# Special Relativity

Asaf Pe'er[1]

December 4, 2013

## 1. The Michelson - Morley experiment

[1]

By the late 1800's, the **wave nature of light** was well established, with diffraction experiments carried by *Thomas Young*, and polarization experiment carried by *Augustin-Jean Fresnel*. However, a question was raised: as sound waves, or pressure waves propagate in *medium* (e.g., air, water), **what is the medium in which light waves propagate ?** While there was no observational evidence for such a medium, in analogy to sound or pressure waves, it was **hypothesized** that such a medium must exist. This medium was called **luminiferous aether**.

Clearly, as we see light coming from the sun and the stars, as well as in vacuum chambers, this medium must exist *everywhere*. Moreover, as we observer no friction in the motion, e.g., of the earth through the aether, it must be very light, with unique properties that interest late 19th century physicists.

The question at hand was not whether aether existed; this was considered to be "obvious". The question was what is the "drag" between the earth and the surrounding aether. It was concluded that as the earth rotates around the sun, there must be **relative motion** between earth and aether, which would look to an observer on earth as **"aether wind"**. At any given point on the Earth's surface, the magnitude and direction of the wind would vary with time of day and season (see Figure 1).

By mid-19th century, it was known that the earth moves around the sun at $\approx 30\,\mathrm{km\,s^{-1}}$, and that the speed of light is finite and equals to $c \approx 300,000\,\mathrm{km\,s^{-1}}$; this was measured by *Fizeau* and *Foucault* (the same guy from Foucault pendulum !). The measurement, using Fizeau-Foucault apparatus is schematically illustrated in Figure 2.

As the light (is assumed to) propagate through the aether, and as earth moves through the same aether, it was thought that the relative motion of the earth and aether could be

[1]Physics Dep., University College Cork

[1]This section is from **Berkeley Physics Course Vol I - Mechanics**, by Kittel, Knight and Helmholz.

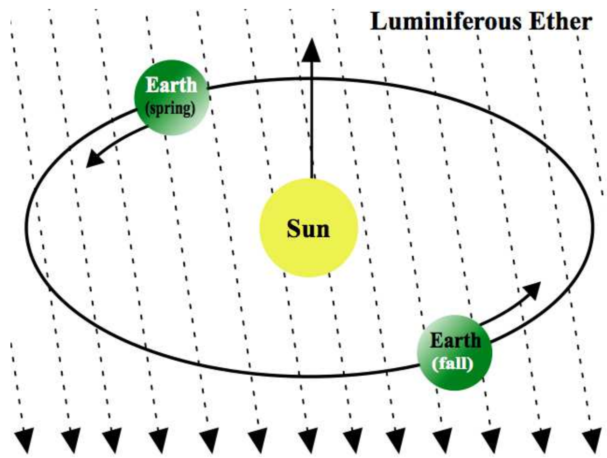


Fig. 1.— As the earth rotates around the sun, it was thought that it should be exposed to aether "wind", who's magnitude and direction vary with time of day and season.

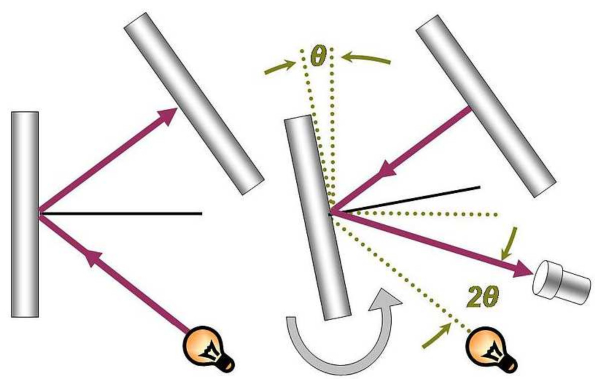


Fig. 2.— Fizeau-Foucault apparatus was used to measure the speed of light in 1850. The apparatus involves light reflecting off a rotating mirror, toward a stationary mirror about 35 kilometers away. As the rotating mirror will have moved slightly in the time it takes for the light to bounce off the stationary mirror (and return to the rotating mirror), it will thus be deflected away from the original source, by a small angle. Since the deflection angle is $\theta = (d\theta/dt)(2h/c)$ where $h$ is the distance between the mirrors, the speed of light $c$ is given by $c = (d\theta/dt)(2h/\theta)$.

measured by analyzing the return speed of light in different directions at various different times. However, until the end of the 19th century, no experience was accurate enough to measure such tiny differences, as expected from Galilean transformation ($c + v_{earth-aether}$ vs $c - v_{earth-aether}$).

*Albert Michelson* invented the **Michelson interferometer**, the first device that was sensitive enough to measure these differences. An interference pattern is produced by splitting a beam of light through a half-silvered mirror into two paths, bouncing the beams back and recombining them (see Figure 3). The recombined beam shows a pattern of constructive and destructive interference whose transverse displacement would depend on the relative time it takes light to transit the longitudinal vs. the transverse arms.

