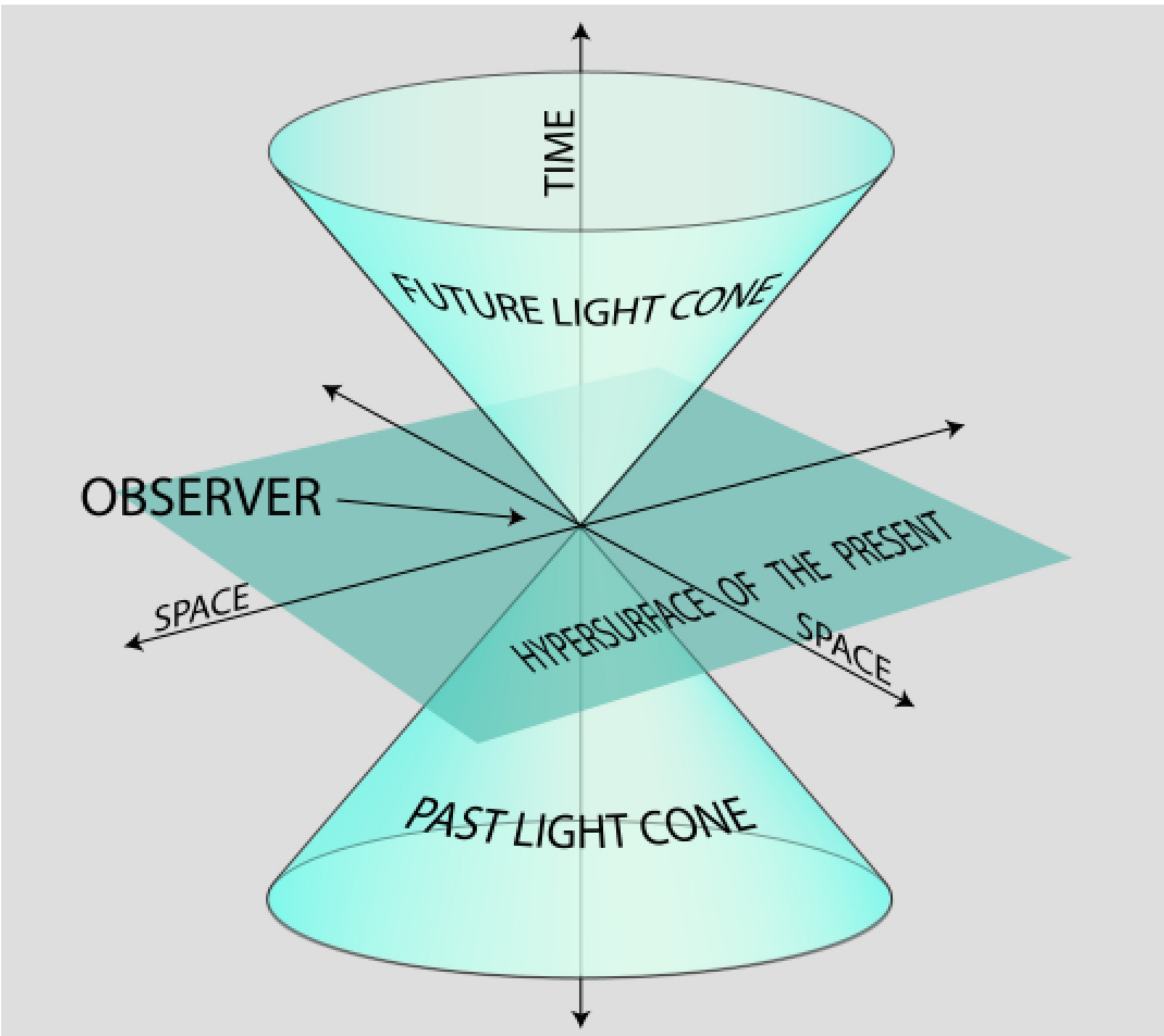
Introduction to special relativity

Panos Christakoglou[1,2]



[1] P.Christakoglou@uu.nl - Utrecht University, Leonard S. Ornsteinlaboratorium Princetonplein 1, 3584CC Utrecht, The Netherlands
[2] Panos.Christakoglou@nikhef.nl - Nikhef, Science Park 105, 1098XG Amsterdam, The Netherlands

Contents

Chapter 1

Introduction to four vectors

Let us consider one event A that happens in one coordinate system S at (t,x). Assume now that we have another system S' that moves with velocity u along the x-axis to the right. This is schematically depicted in fig. 1.1. For this new reference system, the coordinates of A are (t',x'). We assumed here that when the origin of the two systems coincided, the clocks synchronised at $t = t' = 0$.

