GR Marathon Notes

Aiman Al-Eryani

October 28, 2018

Contents

1 Events, Inertial Frames of Reference and Galilean Relativity

- An *event* is defined a point $\vec{x}$ in space at a certain time t. We write them together $(t, \vec{x})$
- An *inertial frame of reference* is a non-accelerating frame of reference. Alternatively, it's a reference frame where Newton's laws holds. In particular, it's usually Newton's second law that gets violated in non-inertial frames of reference (acceleration appears without a force).

1.1 Example of an inertial frames of reference

- If you set up a lab on an airplane, Newton's second law will still hold; moving objects keep moving without accelerating and stationary objects remain stationary unless an external force intervenes (friction, pushing ..etc). So the frame of the lab is an inertial frame of reference. You can place a cup of water on your table and even play a game of pool with your friend.
- If you set up a lab on the surface of the earth (neglecting that the earth rotates), Newton's second law will also apply.

1.2 Non-example of inertial frames of reference

- If you set up your small lab on a roller coaster, then you attempt to place your pool balls stationary on the table, your game will quickly be interrupted everytime the roller coaster takes a turn or go through a loop. You'll also quickly find that cup of water has also moved and that the water is no longer in the cup. Stationary objects don't remain stationary despite you not applying any forces on them. This violates Newton's second law, and so this frame is not inertial.

1.3 Galilean transformations

- For the two inertial frames of reference: the earth's and the plane's. Is the plane's lab moving and the earth is stationary, or is the earth moving and the plane is stationary?
- If you have your axes such that the origin is always moving with the lab on the plane, you're doing physics *in the reference frame of the plane*
- Similarly, if you have your axes such that the they're always moving in the center of your lab on earth, then you're doing physics *in the reference frame of the earth*
- There's **no absolute inertial frame** or one that is special; all inertial frames are just as good.
- Suppose a ball is stationary in the plane's frame (with respect to the origin on the plane's lab) at position x_P and time t_P. What would be the event of the ball (t'_P, x'_P) at a later time t'_P?

$$\begin{aligned} x'_P &= x_P && \text{;ball is stationary, so new position equals old position} \\ t'_P &= t'_P && \text{;the new time is the new time} \end{aligned}$$

- What about in the reference frame of the earth? What would be the event of the ball at a later time t'_E if the plane is moving away from the lab on earth at velocity u?

$$x'_E = x'_P + ut'_P \qquad \text{;position evolves as ball is moving in this frame}$$
$$t'_E = t'_P \qquad \text{;time on the plane equals the time on earth later}$$

- If the plane is stationary with respect to the lab on earth, $u = 0$ and:

$$x'_E = x'_P + ut'_P = x'_P$$
$$t'_E = t'_P$$

which tells us that the two frames are exactly the same.

- These transformations from one reference frame to another are called *Galilean transformations*.
- We can see that the relative velocity between the two reference frames is what differentiates them and is what determines what the galilean transformation.
- We'll denote a galilean transformation from frame A to another frame B moving with velocity u relative to A as G_u.
- Using fancy vector notation, we could have written the transformations above from frame of the plane P to frame of the earth E with G_u being a matrix and the event as a vector:

$$\begin{pmatrix} t'_E \\ x'_E \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix} \begin{pmatrix} t'_P \\ x'_P \end{pmatrix} = \begin{pmatrix} t'_E \\ ut'_P + x'_P \end{pmatrix}$$

so G_u can be expressed as a matrix:

$$G_u := \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}$$

which you multiply the event vector by.

1.4 Three reference frames

- Suppose we have three inertial frames of reference. One reference frame T attached to a tree, another reference frame C attached to a canoe (or a boat), and one more reference frame B attached to a ball.
- The ball - and therefore the frame B attached to it - is moving with velocity v relative to the canoe - or to its frame C.
- The canoe with its frame C is moving with velocity u with repsect to the frame tree with its frame T.
- The ball is always stationary in the reference frame B since it's attached to it. Therefore in B, the event of the ball at a later time t'_B is:

$$x'_B = x_B$$
$$t'_B = t'_B$$

- What about in the reference frame of the canoe C? We know the velocity of the ball in this frame is v, so we just use a galilean transformation to get the evolved coordinates in the frame C, just like before:

$$
\begin{aligned}
x'_C &= x'_B + ut'_B \\
t'_C &= t'_B
\end{aligned}
$$

- What about in the reference frame of the tree T? We only know that the canoe is moving with velocity u relative to the tree, but we don't know yet how the ball is moving relative to the tree (you do know it'll be $u + v$, but that's what we're trying to prove). What would be (t'_T, x'_T)?

- Idea: We do a two-way galilean transformation

$$
(t'_B, x'_B) \xrightarrow{G_v} (t'_C, x'_C) \xrightarrow{G_u} (t'_T, x'_T)
$$

Let's compute that manually, and use the fancy vector notation:

$$
\begin{aligned}
\text{We know that:} \quad \begin{pmatrix} t'_C \\ x'_C \end{pmatrix} &= G_v \begin{pmatrix} t'_B \\ x'_B \end{pmatrix} \\
\text{and that:} \quad \begin{pmatrix} t'_T \\ x'_T \end{pmatrix} &= G_u \begin{pmatrix} t'_C \\ x'_C \end{pmatrix} \\
\text{substituting:} \quad \implies \begin{pmatrix} t'_T \\ x'_T \end{pmatrix} &= G_u G_v \begin{pmatrix} t'_B \\ x'_B \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} t'_B \\ x'_B \end{pmatrix} \\
\begin{pmatrix} t'_T \\ x'_T \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ u+v & 1 \end{pmatrix} \begin{pmatrix} t'_B \\ x'_B \end{pmatrix} = \begin{pmatrix} t'_B \\ (u+v)t'_B + x'_B \end{pmatrix}
\end{aligned}
$$

The main result here is that applying G_v to go from B to C then G_u to go from C to T is equivalent to applying G_{u+v} to go immediately from B to T. This tells us that *velocities add up*. This is called the **Galilean law addition of velocities**. It's something that we all know very well; that if you can pitch a ball at 100 mph, then the catcher would recieve it at 100 mph. But if you do it while on top of a skateboard moving at 10 mph, then the catcher will recieve a 100 mph + 10 mph = 110 mph fast ball.

However this will turn out not to be the case in great generality. In particular, with speeds near the speed of light, velocities don't add up like that.

2 Special Relativity

2.1 Motivation

- What if you shine light in reference frame B, and then view it in frame A that moving with velocity u w.r.t frame B? Galilean law of addition of velocities tells you that the velocity of light would be $v + c$.

- Problem: Maxwell's theory of electromagnetism suggests that light is an electromagnetic wave, and that all electromagnetic waves travel at speed $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$.

- Both the magnetic constant μ_0 and the electric constant ε_0 are constants that have the same value in all reference frames, and c should there just be $\frac{1}{\sqrt{\mu_0 \varepsilon_0}}$ in all frames. Contradiction?
- People concluded that light, like any wave, must be travelling in a medium that they called 'ether'. Ether can be dragged around with a reference frame so Maxwell's theory still holds.
- Michelson-Morley would then in 1887 do an experiment involving mirrors and making use of interference properties and show that, contrary to what they would have liked, that light does not travel in a medium.
- So either Maxwell or Newton and his company did some mistake. Einstein sided with Maxwell.

