Introduction to General Relativity

Barsam Rahimi

March 2024

1 What To Take For Your Own Paper

The most important part of the paper is your model, your equation, the mathematical expression that supports your finding. In physics, all claims must be supported by proper mathematical derivations and the interpretations of the theories arise from analyzing those mathematical models. Unless your work is extremely theoretical and cannot be tested through any type of experimentation (which is a very rare case), you need some sort of experimental data to support your theory. If the data that you find from an experiment approximately matches up with your model, then your theory is both mathematically and experimentally consistent. However, in the case where it does not, there are 3 possibilities; an error in your experimentation such as not taking into account all there is to be thought of, improper lab procedure, incorrect data processing , etc. Another possibility is having an inconsistent mathematical model. If you are sure that your experimentation has been performed correctly and you have considered everything, then your model is probably inconsistent with the previously discovered laws and must be adjusted or re-derived. The third case is when you have both experimental errors and an inconsistent mathematical model. Hopefully this will not happen! Another thing to keep in mind when writing a physics paper is knowing when to use what. If you are talking about particles at a subatomic level such as electrons or particle accelerators, you need to use quantum mechanics and classical mechanics will not be consistent with the experimental findings! In the case of using astronomical data and massive objects such as stars, planets and black holes, you need to use the theory of general relativity and again, cannot use classical mechanics. This is due to the fact that these minor adjustments make all the difference in these two fields. A very small decimal place could make the difference between a particle existing or not existing in the quantum world. Good luck!

2 Classical Mechanics

Throughout this handout, we will make great use of calculus so let us do a quick review of some calculus concepts.

2.1 Interlude: Calculus Review

If we want to know the rate of change of a function $f(x)$, we are seeking the slope of its tangent lines. This is defined as the derivative of the function, $\frac{df}{dx}$ or $f'(x)$, and is calculated through:

$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

If we wish to find the area under the same curve, the area made between the curve and the x-axis, we use the definite integral:

$$\text{Area}_{f(x)} = \int_a^b f(x)\,dx = \lim_{n \to \infty} \sum_{k=0}^{n} f(x_k)\Delta x_k$$

And from the fundamental theorem of calculus we know that:

$$\frac{d}{dx} \int f(x)\,dx = f(x)$$

and that if $F(x)$ is the anti-derivative of $f(x)$, then:

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

When we have a function of several variables, such as $f(x, y, z)$, then the derivative of the function with respect to one variable is a *partial derivative* of the function:

$$\frac{df}{dx} \to \frac{\partial f}{\partial x}$$

and the definition is:

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x}$$

Similar to the fundamental theorem for one variable, in multiple variables;

$$\iiint \frac{\partial^3 f}{\partial x\, \partial y\, \partial z}\, dx\, dy\, dz = f(x, y, z)$$

and the gradient of the function, $\vec{\nabla} f$ is defined as:

$$\vec{\nabla} f = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}$$

Given 2 vectors:

$$\vec{u} = \langle u_1, u_2, u_3 \rangle$$

$$\vec{v} = \langle v_1, v_2, v_3 \rangle$$

Then:

$$\vec{u} \cdot \vec{v} = (u_1 v_1)\hat{i} + (u_2 v_2)\hat{j} + (u_3 v_3)\hat{k}$$

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Which can be calculated using Cramer's rule. The transformations of these vectors from some coordinate system (x, y) to (u, v) is given by the Jacobian, $\mathcal{J}$, of the vectors:

$$(V')^m = \mathcal{J}(V^n) = \frac{\partial U^p}{\partial x^q}(V^n)$$

The same is true for vectors with covariant components.

If we are going form the coordinate system $x^m \to y^n$, for a tensor, the Jacobian transformation must be applied to each index once.

Green's and Stokes' theorem for a vector field $\vec{F}(x, y)$ where $\vec{n}$ is the normal vector, indicates for some curve C and region in space R:

$$\oint_C ds\, \vec{F} \cdot \vec{n} = \iint_R dx\, dy\, \frac{\partial \vec{F}}{\partial x} + \frac{\partial \vec{F}}{\partial y} = \iint_R d\sigma\, \nabla \cdot \vec{F}$$

$$\oint_C \vec{F} \cdot d\vec{r} = \iiint_R d\sigma\, \nabla \times \vec{F} \cdot \hat{k}$$

2.2 Lagrangian Mechanics

In Newtonian mechanics, we often use the famous second law:

$$\sum \vec{F} = m\vec{a}$$

However, this equation is very difficult to work with and hard to adjust for different coordinate systems. In General Relativity, one of the biggest issues is to adjust our measurements based on different reference frames. The problem of telling which reference frame we are working with is mathematically similar to using different coordinate systems to make the same measurements. One issue that arises from this is that Newton's second law is only true when we work with Cartesian coordinates. The other issue with it is that it is a vector equation which is harder to work with compared to scalar equations. So, what we need is an equation of motion that only consists of scalar values and is true in all sorts of different coordinate systems.

To derive this equation let us first re-write the second law as:

$$F = m\ddot{x} = m\frac{d^2x}{dt^2}$$

Now let us recall, the equation for the potential energy:

$$\mathcal{U} = mgh$$

Since the force on an object due to gravity (in classical mechanics) is defined as $F_g = -mg$, we can substitute this into potential energy to get:

$$\mathcal{U} = F_g h$$

which immediately implies that:

$$\mathcal{U} = -\int_{x_0}^{x_1} dh\, F_g \quad \text{or} \quad F = -\frac{d}{dx}\mathcal{U}$$

Expanding the second part into n dimensions we get:

$$\sum_{i \to n} m\ddot{x}_i = \sum_{i \to n} \frac{\partial \mathcal{U}}{\partial x_i}$$

Now, we have solved the issue of having a scalar equation instead of a vector. However, we have multiple scalar functions involved in our equation. How can make it such that it only consists of 1 scalar function ?

To do so, let us recall the equation for the kinetic energy of an object:

$$\mathcal{T} = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{x}^2$$

If we differentiate this on both sides with respect to $\dot{x}$, we get:

$$\frac{d}{d\dot{x}}\mathcal{T} = m\dot{x}$$

from this we get a shocking fact; the time derivative of this expression is equal to the force and that is equal to the position derivative of potential energy!

$$\frac{d}{dt}\left(\frac{\partial \mathcal{T}}{\partial \dot{x}}\right) = m\ddot{x} = \frac{\partial \mathcal{U}}{\partial x}$$

We can also see that if we differentiate the potential energy with respect to $\dot{x}$, we get 0. The same is true fro the position derivative of kinetic energy:

$$\frac{\partial \mathcal{T}}{\partial x} = \frac{\partial \mathcal{U}}{\partial \dot{x}} = 0$$

So if we subtract both of these from the previous equation, we get:

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}}(\mathcal{T} - \mathcal{U})\right) = \frac{\partial}{\partial x}(\mathcal{T} - \mathcal{U})$$

Now, we call the function $\mathcal{T} - \mathcal{U} = \mathcal{L}$, where $\mathcal{L}$ is known as the *Lagrangian* of our system. And from this we get a beautiful equation known as the *Euler-Lagrange Equation*:

$$\boxed{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = 0}$$

and in multiple dimensions, this becomes:

$$\sum_{i\to n} \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0$$

The great things about this equation is that it is precise, and it is everywhere! It does not only apply to one scenario but almost any scenario you may encounter in classical mechanics. When searching for your mathematical model in your paper, it is important for it to have these properties. An equation that has too many exceptions or is only fit for very specific conditions is of no use in physics. Now, let us analyze our Lagrangian and see what we can learn about it. Well, if we write the function out, it will be:

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz$$

Since

$$m\ddot{x} = m\ddot{y} = 0 \quad \text{and} \quad m\ddot{z} = -mg$$

Now, we can see an interesting property of the Lagrangian:

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} = p \quad \implies \quad \frac{d}{dt}p = \frac{\partial \mathcal{L}}{\partial x}$$

And right from this, we know that if

$$\frac{dp}{dt} = constant \implies \frac{\partial \mathcal{L}}{\partial x} = constant$$

which implies that momentum is conserved in our system!

Now, why do we want to use the Lagrangian and why did I introduce this in a General Relativity lecture ? When discussing geodesics, we will come back to this. The reason we want to use the Lagrangian is because it is compatible with different coordinate systems and it is a scalar function meaning it is much easier to work with than vector functions. Let us look at these properties. Let us imagine an object with mass m orbiting the earth with mass M. Given that r is the radial distance between the two and φ is the angle from 0^o, then the potential energy of the object is:

$$\mathcal{U} = -\frac{GMm}{r}$$

and the force on the object is:

$$F = -\frac{GMm}{r^2}\vec{e}$$

Where $\vec{e}$ is some directionality vector. Now let:

$$x = x(r,\varphi) \quad \text{and} \quad y = y(r,\varphi)$$

Then

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) + \frac{GMm}{r}$$

Now, we know that the Euler-Lagrange Equation is true for this system as well:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} - \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{r}} - \frac{\partial \mathcal{L}}{\partial r} = 0$$

which implies that $\mathcal{L} = \mathcal{L}(\dot{r}, \dot{\varphi}, r)$ meaning it does not depend on φ!

