Relativity

W$\hbar$Yβmaths

1 Relativity

Now that we are well versed with tensors and manifolds we have all of the tools required to formally study special and general relativity. Some of you may be entering this series without having seen my differential geometry videos, so I would encourage you to watch those first in order to get the most out of these videos, however, it would be possible to watch these videos and fill in any gaps as necessary when they arise. I will do my best to point out where I am using concepts I have covered in more detail previously, but may end up doing it without realising.

To begin studying relativity, we will first review special relativity (hoping that you have at least seen or heard some of the basics), although no prior knowledge is required as we will formalise the construction of the subject in entirety. I am going to focus more on the mathematical structure of the theory, and focus less on physical consequences and intuition, although I will comment where certain concepts that may seem unclear from a physical perspective are given concrete meaning once differential geometry is introduced.

2 Spacetime

One of the key concepts underlying the transition from classical to relativistic mechanics is the unification of space and time into a single object, spacetime. Previously, classical mechanics relied on specifying the location of an object within $\mathbb{R}^3$ (a set of locations), and time was a universal parameter independent of the space of locations. The trajectory of any particle was then given by $x^i(t)$, a curve paramaterised by the global time coordinate. Since the time coordinate is a global quantity, all observers will agree on the time measured by the particle.

Spacetime realises the universe as a four dimensional manifold, $\mathbb{R}^4$, which is now a set of *events*, given by the 4-tuple (ct, x, y, z). We will frequently refer to these coordinates as x^μ, where latin indices run from 0 to 3, with 0 being the 'time' component and $i = 1, 2, 3$ the space components. The constant c is required for dimensional consistency, such that all spacetime coordinates have dimension length. Since it is arbitrary it can be set equal to one by an appropriate choice of time coordinate, this simplification is frequently chosen in the literature however we will leave the c explicit for now in order to keep track of its influence. For now, c is just an arbitrary constant with units of velocity, and we will discover the physical consequences it has shortly.

Notice now that time appears as an independent coordinate, and so we can no longer use it to paramaterise trajectories through this 4D spacetime. This has dramatic consequences as we will formalise eventually, but I will just briefly summarise now:

Time is no longer universal, and rather depends on the coordinate system used (coordinate frame/lorentz frame/reference frame). Coordinate systems are related by lorentz transformations (rotations + boosts to a certain velocity) and we will see that boosted observers travelling relative to one another (in different coordinate systems) will measure time differently. Therefore, time is no longer a global quantity agreed by all observers, rather it is relative to the particular frame. Stated another way, each coordinate frame defines a unique time coordinate, measured by all observers in that frame. Observers in a relative frame will measure an entirely different time coordinate, based on the frame they are in. This gives a notion of proper time, τ, the time experienced within your own coordinate frame. Then coordinate time, t, is the time in some other particular lorentz frame. All of this will be formalised shortly so if it doesn't make sense now come back and read this again later!

3 Worldlines

We should realise that spacetime is now a set of events, not just locations as in classical mechanics. The trajectory of a particle through spacetime now specifies the past, present and future trajectory of such a particle and is usually referred to as a 'worldline'. The worldline of a particle will be a map from some parameter into $\mathbb{R}^4$, specifying the trajectory of a particle as seen in some particular frame, dependent on some parameter that will be discussed in detail later but is arbitrary for now. Hence, a worldline is the map

$$x^\mu : \mathbb{R} \longrightarrow \mathbb{R}^4 : \lambda \mapsto x^\mu(\lambda) \tag{1}$$

where the parameter λ is known as an affine parameter, one that ranges over all of $\mathbb{R}$ to map the entire worldline. This parameter should not yet be interpret as a time parameter, and it should be noted that coordinate time is now a paramaterised function of λ.

We can also consider vector quantities such as the 4-velocity (4-vector is terminology for a vector in Minkowski space), which is given by the parameter derivative of the worldline. This should not be confused with the coordinate velocity (3-vector), which is measured w.r.t the time coordinate in some particular frame. the differences between the two will be discussed in detail shortly.

$$\text{4-velocity} \quad u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \qquad\qquad \text{Coordinate velocity} \quad v^i = \frac{\mathrm{d}x^i}{\mathrm{d}t} \tag{2}$$

Since spacetime coordinates refer to one particular 'Lorentz frame' (inertial frame), this worldline will appear differently in relative frames. Each of these frames represents an inertial observer, either stationary or travelling at constant velocity. These reference frames are related to each other by Lorentz transformations; coordinate transformations of spacetime that relate observers in different Lorentz frames. From now on, you should realise that observers and frames are synonymous; since every (inertial) observer defines a particular set of coordinates (lorentz frame) within which they measure the events in spacetime. Within their reference frame, they assign particular coordinate values (e.g the time and location) of an event. Since observers in different inertial frames must be using different coordinates, they will inherently be measuring the events in spacetime differently, i.e they will assign different time values to events than any other inertial observer, with these values being related by lorentz (coordinate) transformations.