2.2 Postulates of Special Relativity

Einstein set himself two postulates which his new theory should follow:

1. The laws of physics are the same in all inertial frames of reference.
2. The speed of light c is the same in all inertial frames of reference.

We can see that both postulates were taken one step further from what we've taken previously. We previously only wanted Newton's laws to be the same in all inertial frames, but now it's all the laws of physics. In the second postulate, Einstein did not only postulate that *electromagnetic waves* in particular travel in c in all reference frames; rather he took the speed c as special, and anything that travels at c must travel at c in all inertial frames.

2.3 The new transformation

- Before we start thinking of a new transformation that would keep c constant in all inertial frames, we must consider that for $\frac{v}{c} << 1$, this transformation must correspond to the galilean transformation. This is known as *the correspondence principle.*
- The plot of the position of object x against time t is called its *spacetime diagram.*
- It's clear that the spacetime diagram of an object moving at velocity u is a straight line, as long as there are no forces acting on it. This is just Newton's second law, as any accleration will result in a curve.
- For the Newton's second law to hold, it must mean that going from one reference frame to another should use a transformation that keeps the straight lines in spacetime diagrams straight, since any curve implies acceleration (change in slope is a change in velocity).
- Geometrically, a galilean transformation 'shears' (sort of "stretches") the spacetime diagram in the direction of x. Shearing can never make straight lines in spacetime diagrams curve.
- We also know that the new transformation we're looking for must also keep the straight lines straight for Newton's second law to hold in all reference frames. A change between inertial frames shouldn't introduce any acceleration.
- In other words, the transformation is linear. Any linear transformation can be represented by a matrix.

- Let's call denote this transformation matrix by L_u, where u is a velocity characterising it. The most general form of this matrix is:

$$L_u = \begin{pmatrix} \theta(u) & \tau(u) \\ \alpha(u) & \gamma(u) \end{pmatrix}$$

where $\theta(u)$, $\tau(u)$, $\alpha(u)$ and $\gamma(u)$ are unknown functions of u. What we'll do next is find these functions, and therefore determine the matrix L_u.

2.4 Finding the Lorentz transformation

- Letting go of any previously held notions of galilean transformations of how velocities add up, the following are things we least expect from this new transformation L_u:

$L_u \begin{pmatrix} t \\ -ut \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ 0 \end{pmatrix}$; in a frame where your frame is moving at u, an object that is moving at $-u$ in your frame should look stationary

$L_{-u}L_u = L_uL_{-u} \stackrel{!}{=} I$; this should be identity for $u \neq c$

$L_u \begin{pmatrix} t \\ ct \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ ct' \end{pmatrix}$; the second postulate: the speed of light c is the same in any frame

$L_0 \stackrel{!}{=} I$; going from a frame to another comoving frame must be identity

$L_cL_u \neq L_{c+u}$; velocities should not add up in the galilean sense

- Only the first three conditions are needed to find L_u completely. The last two will just be used to test that we have arrived at the correct result.

- Note that with galilean transformations, time never transformed; it would always be the same no matter what reference frame we pick. However, here, we do not exclude that the time component gets transformed. And indeed, as we shall see, t is not always equal to t'.

2.4.1 Finding $\alpha(u)$ in terms of $\gamma(u)$ and showing that $\gamma(u) = \gamma(-u)$:

- Starting from the first condition:

$$\begin{pmatrix} \theta(u) & \tau(u) \\ \alpha(u) & \gamma(u) \end{pmatrix} \begin{pmatrix} t \\ -ut \end{pmatrix} = \begin{pmatrix} t\theta(u) - ut\tau(u) \\ t\alpha(u) - ut\gamma(u) \end{pmatrix} p \stackrel{!}{=} \begin{pmatrix} t' \\ 0 \end{pmatrix}$$

$t\alpha(u) - ut\gamma(u) = 0 \implies \boxed{\alpha(u) = u\gamma(u)}$

- Let's update the matrix L_u so that it no longer has $\alpha(u)$:

$$L_u = \begin{pmatrix} \theta(u) & \tau(u) \\ u\gamma(u) & \gamma(u) \end{pmatrix}$$

- The second condition is that the inverse of all of our transformations L_u must be L_{-u}. Let's try to hit the expression of the first condition with L_{-u} on both sides to see if we can squeeze out any more info from it:

$$\cancel{(L_{-u})}L_u \begin{pmatrix} t \\ -ut \end{pmatrix} \stackrel{!}{=} (L_{-u}) \begin{pmatrix} t' \\ 0 \end{pmatrix} = \begin{pmatrix} \theta(-u) & \tau(-u) \\ -u\gamma(-u) & \gamma(-u) \end{pmatrix} \begin{pmatrix} t' \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} t \\ -ut \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t'\theta(-u) \\ -ut'\gamma(-u) \end{pmatrix}$$

we have $-ut = -ut'\gamma(-u) \implies t = t'\gamma(-u)$. By the isotropy of space (time should transform the same way regardless of what direction u is), we should also have that $t = t'\gamma(u)$. This implies that $\boxed{\gamma(u) = \gamma(-u)}$. The same argument can be repeated with the equation between elements of the first components, and we get that $\theta(-u) = \theta(u)$.

2.4.2 Finding $\theta(u)$ and $\tau(u)$ in terms of $\gamma(u)$

- Now let's see what we can get if we expand the expression of the second condition:

$$L_{-u}L_u \stackrel{!}{=} I$$

$$\begin{pmatrix} \theta(-u)\theta(u) + \tau(-u)(u\gamma(u)) & \theta(-u)\tau(u) + \tau(-u)\gamma(u) \\ (-u\gamma(-u))\theta(u) + \gamma(-u)(u\gamma(u)) & (-u\gamma(-u))\tau(u) + \gamma(-u)\gamma(u) \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

looks very scary but its just ordinary matrix multiplication. These are four equalities. Let's look at the equality in the left-bottom corner of the matrix:

$$(-u\gamma(-u))\theta(u) + \gamma(-u)(u\gamma(u)) \stackrel{!}{=} 0$$

Recall that $\gamma(u) = \gamma(-u)$. We can then solve for $\theta(u)$ in terms of $\gamma(u)$:

$$-u\gamma(u)\theta(u) + u\gamma^2(u) \stackrel{!}{=} 0 \implies \boxed{\theta(u) = \gamma(u)}$$

- Now let's look at the bottom-right corner and try to solve for $\tau(u)$. We find that:

$$(-u\gamma(-u))\tau(u) + \gamma(-u)\gamma(u) \stackrel{!}{=} 1 \implies -u\gamma(u)\tau(u) + \gamma^2(u) = 1$$

$$\implies \boxed{\tau(u) = \frac{\gamma^2(u) - 1}{u\gamma(u)}}$$

- We can now express the entire matrix in terms of $\gamma(u)$!

$$L_u = \begin{pmatrix} \gamma(u) & \frac{\gamma^2(u)-1}{u\gamma(u)} \\ u\gamma(u) & \gamma(u) \end{pmatrix}$$

now all is left to determine L_u is to determine the function $\gamma(u)$.

