a tensor. So we define

$$E_{ijk} \equiv \sqrt{g}\,\varepsilon_{ijk}$$

$$E^{ijk} \equiv \frac{1}{\sqrt{g}} \varepsilon^{ijk}$$

Recall that making indices upstairs does not change the value of the Levi-Civita. Using this tensor, we can define the ith component of the covariant curl, C^i, as

$$C^i = E^{ijk} \nabla_j A_k = \frac{1}{\sqrt{g}} \varepsilon^{ijk} \nabla_j A_k$$

Suppose ijk is cylic. Then;

$$C^i = \frac{1}{\sqrt{g}} (\varepsilon^{ijk} \nabla_j A_k + \varepsilon^{ikj} \nabla_k A_j)$$

Where

$$\varepsilon^{ijk} = 1, \varepsilon^{ikj} = -1$$

$$\implies C^i = \frac{1}{\sqrt{g}} (\nabla_j A_k - \nabla_k A_j) = \frac{-1}{\sqrt{g}} (\nabla_k A_j - \nabla_j A_k)$$

On the left, we have a contravariant vector, on the right we have covariant ones. So we need to fix this.

$$\begin{aligned}
C^i = \frac{1}{\sqrt{g}} C_j k &= \frac{1}{\sqrt{g}} (\nabla_j A_k - \nabla_k A_j) \\
&= \frac{1}{\sqrt{g}} (\partial_j A_k - \Gamma^l_{jk} A_l - \partial_k A_j + \Gamma^l_{kj} A_l) \\
&\text{Since metric is symmetric, } \Gamma \text{ is symmetric} \\
&\therefore C^i = \frac{1}{\sqrt{g}} (\partial_j A_k - \partial_k A_j)
\end{aligned}$$

And that is the end of our pursuit for the covariant curl! Note that here;

$$C^i = \frac{1}{h_i} \tilde{C}_i$$

Where the Tilde vector components belong to the "ordinary vector".

7.4 Divergence For Spherical Coordinates

Take the spherical coordinate system in flat space. The metric of this space is

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}$$

$$\sqrt{g} = \sqrt{r^4 \sin^2 \theta} = r^2 \sin \theta$$

$$\implies \nabla_\mu A^\mu = \frac{1}{r^2 sin\theta} \left[\frac{\partial}{\partial r} (r^2 \sin \theta A^r) + \frac{\partial}{\partial \theta} (r^2 \sin \theta A^\theta) + \frac{\partial}{\partial \phi} (r^2 \sin \theta A^\phi) \right]$$

But we know that $A^\mu = \frac{\tilde{A_\mu}}{h_\mu}$ so we can substitute it in:

$$\nabla_\mu A^\mu = \frac{1}{r^2 \sin \theta} \left[\frac{\partial}{\partial r} \left(r^2 \sin \theta \frac{\tilde{A}_r}{1} \right) + \frac{\partial}{\partial \theta} \left(r^2 \sin \theta \frac{\tilde{A}_\theta}{r} \right) + \frac{\partial}{\partial \phi} \left(r^2 \sin \theta \frac{\tilde{A}_\phi}{r \sin \theta} \right) \right]$$

Which after cancelling the sine term gives us the ultimate form of the covariant divergence;

$$\nabla_\mu A^\mu = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \tilde{A}_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \tilde{A}_\theta) + \frac{1}{r \sin \theta} \frac{\partial \tilde{A}_\phi}{\partial \phi}$$

7.5 Curl For Spherical Coordinates

For spherical coordinates, we will take the θth component of the curl. Note that here, a cyclic permutation would be r, θ, ϕ.

$$
\begin{aligned}
C^\theta &= \frac{1}{r^2 \sin\theta}\left[\frac{\partial}{\partial\phi}A_r - \frac{\partial}{\partial r}A_\phi\right] \\
A_\mu &= h_\mu \tilde{A}_\mu, \quad h_\mu = (1, r, r\sin\theta) \\
C^\theta &= \frac{1}{r^2 \sin\theta}\left[\frac{\partial}{\partial\phi}\tilde{A}_r - \frac{\partial}{\partial r}(r\sin\theta\tilde{A}_\phi)\right] \\
&= \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial\tilde{A}_r}{\partial\phi} - \frac{\partial}{\partial r}(r\tilde{A}_\phi)\right] \\
\implies \tilde{C}_\theta = rC^\theta &= \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial\tilde{A}_r}{\partial\phi} - \frac{\partial}{\partial r}(r\tilde{A}_\phi)\right]
\end{aligned}
$$

8 Curvature and the Einstein Field Equations

So far, we have been purely developing tools. Now, it is time to take a step back, look our creation, and apply it. There are many many applications of tensors. We, however, will use them to derive the equations of motion in general relativity, namely the Einstein Field Equations.