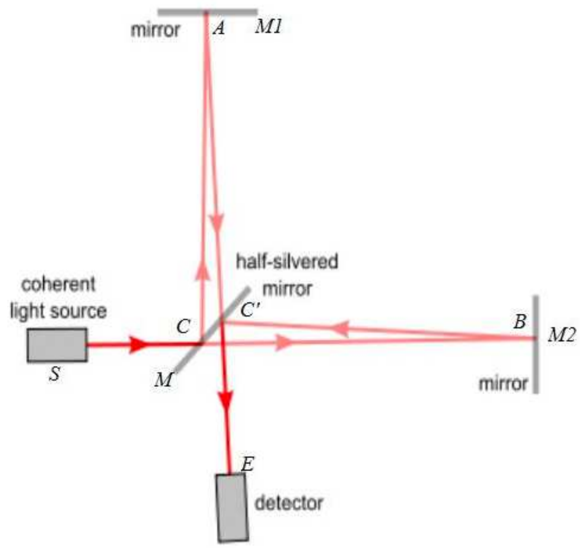


Fig. 3.— Michelson interferometer is composed of a beam splitter and two reflective mirrors. The recombined beam shows interference pattern, which depends on the relative time it takes the two light beams to transit the two arms.

Let us assume that the distance between the splitting mirror and each of the beams is $L$. According to Galilean transformation, the time it takes the longitudinal beam to cross back and forth is

$$T_1 = \frac{L}{c-v} + \frac{L}{c+v} = \frac{2L}{c}\frac{1}{1-\frac{v^2}{c^2}}, \tag{1}$$

where $v$ is the earth's velocity and $c$ is the speed of light.

In the transverse direction: by the time the light gets to the mirror, it moves perpendicular to the light direction by $vT$. Thus, the length the light ray travels is $\sqrt{L^2+(vT)^2}$. Let us denote the time the light travel to the mirror and back by $T_2$. Therefore, $T = T_2/2$, and the time it takes to travel (back and forth) is $T_2 = (2\sqrt{L^2+(vT_2/2)^2}/c)$, or

$$T_2 = \frac{2L}{\sqrt{c^2-v^2}} = \frac{2L}{c}\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{2}$$

The time difference is thus

$$\Delta t = T_1 - T_2 \approx \frac{2L}{c}\left(1 + \frac{v^2}{c^2} - 1 - \frac{v^2}{2c^2}\right) = \left(\frac{L}{c}\right)\left(\frac{v^2}{c^2}\right). \tag{3}$$

Such a time difference should be enough to produce a noticeable interference pattern. This pattern should vary with the time in day, season, etc. **No such pattern was ever observed. The inevitable conclusion is that the speed of light $c$ is independent on the motion of the source or the observer. The aether does not exist. Galilean transformation fails.**

-##-

## 2. The principle of relativity and its consequences

The null result of the Michelson-Morley experiment implies that **the speed of light is independent of the motion of the light source or receiver**. Namely, the speed of light is the same in all reference frames in uniform motion with respect to the source.

We further note that the universe (on a large scale) is *homogeneous* and *isotropic*, namely has no preferred direction. These facts had led *Einstein* to postulate **the principle of relativity**, which states that:

*The laws of physics have the same form in all inertial reference frames*

This implies that both the *form* of the laws and the *numerical values* of the physical constants are the same in all inertial frames. All the vast consequences of the special theory of relativity follow from this postulate.

Let us consider two observers, denoted by $O$ and $O'$. Assume that the observer $O'$ is moving (at constant velocity) with respect to the observer $O$, so that both frames are inertial. We denote by $x, y, z, t$ the (space-time) coordinate system of observer $O$, and by $x', y', z', t'$ the coordinate system of observer $O'$. We further assume that the coordinate origins, $O$ and $O'$ coincide at $t = t' = 0$.

Consider a light source located at the origin in the frame $O$ that emits light at $t = 0$. At time $t$, the spherical wave front of the light is at location

$$x^2 + y^2 + z^2 = c^2t^2. \tag{4}$$

A similar analysis holds for the observer $O'$, thus

$$x'^2 + y'^2 + z'^2 = c^2 t'^2. \tag{5}$$

We can assume (without loss of generality), that $O'$ moves with respect to $O$ in the $\hat{x}$ direction. Thus, $x' \neq x$, while $y' = y$ and $z' = z$. Equations 4, 5 thus imply that $t \neq t'$. Thus, **time separation is not the same in both frames**. In other words, *time is not an* **invariant** *quantity.*

While the time separation between two events (e.g.: emission of photon and detection of the photon) is not invariant, Equations 4 and 5 imply that one can define an **invariant** quantity, which is called the **interval** between two events. The interval is defined by

$$(interval)^2 \equiv (c\Delta t)^2 - (\Delta r)^2 = c^2(\Delta t^2) - \left((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2\right), \tag{6}$$

where $\Delta t$ is the time interval between two events, and $\Delta r$ is the space interval between the same events.

