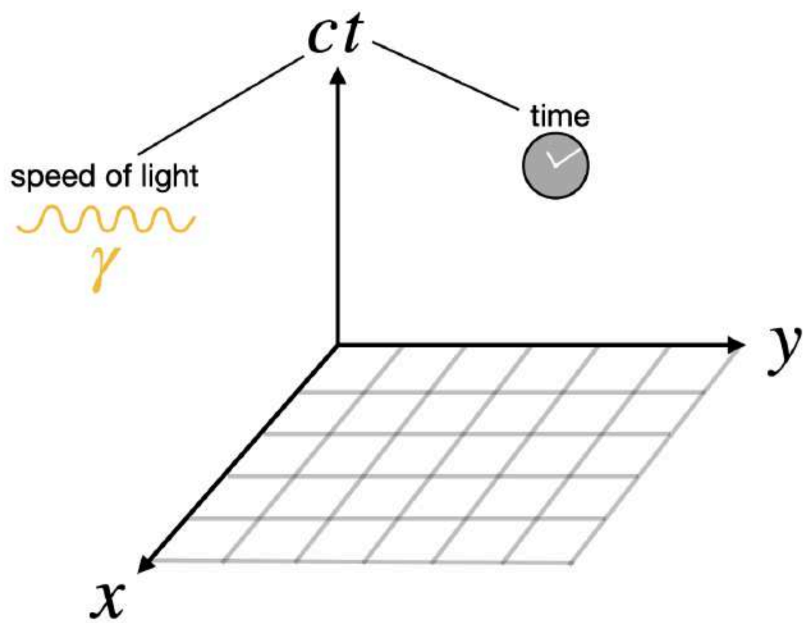
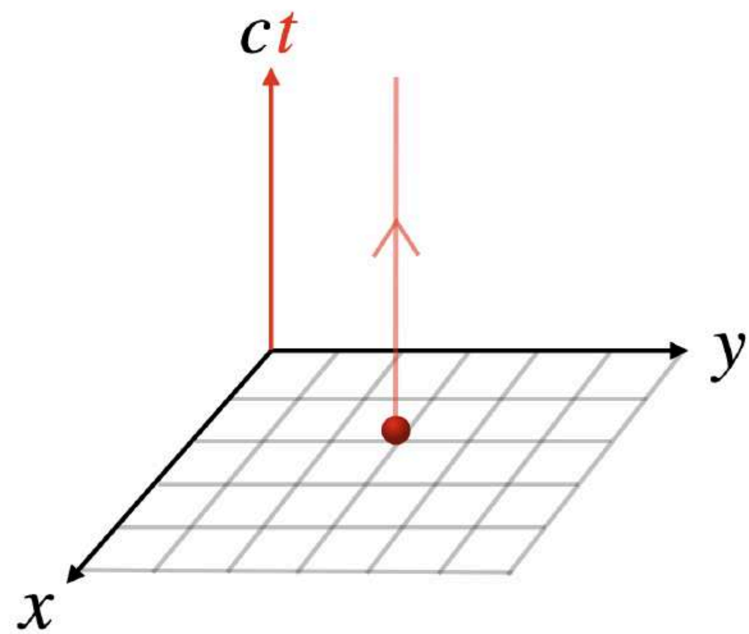
How To Derive Minkowski Geometry From Nothing

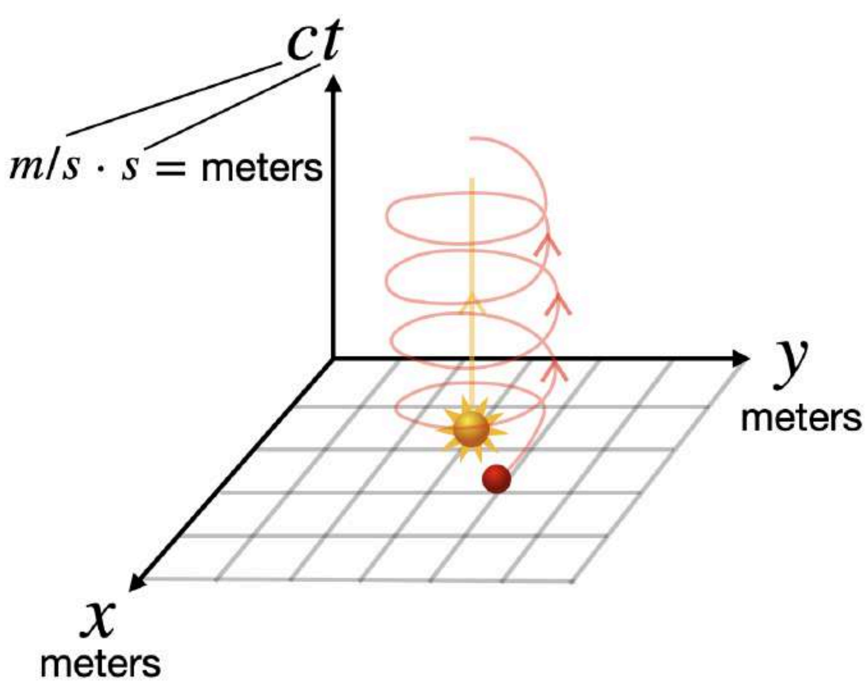
by Dibeos



This graph represents space as a $2D$ xy-plane and the vertical axis is the speed of light c multiplied by the time t. Usually this axis is used to represent just time t, and this way an object at rest in space would move in a straight line in spacetime:

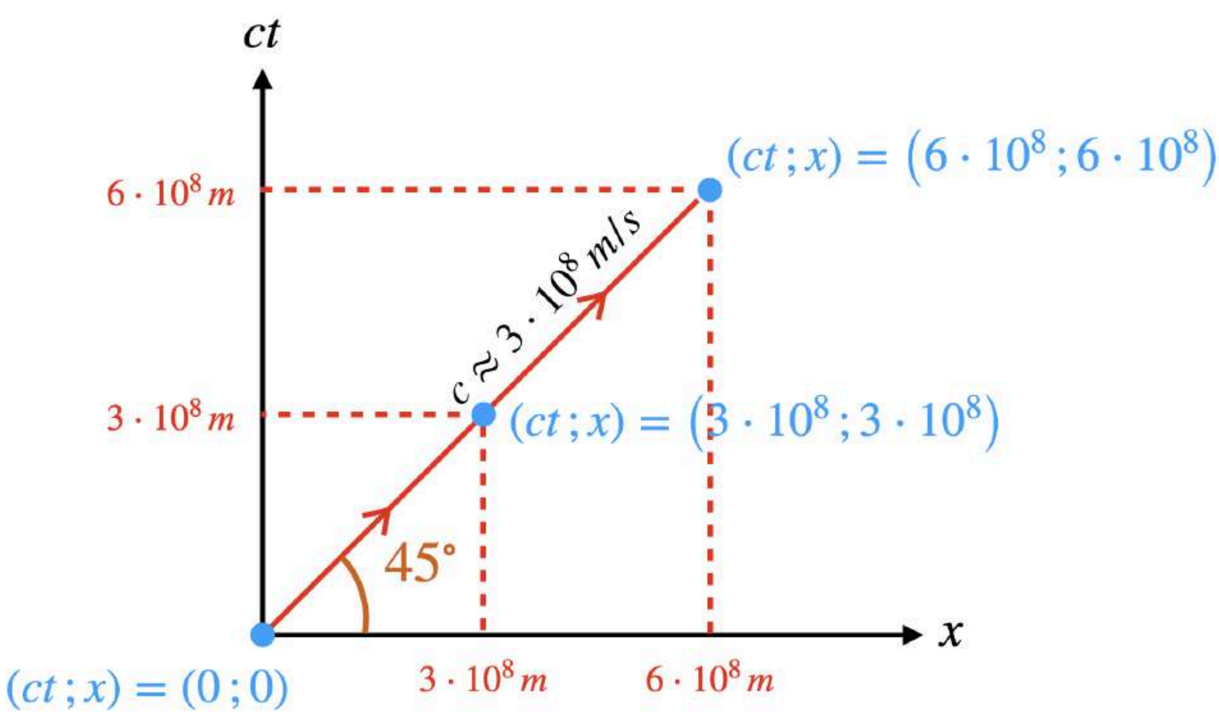


Similarly, a planet orbiting the sun would draw a spiral when moving throughout spacetime:

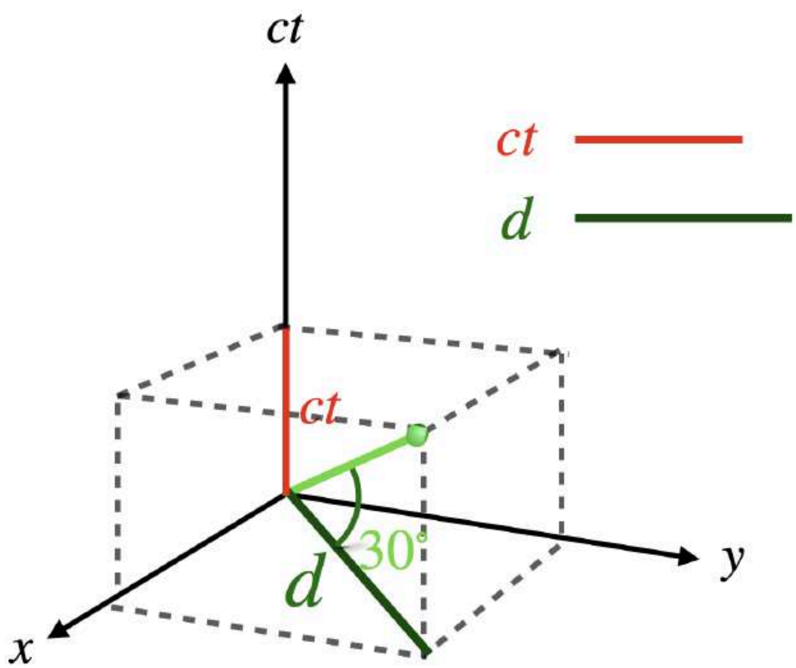


But here we added this c in front of t, and one of the reasons is because this way all axes have the same unit of measurement (let's say meters, for example).

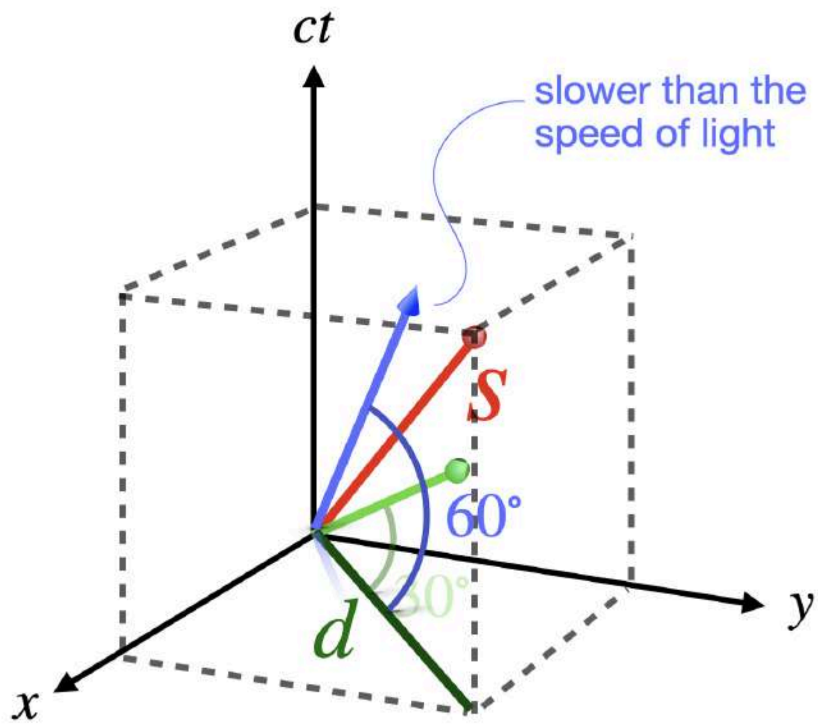
Another reason for using ct instead of t is because this choice gives us a very nice symmetry that will be extremely useful later on. We will see this symmetry through the following example. Imagine we have a point moving at the speed of light ($\approx 3 \cdot 10^8 m/s$) from the origin. This point will trace a spatial distance of $3 \cdot 10^8$ meters after 1 second. Similarly, the distance travelled in the vertical axis would be $3 \cdot 10^8\, m/s$ times 1 second. Now we do the same for 2 seconds and we get the point $(ct\,,\, x) = \left(6 \cdot 10^8\,,\, 6 \cdot 10^8\right)$. In other words, this specific choice of axis (ct) allows us to trace a line with a 45° angle that defines the path of any point moving at the maximum possible speed, i.e. the speed of light.