8.1 The Riemann Curvature Tensor

Consider a sphere and a cylinder. The question we have is "are these objects curved ?" The answer, is that the question is ill defined. What is it that we mean by "curvature" ? There are 2 types of curvature: *extrinsic curvature* and *intrinsic curvature*.

Extrinsic curvature is when we embed a surface to a higher dimensional space, to tell if it is curved. The way we embed this surface matters. I will not go into the topology of this as that is a long and well-detailed discussion which honestly deserves multiple handouts on its own.

Intrinsic curvature is a bit different. Imagine being a bug on the surface of a sphere. Is the bug able to tell if the surface it is on is curved or not ? One way is measuring the sum of the angles of a triangle. There is also the geodesic deviation. The approach that we will take is the most common approach taken by General Relativity textbooks; Parallel Transporting. Parallel transporting a vector means going back to the notion of covariant derivative; taking a vector and moving it along a path and back to where it was in multiple paths and see the changes in its direction. On a flat spacetime, parallel transporting a vector does not change anything no matter which path you take. On the other hand, on a curved surface they cause changes to the components ratios of a vector and different paths may change these ratios differently. So, let us observe this by defining some vector $\vec{a}$. Now, let us take each movement of the vector and show it by applying a operator on our vector. If we define 4 operators for each path, A, B, C, D then;

$$\vec{a} - \hat{D}\hat{C}\hat{B}\hat{A}\vec{a} = 0$$
$$\hat{A}\hat{B}\vec{a} - \hat{B}\hat{A}\vec{a} = 0$$

We want to do this for infinitesimally small distances because for surfaces with varying curvature, we miss the local curvature. Our operators will turn to derivatives when we work in infinitesimally small distances. Recall:

$$\frac{df}{dx} = \frac{f(x+dx) - f(x)}{dx}$$
$$\implies f(x+dx) = f(x) + dx\frac{df}{dx}$$

for infinitesimally small differentials of x, this works. Derivatives are the generators of infinitesimal translations.

We can think of covariant derivatives in terms of generators of infinitesimal parallel transport. So we can replace the operators with covariant derivatives;

$$\hat{A}\hat{B}\vec{a} - \hat{B}\hat{A}\vec{a} = \nabla_\mu\nabla_\nu A^\lambda - \nabla_\nu\nabla_\mu A^\lambda$$
$$= [\nabla_\mu, \nabla_\nu]A^\lambda, \quad [\nabla_\mu, \nabla_\nu] \equiv \nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu$$

So the problem is, what is the value of $[\nabla_\mu, \nabla_\nu]A^\lambda = ?$ We know that if it is 0, then space is flat. If it is not 0, then the space must be curved. Now, if we expand the commutator brackets with the definition of the covariant derivative we get

$$\nabla_\mu\nabla_\nu A^\lambda - \nabla_\nu\nabla_\mu A^\lambda$$
$$= \nabla_\mu(\partial_\nu A^\lambda + \Gamma^\lambda_{\alpha\nu}A^\alpha) - \nabla_\nu(\partial_\mu A^\lambda + \Gamma^\lambda_{\alpha\mu}A^\alpha)$$
$$\text{let} \quad T^\lambda_\nu = \partial_\nu A^\lambda + \Gamma^\lambda_{\alpha\nu}A^\alpha \qquad\qquad \text{let} \quad T^\lambda_\mu = \partial_\mu A^\lambda + \Gamma^\lambda_{\alpha\mu}A^\alpha$$

This gives us

$$\implies = \nabla_\mu T^\lambda_\nu - \nabla_\nu T^\lambda_\mu$$

$$\nabla_\mu T^\lambda_\nu = \partial_\mu T^\lambda_\nu + \Gamma^\lambda_{\sigma\mu} T^\sigma_\nu - \Gamma^\gamma_{\mu\nu} T^\lambda_\gamma$$

$$= \partial_\mu(\partial_\nu A^\lambda + \Gamma^\lambda_{\alpha\nu} A^\alpha) + \Gamma^\lambda_{\sigma\mu}(\partial_\nu A^\sigma + \Gamma^\sigma_{\alpha\nu} A^\alpha) - \Gamma^\gamma_{\mu\nu}(\partial_\gamma A^\lambda + \Gamma^\lambda_{\alpha\gamma} A^\alpha)$$

$$= \partial_\mu \partial_\nu A^\lambda + A^\alpha \partial_\mu \Gamma^\lambda_{\alpha\nu} + \Gamma^\gamma_{\mu\nu} \partial_\gamma A^\lambda + \Gamma^\lambda_{\sigma\mu} \partial_\nu A^\sigma + \Gamma^\lambda_{\sigma\mu} \Gamma^\sigma_{\alpha\nu} A^\alpha + \Gamma^\lambda_{\alpha\nu} \partial_\mu A^\alpha - \Gamma^\gamma_{\mu\nu} \Gamma^\lambda_{\alpha\gamma} A^\alpha$$