We draw worldlines of particles on a 'spacetime diagram', which for simplicity often supresses one or more space dimension. The most common diagram is of a (1+1) dimensional spacetime, with the time and one space axes shown. Time flows upwards in such a diagram, and the worldline of a particle is a continuous path drawn through spacetime. A straight line would therefore correspond to a particle travelling at a constant velocity, given by the gradient of the line. It is important to stress that such a spacetime diagram is not inherently relativstic, we can draw these diagrams in classical mechanics. However, as we will shortly realise, the geometry of spacetime in special relativity is NOT the same as classical mechanics (Euclidean geometry), and hence such pictures must be treated with caution in the relativistic setting to avoid errors based on Euclidean reasoning.

4 Minkowski Space

Spacetime as a manifold is globally $\mathbb{R}^4$, however, rather than the Euclidean metric it is endowed with the pseudo-Riemannian metric $\eta_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$, hence it is customary to refer to it as the space $\mathbb{R}^{1,3}$. Consequently, the line element is

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -c^2\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \tag{3}$$

This then fully characterises the (1+3) dimensional spacetime known as Minkowski space. The consequences of the differing sign for the time and space components are significant and difficult to grasp without more details, so we will frequently revisit this point. For now, we take it for granted and study the consequences of such a metric (geometry).

5 Lightcone structure

We can use the Minkowski metric to probe the geometric structure of spacetime. The line element gives us the infentesimal spacetime distance between two events separated by $\mathrm{d}x^\mu$, this can equivalently be integrated to express finite displacements Δs. Consider a (1+1) dimensional spacetime for simplicity. The spacetime (finite) interval is equivalently expressed

$$\Delta s^2 = -c^2\Delta t^2 + \Delta x^2 \qquad \Longrightarrow \qquad \Delta s = \sqrt{-c^2\Delta t^2 + \Delta x^2} \tag{4}$$

The presence of the minus sign in the root means that the space time distance can be zero or even imaginary, depending on the values of Δx and Δt; this has dramatic consequences as we will now begin to unravel. Every Lorentz frame is now split into three separate regions, dependent on the sign of the line element. If we consider for now the distance measured from the origin in a particular Lorentz frame, $\Delta x = x - 0 = x$. Hence, measured in this frame, distances in spacetime must satisfy $\Delta s^2 = -c^2t^2 + x^2$. Consider the case where $\Delta s^2 = 0$, which occurs when

$$x^2 = c^2t^2 \quad \Longrightarrow \quad x = \pm ct$$

These correspond to lines drawn at 45 degrees on a spacetime diagram (ct, x), and define what is known as the *lightcone* of that particular reference frame (observer). Within the reference frame with coordinate time t, the lines $x = \pm ct$ would correspond to a worldline with velocity $\frac{\mathrm{d}x}{\mathrm{d}t} = c$, i.e that of a light ray emitted at the origin - more on velocity later. The lightcone hence separates Minkowski space into three regions, summarised below, dependent on the sign of ds^2.

$$\mathrm{d}s^2 : \begin{cases} \mathrm{d}s^2 < 0 : \text{timelike} \\ \mathrm{d}s^2 = 0 : \text{null} \\ \mathrm{d}s^2 > 0 : \text{spacelike} \end{cases} \qquad \Longrightarrow \qquad \mathrm{d}s : \begin{cases} \text{imaginary} \\ \text{null} \\ \text{real} \end{cases} \tag{5}$$

These sign conventions hold for the norm of any quantity (i.e contracting any two vectors with the metric). For example the 4-velocity vector of a particle, real particles are bound to travel less than the speed of light and hence they must follow timelike trajectories (with a timelike velocity vector). Light, which is massless, can travel on a null trajectory, and nothing can travel on a spacelike trajectory.

It should be stressed that a spacetime diagram, whilst it resembles $\mathbb{R}^2$, the geometry is NOT that of traditional Euclidean $\mathbb{R}^2$ that would be used in say classical mechanics. Most importantly that points on the lines $x = \pm ct$ have zero separation in spacetime from the origin - in stark contrast to if the metric had been plus.

6 From Euclidean to Minkowski Geometry

As I aluded to in the previous section, whilst the spacetime diagrams we have been drawing resemble $\mathbb{R}^2$, it is important and illuminating to explore the differences between the standard Euclidean geometry and our newly discover Minkowski geometry. For simplicity I have set $c = 1$, and for further simplicity let us work with a two dimensional spacetime (t, x). As a manifold, this spacetime is $\mathbb{R}^2$, and we define the Euclidean/Minkowski geometries respectively by the metrics (line elements)

$$\mathrm{d}s_{\mathrm{E}}^2 = \mathrm{d}t^2 + \mathrm{d}x^2 \qquad \mathrm{d}s_{\mathrm{M}}^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 \tag{6}$$

Let us now consider non-infentesimal distances measured from the origin as previously, we obtain the following two equations for each geometry

$$\Delta s_{\mathrm{E}}^2 = t^2 + x^2 \qquad \Delta s_{\mathrm{M}}^2 = -t^2 + x^2 \tag{7}$$

We can use these two equations to probe the geometry, by considering the curves along which Δs is constant. As we have seen previously, the curves $\Delta s_{\mathrm{M}}^2 = 0$ define the lightcone given by the straight lines $t = \pm x$ in the Minkowski geometry. However, $\Delta s_{\mathrm{E}}^2 = 0 = t^2 + x^2$ is only true for $x = t = 0$, already a stark difference. We are familiar with the Euclidean notion of $\Delta s_{\mathrm{E}}^2 = 0$, as this would simply correspond to zero distance in separation between two points, which as measured from the origin is simply the origin in our diagram. The notion of zero 'distance' in Minkowski space is vastly different however, physical distances now corresponds to intervals in the space dimensions, however, spacetime distance is both a function of the space and time interval. Since the spacetime interval (line element) is an invariant quantity in all coordinate systems, space and time intervals can be measured differently depending on the coordinates you use (the reference frame you are in). This will become evident when we consider Lorentz transformations and relativistic physical effects such as time dilation.

