

Special Relativity

Wednesday, 6 September 2023 4:23 PM

In the lecture, we have used the energy-momentum conservation and 4-vectors. While the full theory of special relativity is beyond the scope of this module, this note serves to show where they come from for those who are interested.

Postulates of Special Relativity

The entirety of special relativity comes from the following 2 postulates:

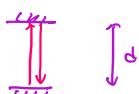
- ① The speed of light is the same in all inertial frames.
- ② The laws of physics are the same in all inertial frames.

Everything else follows from these 2. Postulate ① may sound weird, but it has been verified experimentally through the Michelson-Morley experiment

PLS: Think of inertial frames as frames that move at constant velocity

Time Dilation

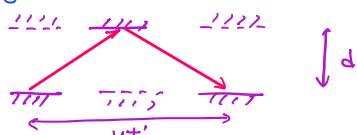
To keep the speed of light constant in all inertial frames, time will have to alter to compensate for the motion of the inertial frame. This can be observed from the following thought experiment. Consider a light travelling vertically upwards from one mirror to another then returned



The distance between the 2 mirror is

$$d = \frac{ct}{2}$$

On the other hand, in an inertial frame that travels horizontally with velocity \bar{u}



The mirror would have travelled horizontally by ut' .

By Pythagoras' theorem :

$$\textcircled{1} \quad d = \sqrt{\left(\frac{ct'}{2}\right)^2 - \left(\frac{ut'}{2}\right)^2}$$
$$\frac{ct}{2} = \sqrt{c^2 - u^2} \frac{t'}{2}$$
$$\therefore t = \sqrt{1 - \left(\frac{u}{c}\right)^2} t'$$

Q: Why did we consider 2-way path rather than a 1-way path when they would have given the same answer?

By convention, we define

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

So, we have

$$t' = \gamma t$$

Since $u < c \Rightarrow \beta < 1$

$$\Rightarrow \gamma > 1$$

$$\Rightarrow t' > t$$

This is called time dilation. The time measured in the rest frame is the shortest. This is why particles that are travelling fast have longer lifetimes.

Length Contraction

In the previous example, since the frame travels in a direction perpendicular to the light direction, the vertical length is not affected. What if the frame travels along the same direction of the light? In the frame where the mirrors are at rest:

$$d = \frac{ct}{2}$$

In the frame where the mirror is moving

① $d' + ut_1' = ct_1'$

② $d' = ct_2' + ut_2'$

From time dilation:

$$\begin{aligned} t_1' + t_2' &= \gamma t \\ \frac{d'}{c-u} + \frac{d'}{c+u} &= \gamma \frac{2d}{c} \\ d' - \frac{2uc}{c^2-u^2} &= \gamma \frac{2d}{c} \end{aligned}$$

Q: What happens if we only take half of the journey to derive length contraction?
Do you get the same thing?

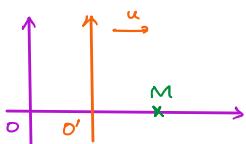
$$d' = \gamma d (1-\beta^2) \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

$$d' = \frac{d}{\gamma}$$

Since $\gamma > 1 \Rightarrow d' < d$, the length in the rest frame d is the longest, and hence length contraction in the moving frame.

Lorentz Transformation

We have derived time dilation and length contraction for the specific situation above. How does the coordinates transform in general? Consider the same event observed from two frames of reference



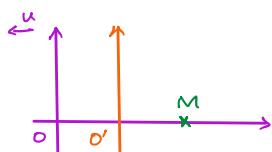
If we are an observer in the O frame

$$\begin{aligned} OM &= OO' + O'M \\ x &= ut + \frac{x'}{\gamma} \end{aligned}$$

length contraction because x' is measured in the O' frame

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

If we are an observer in the O' frame



$$\begin{aligned} OM &= OO' + O'M \\ \frac{x}{\gamma} &= ut' + x' \end{aligned}$$

