

Phys 514

Relativity

April Sada Solomon

Winter 2021

Course Description

The Principle of Relativity, Special Relativity, Manifolds, Spacetime Curvature, Gravitation, Schwarzschild Solution, Black Holes, Perturbations, Radiation, Introduction to Cosmology, Introduction to QFT in Spacetime.

Contents

1	The Principle of Relativity	1
1.1	Velocity of Propagation of Interaction	1
1.2	Intervals	3
1.3	Proper Time	5
1.4	Lorentz Transformations	8

1 The Principle of Relativity

1.1 Velocity of Propagation of Interaction

In nature we often describe the processes and phenomena that occurs using frames of reference, that is, defined spatial and temporal coordinates which are considered to be both homogeneous and isotropic. When a free body propagates in a frame of reference at a constant velocity through time, that frame of reference is understood to be inertial. Extending this notion to multiple systems, a couple of frames moving uniformly relative to each other, where one is said to be inertial, implies that the other is inertial as well. The same applies for infinitely many frames of reference, as long as they move uniformly relative to one another.

Through many experiments along the course of history, physicists have demonstrated that the laws of physics hold in all inertial frames of reference, implying that the **principle of relativity** is true. This means that the equations of motion are understood to be invariant with respect to transformations of coordinates and time. Furthermore, the interactions that material objects are subjected to are described through potential energy functions with respect to the position of the particles in the system. We assume this to be the case as we also assume these interactions to occur instantaneously, and to a degree, they do.

However, recent experiments during the XXth and XXIst centuries have demonstrated that instantaneous propagation of interaction is indeed not the case in general. Our worldview is distorted by our perception of both time and space. In the scale of the universe and the cosmos, both time and space are unfathomably large in comparison to our own human scale. Thus, if we do not pay careful attention, many of the true secrets of the universe will remain hidden, as they did for thousands of years.

In actuality, any and all changes among the interactions of free bodies will occur after a period of time, even if it is really, really, really small. We can denote this time as Δt , such that when we divide the distance Δx between these bodies by this time, we obtain the **velocity of propagation of interaction** v :

$$\boxed{v = \frac{\Delta x}{\Delta t}} \quad (1.1)$$

This limit to the velocity or propagation very clearly implies that any material object cannot have a velocity higher than this. In particle physics, we assume elementary particles propagate signals of their dynamics to other particles via units of information. In this case, the velocity of propagation is called **signal velocity**. With this limit and the principle of relativity, it becomes evident that the maximum velocity of propagation will remain the same in all inertial frames, meaning it is a universal constant and denoted

$$\boxed{c = 2.998 \times 10^8 \text{ [m s}^{-1}\text{]}} \quad (1.2)$$

This is the **speed of propagation of light signals in a vacuum**. It very much so emphasizes and underlines the reasoning for why it is apparently instantaneous for interactions to occur in our eyes, yet not the case in reality. The speed at which signals are propagated is so fast that we are too small to notice it

relatively.

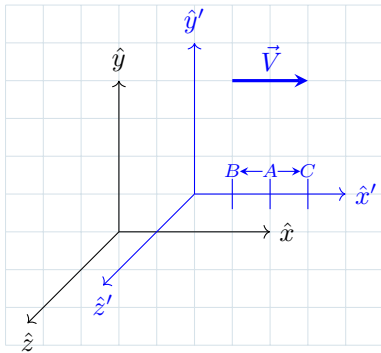
Albert Einstein later refined and generalized Galileo's principle of relativity to what is today known as the Theory of General Relativity. This refinement subsequently led to the modification of mechanics themselves into Relativistic Mechanics. In this case, the propagation of interaction occurs with respect to time as well as position of the bodies in a reference frame. In Newtonian or Classical Mechanics, we assume that interactions only occur with respect to position of the bodies in a frame, as the propagation of interaction is considered instantaneous. In more broader terms, Newtonian Mechanics assumes that the relations of different events in space depend solely on the reference system defined, such that two events happening simultaneously can be well-defined over every reference frame in the system. Contrary to Relativistic Mechanics, Classical Mechanics assumes time to be absolute and independent of the reference frame.