Example. Consider a rocket that moves at velocity $\vec{v} = v\hat{x}$ in the lab frame ($O$). At time $t = t' = 0$, a flash of light is emitted in the rocket, in the $y'$ direction; this is *event A*. It travels a distance $d$, reflected by a mirror and detected at the origin - this is *event B*.

In the *rocket* frame, the origin $x'$ is at rest; thus $\Delta x' = 0$. The time it takes the light ray to travel is $t' = 2d/c$. Thus the *interval* between events $A$ and $B$ is $\sqrt{c^2 t'^2 - \Delta x'^2} = 2d$.

In the *lab* frame, the rocket moves from $x = 0$ to $x$ by the time the light ray was detected. The distance traveled by the light ray to the mirror is thus $\sqrt{d^2 + (x/2)^2}$, and the time it took it is to travel back and forth is $t = (2/c)\sqrt{d^2 + (x/2)^2}$. The interval is $\sqrt{4\left(d^2 + (x/2)^2\right) - x^2} = 2d$.

|One can discriminate between three cases: (i) *spacelike intervals*, for which $x > ct$; (ii) *lightlike intervals*, for which $x = ct$; and (iii) *timelike intervals*, for which $x < ct$.

All massive particles travel at velocities $v < c$, hence events along their path are separated by timelike intervals.

One can define **proper time**, $\tau$ as the time measured by a clock that passes through both events (namely, as measured in the moving frame, or the *rest frame* of the studied object).

$$\Delta\tau = \sqrt{(\Delta t)^2 - \frac{(\Delta r)^2}{c^2}}. \tag{7}$$

Note that:
(i) in timelike intervals, $\Delta\tau$ is always positive.

(ii) In all other frames, the time interval between events is always **longer** than the proper time, since $\Delta t = \sqrt{(\Delta\tau)^2 + (\Delta r)^2/c^2}$, and $\Delta r > 0$.
(iii) Time and space are connected: if one of the coordinates, $(x, y, z$ or $t)$ has a different value when changing frames, the other coordinates must also have different values in the different frames, so that the invariance of the interval is maintained. Thus, one needs to think of problems in terms of *four dimension* **space-time coordinates**.

-##-

## 3. Lorentz transformation

As we have seen in §1, Galilean transformation, which assumes $t = t'$ fails, being inconsistent with the principle of relativity.

We thus seek a new transformation. This transformation should conserve the *interval*, defined by Equation 6. As before, we assume, without loss of generality, that the frame $O'$ moves in the $x$ direction as seen by an observer in the frame $O$, at constant velocity, $v$. At $t = t' = 0$, $x = x' = 0$. We know that in the limit $v \ll c$, this transformation should reduce to the Galilean transformation. In the most general form, the (linear) transformation equations are

$$\begin{aligned} t' &= Ax + Bt; \\ x' &= Cx + Dt, \end{aligned} \tag{8}$$

where $A$, $B$, $C$ and $D$ are constants for a given velocity $v$; moreover, $y = y'$ and $z = z'$.

We know that for $x' = 0$, $dx/dt = v$. Thus, $Cx = -Dt$, or $v = -D/C$. Similarly, for $x = 0$, $dx'/dt' = -v$. Eliminating $t$ from equation 7 by multiplying the upper Equation 8 by $D$ and the lower Equation by $B$ and subtracting, gives

$$\begin{aligned} Dt' &= ADx + BDt; \\ Bx' &= BCx + BDt; \\ \rightarrow Dt' - Bx' &= (AD - BC)x \end{aligned} \tag{9}$$

For $x = 0$, Equation 9 implies $D/B = -v$. Combined, we get

$$C = B. \tag{10}$$

We can now plug these results in Equation 5 and require a form identical to Equation 4,

$$\begin{aligned} x'^2 + y'^2 + z'^2 &= c^2 t'^2 \\ (Bx + Dt)^2 + y^2 + z^2 &= c^2 (Ax + Bt)^2 \\ B^2x^2 + 2BDxt + D^2t^2 + y^2 + z^2 &= c^2(A^2x^2 + 2ABxt + B^2t^2). \end{aligned} \tag{11}$$

We see that Equation 11 can become identical to Equation 4, if

$$\begin{array}{lcl} 2BD & = & 2c^2AB, \\ B^2 - c^2A^2 & = & 1, \\ c^2B^2 - D^2 & = & c^2. \end{array} \quad (12)$$