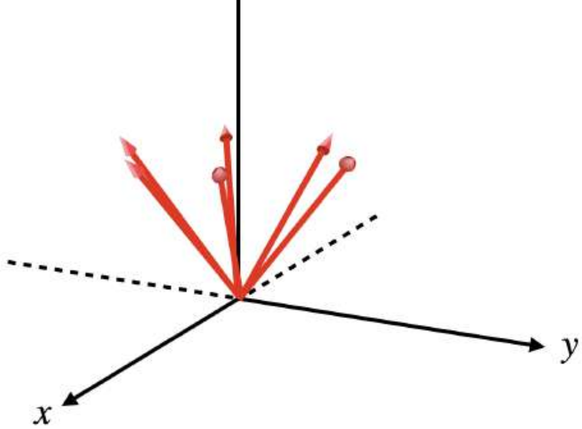
Anything *faster* that the speed of light will trace a line, or a curve with tangent line, that has inclination *less* than 45°. This can be intuitively understood through the fact that the body moves spatial distances that are longer than c-time distances, in the spacetime diagram.

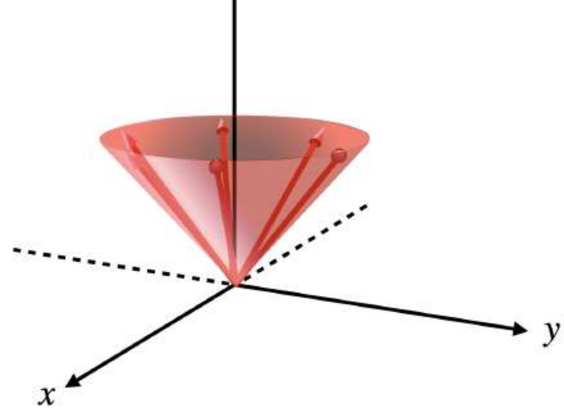


Anything *slower* that the speed of light will trace a line, or a curve with tangent line, that has inclination *greater* than 45°.

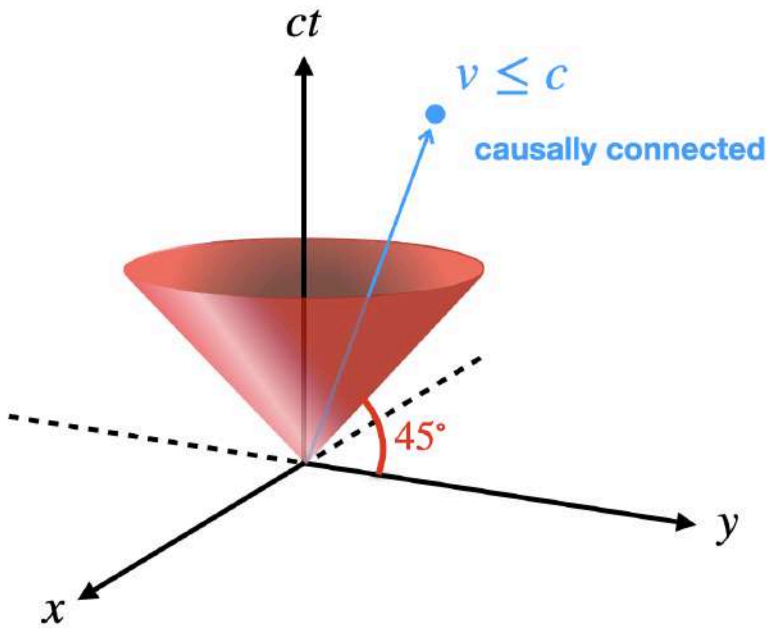


This is not only true for points travelling to the right in x and to the right in y, but also to the left in x and to the right in y, or any permutations.

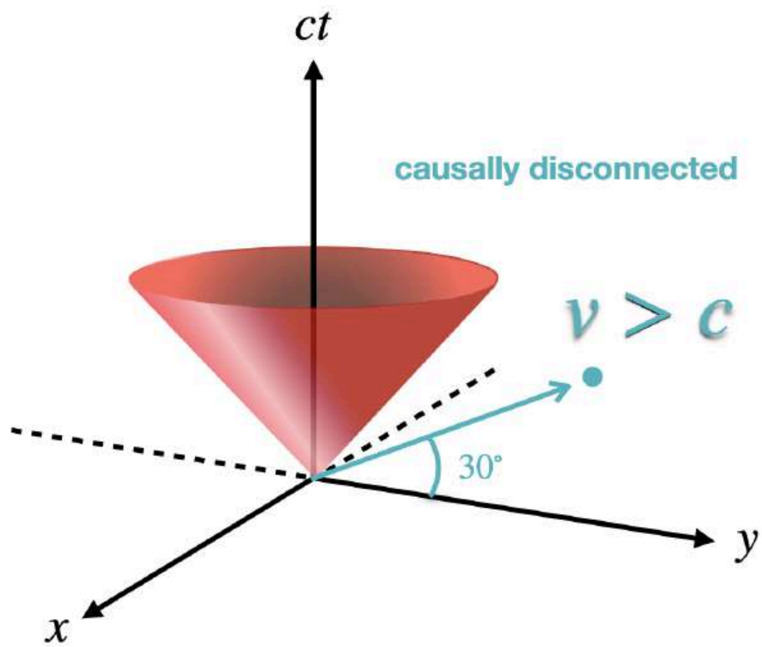




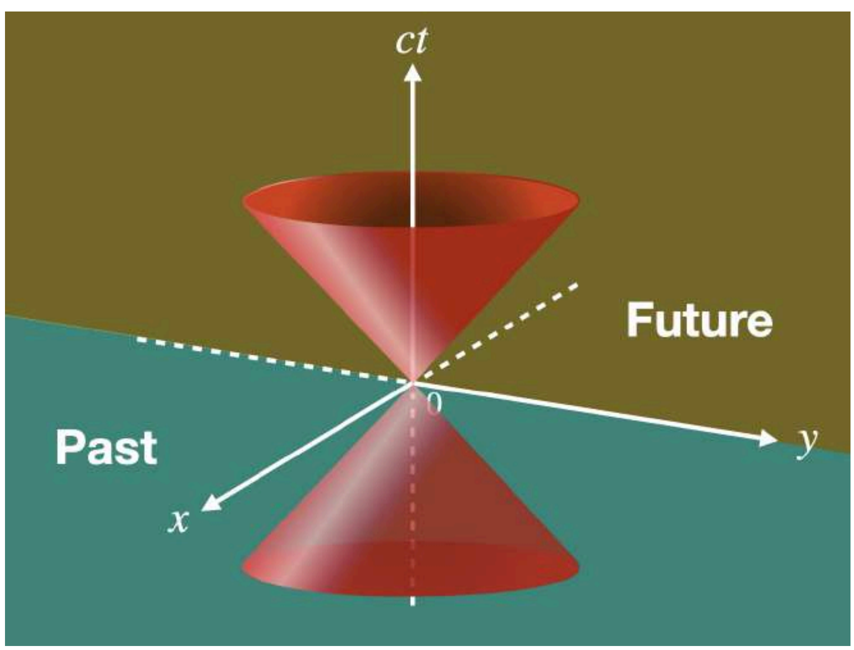
This creates a cone, such that everything that is inside of the cone is said to be *causally connected* to the origin, because it can be reached with a speed $v \leq c$.



Anything that is located outside of the cone is said to be *causally disconnected* to the origin, because it could be reached only with a speed $v > c$.



This is not only true for future events, but for past events as well.



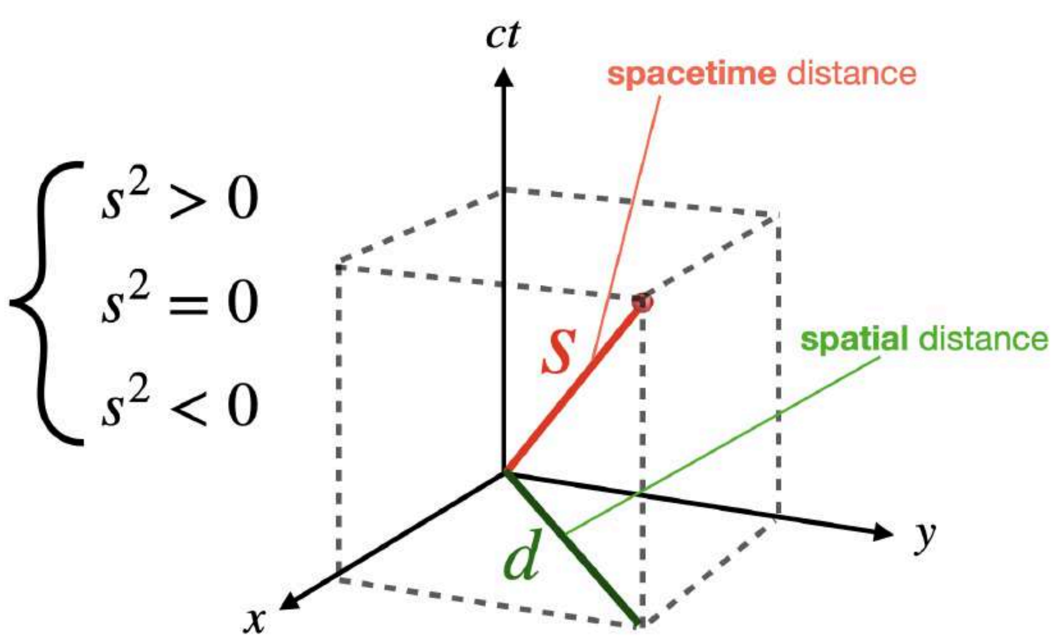
The elegance of Minkowski geometry emerges from the fact that all of it, and plus a few other things that we will see in a bit, comes from just 3 simple assumptions. In this spacetime diagram there are 3 kinds of regions. The existence of these 3 kinds of regions is a direct and exclusive consequence of the following assumptions:

1. The speed of light (c) is finite ;
2. The speed of light (c) is constant ;
3. The speed of light (c) is the maximum limit speed in the universe ($v \leq c$, $\forall v$) .