Recall that in classical mechanics, a reference frame A moving with relative velocity \vec{v}_A to another moving reference frame B with velocity \vec{v}_B will have a well defined relative velocity

$$\vec{v} = \vec{v}_a + \vec{v}_b$$

over the overall mechanical system. However, this is clearly a contradiction to the velocity of propagation of interaction, which will be different in different inertial frames of reference. It has been shown that the velocity of light is completely independent of its direction, for example. This means that time is actually not absolute, as assumed in Classical Mechanics, and so simultaneous events in one frame of reference may not be simultaneous in another frame of reference.

Example 1.1. Take a look at the following figure.



We have two inertial frames K , where the observer lies, and K' , where an observation occurs. We have coordinates defined by $(\hat{x}, \hat{y}, \hat{z})$ and $(\hat{x}', \hat{y}', \hat{z}')$ correspondingly. The frame K' moves relative to K at a constant velocity \vec{V} in the \hat{x} axis direction. Suppose a signal propagates from a source point A in the K' frame in both directions of the \hat{x}' axis. The velocity of propagation of interaction on the frame K' is constant and equal to c , as in all inertial frames. Therefore, the signal propagating an interaction from a source A to equidistant sinks B and C in opposing directions will reach the sinks at the same time relative to this inertial frame.

Now shift our perspective to the frame of reference K . In this case, the propagation may not occur simultaneously from source A to the sinks B and C . The limiting aspect of the velocity of propagation to the value c implies that the maximum velocity of propagation in the K frame must be c as well. Thus, as

the sink B moves towards the source A relative to the frame K while sink C moves away from the source A , the signal will reach the sink B before it reaches the sink C .

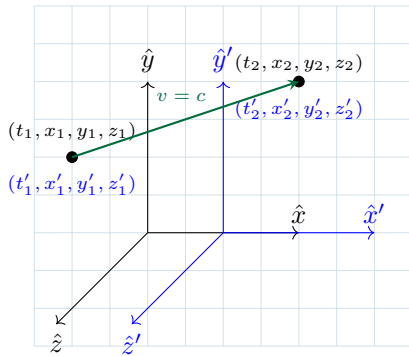
1.2 Intervals

Definition 1.1. **Minkowski spacetime** \mathcal{M} is defined as a 4-dimensional manifold¹ of 3 coordinates of Euclidean space and 1 coordinate for time.

Definition 1.2. An **event** or **world point** is defined as a point in \mathcal{M} .

Definition 1.3. A **world line** is a parametrized, one-dimensional continuous curve in \mathcal{M} defined as a set of events. For instance, the worldline for an elementary particle determines the coordinates of the particle at all moments in time.

We will now express the **Principle of Invariance** of c . We again consider two frames of reference K and K' moving relative to each other with constant velocities.



We assume the coordinate of each frame to be $(\hat{x}, \hat{y}, \hat{z})$ and $(\hat{x}', \hat{y}', \hat{z}')$ respectively such that \hat{x} and \hat{x}' coincide while \hat{y} and \hat{y}' , and \hat{z} and \hat{z}' , are parallel correspondingly. We define a first event on the frame K as (t_1, x_1, y_1, z_1) as the source point of a signal propagating onto another event denoted (t_2, x_2, y_2, z_2) with speed c . We thus arrive at the following conclusion after defining the relative length of propagation as $\ell^2 = c^2(t_2 - t_1)^2$:

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = c^2(t_2 - t_1)^2 = \ell^2$$

We can now rearrange the above to denote a relationship called the **interval** s^2 of these two events in K :

$$0 = -c^2(t_2 - t_1)^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = s^2 \quad (1.3)$$