Using $D = -vB$ in the last line of Equation 12, we get

$$c^2B^2 - v^2B^2 = c^2 \rightarrow B^2 = \frac{c^2}{c^2 - v^2} = \frac{1}{1 - \frac{v^2}{c^2}}; \quad (13)$$

from which we get

$$\begin{array}{lcl} B & = & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}; \\ C = B & = & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}; \\ D = -vB & = & \frac{-v}{\sqrt{1-\frac{v^2}{c^2}}}; \\ A = D/c^2 & = & \frac{-v/c^2}{\sqrt{1-\frac{v^2}{c^2}}} \end{array} \quad (14)$$

It is common to define

$$\begin{array}{l} \beta \equiv \frac{v}{c} \\ \gamma \equiv \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \end{array} \quad (15)$$

With these definitions, using Equations 14 in Equations 8, we finally derive the **Lorentz transformations**,

$$\begin{array}{lcl} t' & = & \gamma \left(t - \frac{vx}{c^2}\right) \\ x' & = & \gamma \left(x - vt\right) \\ y' & = & y \\ z' & = & z \end{array} \quad (16)$$

Equation 16 describe the relativistic transformation law of position and time between two frames.

Note that for every velocity $0 < v < c$, $\gamma > 1$. For $v \rightarrow c$, $\gamma \rightarrow \infty$. However, when $v \rightarrow 0$, $\gamma \rightarrow 1$, and Galilean transformation is retrieved. This can be seen by expanding $\gamma$ in a Taylor series,

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \mathcal{O}(v/c)^6. \quad (17)$$

Thus, for velocities $v \ll c$, the modification to the Galilean transformation law is of the order $(v/c)^2$.

-##-

## 4. Fitzgerald-Lorentz contraction

Consider a ruler that lies in the $x$ direction, attached to a rocket that moves in velocity $\vec{v}$ along the $x$ axis, as seen in the lab frame. The length of the ruler in the lab frame is $l = x_2 - x_1$. Its length in the rocket frame is

$$l' = x'_2 - x'_1. \tag{18}$$

Using Lorentz transformation (Equation 16), we find

$$l' = \gamma(x_2 - vt_2) - \gamma(x_1 - vt_1). \tag{19}$$

If the length measurements of $x_1$ and $x_2$ are made at the same time *in the lab frame*, then $t_2 = t_1 = t$, and one gets

$$\begin{aligned} l' = \gamma(x_2 - x_1) &= \gamma l, \\ l = \tfrac{l'}{\gamma}. \end{aligned} \tag{20}$$

Since $\gamma \geq 1$, we find $l \leq l'$. Thus, **the length of a ruler as measured in a frame in which it is moving is *less* than its length as measured in a frame in which it is at rest.** This is known as *Fitzgerald-Lorentz (length) contraction.* Note that $y = y'$, $z = z'$, namely there is no contraction in direction perpendicular to the motion.

## 5. Time dilation

Consider a clock attached to the moving rocket. The clock ticks at $t'_1$ and $t'_2$, so that $\Delta t' = t'_2 - t'_1$. The time difference in the lab frame is

$$t_2 - t_1 = \gamma\left(t'_2 + \frac{vx'_2}{c^2}\right) - \gamma\left(t'_1 + \frac{vx'_1}{c^2}\right) = \gamma(t'_2 - t'_1), \tag{21}$$

since $x'_2 = x'_1 = 0$. Thus, $\Delta t = t_2 - t_1 > \Delta t'$. Thus , time is 'stretched' in the lab frame. This is known as **time dilation**. Clocks in moving frames runs quicker. Alternatively, the observed time is always longer than the **proper time**.

Observational evidence for that is the arrival of energetic muons to the earth surface. Muons are produced in the upper atmosphere by interactions of cosmic rays with particles in the earth's atmosphere. The muons have half lifetime of $\sim 1.5 \times 10^{-6}$ s, and thus we would not expect them to arrive to earth. However, having $\gamma \sim 30$, their observed decay time is a factor $\gamma$ longer, and they are observed.

### Example: the twin paradox.

The fact that two clocks don't show the same time can lead to seemingly paradoxes. One of the well known paradoxes is *the twin paradox*. Suppose one of a pair of twins becomes an astronaut, and travels to a nearby star (and back), at a velocity close to the speed of light, while his sister stays on earth. She sees his clock slowing, so by the time he returns, she is older than him ($t = \gamma t'$). However, a similar analysis from the astronaut's perspective shows that the clocks on earth slow, and thus the astronaut must be older.

A *qualitative* explanation is that as the astronaut has to switch directions when reaching the star, his frame is not inertial. Thus, the calculation as done on earth (assuming an inertial frame) is the correct one: the person on earth becomes older than the astronaut.