Now we know that the other second covariant derivative looks pretty much the same way but with switching every mu with a nu and vice versa. For the last two terms, since the christoffel symbols are symmetric, they will cancel. So

$$\begin{aligned}[\nabla_\mu, \nabla_\nu] A^\lambda = {} & \partial_\mu \partial_\nu A^\lambda - \partial_\nu \partial_\mu A^\lambda + \Gamma^\lambda_{\alpha\nu} \partial_\mu A^\alpha - \Gamma^\lambda_{\alpha\mu} \partial_\nu A^\alpha \\ & + \Gamma^\lambda_{\sigma\mu} \partial_\nu A^\sigma - \Gamma^\lambda_{\sigma\nu} \partial_\mu A^\sigma + \Gamma^\lambda_{\sigma\mu} \Gamma^\sigma_{\alpha\nu} A^\alpha - \Gamma^\lambda_{\sigma\nu} \Gamma^\sigma_{\alpha\mu} A^\alpha\end{aligned}$$

Going back to our index reordering rules, we see that in the 4th term, α is summed over and in the 5th term σ is summed over. The same thing is happening in the 3rd and the 6th term. So a relabelling will cancel these terms;

$$\sigma \to \alpha$$

$$\implies [\nabla_\mu, \nabla_\nu] A^\lambda = (\partial_\mu \Gamma^\lambda_{\alpha\nu} - \partial_\nu \Gamma^\lambda_{\alpha\mu} + \Gamma^\lambda_{\sigma\mu} \Gamma^\sigma_{\alpha\nu} - \Gamma^\lambda_{\sigma\nu} \Gamma^\sigma_{\alpha\mu}) A^\alpha$$

Now, since σ is being summed over, it is just a dummy index. So this operator, has 4 indices. We define the *Riemann Curvature Tensor* or simply the Riemann Tensor as

$$\mathcal{R}^\lambda_{\alpha\nu\mu} \equiv [\nabla_\mu, \nabla_\nu]$$

Or substituting in the expression in terms of christoffel symbols;

$$\mathcal{R}^\lambda_{\alpha\nu\mu} = \partial_\mu \Gamma^\lambda_{\alpha\nu} - \partial_\nu \Gamma^\lambda_{\alpha\mu} + \Gamma^\lambda_{\sigma\mu} \Gamma^\sigma_{\alpha\nu} - \Gamma^\lambda_{\sigma\nu} \Gamma^\sigma_{\alpha\mu}$$

Recall that christoffel symbols were defined in terms of derivatives of the metric and the Riemann tensor involves derivatives of christoffel symbols. Thus, the Riemann tensor involves second derivatives of the metric and is quadratic in first derivatives of the metric, linear in second derivatives. For flat space, the christoffel symbols must be 0 and that makes sense because then the curvature tensor will be zero, meaning that parallel transporting a vector infinitesimally does not affect its components no matter what path you take. Hence, if all components of the Riemann tensor are 0, then the space is flat.

For example, in 2 dimensions, a cylinder's metric involves 1 and the radius (which is a constant) squared. Implying that the christoffel symbols are 0, meaning the Riemann tensor is 0. On the other hand, for a sphere the metric depends on the angle, θ which varies along paths and hence, the derivative of the metric (which is the christoffel symbol) will not be 0 and hence the Riemann tensor will not be 0 meaning the surface is curved which is true with our knowings. Note that expressing something in spherical coordinates does not affect curvature, what matters is the surface, not the coordinates.

8.2 Relating The Metric To Gravity

The purpose of this part, is learning how we can relate $g_{\mu\nu}$ to a gravitational potential to come up with equations of motion ? We will start with classical gravity and then start building from there. Classical gravity is described by Poisson's equation:

$$\nabla^2\phi = 4\pi G\rho(x, y, z)$$

It has been established that the first component of the energy-momentum tensor, $T_{\mu\nu}$ is defined as the Hamiltonian density:

$$T_{00} = \rho(x, y, z)$$

What about the right hand side ? Obviously, $\nabla^2 \to \nabla_\mu\nabla^\mu$ but what about the potential ϕ ? Let us recall the geodesic equation for free-fall:

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0$$

For $g_{\mu\nu} = \eta_{\mu\nu}$ the christoffel symbol is 0 meaning $\frac{d^2x^\lambda}{d\tau^2} = 0$ Generalizing the classical limit gives

$$\frac{d^2x^i}{dt^2} - \vec{g} = 0$$

How are these two related to each other ?