We have just proven that the Euler-Lagrange Equation is consistent with polar coordinates and you will find that it is also consistent in spherical, cylindrical, curvilinear, etc. coordinates. Whereas $F = ma$ only works in Cartesian coordinates. Quick note also that

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{\partial \mathcal{L}}{\partial \varphi} = 0 \implies \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = constant$$

$$\therefore \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{\partial}{\partial \dot{\varphi}}\left(\frac{1}{2}mr^2\dot{\varphi}^2\right) = mr^2\dot{\varphi} = \omega$$

Since the angular momentum is constant, then momentum is conserved in our system which is consistent with Cartesian coordinate results! There are 2 postulates that must be true which imply the Euler-Lagrange equation:

1. *The system is defined on a scalar function $\mathcal{L}$ depending on three functions of the coordinate*

system: $q(t), \dot{q}(t), t$ *and is closed.*

2. *The path the system takes between two points, S, is the path that extremizes the action meaning if the path is slightly "wiggled", the change is approximately zero. Then:*

$$\boxed{\delta S \approx 0 \implies S(q) = \int_{t_0}^{t_1} dt\, \mathcal{L}(q, \dot{q}, t)}$$

2.3 Hamiltonian Mechanics

If you understand Lagrangian mechanics, it is very easy to understand Hamiltonian mechanics. It is based on the same idea, however, now instead of subtracting potential energy from kinetic energy, we take the sum of all the energies in our system. This is the definition of the Hamiltonian:

$$\mathcal{H} = \mathcal{T} + \mathcal{U}$$

However, we can also re-write this in terms of the Lagrangian through an analysis of its time variance:

$$\frac{d}{dt}\mathcal{L}(q, \dot{q}, t) = \sum_{i \to n} \left(\frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i \right) + \frac{\partial \mathcal{L}}{\partial t}$$

Let us analyze this, term by term. The first term can be written as:

$$\frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i = \dot{p}_i\dot{q}_i$$

And the second term can also be re-written as:

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i = p_i\ddot{q}_i$$

Now we can replace the terms in the parenthesis with:

$$\sum_{i \to n} \frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i = \frac{d}{dt}\sum_{i \to n} p_i\dot{q}_i$$

And then we get a beautiful identity by replacing this in the original equation:

$$\frac{d\mathcal{L}}{dt} = \frac{d}{dt}\sum_{i\to n} p_i\dot{q}_i + \frac{\partial\mathcal{L}}{\partial t}$$

And by definition, the Hamiltonian, $\mathcal{H}$ is defined as:

$$\mathcal{H} = \frac{d}{dt}\sum_{i\to n} p_i\dot{q}_i - \mathcal{L} \implies \boxed{\frac{d\mathcal{H}}{dt} = -\frac{\partial\mathcal{L}}{\partial t}}$$

Now that we have some idea of classical physics, let us start from the dawn of modern physics.

3 Special Relativity

Let us take a break from all that complicated math and use a small brain teaser. I promise you in this section you will see less calculus, mostly just "simple" algebra, and a few thought experiments. It is from here that we leave the good old classical thinking, and get exposed to some of the weirdisms that occur in relativistic cases. Vamonos!

3.1 Postulates

1. The Principle of Relativity

This postulate states that there is no physical way to differentiate an object moving at constant velocity and an object at rest. The object could be moving relative to another object, but it is impossible to determine which one is moving and which one is at rest - *viz.* A and B are moving relative to me, but who says it is not A and I moving relative to B? In other words, the laws of physics will have the same form in all inertial frames of reference (frames in which the net force on the observer is 0) no matter what.

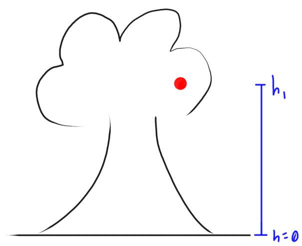
2. The Principle of Invariant Light Speed

This postulate is simple. It states that the speed of light is a constant for all reference frames. Consequently, that means we cannot have a frame of reference at the speed of light synchronizing with a photon so that it is at rest. This brings the idea that the speed of light is the ultimate speed and nothing can go faster than the speed of light.

3.2 Relative positions

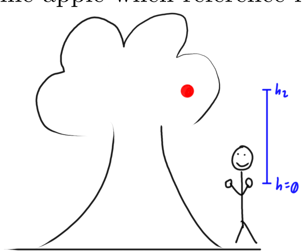
When we make a measurement, we are actually making two: one at zero and one at the measured value. We define the value at zero as the origin. For example, if we are measuring how high an apple is on a tree, we are measuring the value of zero at the ground, and measuring the value of the height at the apple. This value is the height of the apple with respect to (WRT) the ground.

Figure 1: Height of the apple when the reference height is at ground level.



What if we don't want to start measuring at ground level?

Figure 2: Height of the same apple when reference height is someone's hands.



We say that the y-value of the apple is h_1 WRT the ground, and y-value is h_2 WRT the person's hand. To get to h_2 from h_1, we have to subtract the difference between initial height

$$h_2 = h_1 - \Delta h$$

and vice versa

$$h_1 = h_2 + \Delta h$$

This is called the *Galilean Transformation.* We can apply this logic to the remaining two spatial dimensions to obtain a relative *position* WRT to someone else. We are a step closer to understanding things relatively. We are missing another variable - *time.* They say that time is relative, but how?

3.3 Lorentz Transformations

3.3.1 Derivation

An event is something that happens at a definite time and place - say a star exploding. Let us have two observers - you are moving right with a constant velocity u with respect to me. I assign the event with coordinates (x,t) and for you, it is (x',t'). Let us sync our clock, t' = t = 0, as you pass me. Thus the origin of our space-time matches.

Now, using Galilean Transformations, we can find a relation between my coordinates and your coordinates.

$$x' = x - ut \tag{1}$$

$$x = x' + ut \tag{2}$$

This is very easy to understand. As you are constantly moving towards the right (WRT me), the coordinates you assigned to the event at a time t will be offset by the distance you travelled during that time, or ut, relative to the coordinates I assigned.

What about light? From the second postulate, we know that we will agree on its velocity, but that means we must disagree on its length, time, or both. Hence, there should be a modifier factor γ :

$$x' = \gamma(x - ut) \tag{3}$$

$$x = \gamma(x' + ut') \tag{4}$$

Note two things: first, the time variable is no longer guaranteed to be equal between you and me. For me, t time has passed since we synced our clock until the star exploded. For you, *t'* has passed. These two times are not necessarily equal. Second, the modifier factor γ is the same for both you and me. This is from the first postulate, stating that both observers are equivalent.

Now we want to find an equation to describe γ. Suppose back when we synced our clock, we shot a light pulse toward the star and the instance that the pulse reached the star, the star exploded. Hence for me, the light took t time to travel x distance for me, while *t'* time to travel *x'* distance for you. Therefore for this event, we agree that:

$$x = ct \; and \; x' = ct' \tag{5}$$

We multiple the equations together

$$xx' = c^2tt' \tag{6}$$

Onto the algebra. We multiply Eqn 3 and 4 together and replace all x and x' with our Eqn 5 and 6.

$$\begin{aligned}
xx' &= \gamma^2(x - ut)(x' + ut') \\
xx' &= \gamma^2(xx' + xut' - x'ut - u^2tt') \\
c^2tt' &= \gamma^2(c^2tt' + uctt' - uct't - u^2tt') \\
\gamma^2 &= \frac{c^2tt'}{c^2tt' - u^2tt'} \\
\gamma &= \sqrt{\frac{1}{1 - \frac{u^2}{c^2}}} \\
\gamma &= \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}
\end{aligned} \tag{7}$$

Because γ is only dependent on your velocity as you move away from me, u, it doesn't matter we derived it from the light pulse. It can be used as a general case. Let us substitute Eqn 7 back into Eqn 3.

$$x' = \frac{x - ut}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{8}$$

Simple observations shows us that:

$$\lim_{u \to 0} \gamma = 1 \tag{9}$$

$$\lim_{u \to \infty} \gamma = \infty \tag{10}$$

Let us isolate t' from Eqn 4

$$t' = \frac{\frac{x}{\gamma} - x'}{u} \tag{11}$$

Substituting Eqn 7 and 8 into Eqn 11 and simplifying yields:

$$t' = \frac{t - \frac{ux}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{12}$$

And if I wish to compare two events, then the coordinate difference is:

$$\Delta x' = \frac{\Delta x - u\Delta t}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{13}$$

$$\Delta t' = \frac{\Delta t - \frac{u}{c^2}\Delta x}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{14}$$

Note that these are to get your coordinates from mine. If you want to get my coordinates from mine, then simply substitute u to -u (I will be moving towards the left WRT you).

3.3.2 Velocity Law

Let us suppose a particle moves, Δx distance in Δt for me and Δx' in Δt' for you. Let us define the instantaneous velocities:

$$v \equiv \frac{dx}{dt} \qquad \text{according to me} \tag{15}$$

$$w \equiv \frac{dx'}{dt'} \qquad \text{according to you} \tag{16}$$

$$u \equiv \text{your speed according to me}$$

If I think the particle is going at velocity v. What is w? We substitute Eqn 13 and 14 into 16.

$$\begin{aligned} w &= \frac{dx'}{dt'} \\ &= \frac{dx - udt}{dt - \frac{u}{c^2}dx} \\ &= \frac{dt}{dt} \times \frac{\frac{dx}{dt} - u}{1 - \frac{u}{c^2} \times \frac{dx}{dt}} \\ &= \frac{v - u}{1 - \frac{uv}{c^2}} \end{aligned} \tag{17}$$

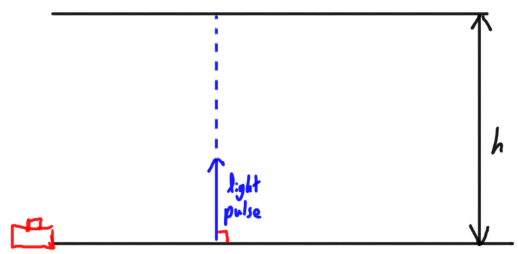
Similarly, if you think the particle is going at velocity w, what is v? We would simply substitute in u as -u.

$$\boxed{v = \frac{w+u}{1+\frac{uw}{c^2}}} \tag{18}$$

3.3.3 Time Dilation

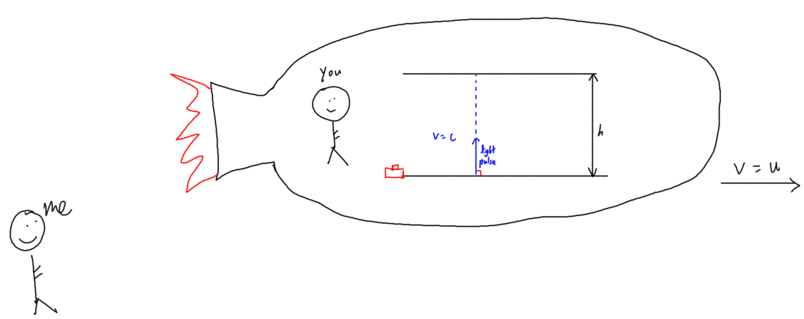
In this section, we will calculate how time is relative to one another. Let us do a quick thought experiment. Imagine two mirrors, parallel to each other and separated by a distance of h. On the lower mirror, there is a button which when pressed, releases a light pulse directly upwards.