## 6. Transformation of velocities

Consider an object which moves at velocity $\vec{u} = u\hat{x}$, as seen in frame $O$. What is the velocity as seen by an observer in frame $O'$, which moves at velocity $\vec{v} = v\hat{x}$ ? We have seen that in Galilean transformation, the answer is simply $u' = dx'/dt' = dx/dt - v = u - v$. However, we have seen that Galilean transformation is incorrect.

Using Lorentz transformation, Equation 16, we derive

$$\begin{aligned} \frac{dx'}{dt} &= \frac{d[\gamma(x-vt)]}{dt} = \gamma\left(\frac{dx}{dt} - v\right) \\ \frac{dt'}{dt} &= \frac{d\left[\gamma\left(t-\frac{vx}{c^2}\right)\right]}{dt} = \gamma\left(1 - \frac{v}{c^2}\frac{dx}{dt}\right), \\ u'_x = \frac{dx'}{dt'} = \frac{\frac{dx'}{dt}}{\frac{dt'}{dt}} &= \frac{\frac{dx}{dt}-v}{1-\frac{v}{c^2}\frac{dx}{dt}} = \frac{u_x-v}{1-\frac{vu_x}{c^2}} \end{aligned} \tag{22}$$

Similarly, since $dt' \neq dt$, the transformation laws of $u_y$ and $u_z$ are

$$u'_y = \frac{dy'}{dt'} = \frac{\frac{dy'}{dt}}{\frac{dt'}{dt}} = \frac{u_y}{\gamma\left(1 - \frac{vu_x}{c^2}\right)}, \tag{23}$$

and a similar equation for $u'_z$.

As expected, we find that for $v \ll c$, the transformation law, Equations 22, 23 reduce to the familiar Galilean transformation.

Note that the transformation law **never** leads to velocities larger than $c$. Further, if $u_x = c$, Equation 22 gives $u'_x = -c$; the velocity of light is $c$ in any reference frame !.

-##-

## 7. Momentum in relativistic mechanics

The "classical" definition of momentum, $\vec{p} = m\vec{v}$, clearly **does not** hold in special relativity. Consider, for example, elastic collision between two balls of similar mass, $m$. We chose the lab frame, in which the balls approach each other at equal and opposite velocities: $\vec{v} = \pm v_x\hat{x} \pm v_y\hat{y}$. The total momentum is zero, both before and after the collision - there is no problem.

Consider now the same collision in a frame which moves at velocity $\vec{u} = v_x\hat{x}$. Assume that ball (1) moves with velocity $\vec{v}(1) = -v_x - v_y$. Using Equations 22, 23, the $x$ and $y$ components of the balls velocities before the collision are given by

$$\begin{aligned} v'_x(1) &= \frac{-v_x - u}{1 + \frac{v_x u}{c^2}} \\ v'_y(1) &= \frac{-v_y}{\gamma\left(1 + \frac{v_x u}{c^2}\right)} \\ v'_x(2) &= \frac{v_x - u}{1 - \frac{v_x u}{c^2}} = 0 \\ v'_y(2) &= \frac{v_y}{\gamma\left(1 - \frac{v_x u}{c^2}\right)} \end{aligned} \tag{24}$$

Since the denominators of the $y$-component of the velocities are not similar, the magnitude of the $y$-component of the velocity is not the same in $O'$, although it was similar in $O$. The change in the $y$-component of the momentum in the collision, $2m(v_y(2) - v_y(1))$ is clearly not zero.

The solution is to re-define the momentum, so that Newton's second law will hold. The 'origin' of the problem is that we used $dr/dt$; this suggests a solution, namely to define the momentum by

$$\vec{p} \equiv m\frac{d\vec{r}}{d\tau}, \tag{25}$$

where $\tau$ is the proper time.

We can write the relativistic momentum as a function of the object's velocity, $v = dx/dt$, as follows. The (inverse) Lorentz transformation (Equation 16) can be written in differential form, $dt = \gamma(dt' + vdx'/c^2)$. In a frame moving with the object, $dx' = 0$, $dt' = d\tau$; thus, we find $dt = \gamma d\tau$, hence $dt/d\tau = \gamma$.

Writing $d\vec{r}/d\tau = (d\vec{r}/dt)(dt/d\tau) = \gamma\vec{v}$, or

$$\vec{p} \equiv m\frac{d\vec{r}}{d\tau} = \gamma m\vec{v}. \tag{26}$$

In the limit $v \ll c$, $\gamma \rightarrow 1$, and we retrieve the "classical" definition of momentum. Furthermore, we can write $p^2 = \gamma^2 m^2 v^2 = \frac{m^2 v^2}{(1 - v^2/c^2)}$, or $v^2 = \frac{p^2/m^2}{1 + p^2/m^2c^2} = \frac{c^2}{1 + m^2c^2/p^2}$. As the denominator is always larger than unity, we get that $u < c$ for any massive particle, and $u = c$ for massless ($m = 0$) particles.