Now consider curves on which the distance as measured from the origin is constant and positive (set = 1 for simplicity). In a Euclidean geometry this is given by $t^2 + x^2 = 1$ which is simply the unit circle, what we would usually realise as the points that lie a fixed distance from the origin. However, in the Minkowski geometry, this corresponds to $-t^2 + x^2 = 1$, which defines a hyperbola as shown in Fig. (REFF). Hence, all points on this hyperbola lie at a constant fixed Minkowski distance from the origin, in stark contrast with our Euclidean intuition. So we see that whilst spacetime diagrams can be realised as an image of $\mathbb{R}^2$, the geometry is not that of the Euclidean plane which is usually associated with $\mathbb{R}^2$.

The differences between the two geometries can therefore be inferred entirely from the line element, and we can even begin to represent various relativistic effects purely by considering geometric arguments within this spacetime diagram.

7 Lorentz Transformations

The core principle of relativity can be stated as requiring that physics should be equivalent regardless of the inertial frame within which you reside. This can be stated using modern terminology as the requirement that physical theories should be *Lorentz* invariant - invariant under Lorentz transformations. Since Lorentz transformations are coordinate transformations, this is realised in practice by the requirement that any physics equation you can write should be written using tensors. Therefore the form of the equation will be unaffected by coordinate (basis) transformations, i.e all observers use the same equation, just with the tensor expressed in their particular basis (coordinate system).

Lorentz transformations are coordinate transformations that preserve the metric, i.e they represent isometries of spacetime. Lorentz transformations are therefore symmetries of spacetime, reflecting the