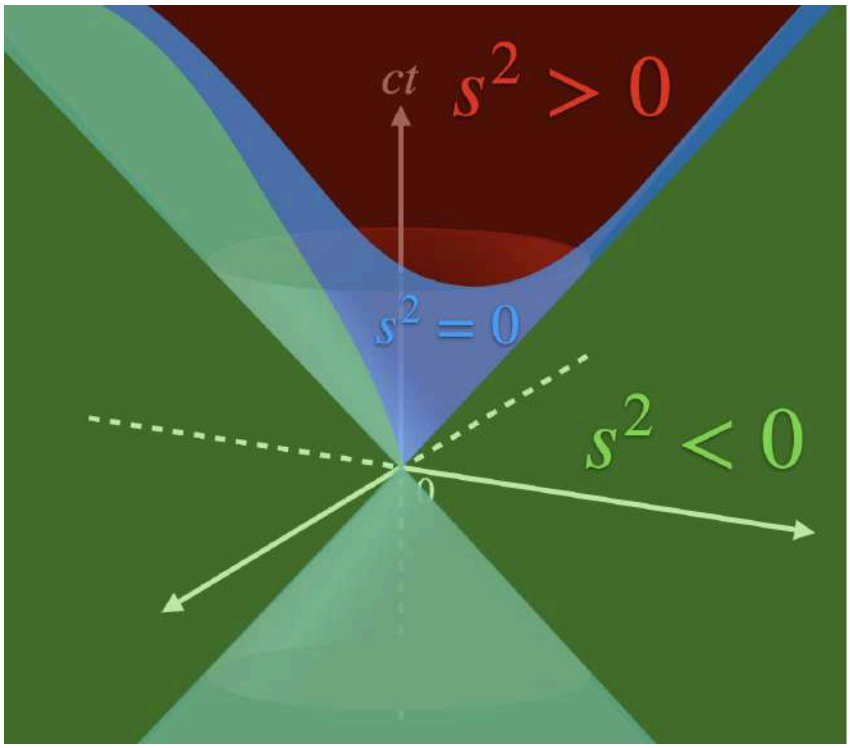
These 3 regions are:

1. $s^2 > 0$;
2. $s^2 = 0$;
3. $s^2 < 0$.

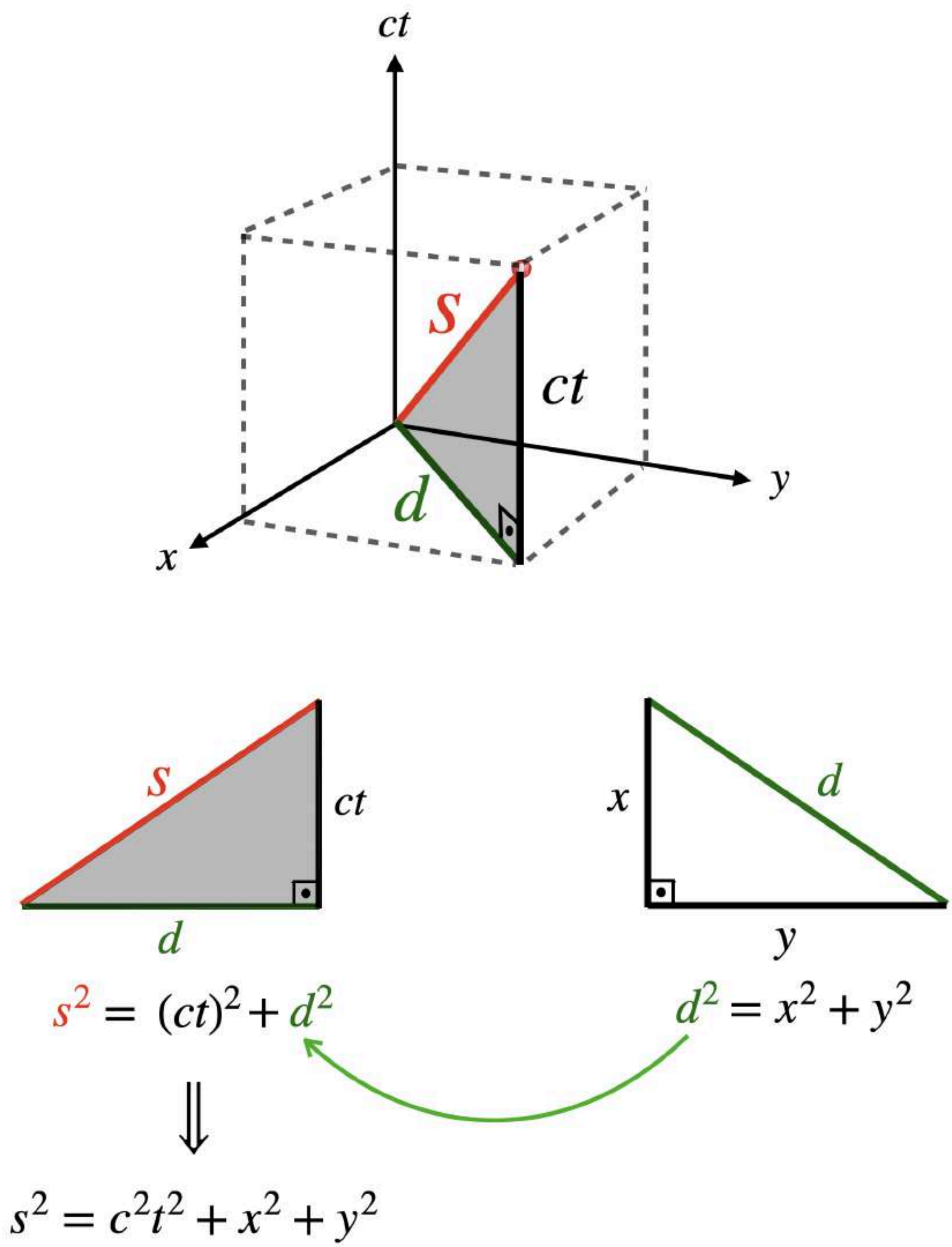
The letter s here represents “spacetime distance” (not spatial distance d).



These are 3 regions are separated by the cone-shaped figure we talked about:



s^2 can be measured using the Pythagorean theorem:

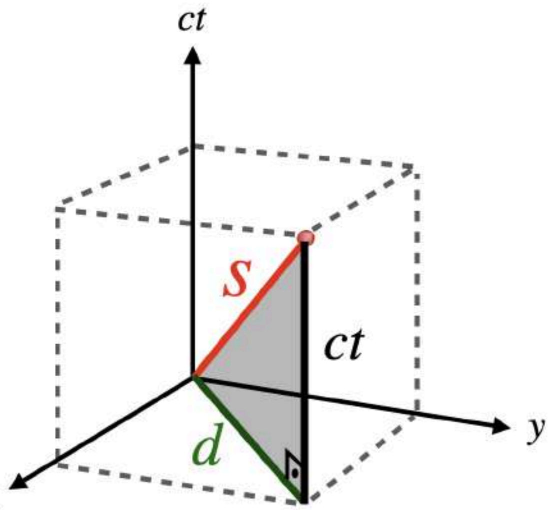


Of course, here we are considering space to be 2-dimensional, but the physically correct assumption would involve a x, y, z structure, plus the ct axis.

It is reasonable that the distance squared from 2 points is *positive* $\left(s^2 > 0\right)$ in spacetime, which means that traveling at any speed $v < c$ I can get to the desired point.

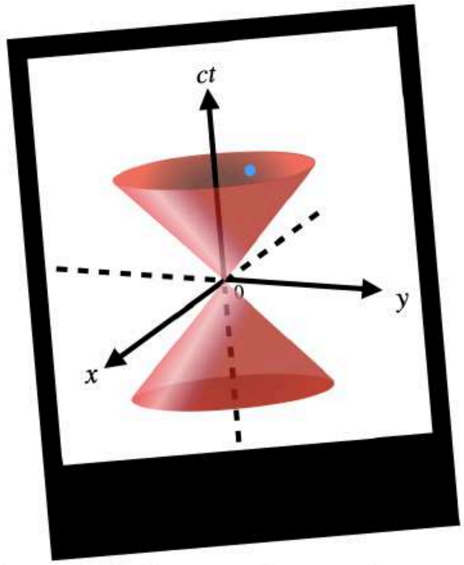
$$s^2 = \underbrace{c^2t^2}_{>0} + \underbrace{x^2}_{>0} + \underbrace{y^2}_{>0} > 0$$

$$s^2 > 0$$



It is not reasonable though that the distance squared is *zero* $\left(s^2 = 0\right)$ in spacetime, when using Euclidean geometry, other than the case in which time is frozen, like a picture.

$$s^2 = c^2t^2 + x^2 + y^2 + z^2 = 0 \Leftrightarrow t = x = y = z = 0$$

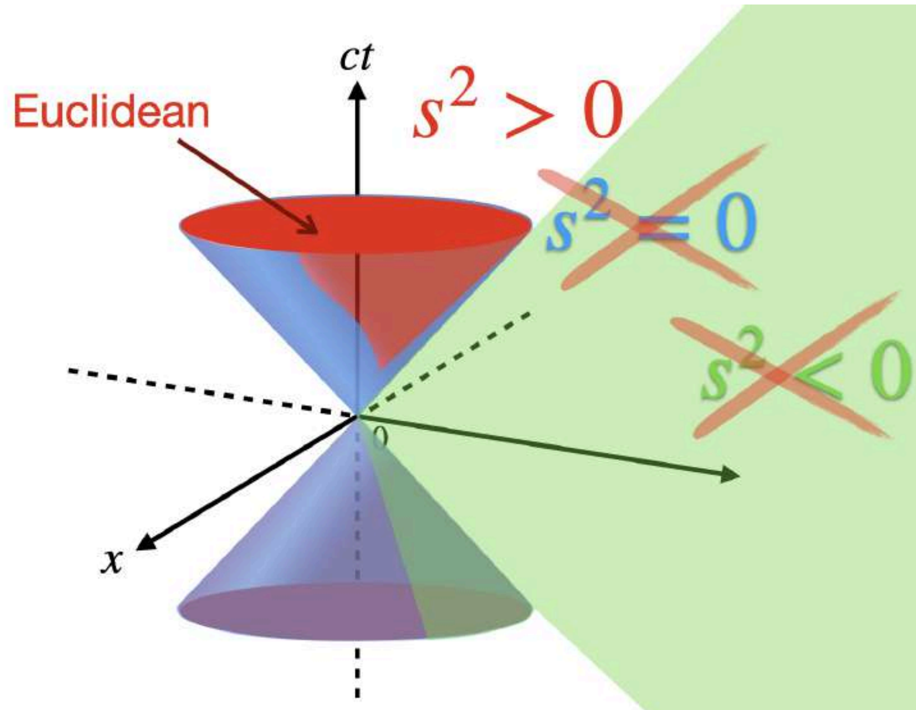


$$s^2 = \underbrace{c^2t^2}_{>0} + \underbrace{x^2}_{>0} + \underbrace{y^2}_{>0} = 0 \quad \lightning$$

And, when using Euclidean geometry, we never have that a distance squared in spacetime is *negative* $\left(s^2 < 0\right)$:

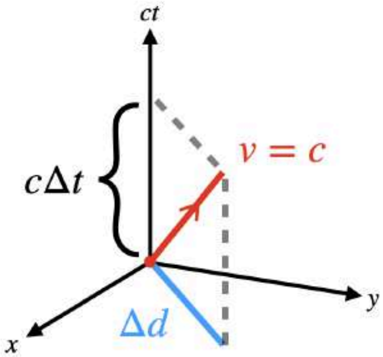
$$s^2 = \underbrace{c^2t^2}_{>0} + \underbrace{x^2}_{>0} + \underbrace{y^2}_{>0} < 0 \quad \lightning$$

So, Euclidean geometry does not accommodate regions $s^2 = 0$ and $s^2 < 0$ of the cone.



The question is: should we change from Euclidean geometry to an alternative geometry? Or should we abandon this understanding of 3 regions of spacetime altogether? Maybe only $s^2 > 0$ is possible after all, and the regions $s^2 = 0$ and $s^2 < 0$ are just nonsense?