Furthermore, we define the coordinates for these events in the reference frame K' as (t'_1, x'_1, y'_1, z'_1) and (t'_2, x'_2, y'_2, z'_2) respectively. Similarly, we have the following interval $(s')^2$ between events in this frame:

$$0 = -c^2(t'_2 - t'_1)^2 + (x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 = (s')^2$$

The principal of invariance of the speed of light c thus states that the speed of light will remain constant in any reference frame, so if the interval $s^2 = 0$ in K , then it will be also 0 in all reference frames. Furthermore, if the two events are infinitely close to each other then we can define the **differential interval** or **line**

¹More on this later.

element ds^2 between them as:

$$\boxed{ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2} \quad (1.4)$$

We can then generalize² \mathcal{M} to be defined in the space spanned by $\{\hat{c}t, \hat{x}, \hat{y}, \hat{z}\}$, where for simplicity we usually assume $c = 1$ such that $\mathcal{M} = \{p : p = -\alpha_t \hat{t} + \alpha_x \hat{x} + \alpha_y \hat{y} + \alpha_z \hat{z}\}$ where α_i for $i = t, x, y, z$ are constants denoting the time and coordinates of a propagation between two events, which may be infinitely close to each other. Recall that if $ds^2 = 0$ in K , then $(ds')^2 = 0$ for every other reference frame K' , and as both are infinitesimally small, we can say they are also proportional to each other:

$$ds^2 \propto (ds')^2 \implies ds^2 = k(ds')^2$$

Where the constant k will depend on the absolute value of the relative velocity between the two reference frames. It will not depend on time nor space as that would contradict the homogeneity of both, meaning that different points in time and space would not be equivalent. Moreover, it cannot depend on the direction of relative velocity as that would violate the isotropy of space.

To examine this in more detail, we consider three different reference frames K , K_1 , and K_2 such that V_1 and V_2 are the relative velocities of K_1 and K_2 due to K correspondingly. Hence,

$$\begin{aligned} ds^2 &= k(V_1)ds_1^2 = k(V_2)ds_2^2 \\ \therefore ds_1^2 &= \frac{k(V_2)ds^2}{k(V_1)} = k(V_{2,1})ds^2 \end{aligned}$$

By the above, we are able to see that the angle between the relative velocities is irrelevant to get the constant $k(V_{2,1})$. As we did with the speed of light, we can then equate $k = 1$ such that

$$ds_1^2 = ds_2^2 \implies s_1^2 = s_2^2$$

Theorem 1.1. *The interval s^2 between two arbitrary events is the same in all inertial frames of reference. The interval is thus invariant under transformations from one inertial system to another.*

Recalling the case of two frames of reference K and K' with two events (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) , as well as the length of propagation $\ell^2 = c^2(t_2 - t_1)^2$, we define $t_{2,1} = t_2 - t_1$ such that the interval in each frame becomes

$$\begin{aligned} s_{2,1}^2 &= -c^2 t_{2,1}^2 + \ell_{2,1}^2 \\ (s'_{2,1})^2 &= -c^2 (t'_{2,1})^2 + (\ell'_{2,1})^2 \end{aligned}$$

²In some textbooks like *The Classical Theory of Fields* by Landau and Lifshitz, where most of my notes came from, the notation is inverted such that $\mathcal{M} = \{p : p = +\alpha_t \hat{t} - \alpha_x \hat{x} - \alpha_y \hat{y} - \alpha_z \hat{z}\}$. The reason for this is because we are taking the square of the interval s^2 , so either notation is allowed. However, care must be taken to avoid confusing which one we are using!