## 8. Four vectors, and the energy-momentum 4-vector

We have seen that an *event* can be described by 4 coordinates, $t, x, y, z$. When transforming to a different frame, one has to use the Lorentz transformation, Equation 16. In Lorentz transformation, the transformation law of each coordinate depends on the values of the other coordinates.

Any quantity that obeys the transformation law given by Equation 16 is called a **four vector**. This can be seen as generalization of the "classical" definition of a vector. Once a quantity is known to be a four-vector, its transformation properties are fully defined.

While the set of space-time coordinates form a four vector, other four-vectors can be found. Consider, for example, the result obtained in Equation 6, namely that

$$(interval)^2 = (\Delta\tau)^2 = (\Delta t^2) - \left(\frac{\Delta x}{c}\right)^2 - \left(\frac{\Delta y}{c}\right)^2 - \left(\frac{\Delta z}{c}\right)^2 \tag{27}$$

is *invariant* under Lorentz transformation. This is analogue to the total length of a vector in 2-dimensional space, $x^2 + y^2 + z^2$, which is invariant under rotations.

We multiply both sides of Equation 27 by the scalar $\left(\frac{mc}{\Delta\tau}\right)^2$, and take the limit $\Delta\tau \to 0$, to get

$$\begin{array}{lll} \left(mc\frac{d\tau}{d\tau}\right)^2 & = & \left(mc\frac{dt}{d\tau}\right)^2 - \left(\frac{mdx}{d\tau}\right)^2 - \left(\frac{mdy}{d\tau}\right)^2 - \left(\frac{mdz}{d\tau}\right)^2 \\ m^2c^2 & = & \left(mc\frac{dt}{d\tau}\right)^2 - {p_x}^2 - {p_y}^2 - {p_z}^2 \end{array} \tag{28}$$

where $p_x$, $p_y$, $p_z$ are the spatial coordinates of the momentum four-vector (see Equation 25).

The first term in the right hand side of Equation 28 can be written as $\frac{1}{c^2}\left(mc^2\frac{dt}{d\tau}\right)^2$. The quantity in parenthesis have dimensions of energy. We thus **define the relativistic energy** by

$$E \equiv mc^2\frac{dt}{d\tau} = \gamma mc^2. \tag{29}$$

Using this definition, the first term on the right hand side of Equation 28 is $\left(\frac{E}{c}\right)^2$.

With this definition, Equation 28 is written as

$$\begin{array}{lll} m^2c^2 & = & \frac{E^2}{c^2} - {p_x}^2 - {p_y}^2 - {p_z}^2 \\ E^2 & = & c^2\left({p_x}^2 + {p_y}^2 + {p_z}^2\right) + m^2c^4 = p^2c^2 + m^2c^4 \end{array} \tag{30}$$

In the limit $m = 0$, a massless particle, e.g., a photon, still has both energy and momentum, connected via $E = pc$. Such a massless particle must travel at speed $c$.

We thus find that the energy and momentum transform in a similar way to the coordinates. They form a four vector, known as **energy-momentum four vector**. This can be seen by multiplying the Lorentz transformation (equations 16) by $m$ and differentiating with respect to $\tau$. One obtains

$$\begin{array}{lcl} m\frac{dx'}{d\tau} = p'_x & = & \gamma\left(m\frac{dx}{d\tau} - mv\frac{dt}{d\tau}\right) = \gamma\left(p_x - v\frac{E}{c^2}\right) \\ p'_y & = & p_y \\ p'_z & = & p_z \\ mc^2\frac{dt'}{d\tau} = E' & = & \gamma\left(mc^2\frac{dt}{d\tau} - mv\frac{dx}{d\tau}\right) = \gamma\left(E - vp_x\right). \end{array} \tag{31}$$

This is the Lorentz transformation of energy-momentum four vector.

We can determine the velocity of a particle, using

$$v_x = \frac{dx}{dt} = \frac{dx}{d\tau}\frac{d\tau}{dt} = \frac{p_x}{m}\frac{mc^2}{E} = \frac{c^2 p_x}{E}, \tag{32}$$

or

$$\vec{p} = \vec{v}\frac{E}{c^2} \tag{33}$$

### 8.1. A few more explanations on vectors

Vectors can be defined, or represented, in 2 ways: (1) a **geometric way** - by specifying the magnitude and direction. (2) An **algebraic** way by specifying the components relative to (Cartesian) coordinate axes.

The definition of a vector as a quantity with magnitude and direction breaks down in advanced work. On the one hand, there are quantities, e.g., elastic constants or index of refraction in anisotropic crystals that have magnitude and direction *but are not vectors*. On the other hand, the geometrical, naive approach is difficult to generalize and extend to more complex quantities such as tensors.

There is also a strong physical motivation for re-defining vectors: while the physical world is described by mathematics, any physical prediction should be **independent** of the mathematical analysis done.

