Special Theory of Relativity

Amaires@Jan 2025

Contents

This write-up is loosely based on Chapters 15 to 17 of *The Feynman Lectures on Physics*, the English translation of Einstein's original Special Relativity paper, a few Wiki pages, and some of my own thoughts. Of course, all mistakes are mine.

1 Galilean Transformation

For over 200 years, Newtonian mechanics was believed to describe nature correctly. In Newtonian mechanics, given a stationary inertial reference frame S and another inertial reference frame S' that is moving with speed u along the x-axis of S, an event observed at location (x, y,z) and time t in S is observed at location (x',y',z') and time t' in S', with these variables subject to the following relationships:

$$\begin{aligned} x' &= x - ut \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned} \tag{1}$$

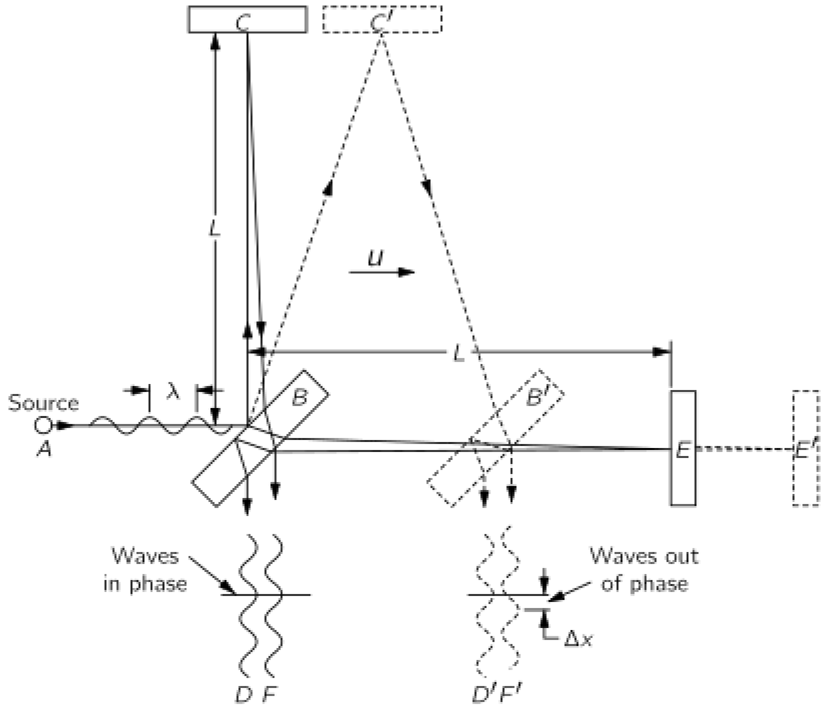

Figure 1: Michelson-Morley Experiment Setup

Such relationships are commonly referred to as the Galilean transformation. The Galilean transformation makes intuitive sense. For example, a car moving at speed $v = dx/dt$ in S is observed to move at speed $v' = dx'/dt = dx/dt - d(ut)/dt = v - u$ in S'. When we are driving on a freeway, other cars driving in the same direction do not appear to be driving fast at all!

2 Failure of Galilean Transformation

The advent of Maxwell's equations and the result of a series of experiments, the most famous of which being the Michelson-Morley experiment, threw the Galilean transformation into question.

2.1 Maxwell's Equations

In the 19th century, the investigations of the phenomena of electricity, magnetism, and light culminated in Maxwell's equations, which describes electricity, magnetism, and light in one unified system. However, if we move from one inertial reference frame to another inertial reference frame and apply the Galilean transformation to (x, y, z, t), the resulting form of the Maxwell's equations changes; therefore in a moving space ship the electrical and optical phenomena should be different from those in the stationary ship. One consequence is that, one could determine the absolute speed of the ship by making suitable optical or electrical measurements. Replace the car in the example in Section 1 with a beam of light traveling at speed $v = c$, v' would be equal to $v' = c - u$. Once v' is measured, u can be derived easily. Unfortunately a bunch of experiments designed along this idea all reach the same conclusion that $u = 0$. The most famous such experiment will be discussed below.