Well, if we think of the last two terms in the geodesic equations as four-velocities

$$v^{\mu} = \frac{dx^{\mu}}{d\tau} = \gamma(c, \vec{v}) = \gamma(1, \vec{v}) \quad c = 1$$

For a classical limit, the velocity factor, $\vec{v}$, is much smaller than the speed of light. So we can think of it as going to 0 and the Lorentz factor going to 1.

$$\implies v^{\mu} = (1, \vec{0})$$

And therefore, if $\mu, \nu \neq 0$ then

$$\Gamma^{\lambda}_{\mu\nu} v^{\mu} v^{\nu} = 0$$

In this case, we only get a contribution from the term $\mu\nu = 00$. So Γ^{λ}_{00} is the only christoffel symbol that matters. Now, we must use the *static-weak field metric* to solve for the christoffel symbol. The static-weak field metric is essentially the Minkowski metric plus some very small perturbation factor:

$$g_{\alpha\beta}(\vec{x}) = \eta_{\alpha\beta} + h_{\alpha\beta}(\vec{x})$$

So what is the christoffel symbol ? we now that $\partial_{\mu} g_{\alpha\nu}$ and $\partial_{\nu} g_{\alpha\mu}$ will just be time derivatives. Since the metric is static, time derivatives are 0. After all, that is what static means; not moving. Hence,

$$\implies \Gamma^{\lambda}_{00} = -\frac{1}{2} g^{\lambda\alpha} \partial_{\alpha} g_{00} = -\frac{1}{2} g^{\lambda\alpha} \partial_{\alpha} h_{00}$$

In the last derivative, since the Minkowski metric is static, its time derivative is just 0. Hence we only need to worry about the perturbation factor. And we also know that the inverse metric times the covariant directional derivative is just the contravariant component

$$\implies = -\frac{1}{2} \partial^{\lambda} h_{00}$$

Note that for $\lambda = 0$, $\partial^0 h_0 0 = 0$. We will only get contributions from the ith components (meaning the spatial components of the derivative, not the time).

$$\implies \Gamma^i_{00} = -\frac{1}{2}\partial^i h_{00}$$

Now, the question is, how is ∂^i related to ∂_i component by component ?

$$g_{ij}\partial^j h_{00} = (\eta_{ij} + h_{ij})\partial^j h_{00}$$

Perturbation states that h is just large enough but h^2 is small enough to be negligible.

$$\implies g_{ij}\partial^j h_{00} \approx \eta_{ij}\partial^j$$

$$\implies \Gamma^\lambda_{00} = \frac{1}{2}\partial_i h_{00} = \frac{1}{2}\nabla h_{00}$$

Before we go ahead and substitute this perturbation into the geodesic equation, let us see what does this mean when we take the classical limit of the geodesic equation, meaning how is $dx^\lambda/d\tau^2$ related to regular time, t ?

$$\begin{aligned} d\tau^2 &= dt^2 - dx^2 \\ \frac{df}{d\tau} &= \frac{df}{dt}\frac{dt}{d\tau} \\ \frac{dt}{d\tau} &= \gamma \approx 1 \quad \text{(in the classical limit)} \\ \implies \frac{d^2x^\lambda}{d\tau^2} &\approx \frac{d^2x^\lambda}{dt^2} \end{aligned}$$

And this shows us that motion and trajectories of objects in general relativity is consistent with both classical mechanics in the classical limit, and it is consistent with special relativity. Now, the problem of whether it is consistent with quantum mechanics or not, I will leave that to you to figure out! Now we substitute our perturbation into the geodesic equation

$$\frac{d^2x^\lambda}{d\tau^2} + \frac{1}{2}\nabla h_{00} = 0$$

if $c \neq 1$, in other words, if we do not use natural units, there will be factors of c^2 all over the place. Now, if we go back to the classical geodesic equation, we see that

$$\frac{d^2x^i}{dt^2} - \vec{g} = 0$$

and using this, we can identify the second terms in the equations to be equal. Hence,

$$\frac{1}{2}\nabla h_{00} = -g$$

Recall that $\vec{F} = m\vec{g} = -m\nabla\phi$ where ϕ is the gravitational potential, not to be confused with potential energy!

$$\begin{aligned}
&\implies g = -\nabla\phi \\
\implies \frac{1}{2}\nabla h_{00} &= \nabla\phi \implies h_{00} = 2\phi \\
\therefore g_{\alpha\beta} &= \eta_{00} + h_{00} = \eta_{00} + 2\phi \\
&= 1 + 2\phi
\end{aligned}$$

So in the classical limit, $g_{00} = 1+2\phi$ connects the metric tensor to gravity, namely the gravitational potential. With $T^{00} \propto \mathcal{H}$, for the velocity being 0,

$$\frac{\mathcal{H}}{c^2} = m$$

Hence, relating the metric and energy-momentum tensor to Poisson's equation will only involve the 00 components.

$$\begin{aligned}
\nabla^2\phi &= 4\pi G\rho \\
\nabla^2 g_{00} = \nabla^2(2\phi) &= 8\pi G\rho \\
\implies \nabla^2 g_{00} &= 8\pi G T_{00}
\end{aligned}$$

And inserting the c factors in, we get the classical limit of Poisson's equation for gravity in terms of the metric and the energy-momentum tensor:

$$\nabla^2 g_{00} = \frac{8\pi G}{c^4} T_{00}$$

Thus, we already know that the Einstein Field Equations must involve second derivatives of the metric on the left hand side and the right hand side must involve the energy-momentum tensor.