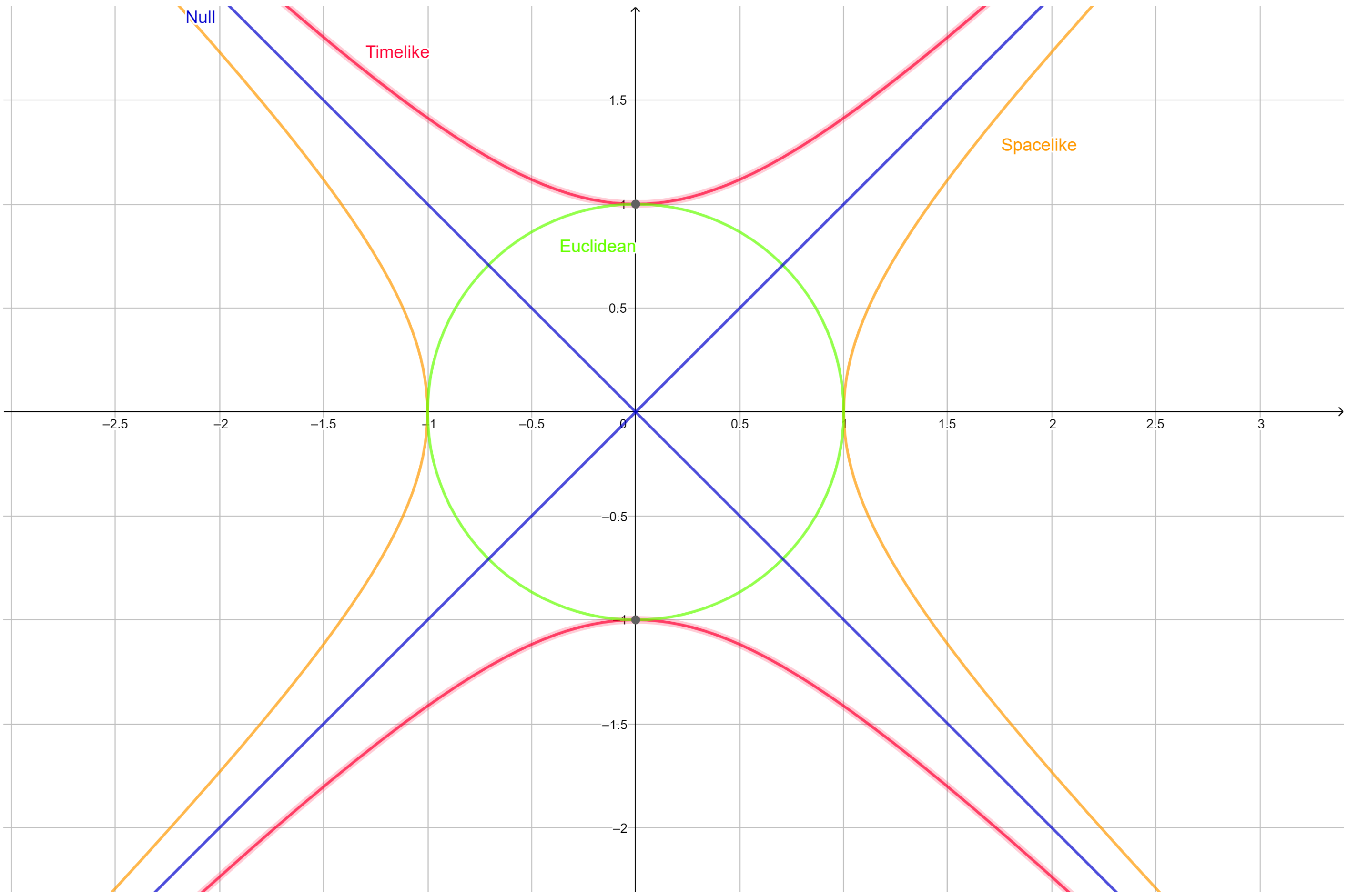

Figure 1: Plot showing points with constant distance from the origin, in the Euclidean and Minkowski geometries respectively

core principle of relativity, that physics should be invariant (symmetric) under Lorentz transformations. The possible symmetries include the three rotations of $\mathbb{R}^3$, along with boosts (to be introduced shortly) along each coordinate direction.

Consider a Lorentz frame, which we will identify using the coordinates $x^\mu = (ct, x^i)$. To explore the possible Lorentz transformations we use the isometry of the metric. Isometry requires that the metric is equivalent when expressed in new coordinates $y^\mu = (c\tilde{t}, y^i)$, simply that

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \eta_{\mu\nu}\mathrm{d}y^\mu \mathrm{d}y^\nu \tag{8}$$

this requirement can then be used to determine the possible coordinate transformations that are isometric (Lorentz transformations). The coordinates y^μ, are related to the x^μ coordinates through an *arbitrary* coordinate transformation

$$y^\mu = \Lambda^\mu{}_\nu x^\nu \tag{9}$$

The Lorentz transformation is a (1,1) tensor, what we would usually associate with a matrix operator that acts on a vector. Hence, it is convenient and frequent to see LT expressed as matrices, although we will opt mostly for tensor notation here.

Now that we have defined the two frames, the Lorentz transformation can be obtained by imposing the isometry requirement, giving

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \eta_{\mu\nu}\mathrm{d}y^\mu \mathrm{d}y^\nu = \eta_{\mu\nu}\mathrm{d}\left(\Lambda^\mu{}_\sigma x^\sigma\right)\mathrm{d}\left(\Lambda^\nu{}_\rho x^\rho\right) = \eta_{\mu\nu}\Lambda^\mu{}_\sigma\Lambda^\nu{}_\rho \mathrm{d}x^\sigma \mathrm{d}x^\rho \tag{10}$$

where in the last step we used the fact that the Lorentz transformation will be a constant matrix. Now using the fact that the metric is a $(0,2)$ tensor (or equivalently the pullback of the one forms $\mathrm{d}y^\mu$), we can equivalently express the Lorentz transformation using the tensor transformation law as

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \eta_{\mu\nu}\mathrm{d}y^\mu\mathrm{d}y^\nu = \eta_{\mu\nu}\frac{\partial y^\mu}{\partial x^\sigma}\frac{\partial y^\mu}{\partial x^\sigma}\mathrm{d}x^\sigma\mathrm{d}x^\rho = \eta_{\mu\nu}\Lambda^\mu{}_\sigma\Lambda^\nu{}_\rho\mathrm{d}x^\sigma\mathrm{d}x^\rho \tag{11}$$

So far all we have done is to realise that the Lorentz transformation is just a coordinate transformation, however, the isometry requirement leads to the following expression

$$\eta_{\mu\nu} = \eta_{\mu\nu}\frac{\partial y^\mu}{\partial x^\sigma}\frac{\partial y^\nu}{\partial x^\rho} = \eta_{\mu\nu}\Lambda^\mu{}_\sigma\Lambda^\nu{}_\rho \tag{12}$$

which is simply the statement that the metric tensor components are invariant when changing coordinates. To now derive the possible Lorentz transformations we must realise the following condition, by taking the determinent of the expression obtained previously

$$\det(\eta_{\mu\nu}) = \det(\eta_{\mu\nu})\det(\Lambda^\mu{}_\sigma)\det(\Lambda^\nu{}_\rho) \qquad \Longrightarrow \qquad \det(\Lambda)^2 = 1 \qquad \Longrightarrow \qquad \det(\Lambda) = \pm 1 \tag{13}$$

It turns out the choice of sign is arbitrary, corresponding to two disconnected sets of Lorentz transformations, the sign representing a reversal of the time axes (we will discuss the disconnected nature of the Lorentz group later) - for consistency and w.l.o.g we can set $\det(\Lambda) = 1$ for convenience.

By taking the partial derivative of the above expression we obtain further constraints as follows,

$$\partial_\alpha\eta_{\mu\nu} = 0 = \eta_{\mu\nu}\left(2\frac{\partial^2 y^\mu}{\partial(x^\sigma)^2}\right)\frac{\partial y^\nu}{\partial x^\rho} \tag{14}$$

Since the Lorentz transformation is non-zero this is realised as the requirement that

$$\frac{\partial^2 y^\mu}{\partial(x^\sigma)^2} = 0 \tag{15}$$

i.e that the Lorentz transformation should be a constant matrix Given the above requirements we now have everything we need to realise the possible Lorentz transformations. We will first simply state the possible transformations, and verify they satisfy the above properties. Fortunately, we can understand the origin of these transformations at a much deeper level as I will present in the next section, for now we simply state. The possible Lorentz transformations include the three orthogonal rotations of $\mathbb{R}^3$, corresponding to rotations in the xy, xz, zy planes. Such transformations are given by orthogonal matrices, which in spacetime act only on the space components, hence can be represented as the block matrices

$$\Lambda_\mathrm{r} = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \tag{16}$$

where R is a 3×3 orthogonal rotation matrix, eg for rotating around the x axis (yz plane)

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{17}$$

where θ is a parameter defining the transformation.