Well, again, imagine a point moving at the speed of light. The amount of time passed and the variation of spatial position are clearly not zero: $\Delta t \neq 0$ and $\Delta d \neq 0$. Indeed, if $v = c$ we have the following:



$$\Delta t \neq 0$$

$$\Delta d \neq 0$$

$$v = c \; : \; c = \frac{\Delta d}{\Delta t} \implies \boxed{\Delta d = c\Delta t}$$

Using Euclidean geometry and this expression ($\Delta d = c\,\Delta t$) , we measure the spacetime distance s :

$$\boxed{\Delta d = c\Delta t}$$

$$s^2 = c^2\,(\Delta t)^2 + (\Delta d)^2 \implies$$

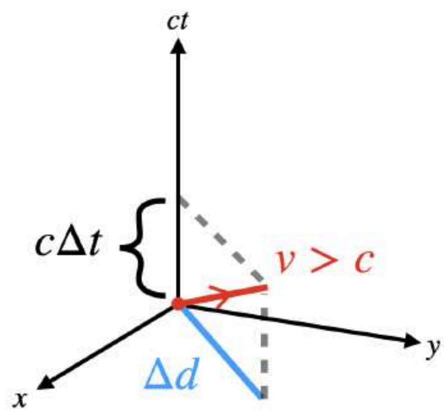
$$\implies s^2 = c^2(\Delta t)^2 + c^2(\Delta t)^2 \implies$$

$$\implies \boxed{s^2 = 2c^2(\Delta t)^2} \neq 0$$

$$\neq 0$$

$$\therefore \; \textbf{Euclidean geometry} \implies \nexists \; s^2 = 0$$

Similarly, in order to reach regions in $s^2 < 0$, we need a velocity $v > c$:



$$v > c : c < \frac{\Delta d}{\Delta t} \implies \boxed{\Delta d > c\Delta t}$$

Using Euclidean geometry and this expression ($\Delta d > c\,\Delta t$) , we can measure the spacetime distance s :

$$\Delta d > c\Delta t$$

$$s^2 = c^2(\Delta t)^2 + (\Delta d)^2 \implies$$

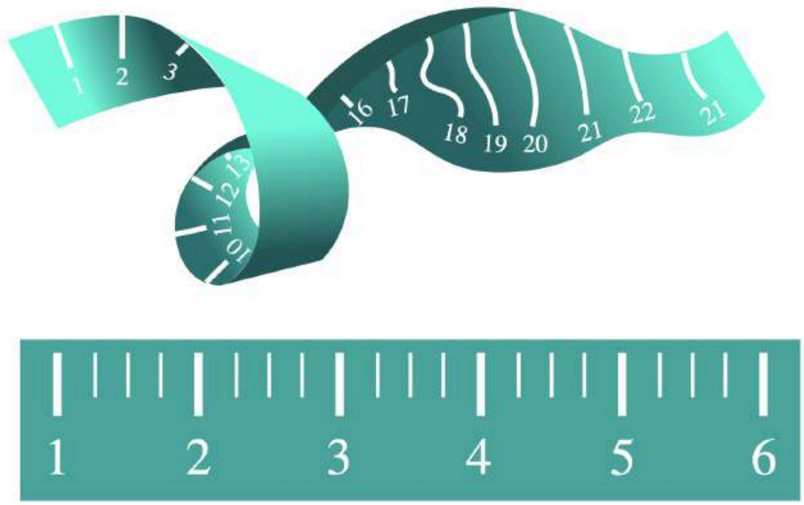
$$\implies s^2 > c^2(\Delta t)^2 + c^2(\Delta t)^2 \implies$$

$$\implies s^2 > 2c^2(\Delta t)^2 \implies s^2 > 0 \implies \neg s^2 < 0$$

$$\therefore \textbf{ Euclidean geometry} \implies \nexists\, s^2 < 0$$

In conclusion, Euclidean geometry allows us to describe points in spacetime only inside of the cone $\left(s^2 > 0\right)$. It does not accommodate for the description of $s^2 = 0$ or $s^2 < 0$, which must be included if we want a complete picture of the spacetime diagram, covering all possible regions.

We need another *metric* (a way of measuring spacetime distances s that will cover all 3 regions). A metric can be thought of as a “special ruler” that will be used to measure spacetime distances. This is not like your regular Euclidean ruler.



Let’s derive this metric.

(1) $s^2 > 0$:

$$v < c \implies \frac{\Delta d}{\Delta t} < c \implies \boxed{\Delta d < c\Delta t} \implies$$

$$\implies \underbrace{x^2 + y^2 + z^2}_{\substack{\textbf{Euclidean metric} \\ \text{(spatial distance } \Delta d)}} < c^2t^2 \implies \underbrace{c^2t^2 - x^2 - y^2 - z^2}_{\text{spacetime distance } = s^2} > 0$$

$$\therefore \boxed{s^2 := c^2t^2 - x^2 - y^2 - z^2}$$

It's ok to use Euclidean geometry for calculating *spatial* distances (there is no contradiction here). But when it comes to "spacetime distance", it turns out that $s^2 = c^2t^2 - x^2 - y^2 - z^2$ will be our new s^2 (or *metric*) rather than the old unsatisfactory Euclidean metric $\left(s^2 = c^2t^2 + x^2 + y^2 + z^2 \right)$.

Let's see how this new metric (which is non-Euclidean) behaves for the other 2 regions:

(2) $s^2 = 0$:

$$\boxed{v = c} \implies \frac{\Delta d}{\Delta t} = c \implies \boxed{\Delta d = c\Delta t} \implies$$

$$\implies x^2 + y^2 + z^2 = c^2t^2 \implies c^2t^2 - x^2 - y^2 - z^2 = 0$$

$$\boxed{\implies s^2 = 0}$$

(3) $s^2 < 0$:

$$\boxed{v > c} \implies \frac{\Delta d}{\Delta t} > c \implies \boxed{\Delta d > c\Delta t} \implies$$

$$\implies x^2 + y^2 + z^2 > c^2t^2 \implies c^2t^2 - x^2 - y^2 - z^2 < 0$$

$$\boxed{\implies s^2 < 0}$$

So we found our metric, the *Minkowski metric*:

$$s^2 = c^2t^2 - x^2 - y^2 - z^2$$

Minkowski metric

The Minkowski metric can also be seen as a matrix transformation.

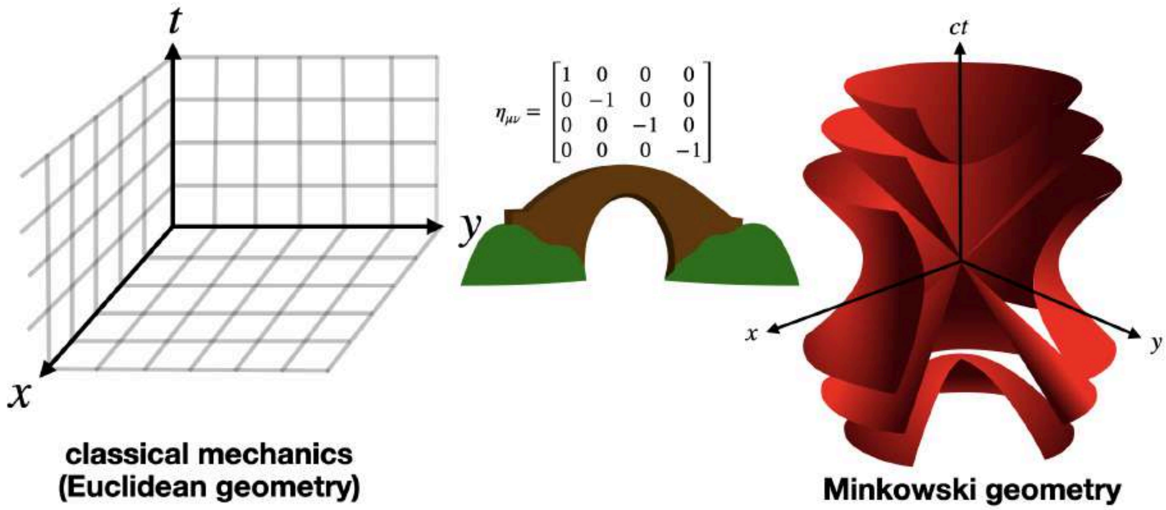
$$s^2 = \begin{bmatrix} ct & x & y & z \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

$$s^2 = c^2t^2 - x^2 - y^2 - z^2$$

$$\therefore \quad \eta_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{matrix} \leftarrow \mu = 0 \\ \leftarrow \mu = 1 \\ \leftarrow \mu = 2 \\ \leftarrow \mu = 3 \end{matrix}$$

$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \nu = 0 & \nu = 1 & \nu = 2 & \nu = 3 \end{matrix}$$

This matrix transformation bridges *classical mechanics* (*Euclidean Geometry*) and *special relativity* (*Minkowski geometry*) by restructuring spacetime.



Let's use *Linear Algebra* in order to understand this new geometry. One of the first things that comes to mind when we find a new matrix transformation is to compute its eigenvectors and eigenvalues.

Check this video (or simply the PDF link below) in order to learn the core of Linear Algebra:

PDF - Core Linear Algebra.pdf

Or this one, about the core of eigenvalues and eigenvectors, in order to understand why and how we do it:

The Core of Eigenvalues & Eigenvectors

PDF - The Core of Eigenvalues & Eigenvectors_compressed.pdf

We want to find the eigenvalues and eigenvectors of the following matrix:

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

In order to accomplish that, we need to solve the following equation:

$$\boxed{\det\left(\eta - \lambda I\right) = 0}$$

$$\det\left(\eta - \lambda I\right) = 0 \implies$$

$$\implies \det \begin{bmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & -1-\lambda & 0 & 0 \\ 0 & 0 & -1-\lambda & 0 \\ 0 & 0 & 0 & -1-\lambda \end{bmatrix} = 0 \implies$$

Eigenvalues

$$\implies (1-\lambda)(-1-\lambda)^3 = 0 \implies \begin{cases} \lambda_1 = 1 \\ \lambda_2 = \lambda_3 = \lambda_4 = -1 \end{cases}$$

For $\lambda_1 = 1$:

$$\eta\vec{u} = \lambda\vec{u} \implies$$

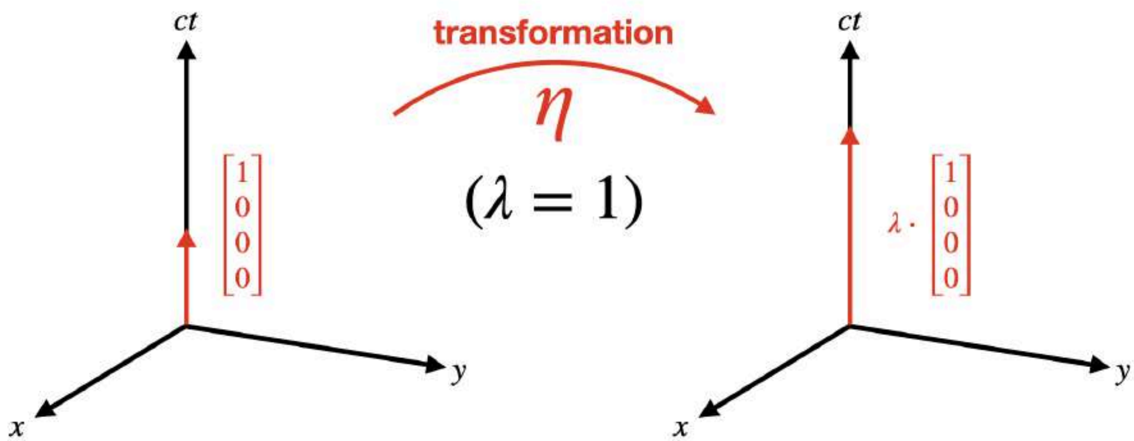
$$\implies \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} u_{ct} \\ u_x \\ u_y \\ u_z \end{bmatrix} = 1 \cdot \begin{bmatrix} u_{ct} \\ u_x \\ u_y \\ u_z \end{bmatrix} \implies$$

$$\implies \begin{cases} u_{ct} = u_{ct} \\ -u_x = u_x \\ -u_y = u_y \\ -u_z = u_z \end{cases} \implies u_x = u_y = u_z = 0 \implies$$

$$\implies \vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Eigenvalues

This eigenvector represents the *c-time direction*. Any vector along this direction is preserved $\left(eigenvalue\ \lambda_1 = 1\right)$ under transformations that respect this metric.



