

## Einstein's Postulates

1. **The principle of relativity.** The laws of physics apply in all inertial reference systems.
2. **The universal speed of light.** The speed of light in vacuum is the same for all inertial observers, regardless of the motion of the source.

## Lorentz Transformations

**Galilean transformations.** If we “start the clock” ( $t = 0$ ) at the moment the origins ( $\mathcal{O}$  and  $\bar{\mathcal{O}}$ ) coincide, then at time  $t$ ,  $\bar{\mathcal{O}}$  will be a distance  $vt$  from  $\mathcal{O}$ , and hence

$$x = d + vt$$

where  $d$  is the distance from  $\bar{\mathcal{O}}$  to  $\bar{A}$  at time  $t$ . Classically,  $d = \bar{x}$  and thus constructed the Galilean transformations:

$$\begin{aligned}\bar{x} &= x - vt \\ \bar{y} &= y \\ \bar{z} &= z \\ \bar{t} &= t\end{aligned}$$

**Lorentz transformations.** In relativity, however, there will be a modification for Galilean transformations. As mentioned earlier,  $d$  is the distance from  $\bar{\mathcal{O}}$  to  $\bar{A}$ , with the addition that it is measured by  $S$ .  $\mathcal{O}$  to  $\bar{A}$  measured by  $\bar{S}$ , however, is  $\bar{x}$  and is at rest. Therefore

$$\bar{x} = \gamma d$$

When this is inserted in Galilean Transformations, we obtain the relativistic version:

$$\bar{x} = \gamma(x - vt).$$

Of course, we could have run the same argument from the point of view of  $\bar{S}$ . The right diagram looks similar, but in this case it depicts the scene

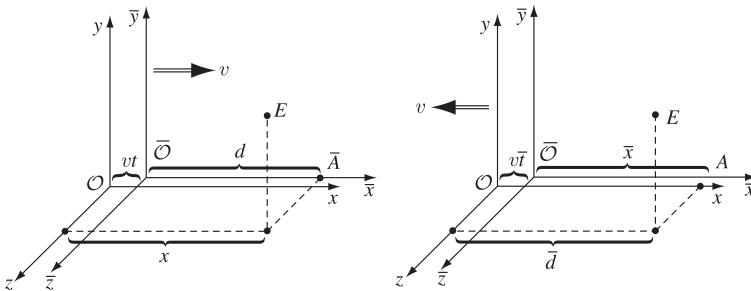


Figure: Reminder that in relativity there is no absolute velocity.  $\bar{S}$  slides along the  $x$  axis at speed  $v$  (with  $S$  being at rest) is the same as  $S$  slides along the  $x$  axis at speed  $-v$  (with  $\bar{S}$  being at rest). Of course that is assuming both systems are inertial.

at time  $\bar{t}$ , whereas left diagram showed the scene at time  $t$ . If we assume that  $\bar{S}$  also starts its clock when the origins coincide, then at time  $\bar{t}$ ,  $O$  will be a distance  $v\bar{t}$  from  $\bar{O}$ , and therefore

$$\bar{x} = \bar{d} - v\bar{t}$$

Similar as before,  $\bar{d}$  is the distance from  $O$  to  $A$  in  $\bar{S}$ , whereas  $x$  is the distance from  $O$  to  $A$  in  $S$  and is at rest. Therefore

$$x = \gamma\bar{d}$$

It follows that

$$x = \gamma(\bar{x} + v\bar{t})$$

This last equation obviously be identical to the formula for  $\bar{x}$ , except for a switch in the sign of  $v$ . Nevertheless, this is a useful result, for if we substitute  $\bar{x}$ , and solve for  $\bar{t}$ , we complete the relativistic “dictionary”:

$$\begin{aligned}\bar{x} &= \gamma(x - vt) \\ \bar{y} &= y \\ \bar{z} &= z \\ \bar{t} &= \gamma\left(t - \frac{v}{c^2}x\right)\end{aligned}$$

These are the famous Lorentz transformations. The reverse dictionary, which carries you from  $\bar{S}$  back to  $S$ , can be obtained algebraically by solving for Lorentz transformations  $x$  and  $t$ , or, more simply, by switching the sign of  $v$ :

$$\begin{aligned}x &= \gamma(\bar{x} + v\bar{t}) \\ y &= \bar{y} \\ z &= \bar{z} \\ t &= \gamma\left(\bar{t} + \frac{v}{c^2}\bar{x}\right)\end{aligned}$$

## Relativity Geometry

**The relativity of simultaneity.** Imagine a freight car, traveling at constant speed along a smooth, straight track; someone then switches the lamp on, and the light spreads out in all directions at speed  $c$ . The two events in question occur simultaneously. However, to an observer on the ground these same two events are not simultaneous. For the beam going to the back end has a shorter distance to travel than the one going forward. According to this observer, therefore, event (b) happens before event (a).