Which directly implies that

$$-c^2 t_{2,1}^2 + \ell_{2,1}^2 = -c^2 (t'_{2,1})^2 + (\ell'_{2,1})^2$$

For the two events to occur at the same point in space inside the K' frame of reference, we would need

$$(\ell'_{2,1})^2 = 0 \implies s_{2,1}^2 = -c^2 t_{2,1}^2 + \ell_{2,1}^2 = -c^2 (t'_{2,1})^2 < 0$$

Theorem 1.2. *A reference frame K' where two events occur at the same point such that $(\ell'_{2,1})^2 = 0$ will exist if and only if an initial reference frame K where the two events were defined has a negative valued interval $s_{2,1}^2 < 0$.*

1.3 Proper Time

Definition 1.4. We say that events are **time-like separated** whenever $s_{2,1}^2 < 0$, implying that the velocity of propagation is less than the speed of light.

Definition 1.5. We say that events are **space-like separated** whenever $s_{2,1}^2 > 0$, implying that the velocity of propagation is more than the speed of light, so these two events cannot be connected due to the principle of invariance.

Definition 1.6. We say that events are **light-like separated** or **null-like separated** whenever the interval of the two events is zero: $s_{2,1}^2 = 0$

Furthermore, we can say for the case of time-like separation that signal propagation between events may occur towards the future or towards the past:

$$s_{2,1}^2 < 0 \wedge t_{2,1} > 0 \implies \text{Signals propagate to the } \mathbf{future}$$

$$s_{2,1}^2 < 0 \wedge t_{2,1} < 0 \implies \text{Signals propagate to the } \mathbf{past}$$

In conclusion, $s_{2,1}^2 < 0$ in some reference frame K will imply that there exists another frame of reference K' in which the two events can occur simultaneously.

Definition 1.7. The **proper time** $\tau_{2,1}^2$ between two time-like separated events in a reference frame K' will be defined as

$$\boxed{\tau_{2,1}^2 = (t'_{2,1})^2 = t_{2,1}^2 - \frac{\ell_{2,1}^2}{c^2} = -\frac{s_{2,1}^2}{c^2}} \quad (1.5)$$

Now, let us change notation to provide more convenience. We will define the coordinates of \mathcal{M} in the following way:

$$x^\mu = \begin{cases} \mu = 0 : & x^0 = ct_{2,1} \\ \mu = 1 : & x^1 = x_{2,1} \\ \mu = 2 : & x^2 = y_{2,1} \\ \mu = 3 : & x^3 = z_{2,1} \end{cases} \implies x^i = \begin{cases} i = 1 : & x^1 = x_{2,1} \\ i = 2 : & x^2 = y_{2,1} \\ i = 3 : & x^3 = z_{2,1} \end{cases}$$

As such, we can thus define a tensor $x^\mu \in \mathcal{M}$ and its own transpose $x^\nu = (x^\mu)^T$. Now, we can make our calculations much more reliable by introducing another concept.

Definition 1.8. A **metric tensor** is defined as a function which takes two tensors \vec{v} and \vec{w} and produces a real scalar value that generalizes properties of the dot product. In the case of \mathcal{M} , the **metric of spacetime** is defined as

$$s_{2,1}^2 = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} t_{2,1} & x_{2,1} & y_{2,1} & z_{2,1} \end{pmatrix} \cdot \begin{pmatrix} t_{2,1} \\ x_{2,1} \\ y_{2,1} \\ z_{2,1} \end{pmatrix} = -c^2 t_{2,1}^2 + x_{2,1}^2 + y_{2,1}^2 + z_{2,1}^2$$

And in the much cleaner **Einstein summation notation**, we get:

$$\boxed{s_{2,1}^2 = \eta_{\mu,\nu} x^\mu x^\nu} \quad (1.6)$$

This allows us to have a much cleaner definition for the **proper time** assuming that $c = 1$:

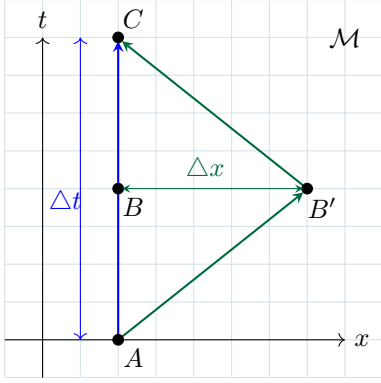
$$\boxed{\tau_{2,1}^2 = -\eta_{\mu,\nu} x^\mu x^\nu = -s_{2,1}^2} \quad (1.7)$$

Theorem 1.3. *The proper time $\tau_{2,1}^2$ between two time-like separated events measures the time elapsed as seen by an observer moving on a straight path between the events on the reference frame K' .*

Notice that if the spatial distance in the initial reference frame K does not change among the two events such that the event is static in space ($\ell_{2,1} = 0$) while displacing in time $t_{2,1}$, then it follows that $\tau_{2,1}^2 = (t'_{2,1})^2 = t_{2,1}^2 \implies \tau_{2,1} = t_{2,1}$. This extrapolates to the fact that the proper time $\tau_{2,1}$ between two events will be invariant regardless of the speed of the inertial frame of the observer, which in this case K' . However, in most general cases where the frames are not inertial, the proper (observed) time $\tau_{2,1}$ and the event time $t_{2,1}$ will not be equal.

$$\tau_{B',A}^2 = \left(\frac{\Delta t}{2}\right)^2 - (\Delta x)^2$$

$$\therefore \boxed{\tau_{B',A} = \frac{\Delta t}{2} \sqrt{1 - v^2}} \quad (1.8)$$



Here, we see the famous *twin's paradox*. We have two observers relative to a frame of reference, usually considered twins, that will go from a point $A(t, x) = (0, 0)$ to a point $C(t, x) = (\Delta t, 0)$ in \mathcal{M} . The first twin will remain static in space while the second will go to point $B'(t, x) = \left(\frac{\Delta t}{2}, \Delta x\right)$ before reaching point C , where $\Delta x = \frac{1}{2}v\Delta t$ for propagation velocity v . The proper time for the trajectory of the first twin from A to B is clearly $\tau_{B,A} = \frac{\Delta t}{2}$, but for the second twin, the proper time from A to B' is given by

By symmetry, it follows that $\tau_{C,B} = \tau_{B,A}$ and $\tau_{B',A} = \tau_{C,B'}$. Considering the entire sequence of movement for both twins, we have

$$\begin{aligned}\tau_{C,B,A} &= \Delta t \\ \tau_{C,B',A} &= \sqrt{1 - v^2} \Delta t\end{aligned}$$

Clearly, $\tau_{C,B,A} > \tau_{C,B',A}$, and so the second twin is younger than the first twin by the time they both reach the point C . This implies that the proper time τ along some world line is related to the interval s^2 traversed through spacetime. Contrary to Classical Mechanics, this implies that non-straight paths are shorter in proper time in Relativistic Mechanics. Because \mathcal{M} is a differentiable manifold, we can redefine the **line element** as

$$\boxed{ds^2 = \eta_{\mu,\nu} dx^\mu dx^\nu} \quad (1.9)$$

We will now get the parametrization for a worldline by defining the spacetime curve $x^\mu(\lambda)$ to be able to calculate derivatives from coordinates of the line element. Thus, we will be able to compute a path integrals to define proper distance and proper time respectively:

$$\begin{aligned}\Delta \mathcal{S} &= \int \sqrt{\eta_{\mu,\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda \\ \Delta \tau &= \int \sqrt{-\eta_{\mu,\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda\end{aligned}$$

This leads us to an important conclusion about elementary particles. Massive particles will propagate with a time-like separation while massless particles will propagate with a space-like separation. However, the distinction between Special and General Relativity only becomes apparent when considering very large masses that can bend spacetime itself. In Special Relativity, bodies that do not bend space time significantly can be assumed to be in flat spacetime at all times.