For $\lambda_2 = \lambda_3 = \lambda_4 = -1$: (threefold degenerate)

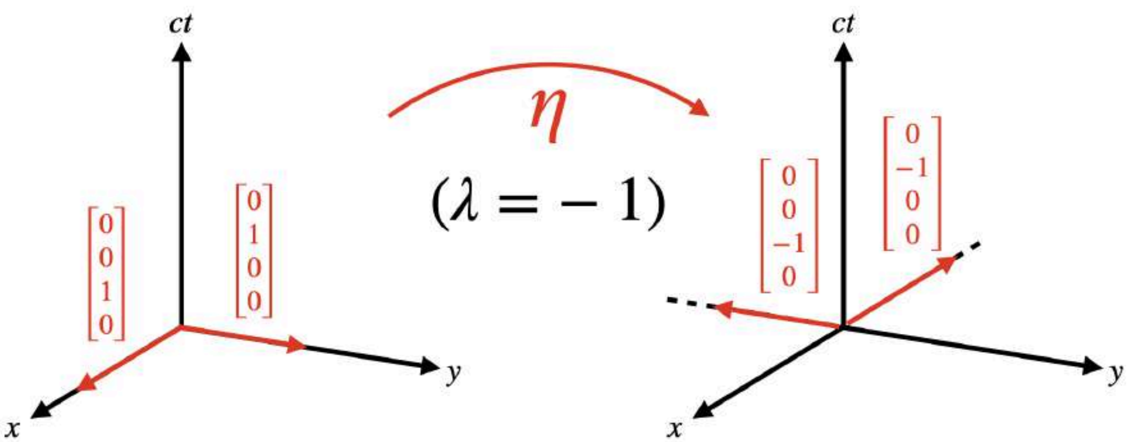
$$\eta \vec{u} = \lambda \vec{u} \implies$$

$$\implies \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} u_{ct} \\ u_x \\ u_y \\ u_z \end{bmatrix} = -1 \cdot \begin{bmatrix} u_{ct} \\ u_x \\ u_y \\ u_z \end{bmatrix} \implies$$

$$\begin{cases} u_{ct} = -u_{ct} \\ -u_x = -u_x \\ -u_y = -u_y \\ -u_z = -u_z \end{cases} \Longrightarrow u_{ct} = 0 \Longrightarrow \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \vec{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

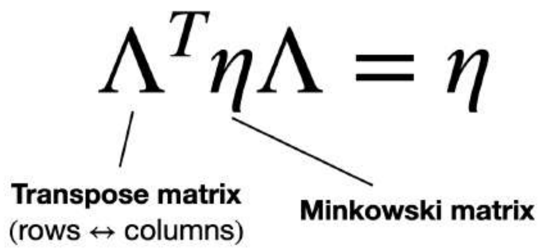
Eigenvectors

These vectors represent the *spatial directions*. Any vector along these directions flips sign $(eigenvalues\ \lambda = -1)$ under this kind transformation.

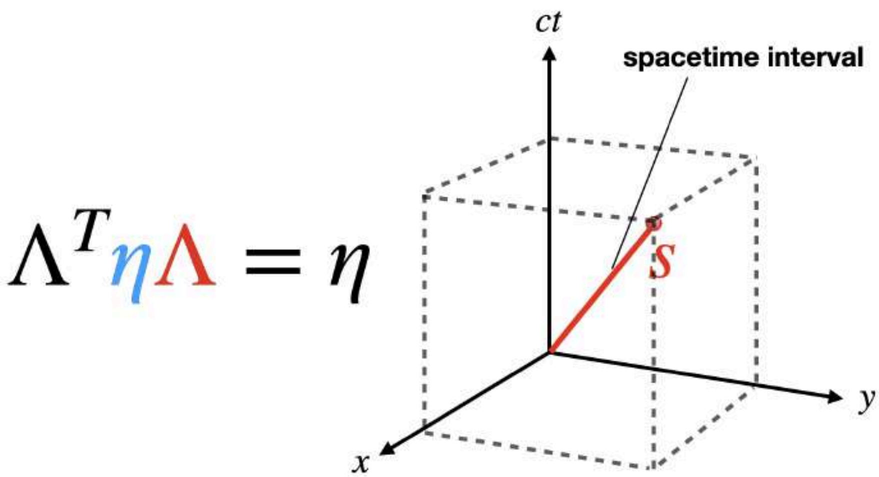


The flipping in sign means that spatial directions reverse under certain transformations, like a *parity transformation* (P). A parity transformation $(x \to -x\,;\ y \to -y\,;\ z \to -z)$ mirrors the spatial coordinates, flipping left and right, up and down, forward and backward. This is especially important in physics because some fundamental laws (like *electromagnetism*) are *symmetric under parity*, while others (like *weak interactions*) *violate parity symmetry*. So, the eigenvalue $\lambda = -1$ for spatial directions shows that these vectors invert under reflections, which is directly related to symmetry operations in physics.

The Minkowski matrix transformation is interesting, but again, it only serves the purpose of opening the door to Minkowski geometry (or special relativity in physics). What we are actually looking for though is a transformation that preserves the most fundamental property of this new geometry, which is the *spacetime interval* s, regardless of the speed v of an observer. In other words, we want to find a matrix that satisfies this equation:



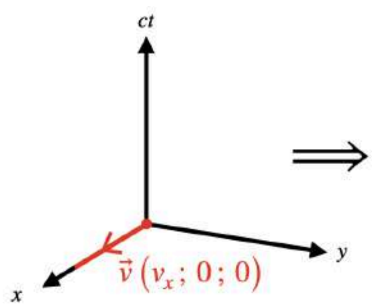
Such a matrix Λ would ensure that the Minkowski metric η remains unchanged, guaranteeing that spacetime intervals are preserved under this transformation. But what kind of transformation could achieve this? We can do some math to find out.



This mysterious matrix acts on all the vectors $(ct\ ;\ x\ ;\ y\ ;\ z)$ in spacetime:

$$\Lambda \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix}$$

For simplicity, let's assume that the point moves with a velocity $\vec{v}$ only in the x-direction $\left(\vec{v} = \left(v_x\ ;\ 0\ ;\ 0\right)\right)$. This implies that Λ must be a linear transformation that mixes ct and x components while leaving y and z unchanged:



$$\implies \Lambda = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} ; a, b, c, d \in \mathbb{R}$$

$$\Lambda^T \eta \Lambda = \eta \implies$$

$$\implies \begin{bmatrix} a & c & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \implies$$

$$\Longrightarrow \begin{bmatrix} a & -c & 0 & 0 \\ b & -d & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \Longrightarrow$$

$$\Longrightarrow \begin{bmatrix} a^2 - c^2 & ab - cd & 0 & 0 \\ ab - cd & b^2 - d^2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \Longrightarrow$$

$$\Longrightarrow \begin{cases} \boxed{a^2 - c^2 = 1} \ (I) \\ \boxed{ab - cd = 0} \ (II) \\ b^2 - d^2 = -1 \ \Longrightarrow \ \boxed{d^2 - b^2 = 1} \ (III) \end{cases}$$

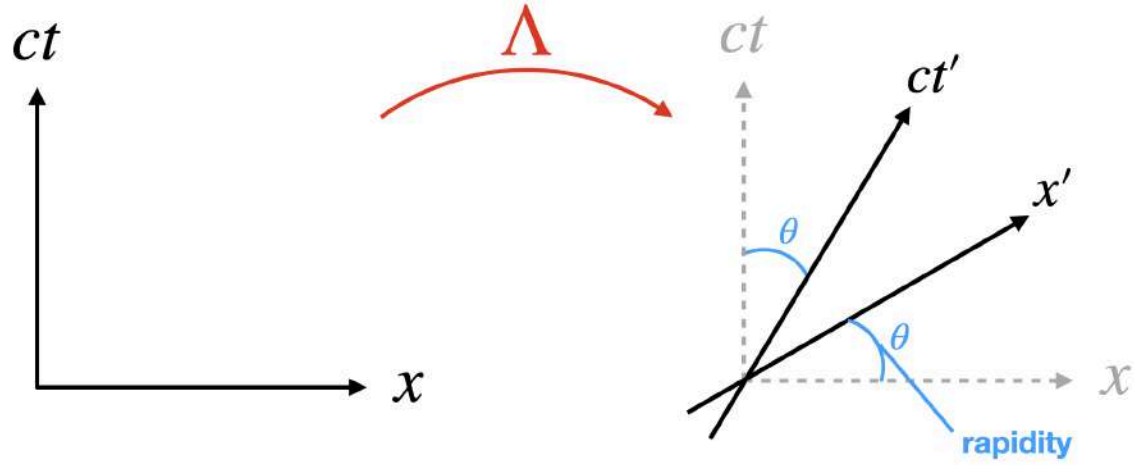
$$(I) \ \wedge \ (II) \ \Longrightarrow \ a^2 - c^2 = d^2 - b^2 \ \Longrightarrow \ \begin{cases} a = d \\ b = c \end{cases}$$

$(I) : \boxed{a^2 - c^2 = 1}$ is a hyperbolic identity:

$$\begin{cases} a := \cosh\theta = d \\ c := \sinh\theta = b \end{cases}$$

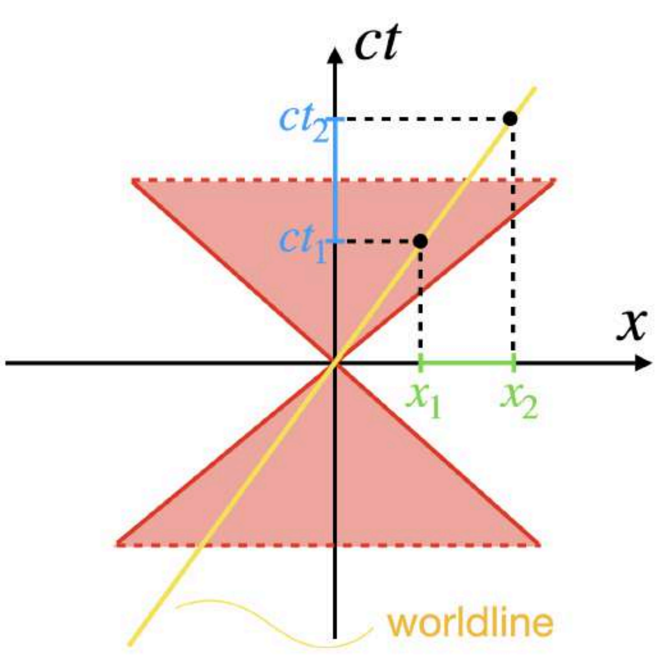
Here, θ is the angle of inclination of the tilt of the new c-time and space axes after the transformation:

$$\begin{cases} a := \cosh\theta = d \\ c := \sinh\theta = b \end{cases}$$



(We will clearly see in a bit why, and by how much, this tilt happens. Btw, this angle θ is usually called *rapidity*.)