Figure 3: After pressing the red button, a light pulse perpendicular to the mirrors is shot upwards.



Let us put this system onto a rocket. you will be on the rocket while I observe you outside the rocket.

Figure 4: Off you go with speed u!

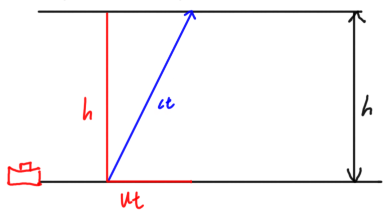


Since you, the system, and the rocket are moving together, everything (except for me) will be relatively stationary WRT you. The system will behave normally. Suppose we sync our clock as you press the button to release the light pulse. Some time has passed, and the light pulse has reached the upper mirror. The time is t for me and t' for you, which is not necessarily equal. From your reference frame we can see that

$$h = ct' \tag{19}$$

But what would the system look like to me? The entire rocket is moving away from me with speed u, including the light pulse. In my reference frame, the light pulse has vertical speed, moving upwards towards the upper mirror. It also has horizontal speed, travelling with the rocket.

Figure 5: Light pulse for me



It is not hard to see that in my reference frame, where t time has passed, the light pulse would be moving diagonally. From my reference frame we can use the Pythagorean theorem and find a relation between t and t':

$$\begin{aligned}
(ct)^2 &= h^2 + (ut)^2 \\
(ct)^2 &= (ct')^2 + (ut)^2 && \text{(Recall Eqn 19)} \\
c^2t^2 - u^2t^2 &= c^2t'^2 \\
t^2(c^2 - u^2) &= c^2t'^2 \\
\frac{t^2}{t'^2} &= \frac{c^2}{c^2 - u^2} \\
\frac{t}{t'} &= \sqrt{\frac{c^2}{c^2(1 - \frac{u^2}{c^2})}} \\
\frac{t}{t'} &= \sqrt{\frac{1}{1 - \frac{u^2}{c^2}}} = \gamma \\
t &= \gamma t' && (20)
\end{aligned}$$

To get from you time in your reference frame to my time in my reference frame, multiply it by a factor of γ. Recall Eqn 9 and 10: as your speed, u, approaches light speed, the γ factor approaches infinity - the faster you move the larger the γ factor is. And using Eqn 20 we just deducted, we can see that the larger the γ factor, the less time has passed for you compared to me. This is time dilation due to speed - the faster you move the slower time is for you.

3.3.4 Length Contraction

Suppose you are carrying a rod with length L_0 as its rest length. Since you are moving away from me at velocity u, how long will I say the rod is? In order to determine its length, I must make measurements of the spatial coordinates of its front and back end *simultaneously*. In this case, $\Delta x = L$ and $\Delta t = 0$, and we use equation 13.

$$L = L_0\sqrt{1 - \frac{u^2}{c^2}} \quad or \quad L = \frac{L_0}{\gamma} \tag{21}$$

Again, recall the that faster you go the larger the γ factor. This means that the faster you go, the shorter the distance will be for you compared to its rest distance.

3.3.5 Order of events

If event 1 causes event 2, then no observer, regardless of which frame of reference they are in, should see the order in reverse. On the other hand, if the events are not causally related, the reverse order will not lead to any contradictions. Recall Eqn 14.

$$\Delta t' = \frac{\Delta t - \frac{u}{c^2}\Delta x}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Suppose that I think event 2 happened after event one. Then, $\Delta t = t_2 - t_1 > 0$. Let us try to find a scenario where you think event 2 happened before event 1, or $\Delta t' = t_2' - t_1' < 0$. The denominator of the equation will always be positive, hence the only way to have a negative $\Delta t'$ is:

$$\begin{aligned} \frac{u}{c^2}\Delta x &> \Delta t \\ \frac{u}{c} &> \frac{c\Delta t}{\Delta x} \end{aligned} \tag{22}$$

Let us look at this equation. If $c\Delta t > \Delta x$, then $\frac{u}{c} > 1$, which is impossible according to our second postulate. What is $c\Delta t$? It is the distance that light travels. What is Δx? It is the spatial separation between the two events. This means that event 2 cannot happen before event 1 for all reference frames if the spatial separation between the events is smaller than the distance light covers during the time between event 1 and 2, *viz.* when they are *causally connected.* This connection works for all reference frames, so if in one reference frame the events are causally connected, you can bet that for all reference frames they will be causally connected.

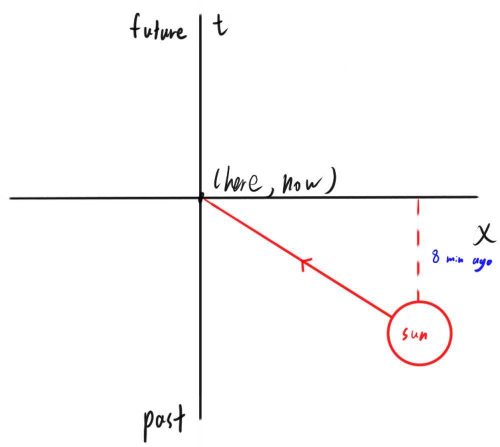
3.4 Spacetime

Now that we have familiarized ourselves slightly with some of the ideas presented in special relativity, we have to learn how it connects to gravity and the theory of spacetime. More importantly, what is spacetime, why should we care about spacetime, and how do we describe spacetime? These are the questions that some of Special Relativity, and most of General Relativity are about.

3.4.1 Minkowski Diagrams

Now we know that time is another dimension, let us graph it. We will ignore the y and z directions and assume only x direction movements.

Figure 6: 8-minute-old Sun ray reaching us.

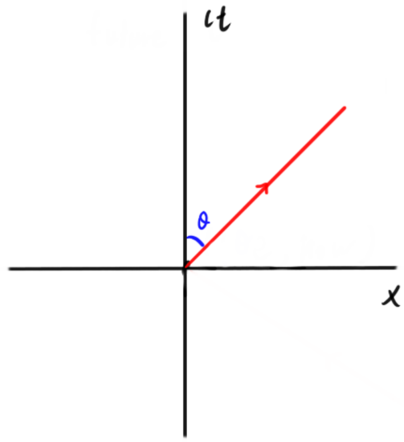


The line represents the path that the light ray took in spacetime, called a *worldline*. You can see that this seems like a displacement-time graph with axes flipped. Do not think of it that way. It is completely different from our diagram. For us time is simply another dimension.
Let us recalibrate our time axes by multiplying the time with the speed of light, c, so that its units

matches the units of other dimensions. Note that since c is a constant for all reference frames, our diagram will still be valid. However, since the speed of light is the cosmological limit, there are limits to what our worldline can be.

Let us suppose there is a photon travelling at the speed of light in the x-direction. It travelled one-light year, which took it, well, a year.

Figure 7: Worldline of a Lightray



Then

$$\tan\theta = \frac{\Delta x}{c\Delta t} = \frac{\textit{one light-year}}{c \times \textit{one year}} = 1$$
$$\theta = 45^\circ$$

This would be the maximum angle a worldline can be. Since particles can never exceed lightspeed, Δx will never be greater than $c\Delta t$.

Let us consider how the graph transfers to different reference frames. Suppose a photon moving away from us at half the speed of light.

$$\tan\theta = \frac{\Delta x}{c\Delta t} = \frac{0.5ct}{ct} = 0.5$$
$$\theta \approx 27^\circ$$

This worldline represents the photon in our stationary reference frame, *S*, and is at an angle from our *ct* axis. However, in the photon's frame of reference, it is stationary, hence its worldline coincides with its reference frame's temporal axis. Let us define the photon's frame to be *S'* and its temporal axis to be *ct'*. Then, the *ct'* axis of the *S'* frame would be equal to the worldline of that photon in our reference frame, *S*. We can prove this by using Lorentz Transformations at $t' = 0$

$$
\begin{aligned}
t' &= \gamma\left(t - \frac{vx}{c^2}\right) \\
0 &= t - \frac{vx}{c^2} \\
t &= \frac{vx}{c^2} \\
\frac{ct}{x} &= \frac{v}{c}
\end{aligned}
\tag{23}
$$

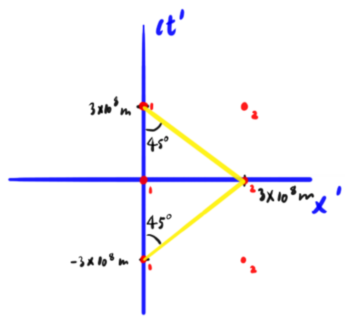
Since the photon is moving at a constant velocity, it does not matter that we used $t' = 0$. The ratio of v/c will always be a constant at any *t'*. Now, we can replace the tan

$$\tan\theta = \frac{ct}{x} = \frac{v}{c} \tag{24}$$

$$\theta = \arctan\frac{v}{c} \tag{25}$$

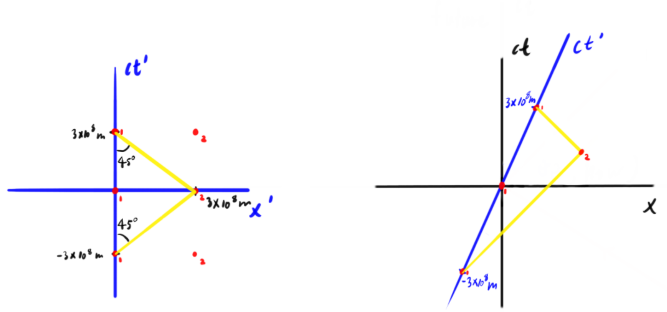
What is the translation of the x-axis? Let us imagine the following scenario: You and your friend are each on a spaceship with a mirror on their back, separated by a light-second away, moving at the same constant velocity on the x-axis. I will be observing you at $x = x' = 0$. Let's say you are closer to me compared to your friend. Now, you shine a flashlight at your friend who is a light-second away.