As we will discuss in the next section, such matrices are elements of the Lie group $SO(3)$, and must be orthogonal $R^T R = \mathbb{I}$ and have unit determinant. It can be easily checked that the above requirements are satisfied, hence rotations are Lorentz transformations.

We can now identify a further set of three transformations, known as boosts. Consider a matrix of the following form

$$\Lambda = \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0 \\ -\sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{18}$$

where ψ is a parameter to be made explicit shortly. It can be easily checked that this is a Lorentz transformation, and we should note its similarity with an orthogonal rotation matrix, albeit with hyperbolic functions replacing trigonometric. Hence, as we will eventually unravel further, Lorentz boosts correspond to *hyperbolic* rotations in spacetime, between the xt, yt and zt planes. This matrix for example, corresponds to a rotation in the xt plane and hence carries a similar structure as in the yz parts of the R_x rotation matrix. The term boost refers to the fact that the transformed frame can be understood to be moving relative to the original frame with some velocity, all of which will be formalised shortly.

The six transformations given previously, three rotations and boosts, are all of the Lorentz transformations possible that preserve the Minkowski metric. Such transformations form what is known as the Lorentz group, $SO(1,3)$, which can be understood as the 'rotations' in spacetime that leave the Minkowski metric invariant (more on this in a later section). Finally, in addition to Lorentz transformations, we also have four (basically trivial) transformations known as Poincare transformations; constant translations along any of the spacetime axes. The full set of all 10 transformations (3, rot, 3 boost, 4 translations) forms what is known as the Poincare group. The inclusion of Poincare transformations is usually not required, since they are so simple as to be almost arbitrary, hence it is customary to only focus on the Lorentz parts of the Poincare group. However we must keep in mind the translational in-variance also present in the full Poincare group.

8 Lorentz Boosts

We will now explore Lorentz boosts in more detail, and hopefully this section should link up to what you may have seen on an introductory relativity course. As we have alluded to previously, a Lorentz boost is a hyberbolic rotation of a space and time axis, we will explore the geometric intuition behind such a rotation, but for now consider it algebraically. For a Lorentz from x^μ, the boosted frame y^ν is given by

$$y^\nu = \Lambda^\nu{}_\mu x^\mu = \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0 \\ -\sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ct\cosh\psi - x\sinh\psi \\ -ct\sinh\psi + x\cosh\psi \\ y \\ z \end{pmatrix} \tag{19}$$

Firstly, we should note that the new boosted coordinates contain a mix of the space and time coordinates of the initial frame, this is the origin of many relativistic effects as we will see shortly.

Consider now the worldline of a stationary observer at rest at the origin in the y^ν coordinates, simply given by $y(\tau) = 0$. Using the Lorentz boost we can obtain that in the x^μ coordinates we must have

$$y^1 = 0 = -ct\sinh\psi + x\cosh\psi \qquad \Longrightarrow \qquad x = ct\tanh\psi \tag{20}$$

Hence, the stationary observer at rest in the origin of the y frame is seen to be moving with velocity $x/t = c\tanh\psi$ in the x^μ coordinates, ie the boosted frame y^ν defines a frame moving relative to x^μ with a constant (frame) velocity

$$v = c\tanh\psi \tag{21}$$

Now we should realise a vital fact that I have avoided stating until this point. The function $\tanh\psi$ takes values in the interval $(-1, 1)$. Hence, the velocity is bounded to the interval $(-c, c)$, i.e no boosted frame can exceed the speed of light. This fact is core to the Lorentz invariance of relativity, that all observers should measure the same speed of light and that you cannot travel faster than light. Here we see it explicitly derived, purely due to the geometric form of the Lorentz boost

To obtain a more familiar form of the Lorentz boost matrix, recall the following identities

$$\cosh\psi = \frac{1}{\sqrt{1-(\tanh\psi)^2}} \quad , \quad \sinh\psi = \tanh\psi\cosh\psi \tag{22}$$

then using $\frac{v}{c} = \tanh\psi$ we obtain the following

$$\cosh\psi = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \quad , \quad \sinh\psi = \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} \tag{23}$$

hence the boost is given by

$$\Lambda = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{24}$$

For brevity, we now identify the Lorentz factor

$$\gamma := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-\beta^2}} \quad : \quad \beta = \frac{v}{c} \tag{25}$$

such that the boost is given by

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{26}$$

We now have the definition of the parameter ψ, known as the rapidity

$$\psi = \cosh^{-1}\gamma = \tanh^{-1}\beta \tag{27}$$

which we will see shortly corresponds to a hyperbolic rotation angle.