1.4 Lorentz Transformations

We will now look at coordinate transformations. In other words, we wish to find a relation between the various inertial frames such that their coordinate systems leave the interval as defined in equation (1.6) invariant.

Definition 1.9. A **translation** is defined as a shift in coordinates on space or time such that

$$x^\mu \rightarrow x^{\mu'} = \delta_{\mu}^{\mu'}(x^\mu + a^\mu) \quad (1.10)$$

where a^μ is a set of 4 coefficients and $\delta_{\mu}^{\mu'}$ is a four dimensional [Kronecker delta](#).

Notice that translations will not alter $x_{2,1}^\mu$ as the metric stays the same, as we are only doing a change in coordinates with respect to a different coordinate system over the same metric, so the interval $s_{2,1}^2$ remains invariant as well.

Definition 1.10. A **boost** transformation is an offset by some constant velocity vector over the general coordinates x^μ . A **spatial rotation** is transformation defined as a change in orientation over the spatial coordinates x^i . Both can be represented by an \mathcal{M} independent matrix given by $\Lambda_{\nu}^{\mu'}$ such that³

$$x^{\mu'} = \Lambda_{\nu}^{\mu'} x^{\nu} \quad (1.11a)$$

$$x'_{2,1} = \Lambda x_{2,1} \quad (1.11b)$$

Contrary to translations, boosts and rotations will change $x_{2,1}^\mu$ by multiplying it by a matrix $\Lambda_{\nu}^{\mu'}$. Also note that matrix transformations are not commutative unless the matrices used are diagonalizable so the order in a matrix transform is important while in a translation it is irrelevant. The interval $s_{2,1}^2$ will hence remain invariant. To illustrate:

$$\begin{aligned} s_{2,1}^2 &= (x_{2,1})^T \eta (x_{2,1}) \\ &= (x'_{2,1})^T \eta (x'_{2,1}) \\ &= (x_{2,1})^T \Lambda^T \eta \Lambda (x_{2,1}) \end{aligned}$$

This leads to the following result in both notations:

$$\eta = \Lambda^T \eta \Lambda \quad (1.12a)$$

$$\eta_{\rho,\sigma} = \Lambda_{\rho}^{\mu'} \eta_{\mu',\nu'} \Lambda_{\sigma}^{\nu'} = \Lambda_{\rho}^{\mu'} \Lambda_{\sigma}^{\nu'} \eta_{\mu',\nu'} \quad (1.12b)$$

So we wish to find matrices $\Lambda_{\nu}^{\mu'}$ such that $\eta_{\mu',\nu'} = \eta_{\rho,\sigma}$.

³Both equations are the same, 1.11b is just a more compact form of 1.11a.

Definition 1.11. The matrices Λ which satisfy $\eta_{\rho,\sigma} = \Lambda_{\rho}^{\mu'} \Lambda_{\sigma}^{\nu'} \eta_{\mu',\nu'}$ are denoted as **Lorentz transformations**. The group of all Lorentz transformations under matrix multiplication is denoted as the **Lorentz group** $SO(3,1)$. The matrices in the group will follow a **Lorentzian metric** where the eigenvalues are of the form $(-, +, +, +)$ or $(+, -, -, -)$. Note that often Lorentzian transformations will not commute, as the group is nonabelian.

Definition 1.12. The **3D Rotations group** $SO(3)$ is the set of all 3×3 matrices R under multiplication that are all orthogonal. This means that $R^T R$ equals a 3×3 identity matrix \mathcal{I} :

$$\mathcal{I} = R^T R = R^T R$$

The matrices in the group will follow an **Euclidean metric** where the eigenvalues are of the form $(+, +, +)$. Like the Lorentz group, the matrices are nonabelian, so they do not commute.