The failure of combining Maxwell's equations with the Galilean transformation first led people to think Maxwell's equations were wrong, so changes were made to Maxwell's questions such that under the Galilean transformation, Maxwell's equations remain unchanged. Unfortunately, the modified equations predicted new electrical phenomena that did not exist in experiments. The implies that the trouble might be with the Galilean transformation.

2.2 Michelson-Morley Experiment

Figure 1 depicts a simplified version of the Michelson-Morley experiment setup. Mounted on a rigid apparatus are a light source A, a partially silvered glass plate B, and two mirrors C and E. The mirrors are placed at equal distance L from B. B splits an oncoming beam of light from A and the two resulting beams continue in mutually perpendicular directions to the mirrors, where they are reflected back to B. It is possible to detect whether the reflected lights arrive at B simultaneously or not by checking whether they reinforce or interfere with each other. The apparatus was oriented such that the line BE was nearly parallel to the earth's motion in its orbit.

At the time of the experiment, light was assumed to only travel in *ether*, a theoretical substance permeating all space, at speed c. The speed of light, observed at different inertial reference frames, according to the Galilean transformation, can be different from c. Let the speed the earth is moving relative to ether be u, we can calculate the time it takes light to travel different paths:

1. $t_{B\to E} = \frac{L}{c-u}$ as the speed of light that travels from B to E, relative to the apparatus, is $c - u$.

2. $t_{E\to B} = \frac{L}{c+u}$ as the speed of the reflected light that travels from E back to B, relative to the apparatus, is $c + u$.

3. $t_{B\to C} = \frac{L/c}{\sqrt{1-u^2/c^2}}$ as $(ct_{B\to C})^2 = L^2 + (ut)^2$.

4. $t_{C\to B} = t_{B\to C} = \frac{L/c}{\sqrt{1-u^2/c^2}}$ due to symmetry.

Using the results above, we see that the time it takes light to travel from B to E and back to B is

$$t_{B\to E\to B} = t_{B\to E} + t_{E\to B} = \frac{L}{c-u} + \frac{L}{c+u} = \frac{2Lc}{c^2-u^2} = \frac{2L/c}{1-u^2/c^2} \tag{2}$$

and the time it takes light to travel from B to C and back to B is

$$t_{B\to C\to B} = t_{B\to C} + t_{C\to B} = \frac{2L/c}{\sqrt{1-u^2/c^2}}. \tag{3}$$

$t_{B\to C\to B}$ is equal to $t_{B\to E\to B}$ iff $u = 0$.

The earth moves along its orbit at the late of 18 miles per second; therefore u (remember that it is relative to ether) should be at least that much at some time of the day or night and at some time during the year. However, the amply sensitive apparatus was not able to detect any difference between $t_{B\to C\to B}$ and $t_{B\to E\to B}$.

3 Lorentz Transformation

In Section 2.1, we discussed that Maxwell's equations and Galilean transformation do not work together, and attempts to revise Maxwell's equations all failed. The Michelson-Morley experiment in Section 2.2 dealt another blow to Galilean transformation by showing results contradictory to it. It was Hendrik Lorentz, based on other people's prior work as well, who noticed that if Galilean transformation was replaced by the following Lorentz transformation, Maxwell's equations would remain the same in different inertial reference frames, and the Michelson-Morley experiment's results could be explained.

$$\begin{aligned} x' &= \frac{x-ut}{\sqrt{1-u^2/c^2}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t-ux/c^2}{\sqrt{1-u^2/c^2}} \end{aligned} \tag{4}$$

There are a few of points worth noting about the Lorentz transformation.

1. When u is far smaller than c, the Galilean transformation is a good approximation of the Lorentz transformation.
2. Though not shown so far, when S and S' do not have their x-axis, y-axis, and z-axis all perfectly aligned, that is, when S has to be rotated to arrive at S', the Galilean transformation tells us that x' is a function of all of x, y, and z, and so are y' and z', but t' is still only a function, in fact an identity function, of only t. A crude interpretation is that in the Galilean transformation, space and time are completely independent. In the Lorentz transformation, space and time are no longer independent. Each of x', y', z', and t' are now a function of all of x, y, z, and t.
3. It is less obvious, but when u is fixed, simple manipulation reveals the Lorentz transformation to be a strictly linear transformation.