2.4.3 Finding the function $\gamma(u)$

- Taking the third condition, which is essentially invoking the second postulate of special relativity (that an object moving at the speed of light should do so in all reference frames):

$$L_u \begin{pmatrix} t \\ ct \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ ct' \end{pmatrix}$$

$$\begin{pmatrix} \gamma(u) & \frac{\gamma^2(u)-1}{u\gamma(u)} \\ u\gamma(u) & \gamma(u) \end{pmatrix} \begin{pmatrix} t \\ ct \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ ct' \end{pmatrix}$$

$$\begin{pmatrix} t\gamma(u) + ct\frac{\gamma^2(u)-1}{u\gamma(u)} \\ tu\gamma(u) + ct\gamma(u) \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ ct' \end{pmatrix}$$

we have two equalities and two unknowns $\gamma(u)$ and t'. Either plug them into Worlfram alpha or solve them. Plugging the expression for t' (the first equality) in the second equality:

$$\cancel{t}u\gamma(u) + \cancel{ct\gamma(u)} = \cancel{c(t\gamma(u))} + c(c\cancel{t}\frac{\gamma^2(u)-1}{u\gamma(u)})$$

$$\implies \boxed{\gamma(u) = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}}$$

we have thus fully determined the new transformation matrix L_u.

- Let's keep using the variable $\gamma(u)$, but let's try to simplify the term in the upper right corner of L_u: $\frac{\gamma^2(u)-1}{u\gamma(u)} = (\frac{1}{u} - \frac{1}{u\gamma^2(u)})\gamma(u) = (\frac{1}{u} - \frac{1-\frac{u^2}{c^2}}{u})\gamma(u) = \gamma(u)(\cancel{\frac{1}{u}} - \cancel{\frac{1}{u}} + \frac{u}{c^2}) = \frac{u}{c^2}\gamma(u)$
- We finally get:

$$\boxed{L_u = \begin{pmatrix} \gamma(u) & \frac{u}{c^2}\gamma(u) \\ u\gamma(u) & \gamma(u) \end{pmatrix}}$$

- The function $\gamma(u)$ is called the *Lorentz factor*, and the transformation L_u is called the *Lorentz transformation*.
- **Excercise:** Check that the last two conditions do indeed hold, and find the new formula for addition of velocities.
- **Excercise:** The fact that c is the same in all reference frames implies that an object that's not already moving at c can not be moving at c in any reference frame. Check that this is indeed the case by trying to compute L_c.

2.5 Sorting the mess, space-interval contraction and time-interval dialation

- **First, we need to sort out one mess**. We have been talking about L_u as the transformation that takes you from your frame to another frame in which according to it you're moving at velocity u. This was easier to calculate, but not so easy to phrase. In other literature, you'll see the parameter of the transform u being the velocity with which that frame is moving relative to your frame. In other words, they define $\Lambda_u = L_{-u}$. We'll switch to that convention from now on.
- We explicitly write Λ_u:

$$L_{-u} = \boxed{\Lambda_u = \begin{pmatrix} \gamma(u) & -\frac{u}{c^2}\gamma(u) \\ -u\gamma(u) & \gamma(u) \end{pmatrix}}$$

- Let's see how arbitrary events (t, x) transform under the lorentz transformation Λ_u that takes us to frame S' that's moving at u:

$$\Lambda_u \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} \gamma(u) & -\frac{u}{c^2}\gamma(u) \\ -u\gamma(u) & \gamma(u) \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} \gamma(u)(t - \frac{u}{c^2}x) \\ \gamma(u)(x - ut) \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ x' \end{pmatrix}$$

- Suppose you have in your frame two events both at the same time t but one at position x and the other at position $x + \Delta x$. One can think the distance between them is:

$$\begin{pmatrix} t \\ x + \Delta x \end{pmatrix} - \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ \Delta x \end{pmatrix}$$

- What would be the distance in the reference frame S'? We need to transform both events and then compute the difference again. By linearity of matrices, it's the same as taking the difference and then computing:

$$\Lambda_u \begin{pmatrix} t \\ x + \Delta x \end{pmatrix} - \Lambda_u \begin{pmatrix} t \\ x \end{pmatrix} = \Lambda_u(\begin{pmatrix} t \\ x + \Delta x \end{pmatrix} - \begin{pmatrix} t \\ x \end{pmatrix}) = \Lambda_u \begin{pmatrix} 0 \\ \Delta x \end{pmatrix}$$
$$\implies (\Delta x)' = \begin{pmatrix} -\frac{u\Delta x}{c^2} \\ (\Delta x)\gamma(u) \end{pmatrix}$$

we have a 'spacital distance' $(\Delta x)\gamma$, but it doesn't help because the events are no longer simultaneous (!). We have a distance in space as well a distance in time in the new frame! Taking the spacial distance wouldn't make sense because it would be the difference between two events at two different times.

- What does that tell us? Two events that are simultaneous in one inerital frame may not be in the other. So "what happened first, a chicken or an egg?" might have different answers in different reference frames.

- In general, there is a trade-off between length contraction and simultaneouity. In most special relativity "paradoxes", the trick is to remember that.

- Now that we know time differences can change, how does the time interval between two events at the same position $(\Delta t + t, x)$ and (t, x) look like in the a different frame S'? In the current frame, the difference is just $(\Delta t, 0)$. Let's Lorentz transform it:

$$\Lambda_u \begin{pmatrix} \Delta t \\ 0 \end{pmatrix} = \begin{pmatrix} \Delta t\gamma(u) \\ -u\Delta t\gamma(u) \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \Delta t' \\ \Delta x' \end{pmatrix}$$

we see that here the spatial distance changes, but that does not matter for time measurement. Therefore, $\Delta t' = \Delta t\gamma(u)$. We have that $\gamma(u) \geq 1$, which means that time dialates when moving; or as they say, *moving clocks tick faster*. This effect is known as *time dialation*.

- A clock at rest is one constructed from events which x component has no velocity term ut for any velocity u involved. If a clock at rest measures the (proper) time interval $\Delta\tau$, then in another moving frame, it would measure $\Delta t = \tau\gamma(u)$ as we have seen above due to time dialation. This implies $\boxed{\Delta\tau = \frac{\Delta t}{\gamma(u)}}$. This is an invariant quantity that's the same in all frames.

- But if time is changing in every frame, what would be the right time to measure? It's simply the time where the clock is at rest. By definition, it's the same in all reference frames. We call it the *proper time* and we define it by $\Delta\tau$ (or sometimes τ).

- With a similar idea, we can define a proper length L in frame that is at rest. One can show in a similar way to proper time that it's $L_0 = \gamma(u)L$ once they have found length contraction.

- **Excercise:** Show that *length contraction*; show that the $\Delta x' = \frac{1}{\gamma(u)}\Delta x$ where $\Delta x'$. Avoid simultaneity issues we encounted before by ensuring $\Delta t' = 0$ in the new frame S'. Hint: Start with Δt arbitrary and then insist that $\Delta t' = 0$. After that, solve for Δt and substitute in the equation for $\Delta x'$.

2.6 Lorentz transformations in three dimensions

- So far, we have only worked with x position and time, but the world has three spatial dimensions.

- Claim: if motion is strictly on the x axis (no velocity term for y and z), transformations for y and z are trivial: $y = y'$, $z = z'$.

- Proof: Since Lorentz transformations are linear, they must map $y = 0$ to $y' = 0$, and $z = 0$ to $z' = 0$. If $y, z \neq 0$, we just translate everything so that they're at zero. Changing the origin of the coordinate system should not change its physics (in general: changing the coordinate system shouldn't affect the physics. That's the principle of *general covariance*).