8.3 The Bianchi Identities

This chapter will be the last one before we finally uncover Einstein's field equations. So just hold on for a few more pages and we will get there. Let us recall the definition of the Riemann tensor,

$$\mathcal{R}^{\lambda}_{\alpha\nu\mu} = \partial_{\mu}\Gamma^{\lambda}_{\alpha\nu} - \partial_{\nu}\Gamma^{\lambda}_{\alpha\mu} + \Gamma^{\lambda}_{\sigma\mu}\Gamma^{\sigma}_{\alpha\nu} - \Gamma^{\lambda}_{\sigma\nu}\Gamma^{\sigma}_{\alpha\mu}$$

Here, must learn about 2 of the contractions of this tensor, the Ricci tensor and the Ricci scalar. The Ricci tensor is given by taking the trace of the contravariant and one covariant index. This repeated with the Ricci tensor gives the Ricci scalar.

$$\mathcal{R}^{\lambda}_{\alpha\lambda\mu} \equiv \mathcal{R}_{\alpha\mu} \quad \text{is the Ricci Tensor}$$
$$\mathcal{R}^{\alpha}_{\alpha} \equiv \mathcal{R} \quad \text{is the Ricci Scalar}$$

Considering the symmetries of the Riemann tensor, we find that it is anti-symmetric for μ and ν

$$\mathcal{R}^{\lambda}_{\alpha\nu\mu} = -\mathcal{R}^{\lambda}_{\alpha\mu\nu}, \quad \mathcal{R}_{\lambda\alpha\nu\mu} = -\mathcal{R}_{\alpha\lambda\nu\mu}$$

Looking at the definition, we find that it involves derivatives and products of christoffel symbols. If we define some local inertial frame:

$$\Gamma \to 0,\ \partial\Gamma \neq 0$$

We can get rid of the $\Gamma\Gamma$ terms. We will form contributions that will end up being valid in all reference frames.

$$\mathcal{R}^{\lambda}_{\alpha\nu\mu} = \partial_{\mu}\Gamma^{\lambda}_{\alpha\nu} - \partial_{\nu}\Gamma^{\lambda}_{\alpha\mu}$$

We are interested in taking covariant derivatives of $\mathcal{R}$. In our reference frame: $\nabla_\mu \to \partial_\mu$. Here, we will introduce some notation. The covariant derivative is also denoted by:

$$\nabla_\sigma \mathcal{R}^\lambda_{\alpha\nu\mu} \equiv \mathcal{R}^\lambda_{\alpha\nu\mu;\sigma}$$

So we are differentiating with respect to what symbol follows the semicolon.

$$\mathcal{R}^\lambda_{\alpha\nu\mu;\sigma} = \partial_\sigma\partial_\mu\Gamma^\lambda_{\alpha\nu} - \partial_\sigma\partial_\nu\Gamma^\lambda_{\alpha\mu}$$

We will write this 2 more times by taking permutations of σ, μ and ν in the orders $(\nu\mu\sigma), (\mu\sigma\nu), (\sigma\nu\mu)$

$$\begin{aligned}
\mathcal{R}^\lambda_{\alpha\nu\mu;\sigma} &= \partial_\sigma\partial_\mu\Gamma^\lambda_{\alpha\nu} - \partial_\sigma\partial_\nu\Gamma^\lambda_{\alpha\mu} \\
\mathcal{R}^\lambda_{\alpha\mu\sigma;\nu} &= \partial_\nu\partial_\sigma\Gamma^\lambda_{\alpha\mu} - \partial_\nu\partial_\mu\Gamma^\lambda_{\alpha\sigma} \\
\mathcal{R}^\lambda_{\alpha\sigma\nu;\mu} &= \partial_\mu\partial_\nu\Gamma^\lambda_{\alpha\sigma} - \partial_\mu\partial_\sigma\Gamma^\lambda_{\alpha\nu}
\end{aligned}$$

Taking the sum of all these, we get the differential Bianchi identity:

$$\mathcal{R}^\lambda_{\alpha\nu\mu;\sigma} + \mathcal{R}^\lambda_{\alpha\mu\sigma;\nu} + \mathcal{R}^\lambda_{\alpha\sigma\nu;\mu} = 0$$

Since this equation is equal to 0, it is true in all frames of reference, even though we calculated it in a specific reference frame. Now, the question is, does this hold if we contract the symbols ? Let us set $\sigma = \lambda$:

$$\mathcal{R}^\lambda_{\alpha\nu\mu;\lambda} + \mathcal{R}^\lambda_{\alpha\mu\lambda;\nu} + \mathcal{R}^\lambda_{\alpha\lambda\nu;\mu} = 0$$