Two events that are simultaneous in one inertial system are not, in general, simultaneous in another.

Using Lorentz transformations we could also come to the sam conclusion. Suppose event A occurs at  $x_A = 0$ ,  $t_A = 0$ , and event B occurs at  $x_B = b$ ,  $t_B = 0$ . The two events are simultaneous in  $S$ . But they are not simultaneous in  $\bar{S}$ , for the Lorentz transformations give  $\bar{x}_A = 0$ ,  $\bar{t}_A = 0$  and  $\bar{x}_B = \gamma x$ ,  $\bar{t}_B = -\gamma(v/c^2)b$ . According  $\bar{S}$  then, B occurs before A, as before.

**Time dilation.** The time it takes the light to leave the bulb and strikes the floor of the car directly below is

$$\Delta\bar{t} = \frac{h}{c}$$

For an observer, however

$$\Delta t = \frac{\sqrt{h^2 + (v\Delta t)^2}}{c}$$

Solving for  $\Delta t$ , we have

$$\Delta t = \frac{h}{c} \frac{1}{\sqrt{1 - v^2/c^2}}$$

therefore

$$\begin{aligned}\Delta\bar{t} &= \sqrt{1 - v^2/c^2} \Delta t \\ &= \frac{1}{\gamma} \Delta t\end{aligned}$$

This is, by the way, how most textbook introduce constant gamma in relativity. Gamma defined as

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

To derive using Lorentz transformations, Suppose that at time  $t = 0$  observer  $S$  decides to examine all the clocks in  $\bar{S}$ . He finds that they read different times, depending on their location

$$\bar{t} = -\gamma \frac{v}{c^2} x$$

Those to the left of the origin (negative  $x$ ) are ahead, and those to the right are behind.  $S$  focuses his attention on a single clock at rest in the  $\bar{S}$  frame, and watches it over some interval  $\Delta t$ . Because  $\bar{x}$  is fixed, the Reverse Lorentz transformations gives

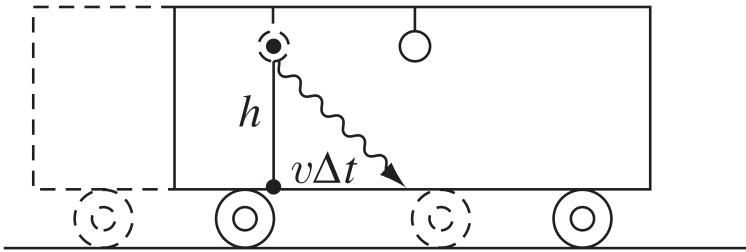
$$\Delta\bar{t} = \frac{1}{\gamma} \Delta t$$

**Lorentz contraction.** To an observer on the train, the time it takes for light signal to be sent down and back

$$\Delta\bar{t} = 2 \frac{\Delta\bar{x}}{c}$$



Figure: Train, freight car even



How long does it take the light to make this trip?

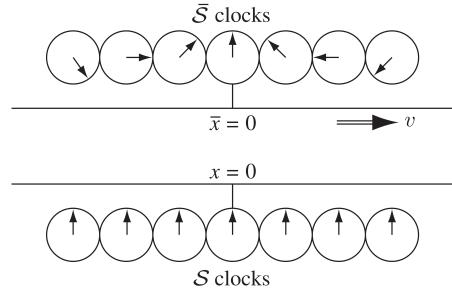


Figure: Clocks that are properly synchronized in one system will not be synchronized when observed from another system