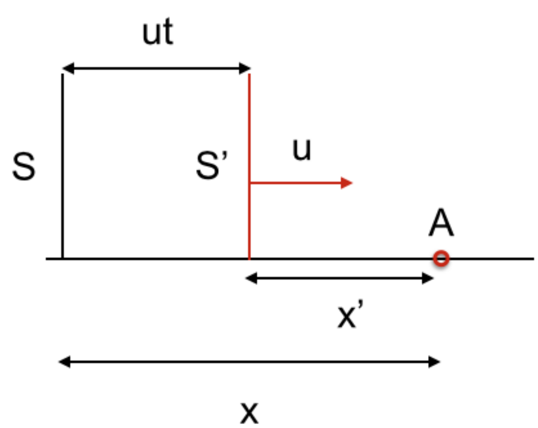


Fig. 1.1: A Galilean transformation between the two reference systems S and S'.

To move from one system to the other, one simply has to perform the relevant transformation that takes the observer from one reference frame to the other. The connection between the two coordinate systems is given by:

$$t' = t$$

$$x' = x - ut$$

If the reference system S' travels with the same velocity as before u but in the opposite direction i.e. to the left, then the spacial coordinates are connected via:

$$x' = x + ut$$

1.1 Special relativity

Let us see now how are these equations altered in the post-Einstein era!

For this we need to first agree that the speed of light is the same in both reference systems i.e. S and $S^{'}$. That means that if we send a light signal, this will travel x (or $x^{'}$) for a time t (or $t^{'}$), such that:

$$x = ct \tag{1.1.1}$$

$$x^{'} = ct^{'} \tag{1.1.2}$$

Now what used to work for the Galilean transformation $x^{'} = x - ut$, should be altered in some way that we are going to try to extract below. Since the way this equation is altered is currently unknown, let us write it as:

$$x^{'} = \gamma(x - ut), \tag{1.1.3}$$

where γ is an unknown fudge factor. The observer in the reference system S will see the object moving in a different way, with the relevant equation now being:

$$x = \gamma(x^{'} + ut^{'}) \tag{1.1.4}$$

At this stage, it is important to note that:

- we assumed here that *gamma*, i.e. the unknown fudge factor, is the same in both systems. This is postulated by the fact that both observers are equivalent.
- The time elapsed from when we synchronised the two clocks and when the light pulse was sent can be different i.e. $t \neq t^{'}$.

Multiplying the two sides of Eq. 1.1.3 and Eq. 1.1.4, we get:

$$xx^{'} = \gamma^2(x - ut)(x^{'} + ut^{'})$$

If one also considers Eq. 1.1.1 and Eq. 1.1.2, the previous relation can be written:

$$c^2tt^{'} = \gamma^2(xx^{'} - utx^{'} - ut^{'}x - u^2tt^{'} = \gamma^2(c^2tt^{'} - \gamma utt^{'} + \gamma utt^{'} - u^2tt^{'} = \gamma^2tt^{'}(c^2 - u^2) \Leftrightarrow$$

$$\gamma^2 = \frac{c^2}{c^2 - u^2} \Leftrightarrow \gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$$

The spatial coordinate thus transforms as:

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

Let us now see how the time coordinate is modified:

$$t^{'} = \frac{1}{u}\left(\frac{x}{\gamma} - x^{'}\right) = \frac{1}{u}\left(\sqrt{1 - \frac{u^2}{c^2}}x - x^{'}\right) = \frac{1}{u}\left(\sqrt{1 - \frac{u^2}{c^2}}x - \gamma x + \gamma ut\right) \Leftrightarrow$$

$$t^{'} = \frac{1}{u}\Big(\sqrt{1-\frac{u^2}{c^2}}x - \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}x + \frac{u}{\sqrt{1-\frac{u^2}{c^2}}}t\Big) = \frac{1}{u\sqrt{1-\frac{u^2}{c^2}}}\Big(x - \frac{u^2}{c^2}x - x + ut\Big) \Leftrightarrow$$

$$t^{'} = \frac{1}{u\sqrt{1-\frac{u^2}{c^2}}}\Big(ut - \frac{u^2}{c^2}x\Big) \Leftrightarrow t^{'} = \frac{t - \frac{u}{c^2}x}{\sqrt{1-\frac{u^2}{c^2}}}$$

Note that similar equations can be written for differences between coordinates e.g. we have two events at A(t_1, x_1) and B(t_2, x_2):

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

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

1.1.1 Time dilation

Let us now consider a clock that in the reference frame S ticks every one second and imagine having two events that happened at the same special coordinate, whose time difference is Δt. The corresponding time difference as measured in a moving reference systems $S^{'}$, as we saw before, would be given by:

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

However, for these two events $\Delta x = 0$, which means that:

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

What does this mean?