Figure 8: Path of the light.



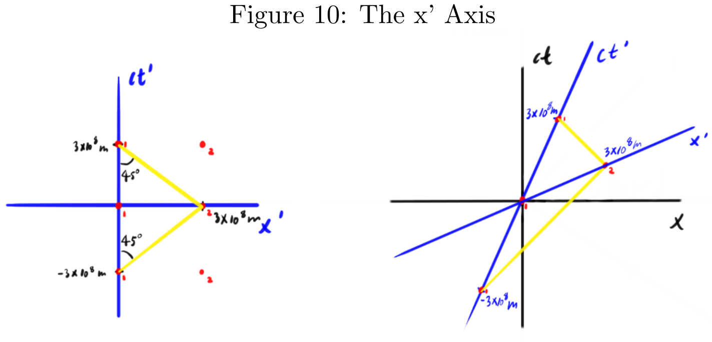
The red dots are the spaceships and the yellow path is the path of light. Since you and your friend have the same velocity according to me, your friend will be stationary according to you. You shined your flashlight a second ago ($3 \times 10^8 m$) in your reference frame, where it took the light a second to reach your friend's ship, then a second to be reflected to you. Now let us look at my frame of reference.

Figure 9: Your Axes in My Frame



We start by marking the position of your ship, which should be directly on your *ct'* axis. We know that the speed of light is a constant and that it must have a slope of 1 or -1. So you fired the light ray forward, and it gets reflected back to you. We know the starting position (a light-second in the past on your axis) and the slope of the forward light ray (1). We can draw that line. We know the ending position (a light-second in the future on your axis) and the slope of the returning light ray (-1). We can draw that line. The interception between these two lines has to be the position of your friend's ship, with the mirror reflecting that light.

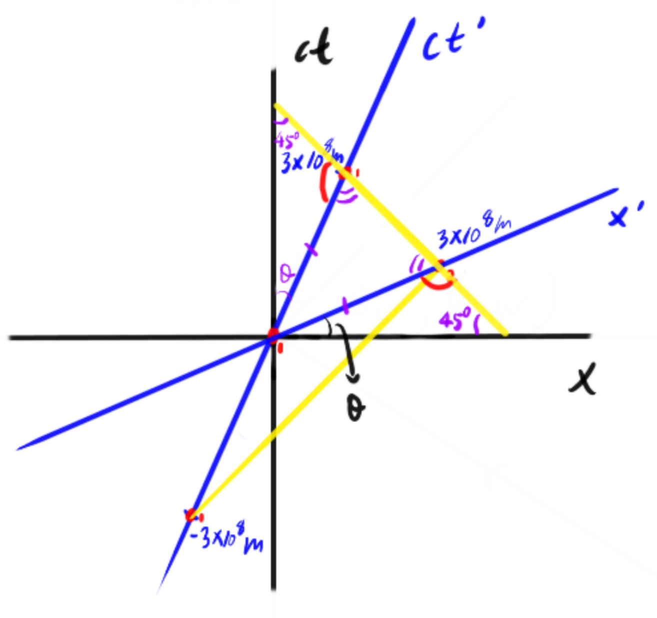
Figure 10: The x' Axis



So that is where in spacetime, I think your friend's ship is at. Where do you think your friend's ship is at? On the x' axis, where $t' = 0$, simultaneously as you reached the origin of your reference frame (which is also my origin). So we can connect these two points and find the x' axis in my reference frame.

We know the angle between the *ct* and *ct'* is θ. What is the angle between x and *x'*? We can find this with some quick geometry.

Figure 11: Extend the Returning Light Ray



After extending the returning light ray, we know that the lightray forms a 45° angle with the *ct* and x axes. We know that the angle between *ct* and *ct'* is θ. We know that the middle triangle, the one formed by lines *ct'*, *x'*, and the returning light ray, is an isosceles, since both the sides are 1 light-second long. This means that its base angles, denoted by purple, is equal. Hence, its supplementary angles, denoted by red, is equal. We now have two triangles, one with $(45°, red\ angle, \theta)$ and the other with $(45°, red\ angle, unknown)$, hence we know the unknown angle between x and *x'* must be equal to θ from the sum of angles in a triangle.

3.4.2 The Four Vectors and Space-time Interval

We have used only two dimensions before this: time and a spatial variable. Let us now consider all the variables in our world.

$$x \equiv (x_0, x_1, x_2, x_3) \equiv (ct, x, y, z) \equiv (x_0, r) \tag{26}$$

Let's say that you are moving away from me in the x axis. Let us define $\beta \equiv u/c$. Using Eqn 13:

$$\begin{aligned} \Delta x' &= \frac{\Delta x - u\Delta t}{\sqrt{1 - \frac{u^2}{c^2}}} \\ x_1' &= \frac{x_1 - \beta x_0}{\sqrt{1-\beta^2}} \end{aligned} \tag{27}$$

And using Eqn 14 for the time:

$$\begin{aligned} \Delta t' &= \frac{\Delta t - \frac{u}{c^2}\Delta x}{\sqrt{1 - \frac{u^2}{c^2}}} \\ x_0' &= \frac{x_0 - \beta x_1}{\sqrt{1-\beta^2}} \end{aligned} \tag{28}$$

The other axes, x_2 and x_3 remain unaffected as you are moving away from me in the x_1 direction only.

Recall that different reference frames can be represented by manipulating the axis of the spacetime diagram. Two observers in different frames may not agree on the coordinates of an event, but they will agree on the distance of that event to the origin of the spacetime graph. Notice the resemblance of Eqn 24 and 25 against the rotation of axes formula:

$$\begin{aligned} x_1' = \frac{x_1 - \beta x_0}{\sqrt{1-\beta^2}} &= x_1 \frac{1}{\sqrt{1-\beta^2}} - x_0 \frac{\beta}{\sqrt{1-\beta^2}} \\ x' &= x\cos\theta + y\sin\theta \end{aligned} \tag{29}$$

$$\begin{aligned} x_0' = \frac{x_0 - \beta x_1}{\sqrt{1-\beta^2}} &= x_0 \frac{1}{\sqrt{1-\beta^2}} - x_1 \frac{\beta}{\sqrt{1-\beta^2}} \\ y' &= y\cos\theta - x\sin\theta \end{aligned} \tag{30}$$

NOTE: if you don't know the axes rotation formula you can derive it really quickly using a little trigonometry and polar coordinates.

Since two observers still agree on the distance from the point to the origin:

$$x^2 + y^2 = x'^2 + y'^2 \tag{31}$$

They would also agree on the dot product of two vectors:

$$A \cdot B = A_x B_x + A_y B_y = AB\cos\theta = A'_x B'_x + A'_y B'_y \tag{32}$$

With A being the position vector $A = x\hat{i} + y\hat{j}$ in the example above, and B being an arbitrary vector.

However, we cannot say that

$$\frac{1}{\sqrt{1-\beta^2}} = \cos\theta \qquad \frac{\beta}{\sqrt{1-\beta^2}} = \sin\theta \tag{33}$$

since it does not obey $\sin^2\theta + \cos^2\theta = 1$. As such, Lorentz Transformation is not exactly a rotation of the axes. But we do have dot product

$$X \cdot X = x_0^2 - x_1^2 = x_0'^2 - x_1'^2 \tag{34}$$

We define this space-time interval as s^2 between the origin and point (x_0, x_1). So if two events are seperated in space by Δx_1 and time by Δx_0, then the square of the space-time interval

$$(\Delta s)^2 = (\Delta x_0)^2 - (\Delta x_1)^2 \tag{35}$$

is the same for all observers. $(\Delta s)^2$ is not positive definite. There are three scenarios:

1. $(\Delta s)^2 < 0$, $\Delta x_1 > \Delta x_0$, is a space-like separation
2. $(\Delta s)^2 > 0$, $\Delta x_0 > \Delta x_1$, is a time-like separation
3. $(\delta s)^2 = 0$, $\Delta x_0 = \Delta x_1$, is a light-like separation.

3.4.3 Momentum and Proper Time

During the classical era, we created a new vector

$$\frac{dr}{dt} = v$$

by dividing a vector by a scalar, we have obtained another vector. We then defined a new vector called momentum by multiplying v by another scalar

$$p = mv$$

Can we use these equations here? $\Delta x = (c\Delta t, \Delta r)$ is the change in coordinates of a point in spacetime. We cannot divide the four components by Δt because time is just another component - it is like dividing x by y. The only thing scalar, and universal between all observers, is the spacetime interval that we defined in the last section:

$$ds = \sqrt{(dx_0)^2 - (dx_1)^2} = cdt\sqrt{1 - \left(\frac{dx_1}{cdt}\right)^2} = cdt\sqrt{1 - \frac{v^2}{c^2}} \tag{36}$$

Where v is the velocity of a particle seen by the observer that assigned the coordinates x to this particle. Note that we can divide dx/dt to receive a velocity specific to a reference frame. we can divide out the constant c, and define a new variable τ:

$$d\tau \equiv \frac{ds}{c} = dt\sqrt{1 - \frac{v^2}{c^2}} = \sqrt{dt^2 - \frac{dx^2}{c^2}} \tag{37}$$

τ is called the "Proper Time", and is universal across all reference frames. Take a look at its units:

$$\sqrt{s^2 - \frac{m^2}{m^2 s^{-2}}} = s$$

We can now use τ, which is universal, as a replacement for t, which is specific to a single reference frame.

$$\begin{aligned} P &\equiv m\left(\frac{dx_0}{d\tau}, \frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau}\right) \\ &= m\left(c\frac{dt}{d\tau}, \frac{dr}{d\tau}\right) \\ &= (P_0, P) \end{aligned} \tag{38}$$

where m is called the (rest) mass of the particle, universal for all inertial reference frames. What exactly did we do? Instead of calculating the momentum using my time or yours (which we do not agree on), we take the time-derivative in the particle's frame of reference, according to its own clock. Let us ignore the y and z directions and assume movements in the x direction.
If you do not like the new variable τ for some reason, then we can simplify it. What is the equation according to my time?

$$\begin{aligned} d\tau &= dt\sqrt{1-\frac{v^2}{c^2}} \\ \frac{df}{d\tau} &= \frac{df}{dt} \times \frac{dt}{d\tau} = \frac{df}{dt} \times \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \end{aligned} \tag{39}$$

where f is an arbitrary function. Let us call it $f \equiv (mx_0, mx_1)$. Then, our momentum equation is now:

$$P = (P_0, P_1) = \left(\frac{mc}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}\right) \tag{40}$$

Note that if the particle is moving slowly, *viz.* $v/c \ll 1$, then $P_0 = mc$ and $P_1 = mv$.
Sometimes you will see other representations of this equations:

$$P_1 = \left(\frac{m}{\sqrt{1-\frac{v^2}{c^2}}}\right) v \equiv m(v)v \tag{41}$$

where $m(v) \equiv \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}$ is a velocity dependant mass, or *relativistic mass*. $m(0)$ is then defined as the *rest mass*. Then the momentum can always be represented as (relativistic) mass times velocity, like the good ol' days.