from above

$$= ut' + \gamma(x - ut)$$

$$\frac{x}{\gamma} = ut' + x' \quad \text{from above}$$

$$= ut' + \gamma(x - ut)$$

$$ut' = \left(\frac{1}{\gamma} - \gamma\right)x + \gamma ut$$

$$\begin{aligned}\frac{1}{\gamma} - \gamma &= \sqrt{1-\beta^2} - \frac{1}{\sqrt{1-\beta^2}} \\ &= \frac{1-\beta^2-1}{\sqrt{1-\beta^2}} \\ &= -\frac{\beta^2}{\sqrt{1-\beta^2}} \\ &= -\gamma\beta^2\end{aligned}$$

$$= -\gamma\beta^2x + \gamma ut$$

$$t' = -\frac{\gamma u}{c^2}x + \gamma t$$

$$ct' = \gamma(ct - \gamma\beta x)$$

Combining the two equations, we have the transformation laws between 2 frames of reference:

$$\begin{cases} ct' = \gamma(ct - \gamma\beta x) \\ x' = \gamma(x - ut) \end{cases}$$

$$\Rightarrow \begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

This is the Lorentz transformation. For completeness, the transverse direction:

$$y' = y$$

$$z' = z$$

Observe that to have a fixed speed of light in vacuum, the time and space coordinates are mixed. This is why in relativity, we consider the 4-vector, where time becomes a coordinate.

To make their units consistent, we consider ct instead of t , which incidentally provides a symmetric form to the Lorentz transformation matrix

Spacetime Interval

In classical physics, we know that when we change from one inertial frame to another, the distance between 2 locations remain the same.

However, in special relativity, we just saw that length contracts in different inertial frames. Then, what is the scalar quantity that measures the "distance" between 2 space-time events? We return to the postulates and consider a spherical wave emitting from the centre in 2 frames:



We have $(ct)^2 = x^2 + y^2 + z^2$ and $(ct')^2 = x'^2 + y'^2 + z'^2$,

which implies

$$0 = (ct)^2 - x^2 - y^2 - z^2 = (ct')^2 - x'^2 - y'^2 - z'^2.$$

which implies

$$0 = (ct)^2 - x^2 - y^2 - z^2 = (ct')^2 - x'^2 - y'^2 - z'^2.$$

This suggests the candidate

$$s^2 = (ct)^2 - x^2 - y^2 - z^2$$

as the "space-time distance." In fact, from the Lorentz transformation, considering only x and t for simplicity:

$$\begin{aligned} ct'^2 - x'^2 &= (ct' - x') \begin{pmatrix} ct' \\ x' \end{pmatrix} \\ &= (ct' - x') \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \\ &= (ct - x) \underbrace{\begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix}}_{\downarrow} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \underbrace{\begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix}}_{\downarrow} \begin{pmatrix} ct \\ x \end{pmatrix} \\ &= \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \underbrace{\begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix}}_{=1 \text{ because just reversing the frame velocity}} \\ &= (ct - x) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \\ &= (ct)^2 - x^2 \end{aligned}$$

The value $(ct)^2 - x^2 = s^2$ is conserved. We call this the spacetime interval. In infinitesimal form:

$$(ds)^2 = (c dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$$

For light, $(ds)^2 = 0$, as seen from above.

For ordinary matter, $(ds)^2 > 0$, so that

$$\begin{aligned} (c dt)^2 - (dx)^2 - (dy)^2 - (dz)^2 &> 0 \\ c^2 &> \frac{(dx)^2 + (dy)^2 + (dz)^2}{(dt)^2} = v^2 \\ \Rightarrow v &< c \end{aligned}$$

The speed of ordinary matter is less than c .

In particular, at the rest frame, where the object is not moving, $dx = dy = dz = 0$, and

$$(ds)^2 = (c dt)^2.$$

We call τ proper time. If you measure time at the rest frame, you are actually measuring the spacetime interval.

Q: Now do you understand why we reflect the light back to the starting point when deriving time dilation and length contraction? Why does the derivation still work for time dilation if we consider half the journey, but not for length contraction?

4-vectors

From Lorentz-transformation and the discussion of spacetime interval, we recognise that time has become a coordinate that needs to be transformed between different frames. Therefore, instead of the usual 3D position vector, in special relativity, we speak of the spacetime 4-vector:

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

the usual 3D position vector, in special relativity, we speak of the spacetime 4-vector:

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

where I have used an underline to denote a 4-vector.