- So we have:

$$\Lambda_u \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma(u) & -\frac{u}{c^2}\gamma(u) & 0 & 0 \\ -u\gamma(u) & \gamma(u) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma(u)(t - \frac{u}{c^2}x) \\ \gamma(u)(x - ut) \\ y \\ z \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}$$

- What if u is not strictly in the x direction? We rotate everything so that u is in the x-axis direction with the rotation matrix R^{-1}, do our physics (lorentz transform), and then rotate back with R. This is equivalent to just hitting L_u by "conjugation" of R (means, hit on the right with inverse, and on the left with itself):

$$\Lambda_{R,u} = R\Lambda_u R^{-1}$$

and then use the new $\Lambda_{R,u}$.

2.7 Lorentz covariance

- In physics, we deal with quantities. But we need good ones that are well defined and consistent. It's a bit vague at the moment, but soon it should be made clear.

- Scalars quantities like the spatial distance between two events Δx are bad. It takes different values in different inertial frames. A good quantity is one that's *Lorentz invariant* a quantity that is the same in all inerital frames. For example, the proper time $\Delta\tau$ is a good quantity and is Lorentz invariant. It's the same in all reference frames. Other examples of Lorentz invariant quantities are the rest mass m_0, the speed of light c, the proper length and so forth.

- Any meaningful vector quantities which involve the physical dimensions Length or Time will have to transform under a lorentz transformation. We base that of the fact that an event that describes a position-time in spacetime transforms by application of the Lorentz transform operator L_u. Velocity for instance is then the derivative of this position, and momentum would be a mass times this position ..etc.

- However, this one disturbing thing with having an event as a quantity of position. A quantity should have one physical dimension (Length, Time ..etc). However, the first component in the event vector has dimension Time, and the rest have dimension Distance/Spactial-Position.

- To make everything consistent, we must make the first components of an event also have a dimension of Spacial-Position. We know the way to make a distance from a time quantity is to multiply it by a velocity quantity.

- In order for the event to still transform similarly, the velocity we multiply by time better be an invariant, and we know that got to be c. This way, it will simply act as a scaling factor.

- We define the *4-position*:

$$X = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

our base position quantity in special relativity. To account for the scaling in time, the Lorentz transformation changes form:

$$\Lambda = \begin{pmatrix} \gamma(u) & -\frac{u}{c}\gamma(u) & 0 & 0 \\ -\frac{u}{c}\gamma(u) & \gamma(u) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

one can easily check that applying this on the the 4-Position vector X gives the same equations when an event without a scaling factor c is hit by the old Lorentz transform.

- One can try to define a velocity out of this. From the formal definition of a derivative, $\frac{dX}{dt} = \lim_{\Delta t \to 0} \frac{\Delta X}{\Delta t}$. The difference between two 4-vectors will transform by multiplication by Λ under a change of inertial frames:

$$\Delta X = X_1 - X_2$$
$$\overset{\text{go to frame } S'}{\rightarrow} \Delta X' = X'_1 - X'_2 = \Lambda X_1 - \Lambda X_2 = \Lambda(X_1 - X_2) = \Lambda \Delta X$$
$$\overset{\text{derivative limit}}{\rightarrow} dX' = \Lambda dX$$

However, Δt is a scalar that's not a lorentz scalar, so will change. Instead, we use the proper time $\Delta\tau$:

$$V = \frac{dX}{d\tau}$$

- Since $d\tau$ is invariant, V transforms like dX which transforms like X: with multiplication by Λ.
- We can express $d\tau$ as follows:

$$\Delta\tau = \frac{1}{\gamma}\Delta t \implies d\tau = \frac{1}{\gamma}dt$$
$$\implies \boxed{V = \gamma\frac{dX}{dt} = \gamma\begin{pmatrix} c \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix}}$$

this is called the *4-velocity*.

- Similarly, we can define other quantities like 4-momentum as $P = m_0 V = m_0\gamma\frac{dX}{dt}$ where m_0 is the rest mass, a Lorentz invariant. And then the 4-force $\frac{dP}{d\tau} = \gamma\frac{dP}{dt}$, which is also a four vector. And the list goes on. These are all quantities that respect Lorentz transformation.
- In general, if all quantities in a law respect lorentz transformations, then the law is lorentz covariant and is compatable with special relativity.
- Maxwell equations, in particular, are lorentz invariant; SR was based to make Maxwell's equations hold afterall. We know how the velocity of light transforms for example; it's a lorentz invariant. One can also check that other quantities form nice four vectors and everything. More will be taught in the second year course "Electrodynamics".
- Newton's law fo gravitation however, does not respect lorentz transofrmations. In particular, it has explicit dependence on the distance between two bodies r, which is not a lorentz invariant nor a lorentz covariant.

2.8 Spacetime diagrams: space and time on equal footing

- So we now have events described by (ct, x, y, z). All components have physical dimension of length. Thus, we can say that an event is a quantity now also known as the *4-position*, the spacetime position of an object.
- Furthermore, 4-position is a lorentz covariant; it transforms "correctly" and consistently under a change of the inertial frame of reference (transforms by multiplication by Λ).
- We saw that after scaling the time component of an event by c to define the $4 - position$, the Lorentz transformation becomes (ignoring the y and z components for simplicity):

$$\Lambda = \begin{pmatrix} \gamma(u) & -\frac{u}{c}\gamma(u) \\ -\frac{u}{c}\gamma(u) & \gamma(u) \end{pmatrix}$$

which looks slightly nicer, as it has now became symmetric if the movement is only on one spacial axis x. Suppose we take the event vector and swap the time axis with the event axis so that it's (x, ct) instead of (ct, x). The lorentz transformation matrix would not be exactly the same!

- This tells us that space and time are to be treated on equal footing. This was not the case in Galilean transformation:

$$G = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}$$
$$G^T = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$$

where the matrix is not symmetric and changing (t, x) to (x, t) would require we take G^T.

- If we were to talk about "Galilean invariants", there are plenty. The speed of light c is not a galilean invariant (light speed is different in different inertial frames). Δx is an invariant in galilean relativity; changing inertial frames doesn't change the distance between two points. To generalize to 3D, the invariant space interval Δs is the galilean sense is given by:

$$\Delta s = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \xrightarrow{\text{infinitisimal}} ds^2 = dx^2 + dy^2 + dz^2$$

- With lorentz', Δx and dx is not an invariant. Going to a boosted frame (a frame that's moving) causes length contraction. But it also causes time dialation. One wonders whether this trade off makes a combination of both a lorentz invariance (spoiler: yes).

2.8.1 Metrics

- There's a name given to an object that takes two objects and gives the distance between them. This object is called the *metric* and can be denoted by $m(a, b) = ||a - b||$. It takes two points $\vec{a}$ and $\vec{b}$ and gives a number that's the distance between them. The metric above can be written as:

$$m(\vec{a}, \vec{b}) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}$$

- Another way to define the distance between two points $\vec{\Delta x}$ is by taking the square root of its dotproduct with itself since:

$$\vec{\Delta x} \cdot \vec{\Delta x} = ||\vec{\Delta x}||^2 = \Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2$$

Metrics defined that way are called "induced metrics". In physics, the dotproduct is so important it's often called the metric, even though the value of the distance is defined by its square root.

- So from now on, when we say the "metric", we mean the "inner product". Afterall, what we care most about is whether its value is invariant or not. If the distance is invariant, then so should be its square. When we want distance, we'll say "the distance" or "the interval" explicitly.