This shows that the time $\Delta t^{'}$ between the two ticks as seen in the frame in which the clock is moving $S^{?}$, is longer than the time Δt between these ticks as measured in the rest frame of the clock S. In other words, moving clocks run slower.

1.1.2 Length contraction

Suppose we carry a rod of length L_0, aligned along the x-axis in the reference system S. To measure the length of this rod in the system $S^{'}$, in which the clock is moving, the two ends of the rod need to be measured at the same time in $S^{'}$. This means that $\Delta t^{'} = 0$. The relation between the lengths in both systems is given by

$$\Delta x = \frac{\Delta x^{'} + u\Delta t^{'}}{\sqrt{1-\frac{u^2}{c^2}}} \Leftrightarrow \Delta x^{'} = L_0\sqrt{1-\frac{u^2}{c^2}}$$

Both time dilation and length contraction, as you can imagine, have quite some implications, the most characteristic of which is the example of the twins: at the age of 20, one of them leaves the earth for a round trip with a spacecraft that travels with a constant velocity $u = 0.5c$. The trip takes him $d = 10$ ly (i.e. light years) away and back to earth. For his brother who stayed back the time elapsed between the start and the end of the trip is:

$$\Delta t = \frac{2d}{u} = 40\text{years}$$

Let us now see what is the time elapsed for the astronaut. The astronaut travels at a constant speed but not for 10 ly. The distance is now:

$$\Delta x^{'} = L_0\sqrt{1-\frac{u^2}{c^2}} = 8.66\text{ly}$$

Hence, the trip lasts for $t = 34.6$years. The astronaut will return to the earth at the age of 54.6 and will meet his 60 year-old twin!

1.1.3 Causality

Assume that we have two events, one being the outcome and the other the cause. We say that these two events are causally related. It makes sense to say that no one should be able to see the outcome before the cause. On the other hand if the two events are not causally related, a reversed order will not lead to logical contradictions.

According to Lorentz transformations:

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

Let's now assume that in S $\Delta t = t_2 - t_1 > 0$. To reverse the order of the two events in another, moving reference system $S^{'}$ we need to have $\Delta t^{'} = t_2^{'} - t_1^{'} < 0$:

$$\frac{\Delta t - \frac{u}{c^2}\Delta x}{\sqrt{1-\frac{u^2}{c^2}}} < 0 \Leftrightarrow \Delta t < \frac{u}{c^2}\Delta x \Leftrightarrow \frac{c\Delta t}{\Delta x} < \frac{u}{c}$$

- If we send a light pulse that travels for Δt, then if $c\Delta t > \Delta x$ (i.e. case A in fig. 1.2), the two events can be reversed in time if

  $$\frac{u}{c} > 1 \Leftrightarrow u > c$$

  If a light signal has enough time to travel between the two events (i.e. to cover their spatial separation) their order can not be reversed.
- On the other hand, if the light signal does not have enough time to travel the entire spatial separation of these two events (i.e. case B in fig. 1.2), then

  $$c\Delta t < \Delta x \Leftrightarrow u < c$$

  That means that we can find an observer for whom the ordering can be reversed. That is harmless though since if a light can not connect two events, nobody can!

In summary, if there is time for a light signal to connect two events which are causally connected, then there is no frame where the time ordering of these two events can be reversed.

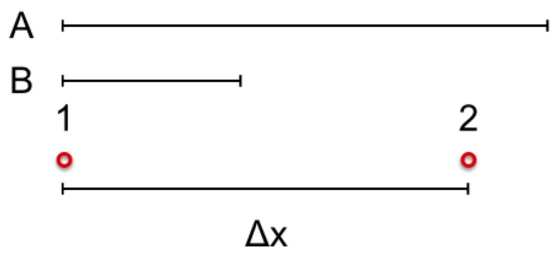


Fig. 1.2: Two events (1) and (2) that causally related in S occur at distance Δx. Depending on whether there is enough time for a light signal to travel (at least) the distance of the two events or not, the time ordering can not or can be reversed in another reference frame.

This now defines the so-called light cone seen in fig. 1.3. The two axis give the spatial (i.e. horizontal axis) and the time (i.e. vertical axis) coordinate. The light cone can be divided in three regions as follows

- The time-like region, with $c\Delta t > \Delta x$, also called as the absolute past (i.e. blue triangle) and absolute future (i.e. yellow triangle),
- the space-like region, with $c\Delta t < \Delta x$,
- and the lines along the diagonals, where $c\Delta t = \Delta x$, known as light-like.