From the spacetime interval, the "distance" between 2 events in spacetime is given by

$$(ds)^2 = (c dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$$

$$= (c dt \ dx \ dy \ dz) \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}_{\equiv \eta} \begin{pmatrix} c dt \\ dx \\ dy \\ dz \end{pmatrix}$$

The quantity η is called the Minkowski metric, which encodes how distances are measured in spacetime.

Pls: In relativity, you may encounter the convention of

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \sim \begin{array}{l} \text{Sometimes called} \\ \text{East Coast metric} \end{array}$$

but in particle physics, the usual convention is

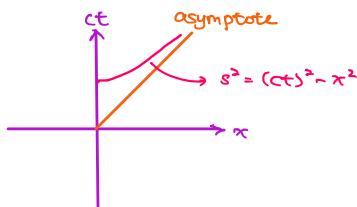
$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \sim \begin{array}{l} \text{Sometimes called} \\ \text{West Coast metric} \end{array}$$

The difference is only at the sign of the spacetime interval.
All physics remain the same.

Since the spacetime interval is the same in all inertial frames,
note that (ignoring y and z for simplicity)

$$s^2 = (ct)^2 - x^2$$

encodes a hyperbola. The underlying geometry is hyperbolic.
By the asymptotes of hyperbola, we observe that the object can never travel faster than the speed of light if it has a rest frame.



Because the geometry is hyperbolic, we can understand Lorentz transformation as a hyperbolic rotation, which can be written as

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh \rho & -\sinh \rho \\ -\sinh \rho & \cosh \rho \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix},$$

where $\cosh \rho = \gamma$ and $\sinh \rho = \gamma \beta$

$$\Rightarrow \boxed{\tanh \rho = \beta} \quad \text{called rapidity}$$

4-momentum

While the 4-position is defined, one still needs to define the 4-momentum to perform relativistic kinematics. It is natural to use the classical expression:

$$\vec{p} = m \vec{v}. \sim \text{use arrow to denote 3-vectors}$$

However, the 3-velocity in different frames are

$$v_x' = v_x \cos \theta + v_y \sin \theta$$

$$\vec{P} = m \vec{v} \rightarrow \text{we want to derive 3-vectors}$$

However, the 3-velocity in different frames are

$$\frac{dx'}{dt'} = \frac{-\gamma u dt + r dx}{\gamma dt - \gamma^2 c^2 dx}$$

$$= \frac{-\gamma u + \gamma \frac{dx}{dt}}{\gamma - \gamma \frac{u}{c^2} \frac{dx}{dt}}$$

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

$$\frac{dy'}{dt'} = \frac{dy}{\gamma dt - \gamma^2 c^2 dx}$$

$$= \frac{\frac{dy}{dt}}{\gamma - \gamma \frac{u}{c^2} \frac{dx}{dt}}$$

$$v_y' = \frac{v_y}{\gamma(1 - \frac{u}{c^2} v_x)}$$

If we naively define

$$\underline{x}' = \frac{dx'}{dt'}$$

$$= \frac{d}{dt'} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix}$$

$$= \begin{pmatrix} c \\ v_x' \\ v_y' \\ v_z' \end{pmatrix}$$

$$= \begin{pmatrix} c(1 - \beta \frac{v_x}{c}) \\ v_x - u \\ v_y/\gamma \\ v_z/\gamma \end{pmatrix} \frac{1}{1 - \beta \frac{v_x}{c}}$$

If we take an inner product following the rules of spacetime interval:

$$\underline{x}' \cdot \underline{x}' \equiv (c \ v_x' \ v_y' \ v_z') \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} v_x' \\ v_y' \\ v_z' \end{pmatrix}$$

$$= c^2 - v_x'^2 - v_y'^2 - v_z'^2$$

$$= c^2 - \frac{1}{(1 - \beta \frac{v_x}{c})^2} [(v_x - u)^2 - \frac{v_y^2}{\gamma^2} - \frac{v_z^2}{\gamma^2}]$$

$$\neq c^2 - v_x^2 - v_y^2 - v_z^2,$$

its norm is not conserved as expected of a scalar quantity. This is because time t changes with different frames of reference. We need a "time quantity" that does not change across different frame of reference.