4 Riemannian Geometry and Tensors

Tensors are crucial mathematical tools to the study of General Relativity and higher levels of Quantum Mechanics. They are also crucial to understanding Riemannian Geometry which is the mathematics of General Relativity. Hence, let us start with understanding what tensors are.

4.1 Tensor Mathematics

In simple words, tensors are a generalization of vectors and scalars. When we think, geometrically, about a scalar value, what it is doing is assigning a value to a specific coordinate. Similarly, a scalar field assigns a scalar value to each point in the coordinate system. Vectors are similar. They assign one or more scalar values, which also correspond to a direction, to a specific point in the space and a vector field assigns a vector to each point in the space.

A tensor assigns one or more vectors to a point in the geometry and similarly, a tensor field assigns a specific tensor to each point in the geometry. The theory of General Relativity makes heavy use of tensors and tensor fields. From this definition, we can see that a vector or a scalar in itself is a type of tensor. Scalars are what we call a *tensor of rank 0* and vectors would be *tensors of rank 1.* To denote a tensor, we use a letter that is the name of that tensor and indices. A tensor of rank 0 would have no indices, a rank 1 would have 1 index, rank 2 would have 2 indices and so on. For example, a vector $\vec{V}$ with contravariant components could be represented by:

$$\vec{V} = V^m$$

And similarly, one with covariant components:

$$\vec{V} = V_m$$

We will learn what contravariant and covariant mean soon. Now let us talk about the transformation of a tensor. Given a rank 2 tensor with an upstairs index and a downstairs index, T^p_q that is being transformed from coordinates x^u to y^u, then the transformed tensor, T^i_j will be given by:

$$T^i_j = \frac{\partial y^i}{\partial x^p} \frac{\partial x^q}{\partial y^j} T^p_q \quad \text{or} \quad T^i_j = \mathcal{J}(x^u, y^u) T^p_q$$

There is one small thing to note here, in the first notation, there is no summation written while there has to be a summation over the p and q indices. This notation is called the Einstein Summation Convention. The formula in standard mathematical notation is:

$$\boxed{T^i_j = \sum_{pq} \frac{\partial y^i}{\partial x^p} \frac{\partial x^q}{\partial y^j} T^p_q}$$

However, the Einstein Summation Convention tells us that every time an index is repeated one time upstairs and one time downstairs, there is no need to write the summation sign as it is assumed that the indices are summed over which makes the formula take the form I initially wrote it in.

Why do we want to use tensors in general relativity ? Why do they matter so much ? Unlike other objects in geometry, tensors have the property that they are invariant under a transformation of reference frame. A tensor will not change when observed by two different observers and this quality allows us to construct frame independent equations that will be true in all reference frames. Even in accelerating ones!

All tensors will follow the above formula when being transformed from one coordinate system to another with additional or fewer transformation coefficients based on the number of their contravariant and covariant indices. This, however, brings back the question of what do contravariant and covariant mean ?

4.2 Contravariant vs. Covariant

Let us imagine some coordinate system x^m in which the basis vectors are $x^m = \{\vec{\mathbf{e}_1}, \vec{\mathbf{e}_2}, \vec{\mathbf{e}_3}\}$ which may or may not be orthogonal to each other. Now, we take some vector V^m which has contravariant components. If we think of this vector in the space we have, it has some components V_1, V_2, V_3. These components are with respect to the basis vectors. The question is what happens to our vector 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 ?

This is where covariant and contravariant components come in. There are different types of vector components that a vector may have. If a vector has covariant components, it means that its size will co-variate with any variation in the basis. If the basis gets smaller, the vector will be smaller and if the basis gets larger then the vector components also get larger. Its variation will be the same as the variations applied to the basis. Likewise, a contravariant vector is one whose components will vary contrary to the basis meaning if the basis shrinks the vector will be larger and if it gets larger the vector will be smaller. There is also invariant meaning it will not change no matter what variation you apply to the basis.

If a vector has covariant components, the index will be a subscript and if it is a contravariant vector, the index will be a superscript. That is why V^m is a contravariant vector whereas V_m is a covariant vector. A tensor may have both covariant and contravariant components or it may only have one type. It varies. In the previous example, the tensor T^i_j has two indices, one covariant index, j, and one contravariant index i.

4.3 The Metric Tensor

When we take a dot product of two vectors (with say, contravariant components although it is true for both) $V_1^m = V_{11}\vec{\mathbf{e_{11}}} + V_{12}\vec{\mathbf{e_{12}}} + V_{13}\vec{\mathbf{e_{13}}}$ and $V_2^n = V_{21}\vec{\mathbf{e_{21}}} + V_{22}\vec{\mathbf{e_{22}}} + V_{23}\vec{\mathbf{e_{23}}}$, we know that we must multiply the corresponding components by each other but what do we do with the direction ? This problem may seem of no importance as it is almost obvious that we are using the same coordinates. But what happens when the directionality vectors are from different coordinate systems ? In this case we will do the same thing as the components, we take the dot product of each directionality vector as well. This gives us the following result:

$$V_1^m \cdot V_2^n = V_{11}V_{21}\vec{\mathbf{e_{11}}}\vec{\mathbf{e_{21}}} + V_{12}V_{22}\vec{\mathbf{e_{12}}}\vec{\mathbf{e_{22}}} + V_{13}V_{23}\vec{\mathbf{e_{13}}}\vec{\mathbf{e_{23}}}$$

Similarly, if we take the cross product of the two vectors, $V_1^m \times V_2^n$, our direction vectors will be:

$$\vec{\mathbf{e}_i} \times \vec{\mathbf{e}_j}$$

where

$$\vec{\mathbf{e}_i} = \{\vec{\mathbf{e_{11}}}, \vec{\mathbf{e_{12}}}, \vec{\mathbf{e_{13}}}\}$$

$$\vec{\mathbf{e}_j} = \{\vec{\mathbf{e_{21}}}, \vec{\mathbf{e_{22}}}, \vec{\mathbf{e_{23}}}\}$$

Now we define the metric tensor, g_{mn} to be:

$$\boxed{g_{mn} = \vec{\mathbf{e}_i} \cdot \vec{\mathbf{e}_j}}$$

So that:

$$\mathbf{V} \cdot \mathbf{V} = V^m \, V^n \, \vec{\mathbf{e}_m} \, \vec{\mathbf{e}_n}$$

And now if we turn the vectors to differentials, we get:

$$dx \, dx = dx^m \, dx^n \, g_{mn}$$

Notice, that the direction vectors and the metric have covariant components not contravariant. However, we know that we can also find the inverse of the metric since the inverse multiplied by the matrix itself must give us the Kronecker delta symbol:

$$g_{mn} \, g^{mi} = \delta^i_n$$

Throughout this lecture we will be making great use of the Kronecker delta so you need to be familiar with it. In General Relativity, the Kronecker delta is treated like the identity matrix of the geometry. Here, its definition would be:

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

Which is by definition the identity matrix in this case but may change based on the context of the problem. The Kronecker delta is very similar to the Dirac delta function but not quite. While the Dirac delta is a function, the Kronecker delta is a matrix (not a tensor) which makes it a much more convenient tool in general relativity than the Dirac delta. We can use the Kronecker delta to find the inverse of any tensor we wish as it can take as many indices as we want it to take. Be careful, however, to not interpret it as a tensor. Not every matrix of numbers is a tensor, we simply use matrices to represent tensors but not vice-versa. Another, very easy to prove, formula that we must know is:

$$V_n = g_{mn}V^m$$

4.4 Generalizing Pythagoras' Theorem with Differential Geometry

A famous equation that we all learn either in elementary school or in high school is the famous Pythagorean theorem:

$$S^2 = x^2 + y^2$$

This formula allows us to calculate the distances between points in a Cartesian coordinate system (x, y). If we extend this to $\mathbb{R}^3$ such that we have (x, y, z), we get:

$$S^2 = x^2 + y^2 + z^2$$

We can also write this formula in different coordinate systems like spherical coordinates (r, θ, φ):

$$S^2 = r^2 + r^2\theta^2 + r^2\sin^2(\theta)\varphi^2$$

We are also allowed to state that the formula can also be written for a change in all these quantities:

$$\Delta S^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \quad \text{or} \quad \Delta S^2 = \Delta r^2 + r^2\Delta\theta^2 + r^2\sin^2(\theta)\Delta\varphi^2$$

And as the size of the interval approaches zero, the intervals become differentials and the formula takes the form:

$$dS^2 = dx^2 + dy^2 + dz^2 \quad \text{or} \quad dS^2 = dr^2 + r^2d\theta^2 + r^2\sin^2(\theta)d\varphi^2$$

Now how can we generalize this theorem ? One of the objectives of mathematics is developing tools needed to generalize cases that are true for specific case and make them work for all possible cases. When observing the pattern we see that the formula is always the product of two dimensions of the coordinate system, multiplied by some coefficient beside it. These coefficients are given to us by the metric tensor. Hence the formula generalizing the Pythagorean theorem which is the central idea of Differential Geometry and General Relativity was formulated by Bernard Riemann as:

$$dS^2 = \sum_{mn} g_{mn}\, dx^m\, dx^n$$

Or using the Einstein Summation Convention:

$$\boxed{dS^2 = g_{mn}\, dx^m\, dx^n}$$

There many ways to calculate the metric tensor. You can even use Pythagoras' theorem itself to calculate the metric! What this formula allows us to do is calculate every infinitesimal distance in the geometry of our space (or space-time). The metric tensor tells us how the curvature of the coordinate system can be accounted for. It is like a matrix of scaling factors that allow for the generalization of Pythagoras' theorem.