We see that the 2nd and 3rd terms are proportional to the Ricci tensor!

$$\implies \mathcal{R}^\lambda_{\alpha\nu\mu;\lambda} - \mathcal{R}_{\alpha\mu;\nu} + \mathcal{R}_{\alpha\nu;\mu} = 0$$

Since we are differentiating with respect to different indices, the Ricci tensors do not cancel.

$$\implies \mathcal{R}^{\lambda\alpha}_{\nu\mu;\lambda} - \mathcal{R}^\alpha_{\mu;\alpha} + \mathcal{R}^\alpha_{\alpha;\mu} = 0$$

Note the derivative in the second term is just a covariant divergence. The third term is the covariant derivative of the Ricci scalar. What about the first term ?

$$\mathcal{R}^{\lambda\alpha}_{\alpha\mu;\lambda} = D_\lambda(\mathcal{R}^{\lambda\alpha}_{\alpha\mu}) = -D_\lambda(\mathcal{R}^{\lambda\alpha}_{\mu\alpha}) = -D_\lambda(\mathcal{R}^{\lambda}_{\mu})$$

Substituting this back in, we get

$$-\mathcal{R}^{\lambda}_{\mu;\lambda} - \mathcal{R}^{\alpha}_{\mu;\alpha} + \mathcal{R}_{;\mu} = 0$$

Since we are summing over λ and α we can rename them. We will relabel $\lambda = \alpha = \nu$

$$-2\mathcal{R}^{\nu}_{\mu;\nu} + \mathcal{R}_{;\mu} = 0$$

If we multiply by a Kronecker delta, we can remove the μ in the second term:

$$\begin{aligned}
\mathcal{R}_{;\mu} &= g^{\nu}_{\mu}\mathcal{R}_{;\nu} \\
\mathcal{R}^{\nu}_{\mu;\nu} - \frac{1}{2}\mathcal{R}_{;\mu} &= 0 \\
\implies \mathcal{R}^{\nu}_{\mu;\nu} - \frac{1}{2}g^{\nu}_{\mu}\mathcal{R}_{;\nu} &= 0 \\
\nabla_\nu\left(\mathcal{R}^{\nu}_{\mu} - \frac{1}{2}g^{\nu}_{\mu}\mathcal{R}\right) &= 0
\end{aligned}$$

And since $\mathcal{R}^{\nu}_{\mu}$ is symmetric, we get the contracted Bianchi identity:

$$\nabla_\mu\left(\mathcal{R}^{\mu\nu} - \frac{1}{2}g^{\mu\nu}\mathcal{R}\right) = 0$$

Recall from physics that if the 4-divergence of any quantity is 0, that quantity is conserved. Note the argument of the derivative is defined as the Einstein Tensor:

$$G^{\mu\nu} \equiv \mathcal{R}^{\mu\nu} - \frac{1}{2}\mathcal{R}g^{\mu\nu}$$

So $\nabla_\nu G^{\mu\nu} = 0$ This concludes every single bit of mathematical tools that we needed to derive the Einstein Field Equations.

8.4 The Einstein Field Equations

We are finally here, the chapter that most physics people look forward too! Let us begin by getting some sort of an interpretation for the energy-momentum tensor. Let us recall Poisson's equation:

$$\nabla^2 g_{00} = 8\pi G T_{00}$$

here are the physical meanings of the some of the components of the tensor:

$$\begin{aligned} T_{0i} &= \vec{p} \quad \text{(momentum density)} \\ T_{ii} &= P \quad \text{(pressure)} \\ T_{i\neq j} &= \tau \quad \text{(shear stress)} \end{aligned}$$

In words, the first row of the energy-momentum tensor is the momentum density, the other diagonal components represent the pressure and the off diagonal components that remain are components of the shear stress.

Now, our goal from all this is to generalize Poisson's equation of gravity. What are the criteria for the left hand side of Poisson's equation ? We know that it must be a second rank tensor because the right hand side involving the energy-momentum tensor is a second rank tensor. We also know that it must be linear in second derivatives of the metric tensor. At the same time, it must satisfy local energy-momentum conservation meaning that the energy-momentum tensor must be divergence-less. Why ? Well, let us look at the divergence of the energy-momentum tensor:

$$\nabla^\mu T_{\mu\nu} = D^0 T_{00} + D^i T_{i0} \implies \partial_t E - \nabla \cdot \vec{p} = 0$$

and this is just the continuity equation. Hence, the left hand side must also have a 0 divergence. We also know that the equation, in the classical limit, must reduce to Poisson's equation. Now what could it be in terms of ?