For example, we assume that space is homogeneous; thus, there is no preferred direction - all directions are equivalent. Thus, any physical system **cannot** and **should not** depend on the arbitrary choice of *orientation* of coordinates.

In a simple example of 2-d rotation, the $x$- and $y$- components of a vector are rotated according to

$$\begin{aligned} x' &= x\cos\phi + y\sin\phi \\ y' &= -x\sin\phi + y\cos\phi. \end{aligned} \tag{34}$$

$x, y$, or $x', y'$ represent the components of the same vector. Alternatively, we **define** any pair of quantities $A_x(x,y)$, $A_y(x,y)$ that transform into $A'_x, A'_y$ by rotation according to

$$\begin{aligned} A'_x &= A_x\cos\phi + A_y\sin\phi \\ A'_y &= -A_x\sin\phi + A_y\cos\phi, \end{aligned} \tag{35}$$

to be the components of a vector $\vec{A}$.

If $A_x$ and $A_y$ are transformed according to Equation 34, we say that they are the components of a vector $\vec{A}$. If they don't - then they do not form a vector.[2]

## 9. Relativistic energy and the mass-energy equivalent

We have seen from the definition of relativistic momentum (Equations 25, 26) that in the limit $v \ll c$, the relativistic momentum approaches the classical (Newtonian) momentum.

What about the energy ? let us use a Taylor expansion to the definition of the energy (Equation 29) in the limit $v \ll c$, to get

$$E = \gamma mc^2 = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} \simeq mc^2\left[1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \mathcal{O}\left(\frac{v}{c}\right)^6\right]. \tag{36}$$

The first term, $mc^2$ is called the **rest energy** of the particle. The second term, $\frac{1}{2}mv^2$ is the classical (Newtonian) kinetic energy. Thus, in the limit $v \to 0$, the particle's energy **is not 0**, but is equal to $mc^2$.

We can use an alternative way to show that **this definition of energy is consistent with the definition of momentum**. We saw that classically, $E = W = \int \vec{F} dx$, where

[2]For further read see the book Arfken & Weber, **Mathematical methods for physicists**.

$\vec{F} = md\vec{p}/dt$. We can generalize that to the relativistic definition of momentum:

$$\begin{aligned} W &= \int F dx = \int \frac{d}{dt}\frac{mv}{\sqrt{1-v^2/c^2}} dx = \\ &= \int \frac{d}{dt}\left(\frac{mv}{\sqrt{1-v^2/c^2}}\right)\frac{dx}{dt}dt \\ &= \int \left[\frac{mv}{\sqrt{1-v^2/c^2}}\frac{dv}{dt} + \frac{mv^3/c^2}{(1-v^2/c^2)^{3/2}}\frac{dv}{dt}\right] dt \\ &= \int \frac{mvdv/dt}{(1-v^2/c^2)^{3/2}} dt \\ &= \int \frac{d}{dt}\left(\frac{mc^2}{\sqrt{1-v^2/c^2}}\right) dt \end{aligned} \tag{37}$$

Taking the upper and lower velocity bounds to be $v$ and 0, we get

$$E_k = W = mc^2(\gamma - 1) \tag{38}$$

We saw that classically, in total plastic (inelastic) collision, Newtonian mechanical energy *is not conserved*, but being lost, mainly to heat. With the new definition of energy in relativistic mechanics, **energy is conserved** (as long as it is not taken from the system) - the loss of kinetic energy is accounted for by a change in the rest-mass energy: **mass is not conserved**.

Example: total inelastic collision of identical particles. A particle of mass $m$ and kinetic (total - rest mass) energy $E_k$ collides with a second particle of mass $m$, initially at rest. Following the collision, the two particles are attached, forming one particle of mass $M$. The combined particle has momentum $p_f$ and energy $E_f$.

**Non relativistically-** we get $M = 2m$; $p_f = p_i \rightarrow v_f = v_i/2$; $E_f = (1/2)Mv_f^2 = mv_i^2/4 = E_i/2$.

**Relativistic treatment.** Using Equation 30, we find

$$M^2c^4 = {E_f}^2 - {p_f}^2c^2, \tag{39}$$

where $E_f$ and $p_f$ are the final (relativistic) energy and momentum of the composite particle of mass $M$.

Conservation laws:
Energy - $E_f = E_i = 2mc^2 + E_k$;
momentum - $p_f = p_i$.