To understand further the action of a Lorentz boost, consider how the coordinates are transformed,

$$y^\nu = \Lambda^\nu{}_\mu x^\mu = \begin{pmatrix} \gamma ct - \beta\gamma x \\ -vt\gamma + \gamma x \\ y \\ z \end{pmatrix} \tag{28}$$

In the limit of low velocities $v \ll c \implies \gamma \simeq 1$ the Lorentz boost reduces to a Galilean boost

$$y^\nu = \begin{pmatrix} ct \\ -vt + x \\ y \\ z \end{pmatrix} \tag{29}$$

However, as the velocity approaches c this is no longer the case, and the space and time axes of the new coordinates are a mix of the spacetime axes of the old coordinates, it is convenient to express this in the hyperbolic language, where the transformation takes the simple form of a hyperbolic rotation as we will now elaborate. Consider again our (1+1) spacetime diagram for simplicity, and consider boosting to new coordinates

$$y^{\nu} = \begin{pmatrix} \cosh\psi & -\sinh\psi \\ -\sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} ct\cosh\psi - x\sinh\psi \\ -ct\sinh\psi + x\cosh\psi \end{pmatrix} \tag{30}$$

9 Relativistic Physical Effects

To understand how such a transformation affects the coordinate axes it is convenient to plot them both on the same diagram, with a common origin. The boosted coordinates appear to scissor inwards dependent on the rapidity parameter. This is the effect of a hyperbolic rotation on the spacetime axes. The benefit

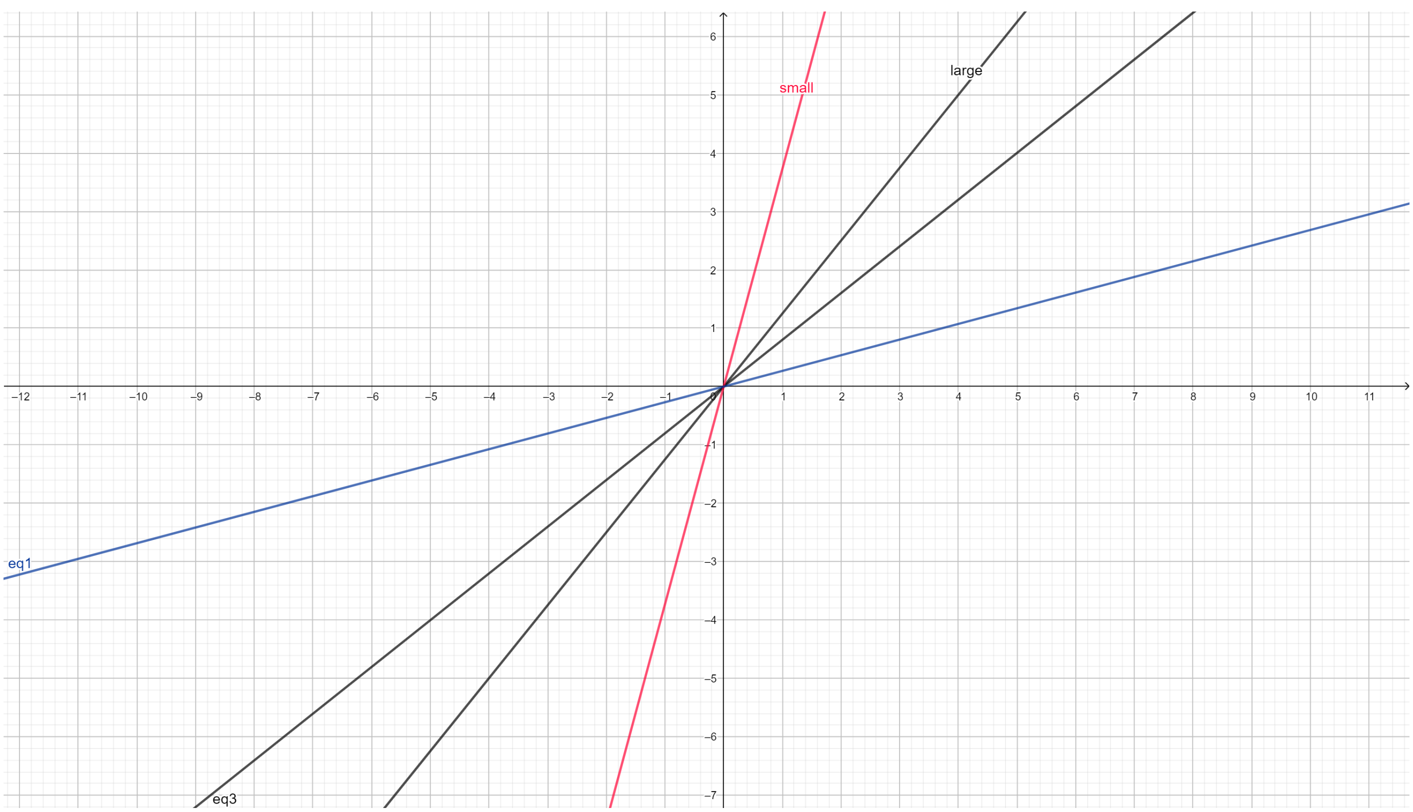

Figure 2: Plot showing boosted Lorentz frames plotted with a common origin

of plotting both frames on such a diagram is that relativistic effects can be visually understood in the following way. Since the line element must be invariant between frames, we can produce a plot of the curves of constant Minkowski separation from the origin as in Fig 1.