A straight *worldline* (i.e., a path in the spacetime diagram) represents constant velocity $\vec{v}$:



$$\frac{\Delta x}{c\Delta t} = \frac{v}{c}$$

The slope of this worldline is $\frac{\Delta x}{c\Delta t} = \frac{v}{c}$, which measures how much space it covers per unit of c-time. Since Minkowski spacetime follows hyperbolic geometry, the natural "angle measure" for this inclination is given by the hyperbolic tangent $\left(tanh\theta = \frac{v}{c}\right)$.

$$\boxed{\tanh\theta = \frac{v}{c}} \implies \frac{\sinh\theta}{\cosh\theta} = \frac{v}{c} \implies \frac{\sqrt{\cosh^2\theta - 1}}{\cosh^2\theta} = \frac{v}{c} \implies$$

$$\implies \frac{\cosh^2\theta - 1}{\cosh^2\theta} = \frac{v^2}{c^2} \implies c^2\cosh^2\theta - c^2 = v^2\cosh^2\theta \implies$$

$$\implies \left(c^2 - v^2\right)\cosh^2\theta = c^2 \implies \cosh^2\theta = \frac{c^2}{c^2 - v^2} \implies$$

$$\implies \boxed{\cosh\theta = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}}} \equiv a \implies$$

$$\implies \sinh\theta = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}} - 1} \implies \sinh\theta = \sqrt{\frac{1 - 1 + v^2/c^2}{1 - \frac{v^2}{c^2}}} \implies$$

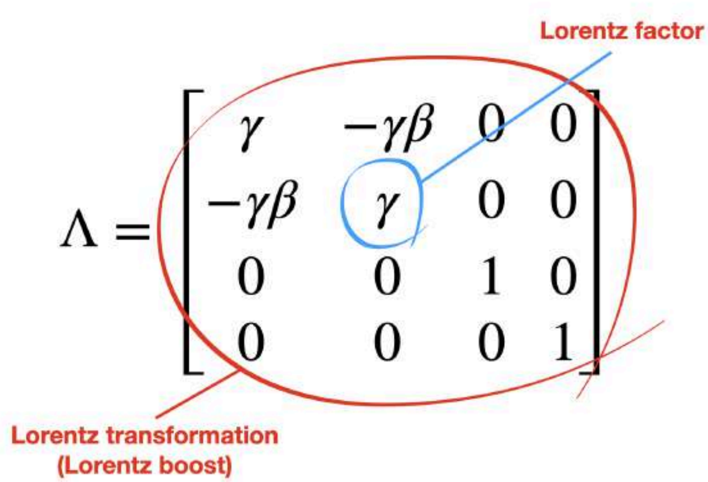
$$\implies \boxed{\sinh\theta = \frac{v/c}{\sqrt{1 - v^2/c^2}}} \equiv c$$

$$\therefore \begin{cases} a = d = \dfrac{1}{\sqrt{1 - v^2/c^2}} \\ b = c = \dfrac{v/c}{\sqrt{1 - v^2/c^2}} \end{cases}$$

Therefore, the matrix transformation that we were looking for is:

$$\Lambda = \begin{bmatrix} \frac{1}{\sqrt{1-\frac{v2}{c^2}}} & \frac{\frac{v}{c}}{\sqrt{1-\frac{v2}{c^2}}} & 0 & 0 \\ \frac{\frac{v}{c}}{\sqrt{1-\frac{v2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Or, more compactly, we can call $\beta := \frac{v}{c}$ and $\gamma := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$:



This is called the *Lorentz transformation matrix* or the *Lorentz boost*, and γ is referred to as the *Lorentz factor*. This matrix is important because it describes how spacetime coordinates transform between observers moving at a constant velocity relative to each other. It preserves the *spacetime interval* $s^2 = c^2t^2 - x^2 - y^2 - z^2$, it ensures that the speed of light remains constant, and it keeps the fundamental structure of Minkowski spacetime unchanged.

Now, this is not the most general version of a Lorentz boost, because we assumed motion only in the x-direction $\left(\vec{v} = \left(v_x ;\ 0 ;\ 0\right)\right)$. If we allow the boost to occur in any random direction $\vec{v} = \left(v_x ;\ v_y ;\ v_z\right)$, then the generalized Lorentz transformation matrix looks like this:

$$\Lambda(\vec{v}) = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1+(\gamma-1)\,\hat{v}_x^2 & (\gamma-1)\,\hat{v}_x\hat{v}_y & (\gamma-1)\,\hat{v}_x\hat{v}_z \\ -\gamma\beta_y & (\gamma-1)\,\hat{v}_y\hat{v}_x & 1+(\gamma-1)\,\hat{v}_y^2 & (\gamma-1)\,\hat{v}_y\hat{v}_z \\ -\gamma\beta_z & (\gamma-1)\,\hat{v}_z\hat{v}_x & (\gamma-1)\,\hat{v}_z\hat{v}_y & 1+(\gamma-1)\,\hat{v}_z^2 \end{bmatrix}$$

, where $\hat{v} = \left(\widehat{v_x} ;\ \widehat{v_y} ;\ \widehat{v_z}\right) = \left(\frac{v_x}{v} ;\ \frac{v_y}{v} ;\ \frac{v_z}{v}\right)$ is the *unit direction* of the velocity, such that $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$ and $\vec{\beta} := \frac{\vec{v}}{c} = \left(\beta_x ;\ \beta_y ;\ \beta_z\right)$.

Of course, this generalized matrix is obtained by combining the boosts in the 3 directions (but not in a linear way):

$$\Lambda_x = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \Lambda_y = \begin{bmatrix} \gamma & 0 & -\gamma\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\gamma\beta & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \Lambda_z = \begin{bmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{bmatrix}$$

From the standard Lorentz transformation for a single axis boost, we know that the time coordinate mixes with all the space coordinates.

$$\Lambda = \begin{bmatrix} \Lambda_{00} & \Lambda_{01} & \Lambda_{02} & \Lambda_{03} \\ \Lambda_{10} & \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{20} & \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{30} & \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{bmatrix}$$

$\underline{\Lambda_{00} = \gamma \ \wedge \ \Lambda_{0i} = \Lambda_{i0} = -\gamma\beta_i}$:

$$\Lambda = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & ? & ? & ? \\ -\gamma\beta_y & ? & ? & ? \\ -\gamma\beta_z & ? & ? & ? \end{bmatrix}$$

spatial-spatial components

We need to derive the transformation of the spatial coordinates x, y, z under a Lorentz boost in an arbitrary direction. For this *spatial-spatial matrix* Λ_{ij} , the expression must be of the following form:

$$\Lambda_{ij} := A \cdot \delta_{ij} + B \cdot v_i v_j \qquad (\star)$$

, where A and B are coefficients that must be found.

δ_{ij} is the *Kronecker delta* (identity matrix) that ensures that when there is no boost $\left(v_i = 0\right)$ in a particular direction, the component is 1.

$$\delta_{ij} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$v_i v_j$ accounts for the effect of velocity direction. I.e., this term introduces the directional dependence of the boost. Each spatial direction (x, y, z) must be affected differently based on how much of the velocity is in that direction. The only way to distribute this effect among the coordinates is through the velocity components $v_i v_j$. A single velocity component is not enough. If we just had v_i , it would define a *vector* , not a *second-rank tensor* (i.e., matrix). We need a matrix that mixes coordinates, and the simplest second-rank tensor that can be built from $\vec{v}$ alone is $v_i v_j$.

Notice that ($\star$) must reduce to the standard Lorentz boost. For a boost along the x-axis, we already know that:

$$\Lambda_{xx} = \gamma \;\; \wedge \;\; \Lambda_{yy} = \Lambda_{zz} = 1$$

If we try any other term than $v_i v_j$, we do not recover this result.

Ok, now let’s move on and find the appropriate coefficients A and B for ($\star$):

$$\Lambda_{xx} = \gamma \implies A \cdot \underbrace{\delta_{xx}}_{1} + B \cdot v_x v_x = \gamma \implies \boxed{\gamma = A + B\, v_x^2} \quad (I)$$

$$\Lambda_{yy} = \Lambda_{zz} = 1 \implies A \cdot \underbrace{\delta_{yy}}_{1} + B \cdot v_y v_y = 1 \implies \boxed{1 = A + B\, v_y^2} \quad (II)$$

$$(I) + (II) \implies \gamma + 1 = 2A + B\underbrace{(v_x^2 + v_y^2)}_{v^2 = v_x^2 + v_y^2 + v_z^2,\ v_z^2 = 0} \implies \boxed{\gamma + 1 = 2A + B\, v^2} \quad (\star)$$

When $v = 0$:

$\gamma + 1 = 2A \implies A = \dfrac{\gamma + 1}{2}$, but if $v = 0$, then $\gamma = \dfrac{1}{\sqrt{1 - \frac{0}{c}}} = 1 \implies$

$\implies \boxed{A = 1}$

$$(\star) \overset{(v \neq 0)}{\implies} B = \frac{\gamma + 1 - 2}{v^2} \implies \boxed{B = \frac{\gamma - 1}{v^2}} \qquad \therefore\ \Lambda_{ij} := \delta_{ij} + (\gamma - 1)\underbrace{\frac{v_i}{v}}_{\hat{v}_i} \cdot \underbrace{\frac{v_j}{v}}_{\hat{v}_j}$$

$$\Lambda = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1 + (\gamma - 1)\,\hat{v}_x^2 & ? & ? \\ -\gamma\beta_y & ? & 1 + (\gamma - 1)\,\hat{v}_y^2 & ? \\ -\gamma\beta_z & ? & ? & 1 + (\gamma - 1)\,\hat{v}_z^2 \end{bmatrix}$$

$$\Lambda_{xy} := \underbrace{\delta_{xy}}_{0} + (\gamma - 1)\underbrace{\frac{v_x}{v}}_{\hat{v}_x} \cdot \underbrace{\frac{v_y}{v}}_{\hat{v}_y} \qquad \Lambda_{yz} := \underbrace{\delta_{yz}}_{0} + (\gamma - 1)\underbrace{\frac{v_y}{v}}_{\hat{v}_y} \cdot \underbrace{\frac{v_z}{v}}_{\hat{v}_z}$$

$$\therefore \quad \Lambda = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1 + (\gamma - 1)\,\hat{v}_x^2 & (\gamma - 1)\,\hat{v}_x\hat{v}_y & (\gamma - 1)\,\hat{v}_x\hat{v}_z \\ -\gamma\beta_y & (\gamma - 1)\,\hat{v}_y\hat{v}_x & 1 + (\gamma - 1)\,\hat{v}_y^2 & (\gamma - 1)\,\hat{v}_y\hat{v}_z \\ -\gamma\beta_z & (\gamma - 1)\,\hat{v}_z\hat{v}_x & (\gamma - 1)\,\hat{v}_z\hat{v}_y & 1 + (\gamma - 1)\,\hat{v}_z^2 \end{bmatrix}$$

And that's the general form of the Lorentz transformation matrix in Minkowski geometry, for a moving reference frame with velocity $\vec{v} = \left(v_x\,;\ v_y\,;\ v_z\right)$.