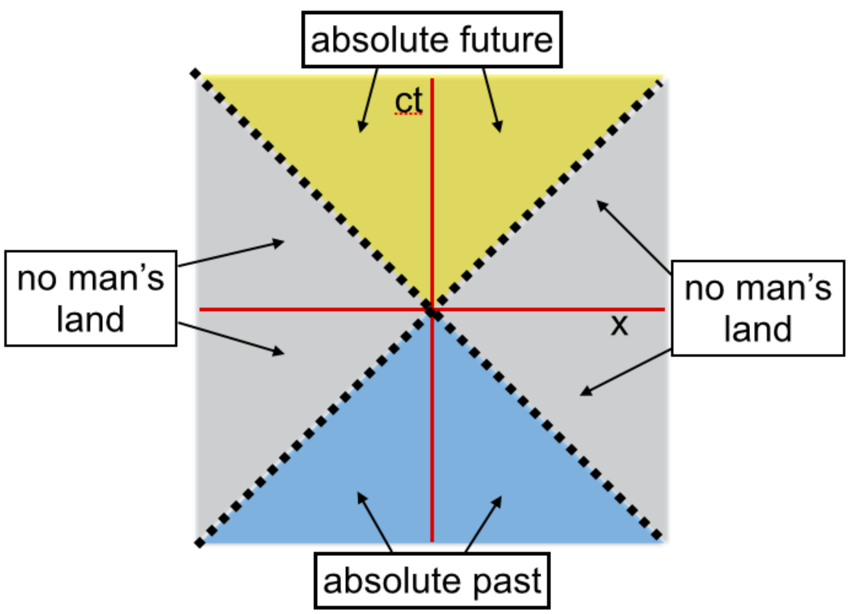


Fig. 1.3: The light cone in special relativity.

1.1.4 Velocity transformations

Let us now consider a particle that moves by Δx in a given time interval Δt measured in a reference system S. The same particle will be seen by an observer who travels along a moving reference system $S^{'}$ as if it traveled a distance $\Delta x^{'}$ for an interval $\Delta t^{'}$. We can now calculate the velocities in the two systems by taking the time derivative:

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

$$v^{'} = \frac{\Delta x^{'}}{\Delta t^{'}}$$

We will now derive the relation between the velocities in these two different systems:

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

$$\frac{\Delta x - u\Delta t}{\Delta t - \frac{u}{c^2}\Delta x} = \frac{\frac{\Delta x}{\Delta t} - u}{1 - \frac{u}{c^2}\frac{\Delta x}{\Delta t}} \Leftrightarrow$$

$$v^{'} = \frac{v - u}{1 - \frac{u}{c^2}v}$$

Going back from $S^{'}$ to S simply means that we have to follow:

$$v = \frac{v^{'} + u}{1 + \frac{u}{c^2}v^{'}}$$

It is interesting to see what kind of velocity one computes for a light pulse in the reference system $S^{'}$:

$$v^{'} = \frac{v - u}{1 - \frac{u}{c^2}v} = \frac{c - u}{1 - \frac{u}{c^2}c} = c\frac{c-u}{c-u} = c$$

We have now established a great part of special relativity, starting from the Lorentz transformations. At this stage, it is important to try to simplify a bit the notation by introducing the following variables:

- The velocity coefficient β:

$$\beta = \frac{u}{c}$$

- The Lorentz factor γ:

$$\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}} = \frac{1}{\sqrt{1-\beta^2}}$$

The Lorentz transformations can thus be written:

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

$$x^{'} = \gamma\left[x - \beta(ct)\right]$$

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

$$(ct^{'}) = \gamma\big[(ct) - \beta x\big]$$

1.2 Space-time four vector

We have seen so far that in special relativity, time is considered as another dimension. It is important to note that this fourth dimension connects with the spatial coordinates via the Lorentz transformations i.e. it is not a completely independent coordinate. Let us now see how we can simplify the way we write down and deal with Lorentz transformations. For this we are going to introduce a beautiful mathematical notion, the one of a four vector. A four vector is an extension of the normal three vectors we all now by now, by adding the time coordinate as the 0-th dimension:

$$\vec{r} = (x, y, z) \to X \equiv (x_0, \vec{r}) = (x_0, x_1, x_2, x_3)$$

In the case of the normal three vectors, we were getting the magnitude (or its length) by taking the Euclidean norm. That means that in the case of $\vec{r} = (x, y, z)$ then:

$$|r| = \sqrt{x^2 + y^2 + z^2}$$

Let's try to do this in the case of a four vector, assuming that the 0-th coordinate is simply t:

$$\sqrt{t^2 + x^2 + ...}$$

You will notice immediately something strange in the previous attempt to write down the formula i.e. the different terms do not have the same units and thus makes little sense to add them directly. Hence we need to change either the time coordinate or the spatial coordinates in a way that they use a common unit system. As you can imagine, it is more convenient to change only one coordinate i.e. the time coordinate, by multiplying it with some velocity. It is natural to multiply t with an invariant velocity i.e. one that does not change when moving from one system to the other and this makes the speed of light the natural choice. By doing this all coordinates have units of length:

$$X \equiv (x_0, \vec{r}) = (x_0, x_1, x_2, x_3) = (ct, x, y, z)$$

In this notation, it is easy to see that the Lorentz transformations can be written as:

$$(ct^{'}) = \gamma\big[(ct) - \beta x\big] \Leftrightarrow x_0^{'} = \gamma\big(x_0 - \beta x_1\big)$$

$$x^{'} = \gamma\big[x - \beta(ct)\big] \Leftrightarrow x_1^{'} = \gamma\big(x_1 - \beta x_0\big)$$

Up to this moment we have not gained much, we just introduced a different notation. We can take advantage though of this notation and extend it a bit further. We can now introduce the indices $\mu : 0, 1, 2, 3$ such that:

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

Note that the index μ is a superscript to the four vector X and this is what we are going to use from now on to indicate that this four vector has a representation in terms of a matrix with μ rows. This representation is known as the **contravariant form**.

We will now see how we can simplify the way we write the Lorentz transformations using this new notation. For this we need to introduce the table

$$\Lambda^{\mu}_{\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where, as in before, the index μ shows the number of rows while ν the number of columns.

With this last addition, the Lorentz transformations can be written as:

$$x^{\mu'} = \sum_{\nu=0}^{3} \Lambda^{\mu}_{\nu} x^{\nu}$$

It turns out that we can simplify even further the previous equation, by using the Einstein's summing convention:

$$x^{\mu'} = \Lambda^{\mu}_{\nu} x^{\nu}$$

1.2.1 Covariant notation

Mathematically, space-time is represented by a four dimensional matrix, known as the Minkowski metric $g_{\mu\nu}$. For the rest of the document, we will be using the (+—) signature[1] The form of $g_{\mu\nu}$ is:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Note that the inverse of this matrix is itself.

With the help of $g_{\mu\nu}$, we can now define the covariant representation of four vectors as:

$$x_{\mu} = g_{\mu\nu} x^{\nu}$$

Note how one can go from the contravariant to the covariant and the other way around: the indices "cancel out"! What is the covariant form of a four vector? Let us try to extract it:

$$x_{\mu} = g_{\mu\nu} x^{\nu} = \left(x_0,\, -x_1,\, -x_2,\, -x_3\right) = \left(x_0,\, -\vec{r}\right)$$

[1] Note that you will find part of the literature following the opposite i.e. (-+++) signature for $g_{\mu\nu}$.

1.2.2 Invariant quantities

We now that although some coordinates change under a given transformation, certain combinations of these quantities remain invariant. A characteristic example is the rotation of the system S to $S^{'}$ by an angle θ. This transforms the vector $\vec{r} = x\hat{i} + y\hat{j}$ to $\vec{r}^{'} = x^{'}\hat{i} + y^{'}\hat{j}$. The connection between the coordinates in these two reference systems is given by:

$$x^{'} = x\cos\theta - y\sin\theta$$

$$y^{'} = x\sin\theta + y\cos\theta$$

If one takes the scalar (i.e. dot) product of each vector with itself, that is if we compute the length of the vector in both systems, then we will realise that it is invariant:

$$\vec{r}^{'}\vec{r}^{'} = (x^{'})^2 + (y^{'})^2 = (x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta)^2 =$$

$$x^2\cos^2\theta + y^2\sin^2\theta - 2xy\cos\theta\sin\theta + x^2\sin^2\theta + y^2\cos^2\theta + 2xy\cos\theta\sin\theta =$$

$$x^2\left(\sin^2\theta + \cos^2\theta\right) + y^2\left(\sin^2\theta + \cos^2\theta\right) = x^2 + y^2 \Leftrightarrow$$

$$\vec{r}^{'}\vec{r}^{'} = \vec{r}\vec{r}$$

The same with four vectors does not work i.e. $(ct)^2 + x^2 \neq (ct)^{'2} + x^{'2}$. On the other hand, what seems to be invariant is the product:

$$I = x_\mu x^\mu = x_0^2 - x_1^2 - x_2^2 - x_3^2$$

$$I = x_\mu^{'} x^{\mu'} = x_0^{'2} - x_1^{'2} - x_2^{'2} - x_3^{'2}$$

Let us try to confirm this, by considering the Lorentz transformations in one direction e.g. on the x-axis, that is taking $x_2 = x_2^{'}$ and $x_3 = x_3^{'}$:

$$x_\mu^{'} x^{\mu'} = x_0^{'2} - x_1^{'2} - x_2^{'2} - x_3^{'2} = \left[\gamma\left(x_0 - \beta x_1\right)\right]^2 - \left[\gamma\left(x_1 - \beta x_0\right)\right]^2 - x_2^2 - x_3^2$$

$$= \gamma^2 x_0^2 + \gamma^2\beta^2 x_1^2 - 2\beta\gamma^2 x_0 x_1 - \gamma^2 x_1^2 - \beta^2\gamma^2 x_0^2 + 2\beta\gamma^2 x_0 x_1 - x_2^2 - x_3^2$$

$$= \gamma^2(1-\beta^2)x_0^2 - \gamma^2(1-\beta^2)x_1^2 - x_2^2 - x_3^2$$

$$= \frac{1-\beta^2}{1-\beta^2}x_0^2 - \frac{1-\beta^2}{1-\beta^2}x_1^2 - x_2^2 - x_3^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2 \Leftrightarrow$$

$$x_\mu^{'} x^{\mu'} = x_\mu x^\mu$$

The quantity $x_\mu x^\mu$ is invariant and it is called the scalar product of x and it is generally given by:

$$\vec{x}^2 = x_\mu x^\mu = x_0^2 - \vec{x}\vec{x}$$

Based on the value of this scalar product, the four vector x^μ can be characterised as:

- time-like if $\vec{x}^2 > 0$,
- space-like if $\vec{x}^2 < 0$,
- light-like if $\vec{x}^2 = 0$,

Let us now summarise:

1. A four vector is an extension of the three vector with the addition of a time related coordinate (i.e. ct) as the 0-th component.
2. A four vector transforms from one inertial reference frame to another via the Lorentz transformation:

$$x^{\mu'} = \Lambda^\mu_\nu x^\nu$$

3. The Minkowski metric is $g_{\mu\nu} = g^{\mu\nu}$, with $g^{-1} = g$.
4. The transition from a contravariant to a covariant your vector happens via:

$$x_\mu = g_{\mu\nu} x^\nu$$

5. The quantity $x_\mu x^\mu$ is invariant and it is called the scalar product of x and it is generally given by:

$$\vec{x}^2 = x_\mu x^\mu = x_0^2 - \vec{x}\vec{x}$$

At this point, the transition to tensors is almost straightforward. Tensors are mathematical objects and they are the generalisation of vectors and scalars. They are usually represented by matrices and describe linear relations between vectors, scalars and other matrices. The order of a tensor is the dimensionality of the matrix e.g. the Minkowski metric $g_{\mu\nu}$ is a second order tensor, a vector is a first order tensor. Tensors transform in the following way:

$$S^{\mu\nu'} = \Lambda^\mu_\kappa \Lambda^\nu_\sigma S^{\kappa\sigma}$$

$$S^{\mu\nu\lambda'} = \Lambda^\mu_\kappa \Lambda^\nu_\sigma \Lambda^\lambda_\tau S^{\kappa\sigma\tau}$$

1.2.3 Proper time

Let us now consider a particle that moves in the (x,t) plane with velocity v. The particle travels Δx at a time interval Δt. The space-time interval is:

$$(\Delta s)^2 = (\Delta x_0)^2 - (\Delta x_1)^2$$

or in a four vector notation

$$\Delta x^\mu = x^\mu_B - x^\mu_A = \begin{pmatrix} \Delta x_0 \\ \Delta x_1 \end{pmatrix} = \begin{pmatrix} c\Delta t \\ \Delta x \end{pmatrix}$$

$$(\Delta s)^2 = \Delta x_\mu \Delta x^\mu = (\Delta x_0)^2 - (\Delta x_1)^2 = (c\Delta t)^2 - (\Delta x)^2$$

$$= (c\Delta t)^2\left[1 - \frac{1}{c^2}\left(\frac{\Delta x}{\Delta t}\right)^2\right] = (c\Delta t)^2\left[1 - \frac{v^2}{c^2}\right] \Leftrightarrow$$