If  $\Delta t_1$  is the time for the light signal to reach the front end, and  $\Delta t_2$  is the return time, then to an observer on the ground

$$\Delta t_1 = \frac{\Delta x + v\Delta t_1}{c} \quad \Delta t_2 = \frac{\Delta x - v\Delta t_2}{c}$$

solving for  $\Delta t_1$  and  $\Delta t_2$

$$\Delta t_1 = \frac{\Delta x}{c-v} \quad \Delta t_2 = \frac{\Delta x}{c+v}$$

So the round-trip time is

$$\Delta t = 2 \frac{\Delta x}{c} \frac{1}{1 - v^2/c^2}$$

Applying time dilation formula and solving for  $\Delta x$

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

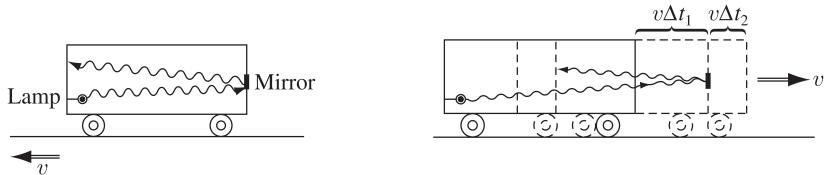


Figure: Railway carriage, if you will

Now, Imagine a stick at rest in  $\bar{S}$ . Its rest length is  $\Delta\bar{x}$ . If an observer in  $S$  were to measure the stick, he would subtract the positions of the two ends at one instant of his time  $t$ . In another word,  $t$  is fixed and Lorentz transformations give

$$\Delta\bar{x} = \gamma\Delta x$$

**Einstein's velocity addition rule.** Suppose a particle moves a distance  $dx$  (in  $S$ ) in a time  $dt$ . Its velocity  $u$  is then

$$u = \frac{dx}{dt}$$

In  $\bar{S}$ , meanwhile, it has moved a distance

$$d\bar{x} = \gamma(dx - vdt)$$

in a time

$$d\bar{t} = \gamma(dt - \frac{v}{c^2}dx)$$

The velocity in  $\bar{S}$  is therefore

$$\bar{u} = \frac{d\bar{x}}{d\bar{t}} = \frac{u - c}{1 - uv/c^2}$$

To recover the more transparent notation, let  $A$  be the particle,  $B$  be  $S$ , and  $C$  be  $\bar{S}$ ; then  $u = v_{AB}$ ,  $\bar{u} = v_{AC}$ , and  $v = v_{CB} = -v_{BC}$ , so

$$v_{AC} = \frac{v_{AB} + v_{BC}}{1 - +(v_{AB}v_{BC}/c^2)}$$

## Structure of Spacetime

**Four-vectors.** The Lorentz transformations take on a simpler appearance when expressed in terms of the quantities

$$x^0 \equiv ct \quad \beta \equiv \frac{v}{c}$$

If, at the same time, we number the  $x, y, z$  coordinates, so that

$$x^1 \equiv x \quad x^2 \equiv y \quad x^3 \equiv z$$

then the Lorentz transformations read

$$\begin{aligned}\bar{x}^0 &= \gamma(x^0 - \beta x^1) \\ \bar{x}^1 &= \gamma(x^1 - \beta x^0) \\ \bar{x}^2 &= x^2 \\ \bar{x}^3 &= x^3\end{aligned}$$

Or, in matrix form:

$$\begin{pmatrix} \bar{x}^0 \\ \bar{x}^1 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

Letting Greek indices run from 0 to 3, this can be distilled into a single equation:

$$\bar{x}^\mu = \sum_{v=0}^3 \Lambda_v^\mu x^v$$

where is  $\Lambda$  the Lorentz transformation matrix in (the superscript  $\mu$  labels the row, the subscript  $v$  labels the column). (3-d) vector defined as any set of three components that transform under rotations the same way ( $x, y, z$ ) do; by extension, we now define a 4-vector as any set of four components that transform in the same manner as  $(x_0, x_1, x_2, x_3)$  under Lorentz transformations:

$$\bar{a}^\mu = \sum_{v=0}^3 \Lambda_v^\mu a^v$$

There is a 4-vector analog to the dot product, but it's not just the sum of the products of like components; rather, the zeroth components have a minus sign

$$\mathbf{A} \cdot \mathbf{B} = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3$$

It has the same value in all inertial systems; just as the ordinary dot product is invariant (unchanged) under rotations, this combination is invariant under Lorentz transformations.