The next logical step would be to calculate the eigenvalues and eigenvectors of this matrix, since they give us information about what directions are preserved in the transformation, and by how much its eigenvectors are stretched or contracted. In order to do so, we will first calculate them for the standard Lorentz boost in the x-direction (ignoring even the existence of the y- and z-directions), and then we will extend the result for the full 4-dimensional Minkowski spacetime.

$$\Lambda_x = \begin{bmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{bmatrix}$$

$$\boxed{\det\left(\Lambda_x - \lambda I\right) = 0} \implies \det\begin{bmatrix} \gamma - \lambda & -\gamma\beta \\ -\gamma\beta & \gamma - \lambda \end{bmatrix} = 0 \implies$$

$$\implies \left(\gamma - \lambda\right)^2 - \left(-\gamma\beta\right)^2 = 0 \implies \left(\gamma - \lambda\right)^2 = \gamma^2\beta^2 \implies$$

$$\implies \sqrt{\left(\gamma - \lambda\right)^2} = \pm\sqrt{\gamma^2\beta^2} \implies \gamma - \lambda = \pm\,\gamma\beta \implies$$

$$\implies \boxed{\lambda = \gamma \pm \gamma\beta} \quad \text{(two eigenvalues)}$$

<u>For $\lambda_1 = \gamma\left(1 + \beta\right)$:</u>

$$\boxed{\Lambda_x\,\vec{u}_1 = \lambda_1\,\vec{u}_1} \implies \begin{bmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{bmatrix} \cdot \begin{bmatrix} u_{ct} \\ u_x \end{bmatrix} = \gamma\left(1 + \beta\right)\begin{bmatrix} u_{ct} \\ u_x \end{bmatrix} \implies$$

$$\implies \begin{cases} \bigstar\; \gamma\,u_{ct} - \gamma\beta\,u_x = \gamma\left(1 + \beta\right)\,u_{ct} \\ \blacklozenge\; -\gamma\beta\,u_{ct} + \gamma\,u_x = \gamma\left(1 + \beta\right)\,u_x \end{cases}$$

The only way to, simultaneously, satisfy both of these equation is if $-u_{ct} = u_x$, because then:

$$\bigstar \implies \gamma\, u_{ct} - \gamma\beta\, u_{ct} = \gamma\left(1+\beta\right)\, u_{ct} \implies$$

$$\implies \left(\gamma + \gamma\beta\right) u_{ct} = \gamma\left(1+\beta\right)\, u_{ct}$$ ✅

$$\blacklozenge \implies \gamma\beta\, u_x + \gamma\, u_x = \gamma\left(1+\beta\right)\, u_x \implies$$

$$\implies \left(\gamma\beta + \gamma\right) u_x = \gamma\left(1+\beta\right)\, u_x$$ ✅

Eingenvector: $\vec{u}_1 = \begin{bmatrix} u_{ct} \\ u_x \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

For $\lambda_2 = \gamma\left(1-\beta\right)$:

$$\boxed{\Lambda_x\, \vec{u}_2 = \lambda_2\, \vec{u}_2} \implies \begin{bmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{bmatrix} \cdot \begin{bmatrix} u_{ct} \\ u_x \end{bmatrix} = \gamma\left(1-\beta\right) \begin{bmatrix} u_{ct} \\ u_x \end{bmatrix} \implies$$

$$\implies \begin{cases} ⬢\ \gamma\, u_{ct} - \gamma\beta\, u_x = \gamma\left(1-\beta\right)\, u_{ct} \\ ◆\ -\gamma\beta\, u_{ct} + \gamma\, u_x = \gamma\left(1-\beta\right)\, u_x \end{cases}$$

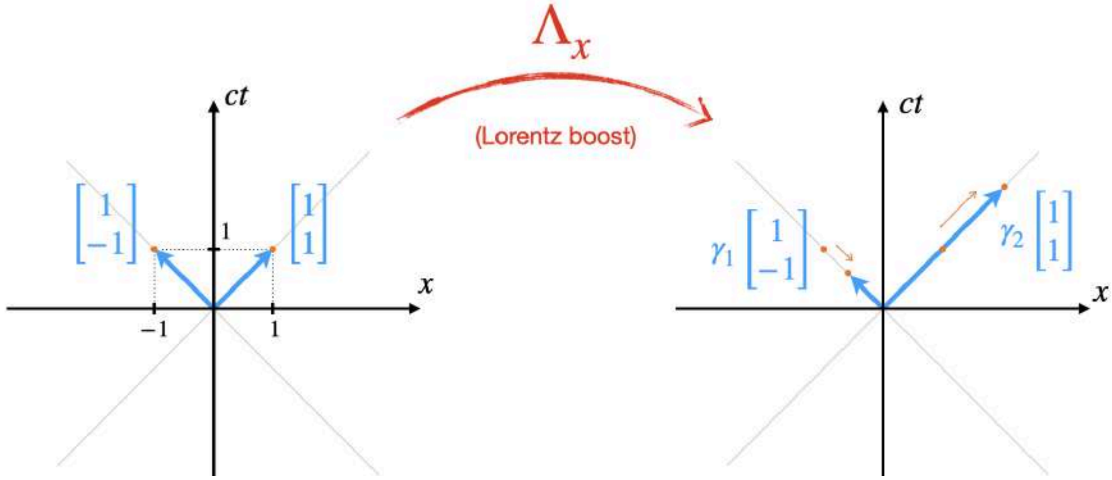
The only way to, simultaneously, satisfy both of these equation is if $u_{ct} = u_x$, because then:

$$\implies \quad \gamma\, u_{ct} - \gamma\beta\, u_{ct} = \gamma\left(1-\beta\right)\, u_{ct}$$

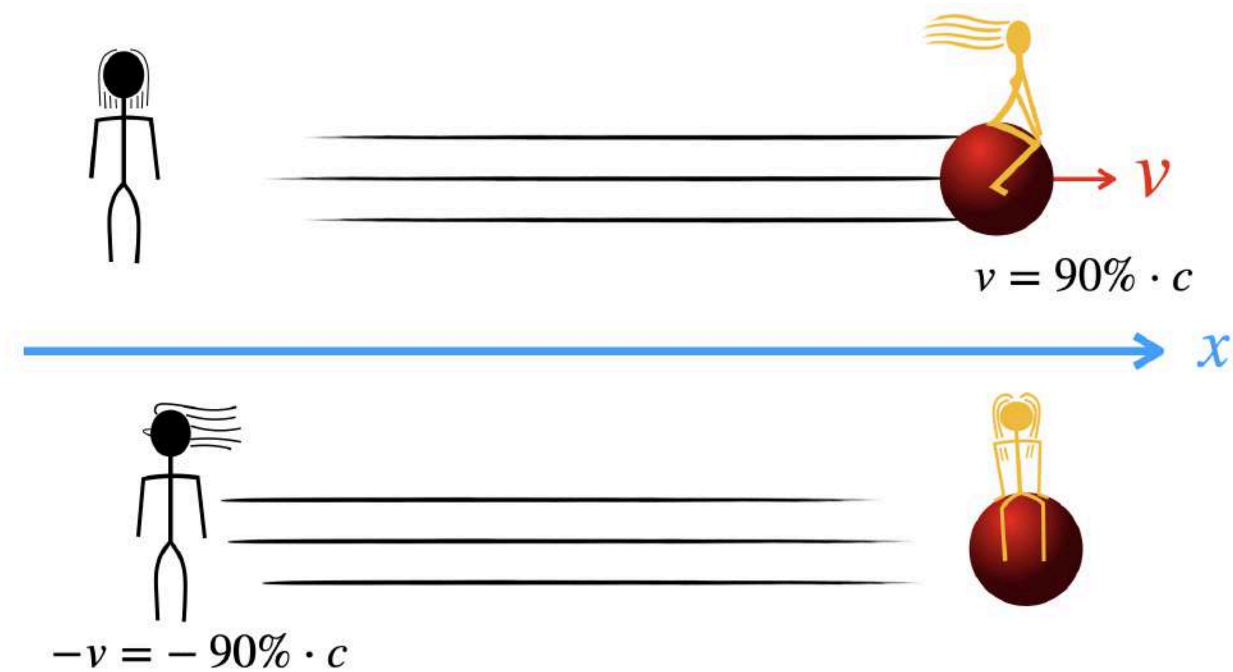
$$\implies \quad -\gamma\beta\, u_x + \gamma\, u_x = \gamma\left(1-\beta\right)\, u_x$$

$$\text{Eingenvector:} \quad \vec{u}_2 = \begin{bmatrix} u_{ct} \\ u_x \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Graphing these results we see that one of the vectors gets *stretched*, while the other gets *contracted*. All of it takes place along these 2 grey lines, which btw form the *light cone*.



As a more concrete example, imagine that this is a Lorentz boost such that the point particle that we are studying moves at 90% of the speed of light ($v = 0.9c$). Notice that the velocity of a stationary observer wrt the particle moving at 90% of the speed of light will be a *negative* number. In this situation, since they are moving *away* from each other with velocity v, we will consider the velocity (measured by the stationary observer) to be $-\ v$.



$$\boxed{\beta = \frac{v}{c}} = \frac{-0.9\cancel{c}}{\cancel{c}} = -\ 0.9$$

Lorentz factor: $\gamma = \dfrac{1}{\sqrt{1-\beta^2}} = \dfrac{1}{\sqrt{1-0.81}} = \dfrac{1}{\sqrt{0.19}} \approx 2.294$

Eigenvalues: $\lambda_1 = \gamma\left(1+\beta\right)$ and $\lambda_2 = \gamma\left(1-\beta\right)$

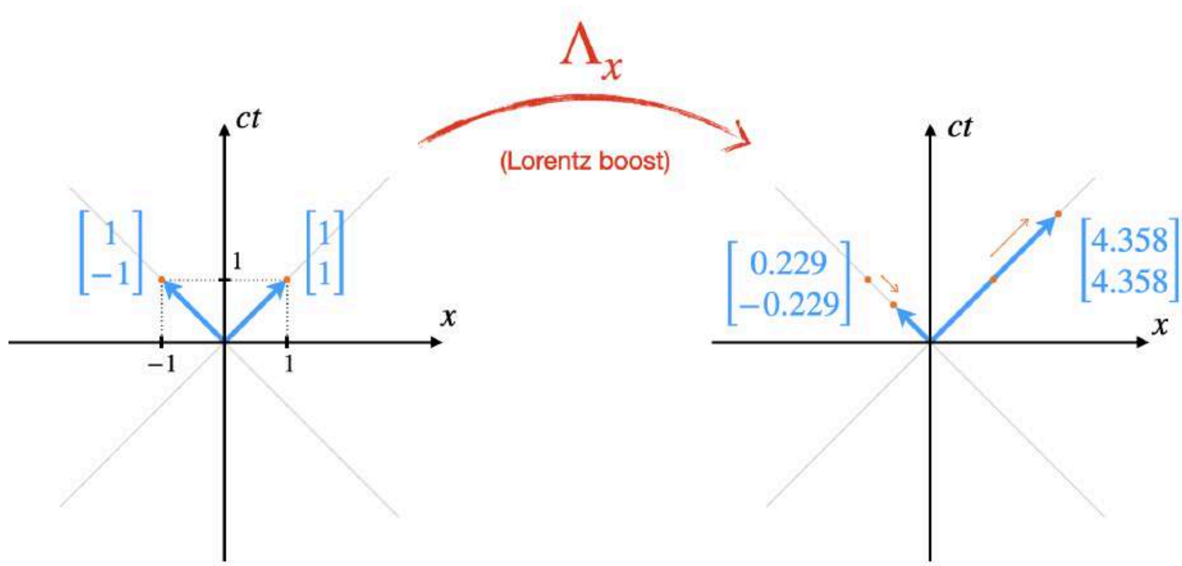
$$\lambda_1 = 2.294\left(1+(-0.9)\right) = 2.294 \cdot 0.1 \approx 0.229$$

$$\lambda_2 = 2.294\left(1-(-0.9)\right) = 2.294 \cdot 1.9 \approx 4.358$$

Eigenvectors: $\vec{u}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

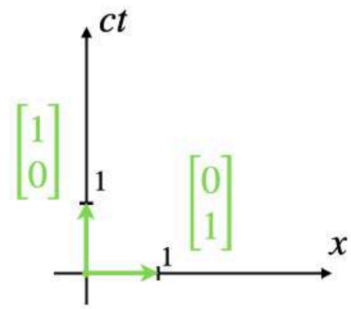
$$\lambda_1 \vec{u}_1 = 0.229 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0.229 \\ -0.229 \end{bmatrix}$$

$$\lambda_2 \vec{u}_2 = 4.358 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4.358 \\ 4.358 \end{bmatrix}$$



Interesting! We are slowly drawing the new diagram we get after the transformation. Let's see what happens with the ct- and x-axes:

ct - axis: $\vec{v}_{ct} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$



$$\boxed{\vec{v}'_{ct} = \Lambda_x \vec{v}_{ct}} \Longrightarrow \begin{bmatrix} ct' \\ x' \end{bmatrix} = \gamma \begin{bmatrix} 1 & -\beta \\ -\beta & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\vec{v}'_{ct} = \begin{bmatrix} ct' \\ x' \end{bmatrix} = \gamma \begin{bmatrix} 1 \\ -\beta \end{bmatrix}$$

x - axis: $\vec{v}_x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\boxed{\vec{v}'_x = \Lambda_x \vec{v}_x} \Longrightarrow \begin{bmatrix} ct' \\ x' \end{bmatrix} = \gamma \begin{bmatrix} 1 & -\beta \\ -\beta & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{v}'_x = \begin{bmatrix} ct' \\ x' \end{bmatrix} = \gamma \begin{bmatrix} -\beta \\ 1 \end{bmatrix}$$

The question is: are $\vec{v}'_{ct} = \gamma \begin{bmatrix} 1 \\ -\beta \end{bmatrix}$ and $\vec{v}'_x = \gamma \begin{bmatrix} -\beta \\ 1 \end{bmatrix}$ orthogonal to each other?