In this figure, the hyperbola represent curves that are a constant Minkowski separation (space and timelike respectively) from the origin. Since $\mathrm{d}s^2$ is an invariant, the Minkowski separation must be equivalent regardless of the reference frame used. This is incredibly useful and can be used to read off many physical effects purely from geometric arguments. Consider a stationary observer in the x^{μ} coordinates, sat at $x = 1$. From the diagram, at $t = 0$ we see this corresponds to a Minkowski separation of $\mathrm{d}s^2 = 1$, hence the stationary observer must have $\mathrm{d}s^2 = 1$ when measured in any coordinate system. Since the origins of our coordinates coincide the hyperbola drawn represent the Minkowski separation

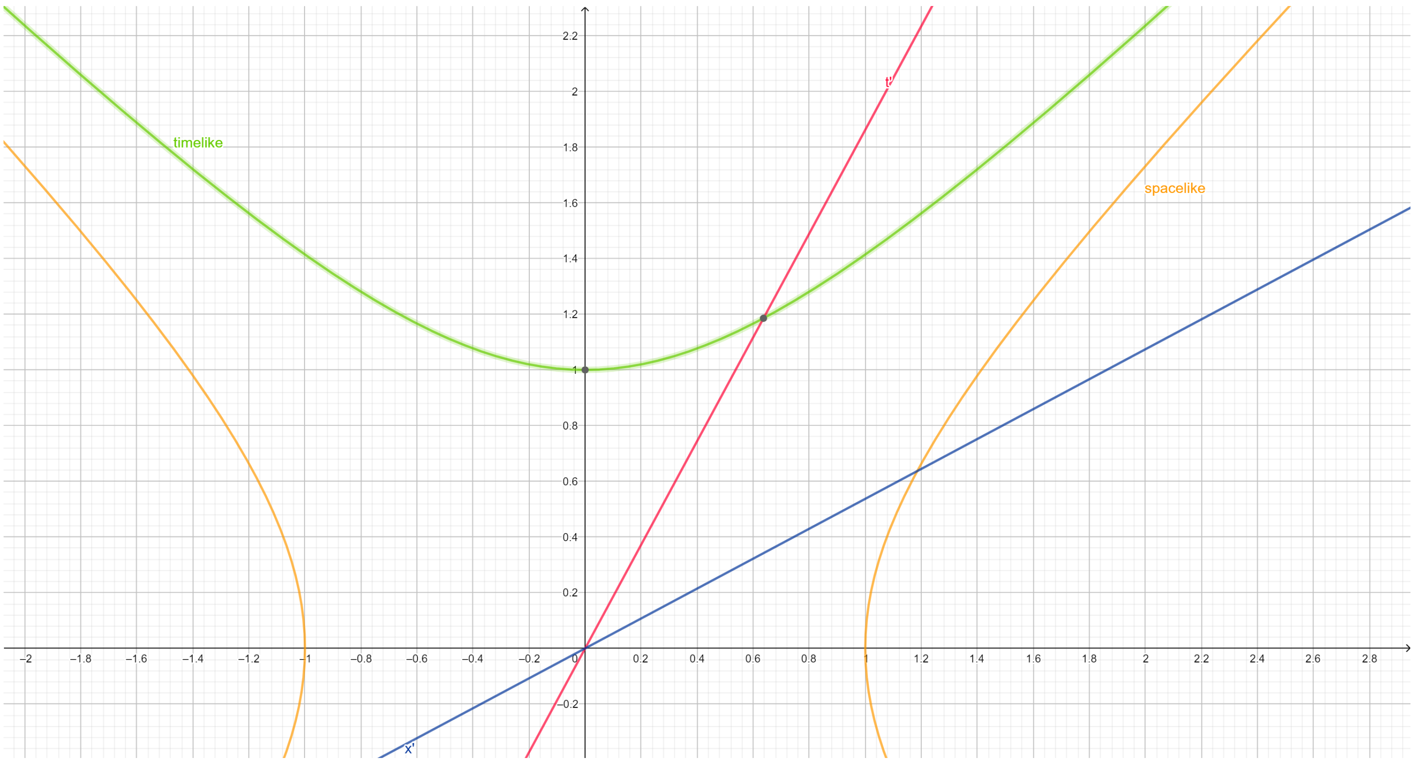

Figure 3: Plot showing boosted Lorentz frames plotted with a common origin

from the origin in both coordinate systems. We can then use this fact to explore how distances appear in each coordinate system.

Consider the purely spacelike interval $\Delta x = 1$, this would specify the Minkowski separation of our stationary observer at the time $t = 0$. Let us now explore how this interval would be viewed by the relatively moving (boosted observer). From the diagram, a spacelike interval of $\Delta x = 1$ is the intersection of the spacelike hyperbola with the x axis (at $x = 1$). Since the interval is an invariant, the boosted observer (scissored axes) must also measure the same interval, hence the intersection of the spacelike hyperbola with the y (blue) axis gives the interval as measured by the y coordinates. Since we chose the convenient time of $t = t' = 0$, this allows us to measure the purely spacelike separation in the y coordinates, which must correspond to $\Delta s^2 = \Delta y^2 = 1$. Now we should follow extreme caution with what follows since I am going to use a Euclidean argument on this Minkowski picture. We should also note, that the red and blue lines corresponding to the (t', y) axes, can equivalently be realised as the lines where y and respectively t' are zero.