$$(\Delta s) = \frac{(c\Delta t)}{\gamma}$$

As we said before, the space-time interval (Δs) is invariant i.e. it does not change between different systems. Now let's calculate the same quantity as seen by the particle itself. For the reference frame that travels with the particle the time interval is $\Delta\tau$ but the space interval is $\Delta x' = 0$ i.e. the space coordinate as seen by the particle did not change. That means that the space-time interval as seen by the particle is:

$$(\Delta s)^2 = \Delta x'_\mu \Delta x^{\mu'} = (c\Delta\tau)^2$$

Now from the invariance of (Δs) one gets:

$$\Delta\tau = \frac{\Delta t}{\gamma}$$

The parameter $\Delta\tau$ is also an invariant and is called **proper time**. One can further write the previous formula in terms of derivatives:

$$\frac{d\tau}{dt} = \frac{1}{\gamma} \Leftrightarrow \frac{dt}{d\tau} = \gamma$$

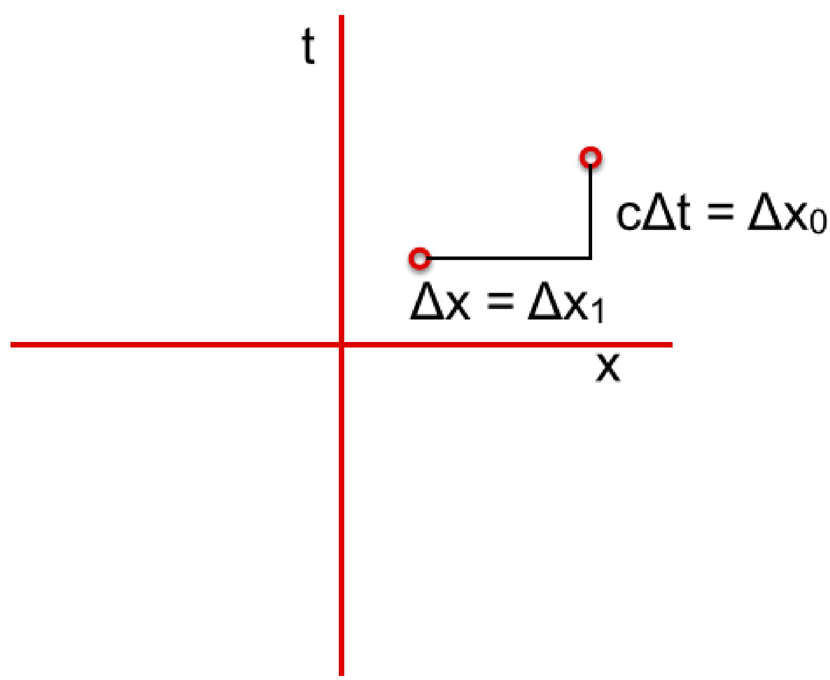


Fig. 1.4: A particle that moves in the (x,t) plane.

1.3 Energy-momentum four vector

In Newtonian mechanics, we used to manufacture new vectors by taking derivatives e.g. $\vec{v} = d\vec{r}/dt$. Then momentum was defined as $\vec{P} = m\vec{v}$. Let us see how things change in the post-Einstein era.

Let us try to generate now more four vectors from x^μ. We can not take the time derivative of x^μ since now time is another coordinate and mixes with x. It is as if we take the derivative of x with respect to y in the three vector notation i.e. it is meaningless! We need to take the derivative with respect to an invariant i.e. the proper time $\Delta\tau$:

$$v^\mu \equiv \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau} = \gamma\frac{dx^\mu}{dt} = \gamma\begin{pmatrix} dx_0/dt \\ dx_1/dt \\ dx_2/dt \\ dx_3/dt \end{pmatrix}$$

The invariant of the four velocity is

$$v_\mu v^\mu = (\gamma c)^2 - (\gamma\vec{v})^2 = \gamma^2 c^2 \big(1 - \frac{v^2}{c^2}\big) = c^2$$

Multiplying the previous with the mass, should give the four momentum:

$$p^\mu = mv^\mu = m\gamma\begin{pmatrix} c \\ \vec{v} \end{pmatrix} = \begin{pmatrix} m\gamma c \\ m\gamma\vec{v} \end{pmatrix}$$

Let us look carefully at the two parts of the new four vector and let us start with the second part i.e. $m\gamma\vec{v}$.

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

This seems as if it is the normal Newtonian mechanics momentum, with the addition of a correction due to velocity. In fact when $u \to 0$ then $m\gamma\vec{v} \to m\vec{v} = \vec{P}$. The momentum is an interesting quantity in relativity: although velocity seems to have an upper limit i.e. the speed of light, momentum does not. This is important for accelerators, where particles are moving very close to the speed of light e.g. $v = 0.9999c$. The next generation of accelerators can add another digit at the end, which makes little difference in the velocity but huge difference in the momentum.