Example 1.2. Lorentzian transformation for a rotation in the xy -plane with θ as a periodic variable with period 2π .

$$\Lambda_{\nu}^{\mu'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Example 1.3. Lorentzian transformation for a boost in the xt -plane with ϕ as a variable with domain $(-\infty, +\infty)$. It can be thought of as a "rotation between space and time":

$$\Lambda_{\nu}^{\mu'} = \begin{pmatrix} \cosh \phi & -\sinh \phi & 0 & 0 \\ -\sinh \phi & \cosh \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Definition 1.13. The nonabelian **Poincaré Group** denotes the set of 10×10 diagonalizable matrices that encompass the 10 degrees of freedom of translations on x^{μ} (4 degrees of freedom), rotations on x^i (3 degrees of freedom), and boosts of both x^0 and some x^i (3 degrees of freedom).

Note. The difference between "boosting" a frame of reference and "accelerating" an object is the same as the difference between "rotating" a frame of reference and "spinning" an object; the first occurs as a change of coordinates, and the second occurs as a consequence of the equations of motion.

Example 1.4. Considering the Lorentzian transformation in Example 1.3, we can get the transformed coordinates as follows:

$$\begin{aligned} t' &= t \cosh \phi - x \sinh \phi \\ x' &= -t \sinh \phi + x \cosh \phi \end{aligned}$$

By setting $x' = 0$, we get the velocity at this point in space.

$$t \sinh \phi = x \cosh \phi \implies v = \frac{x}{t} = \frac{\sinh \phi}{\cosh \phi} = \tanh \phi$$

Hence, we can set $\phi = \tanh^{-1}(v)$ such that

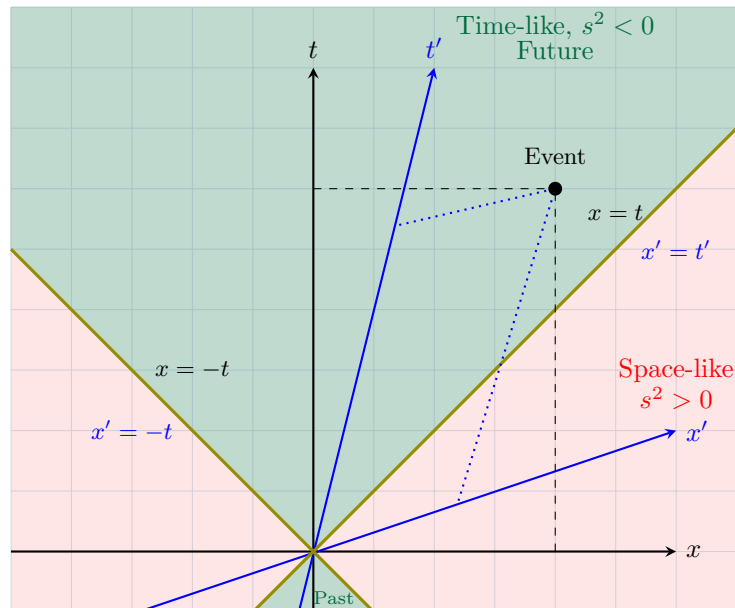
$$\begin{aligned} t' &= t \left(\frac{1}{\sqrt{1-v^2}} \right) - x \left(\frac{v}{\sqrt{1-v^2}} \right) \\ x' &= -t \left(\frac{v}{\sqrt{1-v^2}} \right) + x \left(\frac{1}{\sqrt{1-v^2}} \right) \end{aligned}$$

Let us define $\gamma = (1 - v^2)^{-1/2}$ to get a much cleaner notation of **standard Lorentzian transformations**:

$$t' = \gamma(t - vx) \tag{1.13a}$$

$$x' = \gamma(-vt + x) \tag{1.13b}$$

Graphically an arbitrary Lorentzian transformation that relates coordinates (t', x') to the (t, x) coordinates would look like the following. Note that the event remains unchanged relatively inside the light cone given by the limits $|x| \leq |t|$ and $|x'| \leq |t'|$.



The takeaway here is that, despite the drawing, the axes of the Lorentzian remain orthogonal. The reason being