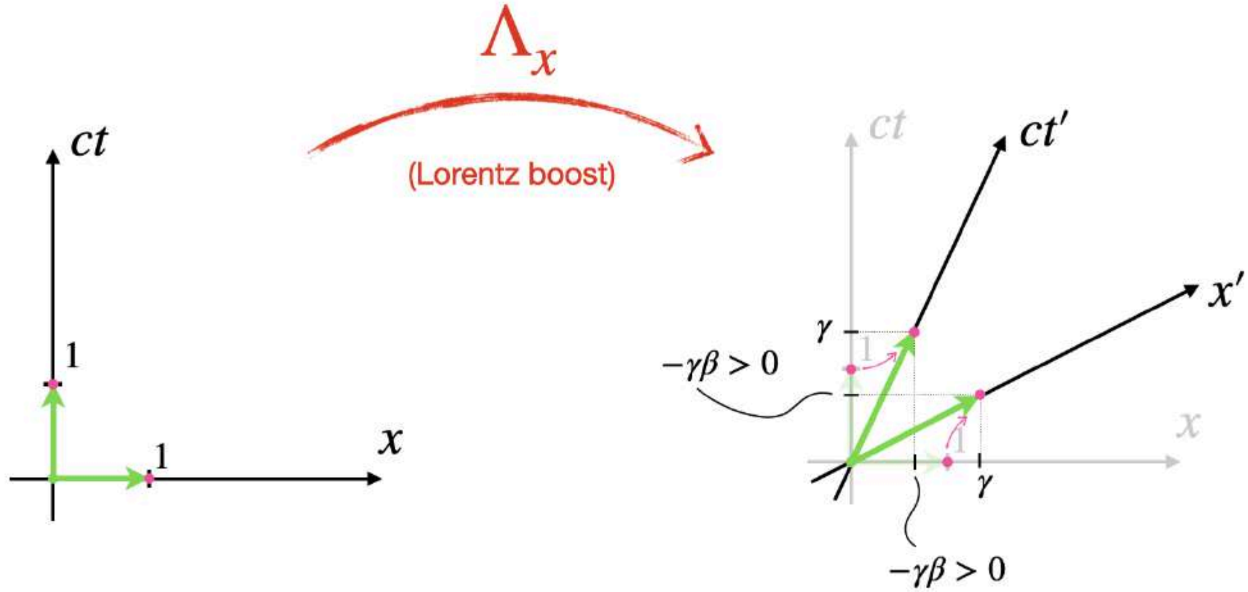
Let's find it out by performing their dot product:

(i.e. if it results in zero, then they are still orthogonal even after the transformation)

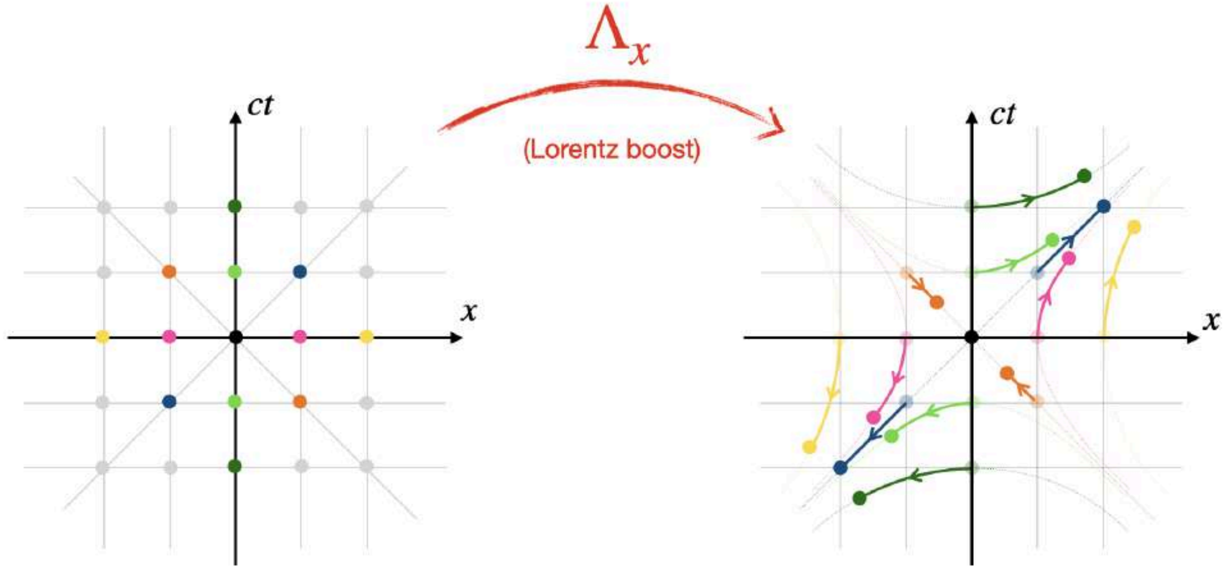
$$\vec{v}'_{ct} \cdot \vec{v}'_x = \gamma \begin{bmatrix} 1 \\ -\beta \end{bmatrix} \cdot \gamma \begin{bmatrix} -\beta \\ 1 \end{bmatrix} = \gamma^2 \left(1 \cdot (-\beta) + (-\beta) \cdot 1\right) = -2\gamma^2\beta$$

Indeed! If $v = 0$ then $-2\gamma^2\beta = -2\gamma^2 \frac{v}{c} = 0$

When $v \neq 0$ $(v < 0)$ the dot product results in a negative number, which implies that the axes ct' and x' are not orthogonal!



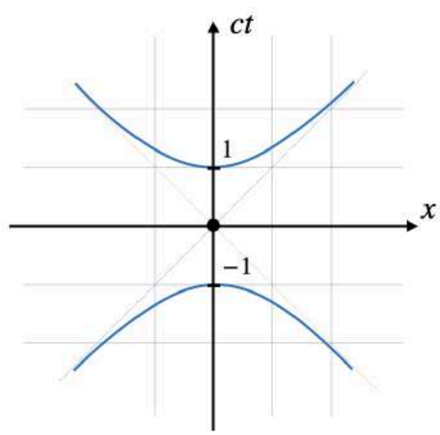
Now, performing the Lorentz boost for many points in the grid of the plane we get a very good grasp of the dynamics of the motion:

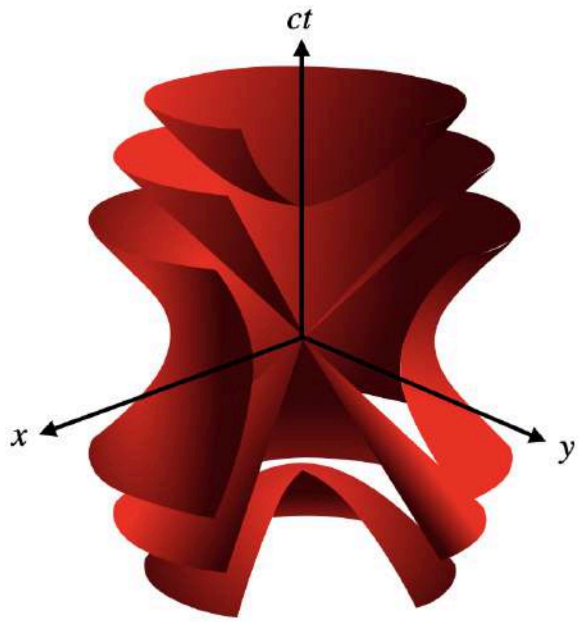


This is, clearly, a *hyperbolic* spacetime. Of course, we can extend it to the 4 -dimensional $(ct\,,\ x\,,\ y\,,\ z)$ Minkowski spacetime.

Formula hyperbola:

$$\underbrace{c^2t^2 - x^2 - y^2 - z^2}_{\text{Minkowski metric}} = 1 \quad \Longleftrightarrow \quad \frac{(ct)^2}{(1)^2} - \frac{(x)^2}{(1)^2} - \frac{(y)^2}{(1)^2} - \frac{(z)^2}{(1)^2} = 1$$





Obviously, we cannot draw the 4-dimensional case, just a 3-dimensional version of it, but you can use your imagination to try to grasp the idea of extending this diagram to 4 dimensions. This is Minkowski geometry.

Extra

As we said before, after applying the Lorentz transformation, the Minkowski metric is kept the same (i.e., it is *invariant* under Lorentz boosts). Let's prove it:

$$\boxed{s^2 = c^2t^2 - x^2} \vdash \boxed{s'^2 = c^2t'^2 - x'^2 = s^2} :$$

Lorentz boost:

$$t' = \gamma\left(t - \frac{vx}{c^2}\right)$$

$$x' = \gamma(x - vt)$$

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

$$s'^2 = c^2\gamma^2\left(t - \frac{vx}{c^2}\right)^2 - \gamma^2(x - vt)^2 =$$

$$= c^2\gamma^2\left(t^2 - \frac{2vxt}{c^2} + \frac{v^2x^2}{c^4}\right) - \gamma^2\left(x^2 - 2xvt + v^2t^2\right) =$$

$$= c^2\gamma^2t^2 - 2vxt\gamma^2 + \frac{\gamma^2v^2x^2}{c^2} - \gamma^2x^2 + 2xvt\gamma^2 - \gamma^2v^2t^2 =$$

$$= \frac{c^2t^2}{1 - v^2/c^2} - \frac{2vxt}{1 - v^2/c^2} + \frac{\frac{v^2x^2}{c^2}}{1 - v^2/c^2} - \frac{x^2}{1 - v^2/c^2} + \frac{2xvt}{1 - v^2/c^2} - \frac{v^2t^2}{1 - v^2/c^2} =$$

$$= \frac{c^4t^2}{c^2 - v^2} - \frac{2vxtc^2}{c^2 - v^2} + \frac{v^2x^2}{c^2 - v^2} - \frac{x^2c^2}{c^2 - v^2} + \frac{2vxtc^2}{c^2 - v^2} - \frac{v^2t^2c^2}{c^2 - v^2} =$$

$$= \frac{c^2t^2}{c^2 - v^2}\left(c^2 - v^2\right) + \frac{x^2}{c^2 - v^2}\left(v^2 - c^2\right) = c^2t^2 - x^2 = s^2 \quad \blacksquare$$

(−1)

Interesting Question

Is the Minkowski spacetime flat or hyperbolic after all?

Minkowski spacetime is flat in the sense of differential geometry, meaning it has zero curvature everywhere. However, it has a hyperbolic structure in how distances (spacetime intervals) are measured.

Why flat? The Minkowski metric $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$ describes a spacetime with zero Ricci curvature and zero Riemann curvature tensor. This means that it has no intrinsic curvature like in General Relativity (where spacetime curvature is linked to gravity).

Why hyperbolic? The Lorentz transformations are analogous to hyperbolic rotations in spacetime. The spacetime interval $s^2 = c^2t^2 - x^2$ defines hyperboloids of constant proper time (rather than spheres as in Euclidean space). Rapidities (used in velocity transformations) are described using hyperbolic trigonometry (e.g., $tanh\,\theta = \frac{v}{c}$)

Please, if you find this document useful, let us know. Or if you found typos and things to improve, let us know as well. Your feedback is very important to us. We're working hard to deliver the best material possible. Contact us at: dibeos.contact@gmail.com