Recall that proper time measures the spacetime interval, which is invariant across different frames. Naturally, we define

$$\underline{x} \equiv \frac{d\underline{x}}{dt}$$

$$= \frac{d}{dt} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

$$= \frac{dt}{dt} \frac{d}{dt} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

time dilation
 $dt = \gamma d\tau$

$$\underline{x} = \gamma \begin{pmatrix} c \\ \vec{v} \end{pmatrix}$$

$$\boxed{\underline{v} = \gamma \left(\begin{array}{c} c \\ \vec{v} \end{array} \right)}$$

Its norm is

$$\begin{aligned}\underline{v} \cdot \underline{v} &\equiv \gamma(c \cdot \vec{v}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \gamma \begin{pmatrix} c \\ \vec{v} \end{pmatrix} \\ &= \gamma^2 (c^2 - \vec{v}^2) \\ &= c^2 \frac{1-\beta^2}{1+\beta^2}\end{aligned}$$

$$\boxed{\underline{v} \cdot \underline{v} = c^2},$$

which is invariant in all frames of reference.

Hence, following classical mechanics, we define the 4-momentum:

$$\begin{aligned}\underline{P} &= m \underline{v} \\ &= m \gamma \left(\begin{array}{c} c \\ \vec{v} \end{array} \right) \\ &\equiv \left(\begin{array}{c} \gamma mc \\ \vec{p} \end{array} \right)\end{aligned}$$

where its norm is

$$\underline{P} \cdot \underline{P} = m^2 \underline{v} \cdot \underline{v}$$

$$\boxed{\underline{P} \cdot \underline{P} = m^2 c^2}$$

The 4-momentum also implies the relativistic momentum

$$\boxed{\vec{p} = \gamma m \vec{v}}$$

What is the "time-component" of the 4-momentum? What is γmc ? To understand it, we consider its non-relativistic limit, where $\beta \ll 1$:

$$\begin{aligned}\gamma mc &= \frac{mc}{\sqrt{1-\beta^2}} \\ &= mc \left(1 + \frac{1}{2}\beta^2 \right) + O(\beta^4) \quad \text{series expansion} \\ &= mc + \frac{1}{2} \frac{mv^2}{c}\end{aligned}$$

The second term is the kinetic energy in classical mechanics, which shows that

$$\gamma mc = \frac{E}{c}$$

Hence, the norm of 4-momentum implies

$$\begin{aligned}m^2 c^2 &= \underline{P} \cdot \underline{P} \\ &= \left(\frac{E}{c} \right)^2 - \vec{p}^2\end{aligned}$$

$$\boxed{E^2 = (mc^2)^2 + (\vec{p}c)^2}$$

This is where the energy-momentum conservation comes from.

Pls: An alternative way to identify E is the following:

We had

$$\underline{v} = \gamma \left(\begin{array}{c} c \\ \vec{v} \end{array} \right).$$

From Newton's equation, we define 4-force

$$\begin{aligned}\underline{f} &\equiv \frac{d}{dt} \underline{P} \\ &= m \frac{d\underline{v}}{dt} \equiv m \underline{a}\end{aligned}$$

$$\text{Since } \underline{v} \cdot \underline{v} = c^2$$

$$\frac{d}{d\tau} (\vec{v} \cdot \vec{v}) = 0$$

$$2\vec{a} \cdot \vec{v} = 0$$

$$\Rightarrow \vec{v} \cdot \vec{a} = 0$$

Then, $\vec{f} \cdot \vec{v} = m\vec{a} \cdot \vec{v} = 0$

$$(c \frac{dp_0}{d\tau} \quad \frac{df}{d\tau}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \gamma \begin{pmatrix} c \\ \vec{v} \end{pmatrix} = 0$$

$$\gamma c \frac{dp_0}{d\tau} - \gamma \vec{v} \cdot \frac{d\vec{p}}{d\tau} = 0$$

$$\frac{dp_0}{d\tau} = \frac{\vec{v}}{c} \cdot \vec{F}$$

$\downarrow \vec{F} \cdot \vec{v}$ is power from classical mechanics,

$$\equiv \frac{1}{c} \frac{d}{d\tau} E$$

$$= \frac{d}{d\tau} \left(\frac{E}{c} \right)$$

Hence, we identify p_0 as E/c