It is convenient to introduce the covariant vector  $a^\mu$ , which differs from the contravariant  $a^\mu$  only in the sign of the zeroth

$$a^\mu = (a_0, a_1, a_2, a_3) \equiv (-a^0, a^1, a^2, a^3)$$

Upper indices designate contravariant vectors; lower indices are for covariant vectors. Raising or lowering the temporal index costs a minus sign; raising or lowering a spatial index changes nothing. Formally,

$$a_\mu = \sum_{v=0}^3 g_{\mu v} a^v \quad q \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The scalar product can now be written with the summation symbol

$$\sum_{\mu=0}^3 a^\mu b_\mu$$

or, more compactly still

$$a^\mu b_\mu$$

This is called the Einstein summation convention, after its inventor, who regarded it as one of his most important contributions. Of course, we could just as well take care of the minus sign by switching to covariant b

$$a^\mu b_\mu = a_\mu b^\mu = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3$$

**The invariant interval.** The scalar product of a 4-vector with itself,

$$a^\mu a_\mu = -(a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2$$

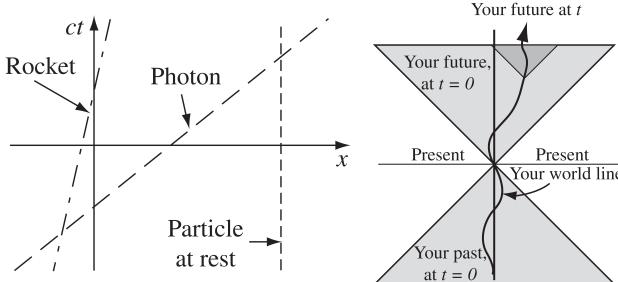


Figure: Minkowski diagrams

can be positive (if the “spatial” terms dominate) or negative (if the “temporal” term dominates) or zero:

- If  $a^\mu a_\mu > 0$ ,  $a^\mu$  is called **spacelike**
- If  $a^\mu a_\mu < 0$ ,  $a^\mu$  is called **timelike**
- If  $a^\mu a_\mu = 0$ ,  $a^\mu$  is called **lightlike**

Suppose event A occurs at  $(x_A^0, x_A^1, x_A^2, x_A^3)$  and event B at  $(x_B^0, x_B^1, x_B^2, x_B^3)$ . The difference

$$\Delta x^\mu \equiv x_A^\mu - x_B^\mu$$

is the displacement 4-vector. The scalar product of  $\Delta x^\mu$  with itself is called the invariant interval between two events:

$$\begin{aligned} I &\equiv (\Delta x)^\mu (\Delta x)_\mu = -(\Delta x^0)^2 + (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 \\ &= -c^2 t^4 + d^2 \end{aligned}$$

If the displacement between two events is timelike ( $I < 0$ ), there exists an inertial system (accessible by Lorentz transformation) in which they occur at the same point. On the other hand, if the displacement is spacelike ( $I > 0$ ), then there exists a system in which the two events occur at the same time. And if the displacement is lightlike ( $I = 0$ ), then the two events could be connected by a light signal.

**Space-time diagrams.** On such a Minkowski diagrams, position plotted horizontally and time (or, better,  $x^0 = ct$ ) vertically. Velocity is then given by the reciprocal of the slope. A particle at rest is represented by a vertical line; a photon, traveling at the speed of light, is described by a 45° line; and a rocket going at some intermediate speed follows a line of slope  $c/v = 1/\beta$ .

The trajectory of a particle on a Minkowski diagram is called a world line. Because no material object can travel faster than light, your world line can never have a slope less than 1. Accordingly, your motion is restricted to the wedge-shaped region bounded by the two 45° lines. We call this your “future,” in the sense that it is the locus of all points accessible to you. As time goes on, and you move along your chosen world line, your options progressively narrow. Meanwhile, the backward wedge represents your “past,” in the sense that it is the locus of all points from which you might have come. As for the rest, this is the generalized “present.” There’s no way you can influence any event in the present; it’s a vast expanse of

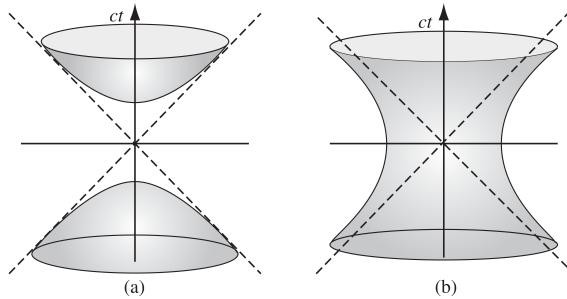


Figure: Hypercones

spacetime that is absolutely inaccessible to you. If we include a  $y$  axis coming out of the page, the “wedges” become cones—and, with an undrawable  $z$  axis, hypercones. Because their boundaries are the trajectories of light rays, we call them the forward light cone and the backward light cone. Your future, in other words, lies within your forward light cone, your past within your backward light cone.