Consider now an object at rest in the boosted coordinates, with length $y = 1$, effectively a rod with one end at the origin and one at $y = 1$. The observer defined by the coordinates x^μ, will see this rod moving with some relative velocity, and will measure the rod with a certain length. What length does x measure? Well in terms of coordinates, they must also simply measure $x = 1$ to be the endpoint of the rod (due to the interval invariance discussed previously). However, the 'physical length' of the rod is going to appear differently in each coordinate system. If we decide that the length of the rod as measured in the frame it appears at rest (y) should be its physical 'proper length' ℓ, we can then consider the length the stationary observer measures observing the moving rod. Furthermore, we should realise that the 'physical' length of the rod as measured along spatial dimensions will be a Euclidean length, so we should be extremely careful to distinguish between Euclidean 'lengths' and spacetime 'distances'.

We can now use the convenient fact that since the length of the rod is measured along a single axis only, we can directly correlate physical distances as drawn on this picture (a Euclidean distance!) with Minkowski distances (since a single axis has a metric equivalent to $\mathbb{R}^1$). This might be confusing and feel unorthodox at first, but we should remember that both of these coordinates are simply maps of some underlying $\mathbb{R}^4$, that we understand as the physical spacetime. If we then decide that a stationary observer can sufficiently measure this physical spacetime using their coordinates, we can use the chart of the stationary observer to represent 'physical' spacetime (recall how the chart for $\mathbb{R}^d$ is just $\mathbb{R}^d$). The two sets of axes I have drawn here simply represent two charts on the same copy of $\mathbb{R}^2$, and since we have designated stationary observers to measure physical distance, physical distances (measured on this picture) are in direct correspondence for both of these frames. Hence, the physical length of the rod is given by its length in the y frame, which can simply be read off in a Euclidean fashion from this picture as the length of the line segment between the origin and the intersection of the blue and orange lines. Hence, as viewed by the stationary observer the rod appears to be shorter in the x frame. The

We should be extremely careful with this argument, in fact it cuts deep to the core of relativity itself. Both of these observers (charts on spacetime) are simply assigning different sets of numbers to points in some underlying topological (physical) space, the numbers they assign are arbitrary as we know any set of coordinates is equally valid on the manifold. Both of these observers would agree that the length of the rod (measured from the origin at time $t = t' = 0$) is given by coordinate value $x = y = 1$, hence they would both conclude in their coordinates that the rod has length 1. However, this is simply a coordinate value and in no way corresponds to a physical length, since proper length is only determined in the frame at which the object is at rest.

10 Orthogonal vs. Hyperbolic Rotations

We should notice a similarity between the form of a Lorentz boost and an (orthogonal) rotation operator. To realise the connection it is convenient again to refer to our spacetime diagram from previous, showing constant values of Δs^2 in the Euclidean/Minkowski geometries, circles/hyperbola respectively. An orthogonal rotation can be understood as a coordinate transformation that leaves the Euclidean metric invariant, intuitively, a transformation that preserves the Euclidean separation between two points. This can be seen from the fact that all points on the unit circle map to the unit circle under an orthogonal rotation. And as can clearly be seen from the diagram, the rotated axes remain orthogonal (hence the name). Such rotations are achieved using orthogonal rotation matrices, that satisfy $R^T R = \mathbb{I}$. The matrices R are elements of the Lie group $O(3)$, the group of orthogonal, 3×3 matrices. As we will discuss in the Lie group part of differential geometry, such groups represent symmetric transformations of the underlying manifold, with the symmetry being isometry of the metric. Orthogonal rotations can be understood as the group of matrices that preserve the Euclidean metric, given by a distance formula of the form $\sum +(x^\mu)^2$, such that distances are invariant under rotations since

$$(Rx^\mu)^2 = (Rx^\mu)^T(Rx^\mu) = (x^\mu R^T)(Rx^\mu) = (x^\mu)^2$$

Furthermore, we specialise to the *special* orthogonal group $SO(3)$, which is the subset of $O(3)$ that contains the identity (and consists of matrices with determinant $+1$). Such a group is found to consist of orientation preserving transformations and throughout we always consider the special orthogonal case.

In the Minkowski setting, we have Lorentz transformations, which are elements of the Lorentz group $\Lambda \in SO(1,3)$. This group contains $SO(3)$, since rotations are three of the six possible LT, however it also contains boosts, which were realised as hyperbolic rotations between space and time dimensions. The Lorentz group elements are characterised by the requirement they preserve the pseudo-euclidean norm, i.e a distance formula of the form $\sum^p -(x^\mu)^2 + \sum^q +(x^\nu)^2$, which for Minkowski space $(p = 1, q = 3)$, returns the Minkowski norm $-(ct)^2 + (x^i)^2$. The Lorentz group elements are so called *pseudo-orthogonal*, satisfying the requirement $\Lambda^T \eta \Lambda = \eta$, which reduces to the orthogonal case when $\eta = \mathbb{I}$.

$$-d\tau^2 = -dt^2 + dx^2 \tag{31}$$

$$d\ell^2 \tag{32}$$