3.1 Change of Unit

If we use c as the unit of length measurement and by extension, the unit of velocity, the Lorentz transformation can be rewritten to have a more symmetric form:

$$\begin{aligned} x' &= \frac{x-ut}{\sqrt{1-u^2}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t-ux}{\sqrt{(1-u^2}}. \end{aligned} \tag{5}$$

Let $\gamma = \frac{1}{\sqrt{(1-u^2}}$, we have

$$\begin{aligned} x' &= \gamma(x-ut) \\ y' &= y \\ z' &= z \\ t' &= \gamma(t-ux). \end{aligned} \tag{6}$$

Our discussion below may use any of the above three forms, the choice of which should be obvious in the context.

3.2 Length Contraction

To measure the length of a object with no motion relative to an inertial reference frame, one can measure the coordinates of the two end points of the object at different times and calculate the differences because the these coordinates do not change with time. Such lengths are called the objects' proper length. In order to measure the length of moving object, however, it is critical that the two end points are measured *at the same time* because under the Lorentz transformation, time and space are intertwined.

Suppose an object's proper length is L, as measured in S, what is its length as measured in S'? Let x_1 and x_2 be the two end points of the object in S, t_1 and t_2 be when x_1 and x_2 are measured, we have the following after applying the Lorentz transformation:

$$\begin{aligned} L &= x_2 - x_1 \\ x'_1 &= \gamma(x_1 - ut_1) \\ x'_2 &= \gamma(x_2 - ut_2) \\ t'_1 &= \gamma(t_1 - ux_1/c^2) \\ t'_2 &= \gamma(t_2 - ux_2/c^2). \end{aligned}$$

If we are to calculate the object's length $L' = x'_2 - x'_1$ in S', we need to have $t'_1 = t'_2$. Plugging this equation in the system of equations above, it follows that

$$L' = L/\gamma = L\sqrt{1 - u^2/c^2}. \tag{7}$$

Clearly $L' \leq L$. This phenomenon is commonly referred to as length contraction.

If S' is moving at speed u along the x-axis of S, from the view of S', S is moving at speed $-u$ along the x-axis of S'. Thus an observer at rest in S will also see the length contraction effect in S'. The following scenario is counter-intuitive, but it is true according to the Lorentz transformation.

> Person P_1 in S uses a meter-long ruler to measure the length of a stick K_1, placed parallel to the x axis and gets 2m. Person P_2 in S' also uses a meter-long ruler to the length of a stick K_2, placed parallel to the x'-axis, and gets 3m. When P_1 observes P_2's action however, he sees P_2 is using a $1/\gamma$m long ruler to measure a $3/\gamma$m long K_2. When P_2 observes P_1's action, he believes P_1 is using a $1/\gamma$m long ruler to measure a $2/\gamma$m long K_1. P_1 and P_2 both believe their own world is normal, but believe somehow the other person's world is shrunk in the x/x' dimension.

The following prediction by the Lorentz transformation is even more counter-intuitive.

> When P_2 finishes measuring, he starts rotating K_2 gradually until the angel between K_2 and the x'-axis grows to 90 degrees. While observing this process, P_1 sees that K_2's length grows gradually from $3/\lambda$m to 3m. Next P_2 decides to measure K_2's length again, and he gets 3m, the same as before. P_1 will also observe that P_2 is using a meter-long ruler to measure a 3m long stick.