- The dot product/metric has two very important properties. The first is that it's symmetric; namely that:

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$$

The second is that it's bilinear; You can split sums and take out/in factors in its two argument "slots":

$$(\alpha\vec{a} + \beta\vec{b}) \cdot \vec{c} = \alpha(\vec{a} \cdot \vec{c}) + \beta(\vec{b} \cdot \vec{c})$$

Since it's symmetric, one can see that if it's linear in one of the arguments, then it's also linear in the other (since we can swap the arguments).

- The generalisation of linear operators with any number of slots that give back numbers are called *tensors*.
- Matrices are *rank 2 tensors* because they can act on two vectors and give a number. So given a matrix A, you can evaluate it on vectors $\vec{a}$, $\vec{b}$ to get a number by taking the first vector transposed, multiplying on the left, and multiplying the other vector on the right: $\vec{a^T} A\vec{b}$.
- Using a matrix, we can write the dotproduct/metric above as:

$$\vec{a} \cdot \vec{b} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a^T I b = a_1 b_1 + a_2 b_2 + a_3 b_3$$

where the matrix in the middle is the identity matrix. So we see that the euclidean metric is really the identity matrix.

2.8.2 The minkowski metric

- Now here is what we're interested in: is there a notion of "space-time" distance that's invariant under lorentz transformations?
- Another way to say it: Is there a metric which if we evaluate on an ΔX we get a lorentz invariant?
- Another way to say it: Is there a metric which the lorentz transformation is an *isometry* of? (an isometry is a map that preserves the distances/evaluations of the metric).
- Let's try to find this metric η. We would use to find the "spacetime" distance between two points δs by applying the metric (our "dot product") on the difference vector of two events $\Delta X = X_1 - X_2$ and taking the square root:

$$\Delta s = \sqrt{(\Delta X)^T \eta (\Delta X)} \implies (\Delta s)^2 = (\Delta X)^T \eta (\Delta X)$$

- Going to a different inertial frame, the difference vector would transform as $\Delta X \to \Lambda \Delta X$. Yet, the spacetime distance should stay the same. This means:

$$(\Delta s)^2 = (\Delta X)^T \eta (\Delta X) \stackrel{!}{=} (\Lambda \Delta X)^T \eta (\Lambda \Delta X) = (\Delta X^T) \Lambda^T \eta \Lambda (\Delta X)$$
$$\iff \eta \stackrel{!}{=} \Lambda^T \eta \Lambda$$

- After staring at the equation for a while, one can find that:

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- η is called the **Minkowski metric**, and it's the metric of spacetime.
- Let's try to compute the distance between X_2 and X_1 denoted by $\Delta X = X_1 - X_2$:

$$\Delta X^T \eta \Delta X = \begin{pmatrix} c\Delta t & \Delta x & \Delta y & \Delta z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} c\Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$$
$$= -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

- Rotations are transformations that keep the euclidean distance the same. Gometrically, it's movement on locus of a circle.
- In this sense, a lorentz transformation is one that keeps the minkowski distance invariant. Geometrically, it's movement on locus of a hyperbola.
- Things to notice: We can apply this to any 4-vector and that Δs^2 can go negative! In mathematics lingo, this makes it a "pseudo-metric". However we shall continue to refer to it by simply "metric".
- **Excercise:** Show that the lorentz invariant magnnitude of the 4-velocity is c ! It means we're moving in c all the time, and that thus $\Delta \tau^2 = -\Delta s^2$

2.9 Physical and geometric interpretation and the light cone

- If one draws a spacetime diagram of ct in the y-axis vs x in the x-axis, one can see that the trajectories $x(t) = ct$ and $x(t) = -ct$ are diagonal lines. These define a *light cone* or the event horizon.
- Any flatter slope would imply a trajectory of something moving faster than c which is not possible. So the boundary of the light cone defines an event horizon outside which nothing can influence anything inside it or viceversa.
- Spacial rotation matrices are the isometries of the euclidean metric. On an x-y diagram, a rotation operation will make the vector $\Delta X = (\Delta x, \Delta y)$ trace contour lines of $\Delta s^2 = \Delta x^2 + \Delta y^2$ with Δs being fixed; that is the locus of a circle of radius Δs.

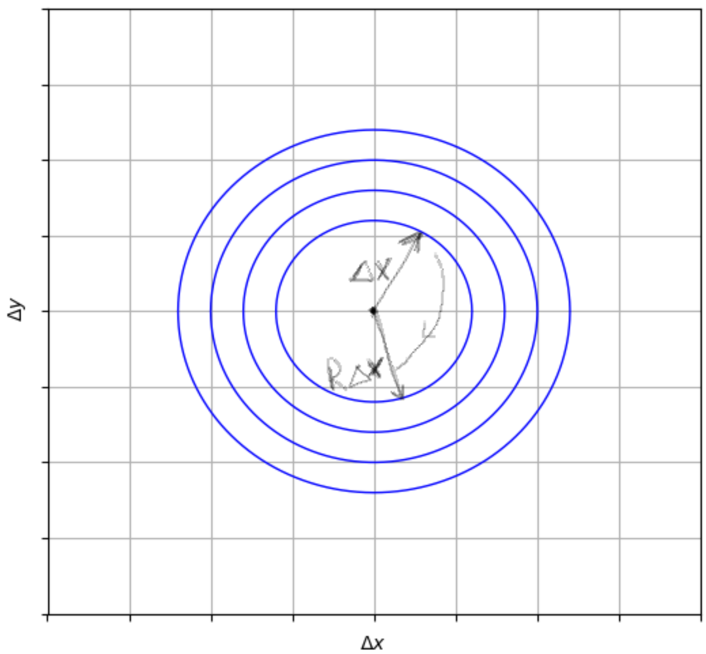

Figure 1: Multiplying ΔX by a rotation matrix R, an isometry of the euclidean metric, is geometrically just sliding the vector on the locus of a circle of fixed radius

- Analogously, Lorentz transformation matrices (boosts) are the isometries of the minkowski metric. On a spacetime diagram with ct against x, a lorentz transformation matrix will make the vector $\Delta X = (c\Delta t, \Delta x)$ trace the contour lines of $\Delta s^2 = -(c\Delta t)^2 + \Delta x^2$. This equation can refer to different shapes depending on the value of Δs^2:
  - If $\Delta s^2 < 0 \implies \Delta x^2 = \pm\sqrt{\pm(c\Delta t)^2 + |\Delta s^2|}$, it will trace the locus of a hyperbola either concaving upwards or downwards. Events with $\Delta s < 0$ are called **timelike separated**. From the geometric picture, it's clear one can always perform a lorentz transformation that makes ΔX point fully up (or fully down); i.e. one can always use a lorentz transformation two make two time-like separated events have zero spatial separation. However, no matter what, you can never "slide" the vector on a locus to make it horizontal (i.e. zero time-sparation or simultaneous) with $\Delta t = 0$.
  - If $\Delta s^2 > 0 \implies \Delta x^2 = \pm\sqrt{\pm(c\Delta t)^2 + |\Delta s^2|}$, it will trace the locus of a hyperbola either concaving leftwards or rightwards.Events with $\Delta s < 0$ are called **timelike separated**. From the geometric picture, it's clear one can always perform a lorentz transformation that makes ΔX point fully up (or fully down); i.e. one can always use a lorentz transformation two make two space-like separated events have zero time separation or be simultaneous! However since this event is outside the light cone, one can never observe this. Conversely to above, you can never "slide" the vector on a locus to make it vertical (i.e. zero space-sparation) with $\Delta x = 0$.
  - If $\Delta s^2 = 0 \implies c\Delta t = \pm\Delta x$, it will trace the locus of a line with slope ± 1. Events with $\Delta s^2 = 0$ are called **null separated**; which means one of the two events is moving at the speed of light. From the geometric picture, since the lines start at the origin, the vector ΔX is always fully intersecting with the line and thus there's nowhere new it can

move; no lorentz transformation can do anything to change it. This is consistent with the second postulate of Special Relativity which states that the speed of light is the same in all inertial reference frames.