Obviously, we need second derivatives of the metric tensor. To satisfy this, we use the uniqueness property of the Riemann tensor; The only 4th rank tensor in terms of linear second and first derivatives of the metric is the Riemann tensor! Thus, since $\mathcal{R}_{\mu\nu}, g_{\mu\nu}\mathcal{R}, \Lambda g_{\mu\nu}$ are all second rank tensors from the Riemann tensor, they could be applicable. Our guess is something of the form:

$$A\mathcal{R}_{\mu\nu} + B\mathcal{R}g_{\mu\nu} = 8\pi GT_{\mu\nu}$$

The rest of this section, we will spend on calculating these coefficients. We can try to use the Riemann tensor for the field equations, but for a vacuum, the energy-momentum tensor is 0 meaning the Riemann tensor will be 0. But if we consider a planet orbiting a star, there is still curvature in the vacuum in-between, so this does not make physical sense. The Riemann tensor includes information that is not needed to calculate the metric tensor and our goal from the Einstein field equations is to solve for metric of spacetime. To find A and B we will use tha law of local electromagnetic and momentum conservation. We begin by taking the divergence of both sides:

$$A\nabla^{\mu}\mathcal{R}_{\mu\nu} + Bg_{\mu\nu}D^{\mu}\mathcal{R} = 0$$

Now, we know from the contracted Bianchi identity, that

$$\nabla^{\mu}\mathcal{R}_{\mu\nu} = \frac{1}{2}g_{\mu\nu}\nabla^{\mu}\mathcal{R}$$

Hence we get that

$$\implies \frac{1}{2}Ag_{\mu\nu}\nabla^{\mu}\mathcal{R} + Bg_{\mu\nu}\nabla^{\mu}\mathcal{R} = 0$$
$$\implies \frac{-A}{2} = B$$

So we know that

$$A\left[\mathcal{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R}\right] = 8\pi GT_{\mu\nu}$$

So what is A ? Let us recall:

$$\mathcal{R}^{\lambda}_{\alpha\nu\mu} = \partial_{\mu}\Gamma^{\lambda}_{\alpha\nu} - \partial_{\nu}\Gamma^{\lambda}_{\alpha\mu}$$
$$\implies \mathcal{R}_{\mu\nu} = \partial_{\nu}\Gamma^{\lambda}_{\mu\lambda} - \partial_{\lambda}\Gamma^{\lambda}_{\mu\nu}$$

Recall that

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}g^{\lambda\beta}\left[\partial_\mu g_{\nu\beta} + \partial_\nu g_{\beta\mu} - \partial_\beta g_{\mu\nu}\right]$$

For the static weak field metric:

$$\begin{aligned} g_{\mu\nu} &= \eta_{\mu\nu} + h_{\mu\nu}(\vec{x}) \\ \mathcal{R}_{00} &= 0 - \partial_\lambda \Gamma^{\lambda}_{00} \\ &= -\partial_i \Gamma^{i}_{00} \end{aligned}$$

Before, we proceed, i must mention that there are 2 metric signatures we can use for the Minkowski metric, $(+---)$ or $(-+++)$. However, since the second one is more commonly used, we will go with that.

$$\begin{aligned} \Gamma^{i}_{00} &= \frac{1}{2}g^{i\beta}\left[\partial_0 g_{\beta 0} + \partial_0 g_{0\beta} - \partial_\beta g_{00}\right] \\ &= \frac{1}{2}g^{i\beta}[-\partial_\beta g_{00}] = \frac{-1}{2}\partial^i g_{00} \\ \implies \mathcal{R}_{00} &= -\partial_i\left(-\frac{1}{2}\partial^i g_{00}\right) = -\frac{1}{2}\nabla^2 g_{00} \\ g_{00} &= \eta_{00} + h_{00} = -1 - 2\phi \quad \text{(with } (-+++) \text{ metric signature)} \\ \implies \mathcal{R}_{00} &= \nabla^2\phi \end{aligned}$$

Now, we will write the Ricci scalar as the sum its components:

$$\begin{aligned} \mathcal{R} = \mathcal{R}^{\alpha}_{\alpha} &= -\mathcal{R}^{0}_{0} + \mathcal{R}^{i}_{i} \qquad \mathcal{R} = g^{\alpha\beta}\mathcal{R}_{\alpha\beta} \\ \implies \mathcal{R} &= -\mathcal{R}_{00} + \mathcal{R}^{i}_{i} \end{aligned}$$

we will use htis in a few moments. In this Newtonian limit, the metric is just the mass density, the rest of the components are 0. Meaning that the equation will reduce to Newton's second law. Hence, for $\mu\nu = ij$, $\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} \approx 0$. So assuming that $\mathcal{R}_{ij} \approx \frac{1}{2}\mathcal{R}g_{ij}$ we get that

$$\begin{aligned} \implies \mathcal{R}^{i}_{i} &= \frac{1}{2}g^{i}_{i}\mathcal{R} = \frac{3}{2}\mathcal{R} \\ \implies \mathcal{R} &= -\mathcal{R}_{00} + \frac{3}{2}\mathcal{R} \end{aligned}$$