Now imagine that we have a tensor field of metric tensors, $g_{mn}(x^p)$, such that each point in our space has its own curvature. This suddenly opens up endless possibilities of having different geometries. There may be some places in which the geometry is hyperbolic, some where the geometry is spherical and some where it is flat/Euclidean. How can we tell this ? If the metric tensor at some point is equal to the Kronecker delta, then Pythagoras' theorem will take the form:

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

Which is only true for flat coordinates. So this equation allows for us to learn what the geometry of our space is like. It is perhaps the second most important equation in general relativity. I know that sounds odd, however, after learning learning the Einstein Field Equations we will see why it is the second most important. Before that, we must familiarize ourselves more with tensors.

4.5 Properties of Tensors and Tensor Algebra

The first formula we learned, which we will be making great use of is the transformation formulas:

$$(V')^m = \frac{\partial y^m}{\partial x^n} V^n$$

$$(V')_m = \frac{\partial x^n}{\partial y^m} V_n$$

Some tensors have the property of being *symmetric*. This means it does not matter in which order we write their indices. For example, the metric tensor is a symmetric tensor. This means that:

$$g_{mn} = g_{nm}$$

Similar to vectors and scalars, tensors with equal number of covariant and contravariant indices can be subtract from each other:

$$T^{lmn}_{pqr} - U^{lmn}_{pqr} = 0$$

$$\implies T^{lmn}_{pqr} = U^{lmn}_{pqr}$$

Tensors are additive if their indices match. Given,

$$T = T^{m\ldots}_{\ldots p} \qquad S = S^{m\ldots}_{\ldots p}$$

then:

$$T + S = (T + S)^{m\ldots}_{\ldots p}$$

We can also multiply tensors of different or similar indices. Take tensors V^m and W_n, then:

$$V^m \, W_n = T^m_n$$

You may also see this notation for multiplying tensors:

$$V^m \otimes W_n = T^m_n$$

In this lecture however, we will denote the tensor product simply by writing no symbols between them and just putting them beside each other. Another extremely important property we will make use of is *tensor contraction*. To understand tensor contraction, we will use a lemma. Consider the quantity:

$$\frac{\partial x^b}{\partial y^m} \frac{\partial y^m}{\partial x^a}$$

Reminder that according to the summation convention, there is a summation to be performed on m. Let us imagine any function, F such that $F = F(y^m)$. Replacing x^b in the above quantity with F, we get:

$$\frac{\partial F}{\partial y^m}\frac{\partial y^m}{\partial x^a}$$

Which is equal to $\partial F/\partial x^a$. What the expression is saying is the rate of change of F as we change x^a little by little. So what happens if F happens to be x^b ? Then the whole shebang becomes a trivial looking expression:

$$\frac{\partial x^b}{\partial y^m}\frac{\partial y^m}{\partial x^a} = \frac{\partial x^b}{\partial x^a}$$

And when we analyze this, we see that after differentiation, we will have 1s where a and b are equal, and 0s everywhere else. This is, in fact, equal to the Kronecker delta!

$$\frac{\partial x^b}{\partial x^a} = \delta^b_a$$

This is how we can find the inverse of tensors. We will see why this matters in the next section where we learn to distinguish between flat and curved space.

5 Flat vs. Curved Spacetime

The main purpose of this section is for us to build tools that help us determine whether a space is flat or curved. The reason of doing this is because in nature, this problem is the same as telling whether a gravitational field is an artifact of the coordinate system we are using or if it is really a gravitational field. First, we must learn about determining whether the geometry at a specific point's surroundings is flat or not. Then we expand the idea to the entire geometry.

5.1 Gaussian Normal Coordinates

To help identify if the geometry around a point is flat, we use a tool called Gaussian Normal Coordinates. To do this, let us imagine some point P on a 2-sphere. A 2-sphere is a sphere which has a fixed r of $r = 1$. The coordinate system on this sphere is simply (θ, ϕ). Now, we construct, on point P, a plane. This plane has flat lines on it which are orthonormal. This means that the angle between them is always 90^o. Essentially, we build a 2D Cartesian Coordinate system on the tangent plane. To describe this flat space we use the Kronecker delta:

If the space is flat, the metric tensor is equivalent to the Kronecker delta.

So for ordinary Euclidean geometry where we are using Cartesian coordinates, Pythagoras' formula takes the form:

$$dS^2 = \delta_{mn}\, dx^m\, dx^n$$

and in curvilinear coordinate systems takes the form:

$$dS^2 = g_{mn}\, dx^m\, dx^n$$

5.2 Tensors Along Curved Coordinates

The most important concepts that you will need to remember from this section are the derivative of a tensor and the Riemann curvature tensor. Let us get into it!

When we discussed generalization of Pythagoras' theorem, we learned that the metric tensor gives us a way to scale the theorem in order for it to be suitable with all sorts of curvy coordinate systems we get under the Jacobin transformation rule. Now, we must ask ourselves, does the metric change as we move along a curve ? The metric tells us how to transform from Cartesian coordinates to any other coordinates that we need but how does the metric itself change as we move along some curve in our curved space ? To answer that we must look at the rate of change of the metric which means, as expected, that we must differentiate the tensor! So how do we go around this ?

Let us start from what our intuition tells us; when we differentiate, there has to be a factor that tells us the transformation of the metric under our coordinate system. The next part we will, for now, denote as ξ. That gives us the following expression for the *covariant derivative of tensor* V^m:

$$D_r V^m = \partial_r V^m + \xi$$

Where ∂_r is a shorthand notation for the expression:

$$\partial_r V^m \equiv \frac{\partial V^m}{\partial x^r}$$

Now we have to figure what ξ is. Again, intuitively, it must have something to do with the variation of metric along a curve. We need a measure of how much has our metric changed with respect to the previous basis vectors. To do this we use an extremely helpful tool in differential geometry known either *Christoffel Symbols* or *Connection Coefficient* denoted using Γ^t_{mn}. The letters we used for the indices are only temporary. We will come back to learn more about them soon, however, this tool allows us to find that the covariant derivative of a tensor is given by the expression:

$$\boxed{D_r V^m = \partial_r V^m + \Gamma^t_{rm} V^t}$$

And for a tensor with a covariant index we get:

$$\boxed{D_r V_m = \partial_r V_m - \Gamma^t_{rm} V_t}$$

Of course, all of these infinitesimally small differentials is calculated over Gaussian normal coordinates. Note that in the above expression there is a summation to be performed over the index t, making the whole thing look like:

$$D_r V_m = \sum_t \partial_r V_m - \Gamma^t_{rm} V_t$$

There is a little bit of a miscommunication here, the term *covariant derivative* does not correspond to the term covariant that we learned about vector components or tensor indices. Really what it means is that the derivative is invariant under a transformation of reference frame. It is a tensor derivative which means the components must be valid in all reference frames.

Now, let us apply this to a vector of contravariant components and then we will see how the rule applies to higher rank tensors. The only difference is that the term with the Christoffel Symbol will have a plus sign beside it. There is nothing else that is different! So the derivative of a rank 1 tensor with a contravariant component, V^m, would be:

$$D_r V^m = \partial_r V^m + \Gamma^t_{rm} V^t$$

Where, again, the index t is summed over. When differentiating tensors with multiple indices, we follow the same logic; the first term is always the coordinate derivative and the next terms would be a christoffel symbol multiplied to the tensor with one common index being summed over. If the index is covariant, the term is subtracted and if it is contravariant, it is added. For example, let us take the derivative of the metric:

$$D_r g_{mn} = \partial_r g_{mn} - \Gamma^t_{rm} g_{tn} - \Gamma^t_{rn} g_{tm}$$

Now, let us dive deeper into Christoffel symbols and what they are.

5.3 Christoffel Symbols

We have met Christoffel symbols now that we have discussed the covariant derivative of tensors. But what are they really ? How do we interpret a christoffel symbol ? I will define this tool for you in 2 ways, one mathematical definition and a geometric interpretation of what it helps us do. Usually, professors in universities like to start with the mathematical definition and then define a geometric interpretation, however, I prefer the other way.

As we discussed in the previous subsection, the connection coefficients first show up when we differentiate a tensor. We said that the first term is to adjust for any new coordinate changes along the curve. The proceeding terms are by definition factors that express the variation of the metric and then are put into the derivative. This is just like a normal function:

$$D\,f(x) = f(x) + dx$$

From this, we learn that the connection coefficients are almost like scaling factors! They are used in order to measure the variation of the tensor at hand with respect to its previous components and basis vectors. They show up even more in Calculus of Variations. Essentially, they measure the rate of variation of the metric tensor along the direction of differentiation in the space at hand. Now let us define the symbol mathematically.