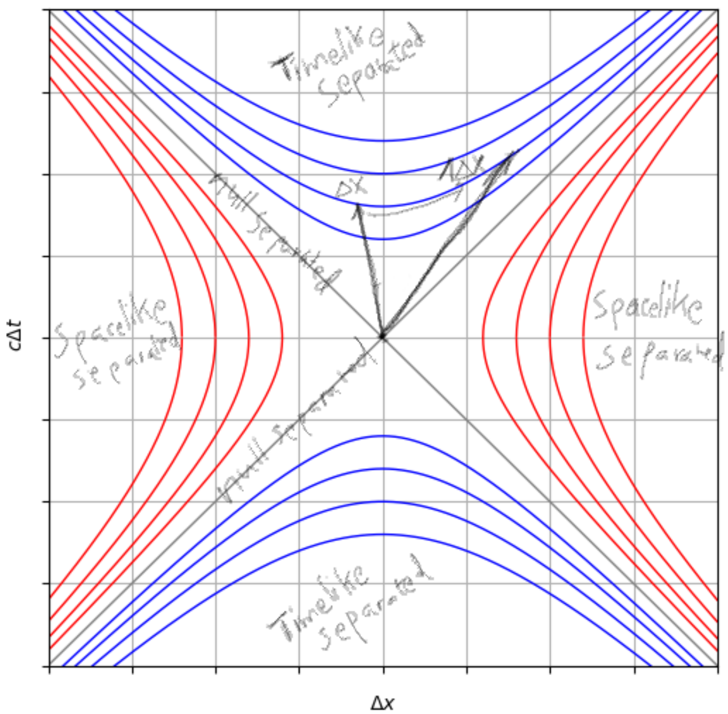

Figure 2: Multiplying the difference between two events ΔX by a lorentz transformation matrix Λ, an isometry of the minkowski metric, is geometrically just sliding the vector on the locus of a hyperbola

- Timelike separated events, i.e. these for which ΔX is inside the light cone, are causally connected and can influence one another. Spacelike separated events with their ΔX outside the light cone are causally disconnected and can never influence one another.
- Suppose you're in a frame where you're at rest. You measure the time Δt, and by definition, since you're at rest, the time you measured is the proper time $\Delta\tau$. So in this case $\Delta t = \Delta\tau$. Note that since you're at rest, the spacial distance you moved is $\Delta x = 0$. The spacetime distance is then

$$\Delta s^2 = -(c\Delta t)^2 + \cancel{(\Delta x)^2} = -(c\Delta t)^2 \implies \boxed{\Delta s^2 = -(c\Delta\tau)^2}$$

- The left handside of the boxed equation is a spacetime distance which is the same in all reference frames by definition (thus a lorentz invariant). The right handside is a product of two lorentz invariants, namely the speed of light c and the proper time $\Delta\tau$. This means that the equation above holds regardless of the inertial frames you're at, and regardless of the two events you're measuring distances between. It gives an expression for the propertime $\Delta\tau$.
- Notice the proper time will only give you a real number between timelike separated events, and is thus only physical for timelike separated events.

2.9.1 Example: The twin paradox

- **Problem:** Alice and Bob each start measuring time at the same point. Bob travels away from Alice at velocity v. After covering a distance d, he makes a sudden turn and returns back to Alice. What would be the time measured for both Alice and Bob?

- **Solution:**

  - Since Bob makes a turn somewhere, the frame attached to him while he took his full trajectory is not inertial. So we should stick to Alice's inertial frame of reference.
  - Since Alice is stationary in her frame, $\Delta x_A = 0$.
  - From Alice's perspective, it takes Bob $\Delta t = \frac{2d}{v}$ seconds to return, so the spacetime interval between the start of the measurement and its end is $\Delta s_A^2 = -(c\Delta t)^2 = -4d^2 \frac{c^2}{v^2}$.
  - Bob, however, will have in addition traveled a distance of $\Delta x = v\Delta t = 2d$. So his spacetime interval would be $(\Delta s_B)^2 = -4d^2 \frac{c^2}{v^2} + 4d^2$.
  - Now the time each measured is their proper time for both. We have that $(\Delta s)^2 = -(c\Delta\tau)^2 \implies \Delta\tau = \sqrt{\frac{-(\Delta s)^2}{c^2}}$

$$\begin{aligned}
\text{So for Alice: } \Delta\tau_A &= \sqrt{\frac{-(\Delta s_A)^2}{c^2}} = \sqrt{\frac{4d^2 \frac{c^2}{v^2}}{c^2}} \\
&= \frac{2d}{v} \\
\text{For Bob: } \Delta\tau_B &= \sqrt{\frac{-(\Delta s_B)^2}{c^2}} = \sqrt{\frac{4d^2 \frac{c^2}{v^2} - 4d^2}{c^2}} \\
&= \frac{2d}{v}\sqrt{1 - \frac{v^2}{c^2}}
\end{aligned}$$

  - So we have that $\Delta\tau_A = \Delta\tau_B\sqrt{1 - \frac{v^2}{c^2}}$ which tells us that $\Delta\tau_A > \Delta\tau_B$ (more time passed for Alice than Bob).

2.10 Algebra of Special Relativity

2.10.1 Groups

- Formally a group is a set with an operation on its elements $K(x, y)$ that satisfies four axioms:

  1. Has an "identity element" I such that $K(x, I) = x$ and $K(I, x) = x$
  2. Each element has an "inverse element" x^{-1} such that $K(x, x^{-1}) = I$ and $K(x^{-1}, x) = I$
  3. Assocativity: $K(K(x, y), z) = K(x, K(y, z))$
  4. Closedness: The operation K always produces an element inside our group.-

- Example: Group of all points in $\mathbb{R}^3$, and let the operator be the addition of the two position vectors (i.e. translation).

- Example: Group of all linear operators that respect the Euclidean metric (preserve pythagorean distances) in the three dimensional space. These are all the invertible matrices A which inverses are themselves transposed; i.e. $A^{-1} = A^T \implies AA^T = A^TA = I$. Its fancy name is the *Special Orthogonal Group* or $SO(3)$. "Special" means invertible, and "Orthogonal" means $A^T = A^{-1}$. These matrices are rotation matrices.

2.10.2 The Lorentz and the Poincare Group

- Group of all matrices which respect the Minkowski metric (preserves the spacetime distance). These are the matrices that obey $A\eta A^T = g$ where η is the minkowski metric.
- We have seen before that Lorentz transformation matrices and matrices that rotate in space obey this relation (in fact, that's how we found η in the first place).
- This group of spacial rotation and boosts (Lorentz transformations) is called *the Lorentz group* and is denoted by $SO(3,1)$.
- One can thing of it as the group of rotations on spacetime, where rotation on spacetime is on a locus of a hyperbola rather than a circle.
- The Lorentz group with the group of operators that translate on spacetime is called *the Poincare group* and is denoted by $TSO(3,1)$.
- Note that $TSO(3,1)$ is not a linear group, however all its elements are isometries of the Minkowski metric.