And this gives us the relationship:

$$\mathcal{R} = 2\mathcal{R}_{00} = 2\nabla^2\phi$$

Now, since we know that the Einstein Field Equations must reduce to Poisson's equation, we will substitute this in to see what coefficient is needed for the equations to match:

$$\begin{aligned} A\left[\mathcal{R}_{00} - \frac{1}{2}g_{00}\mathcal{R}\right] &= 8\pi G T_{00} \\ \implies A\left[\nabla^2\phi - \frac{1}{2}(-2(\nabla^2\phi))\right] &= 8\pi G T_{00} \\ \implies 2A\nabla^2\phi &= 8\pi G T_{00} \\ A\nabla^2\phi &= 4\pi G T_{00} \\ \implies A &= 1 \end{aligned}$$

If we used the other metric signature, we would have gotten that $A = -1$. And now, we know that $B = -\frac{1}{2}$ using the relationship we established earlier. And finally, this gives us the Einstein Field Equations:

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

We have finally made it! If you pay attention, you will see that the left hand side was the Einstein tensor the whole time! So we could also write the field equations as

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}$$

Some other notations are introducing the Einstein gravitational constant:

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

and the cosmological constant, Λ. This constant allows for a non-expanding universe. It was known as Einstein's greatest blunder but it regained importance when we discovered that it did play a role in allowing us to calculate the expansion of the universe. These set of equations connect the curvatures of spacetime to the presence of energy and momentum in that spacetime. Now, I want to show you how great the notations we have introduced in this handout are because if we were to write out the field equations without any of these notation conventions, it would look something like this:

$$\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu}+\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\alpha\partial_\nu g_{\beta\mu}-\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu}-\frac{3}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\mu\partial_\nu g_{\alpha\beta}$$
$$-\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\sum_{\rho=0}^{3}\sum_{\lambda=0}^{3}g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\mu g_{\beta\nu}-\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\sum_{\rho=0}^{3}\sum_{\lambda=0}^{3}g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\nu g_{\mu\beta}$$
$$+\frac{1}{4}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\sum_{\rho=0}^{3}\sum_{\lambda=0}^{3}g^{\beta\lambda}g^{\alpha\rho}\partial_\nu g_{\alpha\lambda}\partial_\mu g_{\rho\beta}+\frac{1}{4g}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\beta g\partial_\nu g_{\mu\alpha}-\frac{1}{4g}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\beta g\partial_\alpha g_{\mu\nu}$$
$$-\frac{1}{4g}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\beta g\partial_\mu g_{\alpha\nu}=8\pi G T_{\mu\nu}$$

This is single handedly the ugliest equation I have ever seen. This goes on to show just how amazingly useful tensor notation and the summation convention are. Now, to make things even worse, imagine having to write out the terms of each sum! It can easily fill up your 360 page notebook. To give you a glance, here is the first term fully written out:

$$\begin{aligned}\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu}=&\frac{1}{2}g^{00}\partial_0\partial_\mu g_{0\nu}+\frac{1}{2}g^{01}\partial_0\partial_\mu g_{1\nu}+\frac{1}{2}g^{02}\partial_0\partial_\mu g_{2\nu}+\frac{1}{2}g^{03}\partial_0\partial_\mu g_{3\nu}\\&+\frac{1}{2}g^{10}\partial_1\partial_\mu g_{0\nu}+\frac{1}{2}g^{11}\partial_1\partial_\mu g_{1\nu}+\frac{1}{2}g^{12}\partial_1\partial_\mu g_{2\nu}+\frac{1}{2}g^{13}\partial_1\partial_\mu g_{3\nu}\\&+\frac{1}{2}g^{20}\partial_2\partial_\mu g_{0\nu}+\frac{1}{2}g^{21}\partial_2\partial_\mu g_{1\nu}+\frac{1}{2}g^{22}\partial_2\partial_\mu g_{2\nu}+\frac{1}{2}g^{23}\partial_2\partial_\mu g_{3\nu}\\&+\frac{1}{2}g^{30}\partial_3\partial_\mu g_{0\nu}+\frac{1}{2}g^{31}\partial_0\partial_\mu g_{1\nu}+\frac{1}{2}g^{32}\partial_0\partial_\mu g_{2\nu}+\frac{1}{2}g^{33}\partial_0\partial_\mu g_{3\nu}\end{aligned}$$

You get the idea. This concludes this handout on tensor calculus. Throughout this handout, I have taken great inspiration from the Tensor Calculus series by Andrew Dotson, Tensor Calculus for Physics by D.Neuenschwander and the lecture series on general relativity by Scott Hughes on MIT Opencourseware.