To do so, first, we must learn one thing abou the christoffel symbols. Similar to how the metric is a *symmetric* tensor, christoffel symbols are also symmetric. This means we can interchange their indices:

$$\Gamma^t_{mn} = \Gamma^t_{nm}$$

Using this property, let us rewrite the derivative of the metric tensor in three different index orders: (rmn), (mnr), (nrm):

$$D_r g_{mn} = \partial_r g_{mn} - \Gamma^t_{rm} g_{tn} - \Gamma^t_{rn} g_{tm}$$

$$D_r g_{mn} = \partial_m g_{nr} - \Gamma^t_{nm} g_{tr} - \Gamma^t_{mr} g_{tn}$$

$$D_r g_{mn} = \partial_n g_{rm} - \Gamma^t_{nr} g_{tm} - \Gamma^t_{nm} g_{tr}$$

Now, to optimize the path (according to Lagrangian mechanics), we consider the scenario in which $D_r g_{mn} = 0$. Then, subtracting the second and third equation from the first equation, we get:

$$\partial_r g_{mn} - \partial_m g_{nr} - \partial_n g_{rm} + 2\Gamma^t_{mn} g_{rt} = 0$$

This operation allows us to have only one christoffel symbol in our equation rather than multiple ones with different indices. To solve for it, we isolate the expression for the last term:

$$\Gamma^t_{mn} g_{rt} = \frac{1}{2} \left[\partial_m g_{nr} + \partial_n g_{rm} - \partial_r g_{mn}\right]$$

To isolate for the christoffel symbol, we cannot just divide by the metric. We must use the property that allows us to convert the metric to the Kronecker delta! This means we need the inverse of the metric since:

$$g_{mn} \otimes g^{rn} = \delta^r_m$$

Hence, after multiplying both sides of the equation by the inverse metric we finally arrive at the mathematical definition of christoffel symbols:

$$\boxed{\Gamma^t_{mn} = \frac{1}{2} g^{rt} \left[\partial_m g_{nr} + \partial_n g_{rm} - \partial_r g_{mn}\right]}$$

When we consider that the metric is a 4×4 matrix and that the christoffel symbols involve multiple derivatives of the metric multiplied by the inverse metric, we learn about, perhaps the greatest problem with general relativity; calculations. These calculations are so rigorous that even the world's best supercomputers take hours to days to calculate these results just to calculate some orbit or the path of some light ray. In regards to setting the derivative equal to 0, it is physically simple to prove that there exists a reference frame in which the derivative of the metric is zero.

5.4 The Riemann Tensor

Let us imagine a cone with a rounded summit. This cone has the property that it is curved around its summit but flat everywhere else. Recall that if a shape that is extrinsically curved, can be re-arranged or unfolded to a flat surface, then it is not intrinsically, and therefore not "really" curved. Now, around this curved summit, let us imagine some vector V^m. How do the components of this vector vary as we move this vector around the summit ? Well, if we move it around the intrinsically flat part, nothing happens; there is simply a linear transformation on the vector's components. But how about the curved section ? The components of the vector vary, but the variations of these changes is what tells us the curvature by definition. Hence, if we can measure the variation of the changes in the vector, we can characterize the curvature of the summit. This, of course, is by definition the second covariant derivative of the vector. However, we must be careful to differentiate with respect to different directions as we do not want to repeat the same derivative as it has the property that it will be constant. Therefore, the second covariant derivative of the vector will be:

$$D_s D_r V_m$$

In ordinary calculus, this is the same thing as saying the order of partial differentiation in a nice multivariate function does not matter. In tensor calculus, however, the order matters! Hence, taking a derivative like this:

$$D_s D_r V_m - D_r D_s V_m$$

Would be equivalent to going around a loop in our curved space. This yields to an important principle:

In flat space, the indices of direction are interchangeable, in curved space they are not. Let us compute this derivative.

$$D_s \left[D_r V_m \right] = D_s \left[\partial_r V_m - \Gamma^t_{rm} V_t \right]$$

We know how to compute this derivative again, just brute force through the differentiation. Note that this gives a tensor. When we take the difference of these different derivatives, it gives us another tensor $\mathcal{R}$ which is even better as it has not 2 but 4 indices!

$$D_s D_r V_m - D_r D_s V_m = \mathcal{R}^t_{srm} V_t$$

$$\text{where} \quad \boxed{\mathcal{R}^t_{srm} = \partial_r \Gamma^t_{sm} - \partial_s \Gamma^t_{rm} + \Gamma^p_{sm}\Gamma^t_{rp} - \Gamma^p_{rm}\Gamma^t_{sp}}$$

This tensor is known as the *curvature tensor* or *the Riemann tensor*. It characterizes the curvature of space and tells us where the geometry is hyperbolic, where it is spherical, and where it is flat. We will be able to use this tensor in the Einstein Field Equations to calculate the curvature of spacetime. But before that, a quick section on geodesics.

Once we learn the kinematics of objects in curved and flat spacetime, we can analyze this behavior to come up with the dynamics on that object. It is just like good old Newtonian mechanics! In summary, the space is flat if and only if the curvature tensor is everywhere zero. The curvature tensor has a complicated form given by the above equation. But when we know the metric, the curvature tensor can be computed at every point of the space. Therefore, it is a practical tool.

5.5 Geodesics

From now on, we will start using spacetime instead of ordinary space. This means, Pythagoras' theroem takes the form:

$$d\tau^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

or applying the spacetime metric given to us by Hermann Minkowski, also known as the *Minkowski metric*, the theorem takes the form:

$$d\tau^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$$

Note that when we use greek indices, it means we are summing over from 0 to 3, with 0 being the time dimension whereas when using Roman indices like English letters, we are only summing over from 1 to 3. So, in spacetime, Pythagoras' theorem takes the form:

$$d\tau^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$$

and the Riemann tensor becomes:

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

Now, let us use the equivalence principle to derive an equation for the motion of particles in spacetime. Before that, however, we must ask what exactly are geodesics ? In short, they are the shortest path from one point to another. They may seem curved when drawn in flat space but for that specific geometry, they are the shortest possible path a particle can take to get there. Airplanes, Space shuttles and fishing ships in the ocean all use geodesics to calculate their paths. To start, let us consider a particle free falling with respect to a free falling coordinate system x^μ. Then:

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

Using the chain rule and then the product rule, we can manipulate this equation to get:

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

$$\implies \frac{d^2x^\nu}{d\tau^2}\frac{\partial x^\mu}{\partial x^\nu} = -\frac{dx^\nu}{d\tau}\frac{\partial^2 x^\mu}{\partial x^\nu\,\partial x^\alpha}$$

Now, we multiply both sides by the expression $\partial x^\lambda/\partial x^\mu$. This gives us:

$$\frac{d^2x^\lambda}{d\tau^2} = -\frac{dx^\nu}{d\tau}\frac{dx^\alpha}{d\tau}\left[\frac{\partial^2 x^\mu}{\partial x^\nu\,\partial x^\alpha}\frac{\partial x^\lambda}{\partial x^\mu}\right]$$

The terms in the bracket is another definition of the christoffel symbol:

$$\Gamma^\lambda_{\nu\alpha} = \frac{\partial^2 x^\mu}{\partial x^\nu\,\partial x^\alpha}\frac{\partial x^\lambda}{\partial x^\mu}$$

Which gives us the geodesic equation:

$$\boxed{\frac{d^2x^\lambda}{d\tau^2} = -\Gamma^\lambda_{\nu\alpha}\frac{dx^\nu}{d\tau}\frac{dx^\alpha}{d\tau}}$$

This can also be derived using the least action principle (also known as *Hamilton's Principle*):

$$S = \int dS$$

Where

$$dS = \sqrt{-g_{\mu\nu}\,dx^\mu\,dx^\nu} \iff dS^2 = -d\tau^2$$

Note the reason why the square root is negative is because we use the signature (-+++) or the opposite and so it has to be time-like.

$$S = \int \sqrt{-g_{\mu\nu}\, dx^\mu\, dx^\nu}$$

$$= \int d\lambda \sqrt{-g_{\mu\nu} \frac{dx^\mu}{dx^\lambda} \frac{dx^\nu}{dx^\lambda}}$$

Now, using the principle of least action (recall the second postulate of lagrangian mechanics), we are going to vary S along some curve x^μ

$$\delta S = \int d\lambda\, \delta \left(\sqrt{-g_{\mu\nu} \frac{dx^\mu}{dx^\lambda} \frac{dx^\nu}{dx^\lambda}} \right) = \int d\lambda \frac{\delta \left(-g_{\mu\nu} \frac{dx^\mu}{dx^\lambda} \frac{dx^\nu}{dx^\lambda} \right)}{2\sqrt{-g_{\mu\nu} \frac{dx^\mu}{dx^\lambda} \frac{dx^\nu}{dx^\lambda}}} = 0$$

Using the product rule and the substitution: $\frac{d\tau}{d\lambda} = \sqrt{-g_{\mu\nu} \frac{dx^\mu}{dx^\lambda} \frac{dx^\nu}{dx^\lambda}}$

$$= \int d\lambda \left(\delta g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\tau} + g_{\mu\nu} \frac{d\delta x^\mu}{d\lambda} \frac{dx^\nu}{d\tau} + g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d\delta x^\nu}{d\tau} \right) = \int d\lambda\, \delta g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\tau} + 2 g_{\mu\nu} \frac{d\delta x^\mu}{d\lambda} \frac{dx^\nu}{d\tau}$$

Now we integrate by parts

$$= \int d\tau \left(g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \partial_\alpha g_{\mu\nu} \delta x^\alpha - 2\delta x^\mu \partial_\alpha g_{\mu\nu} \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} - 2\delta x^\mu \partial_\alpha g_{\mu\nu} \frac{d^2 x^\nu}{d\tau^2} \right)$$

Simplifying this equation and after a small manipulation we get

$$= \int d\tau\, \delta x^\mu \left(-2 g_{\mu\nu} \frac{d^2 x^\nu}{d\tau^2} + \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \partial_\mu g_{\alpha\nu} - \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \partial_\alpha g_{\mu\nu} - \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} \partial_\nu g_{\alpha\mu} \right)$$

Multiplying both sides by $-\frac{1}{2}$ and factoring the term $\frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau}$, we get the form:

$$0 = \int d\tau\, \delta x^\mu \left(g_{\mu\nu} \frac{d^2 x^\nu}{d\tau^2} + \frac{1}{2} \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} (\partial_\alpha g_{\mu\nu} + \partial_\nu g_{\alpha\mu} - \partial_\mu g_{\alpha\nu}) \right)$$

Which helps us establish, according to Hamilton's principle, that the Euler-Lagrange equations is the expression

$$g_{\mu\nu} \frac{d^2 x^\nu}{d\tau^2} + \frac{1}{2} \frac{dx^\alpha}{d\tau} \frac{dx^\nu}{d\tau} (\partial_\alpha g_{\mu\nu} + \partial_\nu g_{\alpha\mu} - \partial_\mu g_{\alpha\nu}) = 0$$