Under Lorentz transformations, the interval  $I$  is preserved, and the locus of all points with a given value of  $I$  is a hyperbola—or, if we include the  $y$  axis, a hyperboloid of revolution. When the displacement is timelike, it’s a “hyperboloid of two sheets”; when the displacement is spacelike, it’s a “hyperboloid of one sheet.” When you perform a Lorentz transformation, the coordinates  $(x, t)$  of a given event will change to  $(\bar{x}, \bar{t})$ , but these new coordinates will lie on the same hyperbola as  $(x, t)$ . No amount of transformation will carry it, say, from the upper sheet of the timelike hyperboloid to the lower sheet, or to a spacelike hyperboloid.

If the displacement 4-vector between two events is timelike, their ordering is absolute; if the interval is spacelike, their ordering depends on the inertial system from which they are observed. In terms of the space-time diagram, an event on the upper sheet of a timelike hyperboloid definitely occurred after  $(0, 0)$ , and one on the lower sheet certainly occurred before; but an event on a spacelike hyperboloid occurred at positive  $t$ , or negative  $t$ , depending on your reference frame. Conclusion: The displacement between causally related events is always timelike, and their temporal ordering is the same for all inertial observers.

Here’s some neat summary

| Timelike separation   | Spacelike separation  | Lightlike separation   |
|---|---|--|
| $I < 0$<br>$a^\mu a_\mu < 0$<br>The two events may or may not have a causal connection. | $I > 0$<br>$a^\mu a_\mu > 0$<br>The two events cannot have a causal connection. | $I = 0$<br>$a^\mu a_\mu = 0$<br>The two events may or may not have a causal connection. (citation needed!) |

|   |  |   |
|---|--|---|
| All observers will agree about the order in which the two events occurred.            | Different observers will disagree about the order in which the two events occurred.  | Different observers will disagree about the distance between them and the time between them, but everyone will agree that the distance equals the time multiplied by c.<br>(source incomplete!) |
| There will be one reference frame in which the two events happened at the same place. | There will be one reference frame in which the two events happened at the same time. |   |

## Proper Time and Proper Velocity

As you progress along your world line, your watch runs slow; while the clock on the wall ticks off an interval  $dt$ , your watch only advances  $d\tau$ :

$$d\tau = \sqrt{1 - u^2/c^2} dt$$

The time  $\tau$  your watch registers (or, more generally, the time associated with the moving object) is called proper time. I'll use  $u$  for the velocity of a particular object—you, in this instance—and reserve  $v$  for the relative velocity of two inertial systems. The velocity

$$\mathbf{u} = \frac{d\mathbf{l}}{dt}$$

relative to ground (displacement divided by the time), both  $d\mathbf{l}$  and  $dt$  are to be measured by the ground observer. However there is also distance covered per unit proper time

$$\boldsymbol{\eta} \equiv \frac{d\mathbf{l}}{d\tau}$$

This hybrid quantity (distance measured on the ground, over time measured by *you*) is called proper velocity.

$$\boldsymbol{\eta} = \frac{1}{\sqrt{1 - u^2/c^2}} \mathbf{u}$$

$\eta$  is the spatial part of a 4-vector

$$\eta \equiv \frac{dx^\mu}{d\tau}$$

whose zeroth component is

$$\eta^0 = \frac{dx^0}{d\tau} = \frac{d}{d\tau} ct = \frac{c}{\sqrt{1 - u^2/c^2}}$$

Thus, for instance, when you go from system  $S$  to system  $\bar{S}$ , moving at speed  $v$  along the common  $x\bar{x}$  axis,

$$\begin{aligned}\bar{\eta}^0 &= \gamma(\eta^0 - \beta\eta^1) \\ \bar{\eta}^1 &= \gamma(\eta^1 - \beta\eta^0) \\ \bar{\eta}^2 &= \eta^2 \\ \bar{\eta}^3 &= \eta^3\end{aligned}$$

More generally

$$\bar{\eta}^\mu = \Lambda_v^\mu \eta^v$$

$\eta^\mu$  is called the proper velocity 4-vector, or simply the 4-velocity. By contrast, the transformation rule for ordinary velocities  $\mathbf{u}$  is quite cumbersome. The reason for the added complexity is plain: we're obliged to transform both the numerator  $d\mathbf{l}$  and the denominator  $dt$ , whereas for proper velocity, the denominator  $d\tau$  is invariant, so the ratio inherits the transformation rule of the numerator alone.