Substituting in Equation 39, we find

$$M^2c^4 = \left(2mc^2 + {E_k}^2\right)^2 - {p_i}^2c^2, \tag{40}$$

where $p_i^2c^2 = E_{i,1}^2 - m^2c^4 = (mc^2 + E_k)^2 - m^2c^4$, where $E_{i,1}$ is the energy (kinetic + rest mass) of particle 1, and thus

$$\begin{aligned} M^2c^4 &= \left(2mc^2 + E_k\right)^2 + m^2c^4 - (mc^2 + E_k)^2 \\ &= 4m^2c^4 + 4mc^2E_k + E_k^2 + m^2c^4 - m^2c^4 - 2mc^2E_k - E_k^4 \\ &= 4m^2c^4 + 2mc^2E_k \end{aligned} \tag{41}$$

or $M = \sqrt{4m^2 + 2mE_k/c^2}$, which implies that for $E_k > 0$, $M > 2m$.

The kinetic energy after the collision is $E_{k,f} = E_f - Mc^2 = E_k + (2m - M)c^2$ which is **less** than the kinetic energy before the collision, $E_k$. The loss of kinetic energy, $\Delta E_k = E_{k,f} - E_k = (2m - M)c^2 = -\Delta mc^2$ - is exactly compensated by the increase of mass (times $c^2$). Thus, **the kinetic energy loss is compensated by gain of mass**. The opposite is also true: in any interaction in which mass is lost, energy is gained, according to

$$\Delta E = \Delta mc^2 \tag{42}$$

Example. Energy production in the sun.

The most important source of energy in the sun and most stars arises from nuclear burning of protons to form a Helium,

$$\begin{aligned} H^1 + p &= H^2 + e^+ + \nu \\ H^2 + H^1 &= He^3 + \text{photon} \\ He^3 + He^3 &= He^4 + 2H^1 \end{aligned} \tag{43}$$

The net result is that 4 hydrogen atoms ($H^1$) are burned (fused) to produce 1 Helium aton ($He^4$). The mass difference is $4M(H^1) - M(He^4) \approx 4.78 \times 10^{-26}\,\text{gr} \approx 52m_e$, which is converted to energy, as $\Delta E = \Delta mc^2$. Note that as neutrinos ($\nu$) are produced, the sun is a very powerful source of neutrinos, which carry up to $\sim 10\%$ of the energy produced in the sun.

## 10. Worked example: acceleration of a charged particle by constant, longitudinal electric field

The equation of motion of a particle of rest mass $m$ and charge $q$ in a uniform electric field $\vec{E} = E\hat{x}$ is

$$\frac{dP}{dt}\hat{x} = F_{\hat{x}} = qE, \tag{44}$$

or

$$\frac{d}{dt}(\gamma mv) = \frac{d}{dt}\left(\frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}}\right) = qE. \tag{45}$$

This can be integrated, to yield

$$\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}} = qEt. \tag{46}$$

(we assume that at $t = 0$, $v = 0$; furthermore, we consider only motion along the $\hat{x}$ direction, thereby assuming $v_y = v_z = 0$).

We note that in the *non-relativistic case*, we would obtain $v = \frac{qEt}{m} = \frac{p}{m}$, namely $v$ grows linearly with time.

Using some algebra, we can now extract $v$ from equation 46:

$$\frac{v}{c} = \frac{\frac{qEt}{mc}}{\sqrt{1+\left(\frac{qEt}{mc}\right)^2}} \tag{47}$$

This is plotted in Figure 4.

Clearly, at short times, $t \ll mc/qE$, the denominator in Equation 47 is replaced by unity, and we retrieve back the classical result. In the other extreme of long times, $t \gg mc/qE$,

$$\frac{v^2}{c^2} = \frac{1}{\left(\frac{mc}{qEt}\right)^2 + 1} \approx 1 - \left(\frac{mc}{qEt}\right)^2. \tag{48}$$

This shows that $v$ approaches $c$ as a limiting velocity (but never exceeds it).

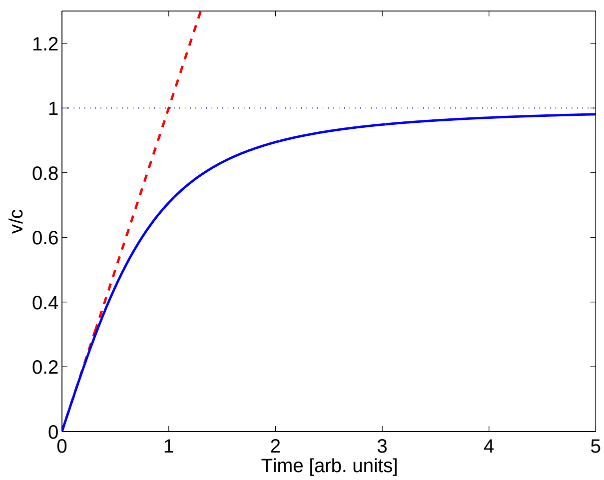


Fig. 4.— The time-dependence of the velocity of a particle subject to a constant force. In the non-relativistic limit (dashed, red), the velocity increases linearly with time. In the relativistic treatment, $v \to c$ at late times, but never exceeds it.