Multiplying by the inverse metric, $g^{\mu\beta}$ gives us

$$\frac{d^2x^\nu}{d\tau^2} + \frac{1}{2}g^{\mu\beta}(\partial_\alpha g_{\mu\nu} + \partial_\nu g_{\alpha\mu} - \partial_\mu g_{\alpha\nu})\frac{dx^\alpha}{d\tau}\frac{dx^\nu}{d\tau} = 0$$

Now, we substitute $\Gamma^\beta_{\alpha\nu}$ instead of its definition and we bring that term to the right. This gives us the same geodesic equation:

$$\boxed{\frac{d^2x^\nu}{d\tau^2} = -\Gamma^\beta_{\alpha\nu}\frac{dx^\alpha}{d\tau}\frac{dx^\nu}{d\tau}}$$

All this calculation just goes on to show how rigorous and difficult it is to calculate outcomes in General Relativity. This means, we have now pretty much figured out the kinematics of particles in a relativistic framework. A good exercise to get more comfortable is to derive this equation from the Lagrangian for a charged particle under a Lorenz force. You should get:

$$\boxed{\frac{d^2x^\mu}{ds^2} = -\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} + \frac{q}{m}F^{\mu\beta}\frac{dx^\alpha}{ds}g_{\alpha\beta}}$$

6 Einstein Field Equations

The Einstein Field Equations is one tensor equation that can be used to model the effect of the presence of Energy and Momentum on Spacetime. To do so, we must first familiarize ourselves with the Energy-Momentum Tensor.

6.1 The Energy-Momentum Tensor

To replace the right hand side of $F = ma$, we must come up with something that models mass in relativity. We know from the famous equation of Einstein, $E = mc^2$ that energy and mass are related quantities. As a matter of fact, if we pick our units so that the scaling is by a factor of c^2, then we simply get $E = m$. Hence, if we relate Energy to spacetime, we have basically related mass to spacetime. To do this, we use the concept of energy density, how much energy there is in an infinitesimally small volume dV. Which brings us to Poisson's equation. We start from Newton's equation for a gravitational field:

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

We must replace the right hand side with Energy density $\rho_E(x, y, z)$. However, we also know that momentum may be involved if there is energy. We can introduce a tensor for this. First we define the field $\vec{P}$ to be:

$$\vec{P} = (E^0, P)$$

And we introduce the tensor in spacetime, $T_{\mu\nu}$ where the first component of the tensor is, by definition, the energy flow:

$$T^{00} = \rho_E(x, y, z)$$

And so after re-scaling the equation of Newton, we find that the Energy-momentum tensor $T_{\mu\nu}$ is fit to the equation by a factor of $2/c^4$. Hence we get for the right hand side of Newton's equation:

$$\nabla^2 g = ?\, \frac{8\pi G}{c^4} T^{\mu\nu}$$

The other components of the energy momentum tensor describe momentum density, energy flux, momentum flux, shear stress and pressure in spacetime. We will not focus on those here as there are plenty of sources that explain the tensor in further detail and this handout is already too long!

6.2 The Einstein Tensor: The Final Mystery!

We have now figured out the right hand side of what will be a relativistically accurate representation of gravity in spacetime. Now, we must come up, for the left hand side, with an expression that describes the curvature of spacetime since we know from the equivalence principle that gravity is simply an artifact of curvature in spacetime. To do so, let us go back to our good old curvature tensor $\mathcal{R}^{\lambda}_{\rho\mu\nu}$. Obviously, the first guess a physicist like Einstein would take is to set the Riemann tensor equal to the right hand side of Newton's modified equation. Unfortunately it does not work as the curvature tensor is a 4 by 4 matrix of 4 by 4 matrices. In other words, the curvature tensor has 4 indices while the energy-momentum tensor only has 2. Hence, we must find a way to get these two tensors to be equal to each other. In order to solve this issue, Einstein temporarily introduced a tensor $G^{\mu\nu}$ and put that in the left hand side. This gives is the initial version, but incomplete, of Einstein's field equations:

$$G^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu}$$

This is our first tensor equation! Now, the mystery is what is $G^{\mu\nu}$? It is known as the Einstein tensor. Now, the only mystery remaining is evaluating the Einstein tensor.

6.3 Evaluating the Einstein Tensor

To find the value of the Einstein tensor, we must first go back to tensor algebra where we learned about tensor contraction. We learned that we can contract a tensor which decreases the number of its indices by 2. Let us apply this technique to the Riemann tensor:

$$\mathcal{R}^{\lambda}_{\rho\mu\nu} R^{\rho}_{\lambda} = \mathcal{R}_{\mu\nu}$$

This new tensor we have is known as the *Ricci tensor*. We can also contract the Ricci tensor or simply, take its trace:

$$\mathcal{R} = \mathrm{tr}_{g_{\mu\nu}}(\mathcal{R}_{\mu\nu})$$

This value is known as the *curvature scalar*. Why does this value matter though ? Well, let us see; first we try putting the Ricci tensor equal to the left hand side of Einstein's equation:

$$\mathcal{R}^{\mu\nu} = \frac{8\pi G}{c^4} T^{\mu\nu}$$

This does not work as there are instances in space where energy is not conserved, momentum is not conserved and other things. We know, intuitively from Newton's equation, that it must involve, some sort of second derivative of the curvature. The Ricci tensor itself is like a derivative of the curvature tensor. Hence, let us see what is it that is missing. After grinding through the pages of calculation you get that:

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

We obviously do not have this term in the equation but it gives us what we need to add to the right side and this brings us to our answer! Hence, we arrive at the final form of Einstein's Field Equations:

$$\boxed{\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R} g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}}$$

Or if we write it with contravariant indices:

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

From this we see 2 things. One that we have finally found an equivalent expression for the Einstein tensor;

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

And the second one is that Einstein's Field Equations have the property that they work with both covariant and contravariant indices! Another form you may see is one in which the unit of scale is set as $c = 1$ which gives a simpler form:

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

There is one other form of Einstein's equation that you may see:

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = \kappa T^{\mu\nu}$$

Here the term Λ is known as the *cosmological constant*. The reason why this formulation is not the best is because of this constant. When Einstein first published his paper with the above formulations of his field equations, he realised that the equations imply that the universe must be expanding. Hence, he added this constant in order to allow for the universe to stay fixed and not contract or expand. He called it the greatest blunder of his life which is ironic because later on, the Hubble telescope showed us that the universe IS in fact expanding and the equations were initially formulated correctly! Hence, I do not suggest using this formulation of the Einstein Field Equations.

Now that we have an equation, we must solve this equation! Are there even any solutions ?

6.4 The Schwarzschild Metric and Black Holes

The goal from the Einstein Field equations is solving for one of the tensors and using that to calculate the metric of the geometry. Obviously, looking at this equation we see that the calculation for a general solution is so complicated that even world's most advanced super computers will need years of calculation. It becomes even more complicated when you remember that the curvature tensor is defined in terms of christoffel symbols which themselves are defined using multiple derivatives of the metric! Hence, we test the equations for the scenarios we need. One example of this was done by Karl Schwarzschild.

During his time at the navy in World War I, Karl Schwarzschild who was an extremely talented astronomer and mathematician who had attended some of Einstein's lectures on General Relativity had spent his time on solving the equations. He tried and tried until he found every students' favorite: the Schwarzschild Metric. This metric describes the dynamics of the curvature of spacetime due to the rotation of the earth around the sun. The metric is given by the expression:

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

This is taking into account that $c^2 = 1$ otherwise all the time dimensions will have a factor of c^2 mutliplied to them and the space-like ones will be multiplied by a factor of $\frac{1}{c^2}$. Note that really it will look like this:

$$d\tau^2 = -c^2\left(1 - \frac{2MG}{r}\right)dt^2 + \left(\sum_{k=1}^{n}\frac{1}{c^{2k}}\right)\left(\frac{1}{\left(1-\frac{2MG}{r}\right)}dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2\right)$$

If we take $c = 1$, we will get the better form:

$$d\tau^2 = -\mathcal{F}dt^2 + \mathcal{F}^{-1}dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2$$

Where $\mathcal{F} = 1 - \frac{2MG}{r}$. This is one solution to Einstein's equation. However, there is one problem. What happens when $r = 2MG$? We all know what it is, it is a black hole.

A small analysis of the equations show that once we reach $r = 2MG$, the whole shebang breaks! Well, one of the researchers that worked on this theory was Oppenheimer! Him and his colleagues at Berkeley showed that when a sphere reaches this radius, known as its *Schwarzschild Radius* for obvious reasons, showed that it goes right through and collapses. For the case of stars, when there is enough gravitational pull and mass, the star gives birth to a black hole. We can spend pages and pages on black holes. However, we will not as we have done a lot. Finally, let us discuss why General Relativity is a successful theory.

7 Success of General Relativity

We can see from our long journey that all these equations are not just symbols for the sake of inscrutability but they have real meanings and allow us to make tons of predictions about the universe. Objects such as black holes, wormholes and concepts such as the expansion and shape of the universe, spinning black holes, interference of two dying stars, etc. all arise from analyzing the field equations of general relativity. This is what makes it such a powerful and successful theory; It allows scientists to make predictions about the universe and they can come up with experiments to constantly check the consistency of the theory. General Relativity was proposed in 1915, yet more than half the things we know about the universe through this theory were discovered after its proposition and formulation by Albert Einstein. To this day, the theory is still being tested and so far has been consistent with our experimental data making it one of the most powerful theories of physics. Two good sources for further studying the topic are online lectures by Prof. Leonard Susskind at Stanford University and Prof. Scott Hughes at MIT, and books; *An Introduction to General Relativity, Spacetime and Geometry* by Prof. Sean Carroll at Johns Hopkins University and formerly California Institute of Technology, and *General Relativity: The Theoretical Minimum* by Prof. Leonard Susskind and Andre Cabannes at MIT. I also suggest studying the entire *The Theoretical Minimum Series* by Prof. Leonard Susskind if you are interested in a thorough study of physics at a high level.