Back to the components of the four momentum now, with the first component which is of particular interest:

$$p_0 = m\gamma c = \frac{mc}{\sqrt{1-u^2/c^2}}$$

which when $u \to 0$, then $p_0 \to mc$ which does not remind us of something. This 0-th component can be written as:

$$p_0 = mc\big(1 - \frac{u^2}{c^2}\big)^{-1/2}$$

We should now remember the fact that

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \ldots$$

for small values of x. This transforms p_0 as follows:

$$p_0 = mc\big(1 + \frac{1}{2}\frac{u^2}{c^2} + \frac{3}{8}\frac{u^4}{c^4} + \ldots\big)$$

$$= mc + \frac{1}{2}\frac{mu^2}{c} + \frac{3}{8}\frac{mu^4}{c^3} + \ldots$$

The first term still tells us nothing, while the second term starts to resemble the kinetic energy. To make the resemblance more clear one has to multiply p_0 by c:

$$cp_0 = mc^2 + \frac{1}{2}mu^2 + \ldots$$

Now the second term is definitely the old known kinetic energy. The first component is called the rest energy of the particle. It shows that even a particle at rest has energy!

The four momentum is now written:

$$p^\mu = \begin{pmatrix} \gamma mc \\ \gamma m\vec{v} \end{pmatrix}$$

Let us now define the relativistic energy

$$E = \gamma mc^2$$

which is connect to the 0-th component of the four momentum as $p_0 = \gamma mc = E/c$. This means that the four momentum can be written as:

$$p^\mu = \begin{pmatrix} E/c \\ \vec{P} \end{pmatrix}$$

The invariant of the four momentum is:

$$p^2 = p_\mu p^\mu = \frac{E^2}{c^2} - \vec{P}^2 = (\gamma mc)^2 - (\gamma mv)^2$$

$$= m^2c^2\gamma^2\big(1 - \frac{v^2}{c^2}\big) \Leftrightarrow p^2 = m^2c^2$$

Being an invariant means that no matter in which frame we will calculate it, its value won't change. In these cases, it is always wise to choose the most convenient (e.g. in terms of complications related to calculations) reference system. In this case, we can try to calculate the dot product in the frame that moves with the particle (i.e. if a particle is what is being studied with the energy-momentum four vector). In this frame, the particle is at rest and the four momentum can be written as:

$$p^\mu = \begin{pmatrix} E/c \\ \vec{0} \end{pmatrix} = \begin{pmatrix} mc \\ \vec{0} \end{pmatrix}$$

It is obvious that the dot product is simply:

$$p^2 = p_\mu p^\mu = m^2c^2$$

This invariant can be written as:

$$\big(\frac{E}{c}\big)^2 - \vec{P}^2 = (mc)^2 \Leftrightarrow$$

$$E^2 = m^2c^4 + P^2c^2$$

This last formula clearly states that the energy of a particle with mass m consists of two parts, one related to its momentum and the other to its rest mass. What happens then for a massless particle e.g. the photon? For this $m = 0$ and $u = c$, which means that both the energy and the momentum part of its energy-momentum four vector is plagued by something that looks as $0/0$. For the case of the photon, its energy-momentum four vector is written as:

$$p^{\mu} = \begin{pmatrix} P \\ \vec{P} \end{pmatrix}$$

since its energy is $E = Pc$. In this case, the energy and momentum is determined by the frequency ν or the wave length λ:

$$E = h\nu = \frac{hc}{\lambda}$$

1.4 Relativistic collisions

Energy and momentum are always conserved. Now that we have defined the energy-momentum four vector, instead of writing four equations, it is obviously more convenient to write everything in a more compact way. Hence, what was up to this moment written as

$$E_1 + E_2 = E_3 + E_4 + ... + E_n$$

$$\vec{P}_1 + \vec{P}_2 = \vec{P}_3 + \vec{P}_4 + ... + \vec{P}_n$$

transforms into

$$P_1^{\mu} + P_2^{\mu} = P_3^{\mu} + P_4^{\mu} + ... + P_n^{\mu}$$

Note that by colliding two incoming particles, we can create at the final state more than two outgoing particles. This is also what is being achieved and studied at accelerator complexes where by colliding e.g. p on p, as done at the Large Hadron Collider (LHC), we can create 3, 4, 5,...,50 particles and more. The equation of the energy momentum four vectors can be generalised to an equality of all incoming four momenta with all outgoing four momenta:

$$P_{\text{in}}^{\mu} = P_{\text{out}}^{\mu}$$