$$\begin{aligned}\bar{u}_x &= \frac{d\bar{x}}{d\bar{t}} = \frac{u_x - v}{1 - vu_x/c^2} \\ \bar{u}_y &= \frac{d\bar{y}}{d\bar{t}} = \frac{u_y}{\gamma(1 - vu_x/c^2)} \\ \bar{u}_z &= \frac{d\bar{z}}{d\bar{t}} = \frac{u_z}{\gamma(1 - vu_x/c^2)}\end{aligned}$$

**Relativistic Energy and Momentum.** Relativistic momentum of an object of mass  $m$  traveling at (ordinary) velocity  $\mathbf{u}$

$$p \equiv m\eta = \frac{m\mathbf{u}}{\sqrt{1 - u^2/c^2}}$$

Relativistic momentum is the spatial part of a 4-vector

$$p^\mu \equiv m\eta^\mu$$

with temporal component

$$p^0 = \frac{mc}{\sqrt{1 - u^2/c^2}}$$

Einstein identified  $p^0c$  as relativistic energy

$$E \equiv \frac{mc^2}{\sqrt{1 - u^2/c^2}}$$

Notice that the relativistic energy is nonzero even when the object is stationary; we call this rest energy:

$$E_{\text{rest}} \equiv mc^2$$

The remainder, which is attributable to the motion, is kinetic energy

$$E_{\text{kin}} \equiv E - mc^2 = mc^2 \left( \frac{1}{\sqrt{1 - u^2/c^2}} - 1 \right)$$

So far, this is all just notation. The physics resides in the experimental fact:

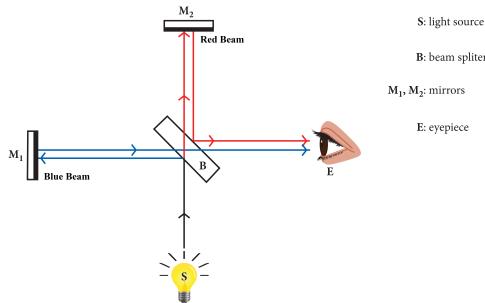


Figure 1: Michelson-Morley Experiment

In every closed system, the total relativistic energy and momentum are conserved.

Mass, however, is not conserved. Specifically , mass is invariant but not conserved; energy is conserved but not invariant; electric charge is both conserved and invariant; velocity is neither conserved nor invariant. Note the distinction between an invariant quantity (same value in all inertial systems) and a conserved quantity (same value before and after some process).

The scalar product of  $p^\mu$  with itself is

$$p^\mu p_\mu = -(p^0)^2 + (\mathbf{p} \cdot \mathbf{p}) = -m^2 c^2$$

In terms of the relativistic energy and momentum

$$E^2 - p^2 c^2 = m^2 c^4$$

For massless and relativistic particle

$$E = pc$$

**Michelson-Morley Experiment.** You already know the story of aether wind and how this experiment disprove that model, so i'll omit that and focus on this experiment itself. The Michelson interferometer is represented in Figure 1. Light from a single source reaches a half-silvered mirror and splits into two beams traveling at right angles to each other. The beams reach mirrors that send them back to the splitter, where they travel together to an eyepiece. There are three possibilities that determine whether or not the two beams arrive in phase at the end of the eyepiece:

1. If the two paths are identical in length, and the two beams travel identical speeds, then the two beams are in phase, and the interference is constructive.
2. If the two paths are different lengths, then the beams may therefore arrive out of phase.
3. If the two paths are identical lengths but are traveled at different speeds, then the two beams may be out of phase.

Now suppose the Earth is traveling along the horizontal axis from  $M_1$  to  $B$ . The aether model predicts that for part of the trip the blue beam is going faster, and for part of the trip the red beam is going faster. But

these two effects do not fully cancel out, and even if the two path lengths are identical, the beams will arrive out of phase.

After measuring the type of interference they observed, they rotated the entire apparatus. A uniform rotation does not change the geometry of the experiment; whatever phase change was caused by differences in path length is still exactly the same as it was. But a phase change caused by differences in speed, due to orientation with respect to the aether wind, will change.

So the experiment was not to see whether the two beams constructively or destructively interfered on the wall, since such interference might have multiple causes. Rather, they looked to see how the interference pattern changed as they rotated the device.

The result is, as you know it, their experiment found no evidence at all that the two beams had traveled at different speeds. Michelson wrote in a letter that the result was “decidedly negative.” Even allowing for measurement errors, his experiment seemed to show that “the relative velocity [of the Earth and the aether] is less than one sixth of the Earth’s velocity.”