3 General Relativity

"Spacetime tells matter how to move. Matter tells spacetime how to curve. "
- John Wheeler

3.1 Motivation

- Newton's law of gravitation says that the magnitude of the force acting on the earth from the the sun is:

$$F_G = \frac{GM_{sun}m_{earth}}{(x_{earth} - x_{sun})^2}$$

  If the sun suddenly disappears, it tells us that since $M_{sun} = 0$ the earth immediately stops feeling any force and leaves orbit. But this is in contradition with special relativity that tells us that it would stop feeling the force at least after ~8 minutes have passed.
- In particular, Newton's law of gravitation is not Lorentz covariant. The three dimensional spacial quantity $x_{earth} - x_{sun}$ is not a four vector nor an invariant.
- Another motivation was the descrepency between Mercury's orbit and the predictions Newton's gravity makes. Newton's gravity predicts orbits follow elliptical paths; the orbits never cross. However in the case of Mercury, the orbit crosses itself and so the perihelion of the ellipse of ist orbit precesses. Therefore a new theory was needed.

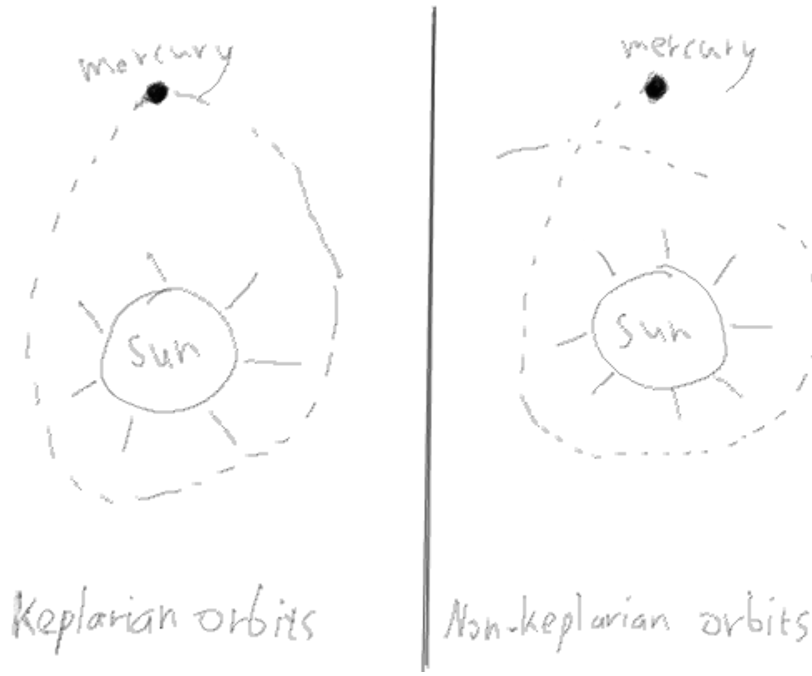

Figure 3: On the left is a keplarian orbit where the planet follows an ellipse. On the right is a non-keplarian orbit that's observed with Mercury

- Einstein's idea: gravity is geometric and gravitional acceleration is just freefalling. Gravitional acceleration is just an emergent property of curved spacetime.
- Special Relativity is a special case of General Relativity where the *spacetime is flat.*
- In General Relativity, spacetime curves in response to the presence of matter. Acceleration then emerges naturally from curvature of spacetime.
- Main concept: Curved space can be covered by multiple patches that are approximately flat. The structure that contains all these patches is called a *manifold.*
- At each point, there's a different flat metric; or one says there's a *metric field* defined on each point on the manifold. The spacetime distance between two events would then be the sum (integral) of all the little distances on the patches you pass by.
- A manifold equipped with a metric field is called a *Riemanian Manifold.* In the case of general relatvitiy, the metric at each point will be pseudoriemanian (we'll show they all can take negative distances like the Minkowski metric), so we will be working with *Pseudo-Riemanian manifolds.* The maths of (Pseudo)Riemanian manifolds is "Riemanian Geometry" which we will have to cover to quantify the ideas of General Relativity.

3.2 The Einstein equivalence principle

3.2.1 The weak equivalence principle

- Inertial mass and gravitational mass are equivalent; i.e. the inertial mass in $F = ma$ and the gravitataional mass in $F = GMm/r^2$ are one and same. This idea is just experimental; there just doesn't seem to be any descripancies, and it seems we can always cancel them out with one another. This was the main motivation behind the idea that gravity must be geometric.

3.2.2 The equivalence principle

- Assume building a little laboratory inside an elevator that's accelerating down uniformly. You take a tiny ball, and let it freefall. The acceleration is indestinguishable from that of gravity.
- But if you take another ball and put it far away, with a high resolution measurement, you may be able to see that the balls are attracting falling into the same center of mass of the earth.
- What you can do: decrease distance between the balls until the measuremnt device can no longer detect it. You can always do this for any resolution of the device.
- So we say **locally** (in a small neighbohrhood), uniform acceleration can't be destinguished from gravity.
- If curvature is responsible for gravity, that's equivalent to saying that locally spacetime is flat in some coordinate system (can involve frames not necessarily inertial).
- In other words, there's a coordinate transformation which makes the metric at point a be $g(a) = \eta$ i.e. a frame that's inertial where acceleration is zero.

3.3 Vectors and the Einstein summation convention

- In special relativity, all our quantities are either (lorentz covariant) 4-vectors or (lorentz invariant) scalars. 4-vectors transform under a change of reference frames by multiplication with a lorentz transformation matrix Λ. Lorentz invariants can be found by taking any four vector X (or two of them, suppose the second is Y) and evaluating $X^T \eta X$ (or $X^T \eta Y$). We have have mentioned before that all of these (scalars, vectors, metric) are tensors of a particular rank (0, 1 and 2) respectively.
- In general relativity, we'll need to use tensors even more heavily. One quantity we have in GR for example is the *Riemann curvature tensor* R (which we'll cover in detail later), which is a tensor of rank **4** (it needs to feed on 4 vectors to give a scalar). There's no obvious way to represent that besides using the function call notation $R(x, y, z, w)$ which can be somewhat inconvenient.
- This motivates Einstein's notation. A tensor T will have a rank equal to the number of its indices.
- A tensor of rank 1 (a vector) will have one upper index T^μ (not to be confused with exponentiation). μ is an index that goes from 0 to 3 and represents the components of T^μ. For example, suppose X^μ is the 4-position. X^0 is its zeroth component (the time component), so $X^0 = ct$, $X^1 = x$, $X^2 = y$ and $X^3 = z$.
- A tensor of rank 2 (e.g. a metric) will have two indices $T_{\mu\nu}$ each going from 0 to 3 and representing a "slot". For example, the minkowski metric η we know has non-zero components only at the diagonals, and the values at the diagonals are $(-1, 1, 1, 1)$. E.g: Ro read off the components, $\eta^{00} = -1$, $\eta^{22} = 1$, $\eta^{21} = 0$ ..etc. The positioning of the indices (up or down) is significant as will be seen in the next point.

- Evaluation of tensors in this this notation assumes the Einstein summation convention: Indices given the same symbol *up* and *down* imply a sum:

$$V^\mu W_\mu = \sum_{\mu=0}^{3} V^\mu W_\mu = V^0 W_0 + V^1 W_1 + V^2 W_2 + V^3 W_3$$

which is a scalar (no free indices).

- Examples:
  - The lorentz transform matrix is a two tensor. But usually we just apply one 4-vector to it to get back another 4-vector. In einstein summation convention we can write it as:

$$\Lambda X = \Lambda^\nu{}_\mu X^\mu = \sum_{\mu=0}^{3} \Lambda^\nu{}_\mu X^\mu = \Lambda^\nu{}_0 X^0 + \Lambda^\nu{}_1 X^1 + \Lambda^\nu{}_2 X^2 + \Lambda^\nu{}_3 X^3$$

there's one free index left, which means the sum is a new vector. We can check the calculation is equivalent to the regular matrix multiplication, assuming $\Lambda = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

$$\Lambda X = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma ct - \frac{v}{c}\gamma x \\ -\frac{v}{c}\gamma ct + \gamma x \\ y \\ z \end{pmatrix}$$

$$\Lambda^\nu{}_\mu X^\mu = \Lambda^\nu{}_0 X^0 + \Lambda^\nu{}_1 X^1 + \Lambda^\nu{}_2 X^2 + \Lambda^\nu{}_3 X^3 = \begin{pmatrix} \gamma \\ -\frac{v}{c}\gamma \\ 0 \\ 0 \end{pmatrix} ct + \begin{pmatrix} -\frac{v}{c}\gamma \\ \gamma \\ 0 \\ 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} y + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} z = \begin{pmatrix} \gamma ct - \frac{v}{c}\gamma x \\ -\frac{v}{c}\gamma ct + \gamma x \\ y \\ z \end{pmatrix}$$

Note that when using einstein's notation, the order of multiplication doesn't matter as it's already dictated by the indices (so the order of indices *does* matter now, unless the tensor is symmetric in two indices, then you can swap them).

- The metric tensor is a rank 2 tensor, and it takes two vectors and gives you the spacetime distance. We can show that $\Delta X^T \eta \Delta X = \eta_{\mu\nu} \Delta X^\mu \Delta X^\nu$:

$$\eta_{\mu\nu} \Delta X^\mu \Delta X^\nu = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} \Delta X^\mu \Delta X^\nu = -(c\Delta t)^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$$

- Co-vectors are operators which take a vector and return a number, and are vectors themselves. In einstein notation, they're simply vectors with a lower index W_μ so that you can apply it on a vector V^μ by just sticking them together to get a number: $W_\mu V^\mu = \sum_{\mu=0}^{3} W_\mu V^\mu$.

- In general relativity in particular, and in physics in general, it's usually useful to identify for each vector V^μ its own dual vector that has its same variable name but with its index position changed: V_μ such that $V_\mu V^\mu$ gives the inner product (evaluation of the metric) on the the two. This means:

$$\Delta X^\mu \Delta X_\mu \stackrel{!}{=} \Delta X^\mu \Delta X^\nu \eta_{\mu\nu} \implies \boxed{\Delta X_\mu = \Delta X^\nu \eta_{\mu\nu}}$$

this identifies the metric as a device that *lowers indices.*

- We know doing $V_\mu = V^\nu \eta_{\mu\nu}$ to lower its index is in matrix notation just the multiplication $V^T\eta$. This tells us that if we need to get V back, we need to multiply by the inverse of the metric η^{-1}. Multiplying it by the inverse is effectively getting its index up again. We thus define η^{-1} to be the tensor that raises the indices and we call it $eta^{\mu\nu}$ (with its indices up).
- So $\eta\eta^{-1} = I \xrightarrow{\text{Einstein notation}} \boxed{\eta^{\alpha\nu}\eta_{\nu\beta} = \delta^\alpha_\beta}$, where δ^α_{beta} is just the identity matrix in einstein's notation (and it gets a special fancy name for some reason "Kronecker delta". But it's really just the identity matrix).
- So now we have that:

$$\text{To lower indices: } \boxed{V_\mu = \eta_{\mu\nu}V^\nu}$$
$$\text{To raise indices: } \boxed{V^\mu = \eta^{\mu\nu}V_\nu}$$

- From now on, we'll use einstein's notation in our equations of general relativity as they're simply much more convenient.

3.4 The Geodesic Equation from the einstein equivalence principle

- Given coordinates x^μ, a new coordinate system is in the most general sense a function of the old coordinates; i.e. $\xi^\mu = \xi^\mu(x^\mu)$.
- So the einstein equilvalence principle tells one that there always exists a coordinate system for a point p where it's locally flat (has the minkowski metric, and thus inertial and thus acceleration is zero). So we have that $\frac{d^2\xi^\alpha}{d^2\tau} = 0$
- Let's work this out further:

$$\begin{aligned}\frac{d^2\xi^\alpha}{d^2\tau} &\stackrel{\text{chain rule}}{=} \frac{d}{d\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial x^\mu}{\partial\tau}\right) \stackrel{!}{=} 0 \\ &\stackrel{\text{product rule}}{=} \frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\frac{\partial x^\mu}{\partial\tau}\frac{\partial x^\nu}{\partial\tau} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial\tau^2} = 0\end{aligned}$$

Now let's multiply through out by $\frac{\partial x^\lambda}{\partial\xi^\alpha}$. Note that this is the inverse of the so called "Jacobian matrix" (matrix of derivatives) of ξ.

$$\begin{aligned}0 &\stackrel{!}{=} \left(\frac{\partial x^\lambda}{\partial\xi^\alpha}\right)\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\frac{\partial x^\mu}{\partial\tau}\frac{\partial x^\nu}{\partial\tau} + \frac{\partial\xi^\alpha}{\partial x^\mu}\left(\frac{\partial x^\lambda}{\partial\xi^\alpha}\right)\frac{\partial^2 x^\mu}{\partial\tau^2} \\ &= \left(\frac{\partial x^\lambda}{\partial\xi^\alpha}\right)\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\frac{\partial x^\mu}{\partial\tau}\frac{\partial x^\nu}{\partial\tau} + \delta^\lambda_\mu\frac{\partial^2 x^\mu}{\partial\tau^2} \\ &= \left(\frac{\partial x^\lambda}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\right)\frac{\partial x^\mu}{\partial\tau}\frac{\partial x^\nu}{\partial\tau} + \frac{\partial^2 x^\lambda}{\partial\tau^2}\end{aligned}$$

Define $\Gamma^{\lambda}_{\mu\nu} := (\frac{\partial x^{\lambda}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\mu} \partial x^{\nu}})$. This is just a group of coeffecients that we call the "connection" or the "Christoffel symbol". Our final equation becomes:

$$\boxed{\frac{\partial^2 x^{\lambda}}{\partial \tau^2} = -\Gamma^{\lambda}_{\mu\nu} \frac{\partial x^{\mu}}{\partial \tau} \frac{\partial x^{\nu}}{\partial \tau}}$$

that's the so called **geodesic equation**, and it's the equation of motion of general relativity.

- On the left handside is the acceleration in the given coordinate system. Notice how it's not equal to zero. It's analogous to the coriolis effect where acceleration can emerge simply by going to a different frame of reference. We will later show that the connection Γ only depends on the metric and its derivatives; and thus that curvature induces acceleration described by this equtaion.