3.3 Interpreting the Michelson-Morley Experiment Results

Length contraction can be used to explain the puzzling result from the Michelson-Morley experiment. In the Michelson-Morley experiment, S is ether, which, if you remember, was believed to be required for light to propagate, and S' is the experiment apparatus. The length measurements for BE and BC are made in S'. To avoid confusion, all the variables in (2) and (3) are rewritten to their "primed" versions as below:

$$t'_{B\to E\to B} = \frac{2L'/c}{1 - u^2/c^2} \tag{8}$$

$$t'_{B\to C\to B} = \frac{2L'/c}{\sqrt{1 - u^2/c^2}}. \tag{9}$$

The final observation of whether the light reflected from C and E reinforces or interferes with each other was made in S because light travels in ether. In S, the length of BC remains to be $L = L'$ because if it is perpendicular to the direction of motion, but the length of BE is $L = L'/\gamma = L'\sqrt{1 - u^2/c^2}$. The derived $t_{B\to E\to B}$ and $t_{B\to C\to B}$ are:

$$t_{B\to E\to B} = \frac{2L/c}{1 - u^2/c^2} = \frac{2L'\sqrt{1 - u^2/c^2}/c}{1 - u^2/c^2} = \frac{2L'/c}{\sqrt{1 - u^2/c^2}} \tag{10}$$

$$t_{B\to C\to B} = \frac{2L/c}{\sqrt{1 - u^2/c^2}} = \frac{2L'/c}{\sqrt{1 - u^2/c^2}}. \tag{11}$$

Now you see that $t_{B\to E\to B} = t_{B\to C\to B}$ which explains why no light interference was observed at B in the Michelson-Morley experiment.

3.4 Velocity Transformation and the Speed of Light

If an object is moving at speed v' in S', along the x' axis. What is the object's speed, as measured by a resting observer in S? Is it simply $v' + u$ as would be concluded by the Galilean transformation?

Before we can find out the answer, it helps to solve the inverse case of (4), which is merely a system of linear equations. The result is

$$\begin{aligned} x &= \gamma(x' + ut') \\ y &= y' \\ z &= z' \\ t &= \gamma(t' + ux'). \end{aligned} \tag{12}$$

The same result can also be arrived at by noticing that if S' is moving at speed u relative to S, S is moving at speed $-u$ relative to S'. Simply replacing u with $-u$ in (4) and renaming the variables get you (12).

By an object is moving at speed v' in S', we really mean that $v' = \frac{x_2'-x_1'}{t_2'-t_1'}$. If x_1', x_2', t_1', and t_2' are translated back to x_1, x_2, t_1, and t_2 using (12), we have:

$$\begin{aligned} v' &= \frac{x_2' - x_1'}{t_2' - t_1'} \\ x_1 &= \gamma(x_1' + ut_1') \\ x_2 &= \gamma(x_2' + ut_2') \\ t_1 &= \gamma(t_1' + ux_1') \\ t_2 &= \gamma(t_2' + ux_2') \end{aligned}$$

and

$$v = \frac{x_2 - x_1}{t_2 - t_1} = \frac{(x_2' - x_1') + u(t_2' - t_1')}{(t_2' - t_1') + u(x_2' - x_1')} = \frac{\frac{x_2'-x_1'}{t_2'-t_1'} + u}{1 + u\frac{x_2'-x_1'}{t_2'-t_1'}} = \frac{v' + u}{1 + uv'}.$$

That is the transformation of velocity. The same result can be derived more easily using calculus. To start, we have

$$v = \frac{dx}{dt} = \frac{d(\gamma(x' + ut'))}{d(\gamma(t' + ux'))} = \frac{\gamma d(x' + ut')}{\gamma d(t' + ux')} = \frac{d(x' + ut')}{d(t' + ux')} = \frac{dx' + udt'}{dt' + udx'}.$$

Dividing both the numerator and denominator by dt', we arrive at

$$v = \frac{\frac{dx'}{dt'} + u}{1 + u\frac{dx'}{dt'}} = \frac{v' + u}{1 + uv'}.$$

The transformation of velocity has a few interesting points:

1. v is a symmetric function of v' and u.
2. When uv' is far smaller than c^2, the velocity transformation degenerates to that derived from the Galilean transformation.
3. If $v' = 1$, which means the speed of light, we have
$$v = \frac{1+u}{1+u} = 1.$$
That is, anything travels at the speed of c in S' will appear to travel at the same speed in S. In particular, *light travels at speed c in any inertial reference frame.*
4. If $|u| \leq 1$ and $|v'| \leq 1$, easy algebraic manipulation gives us$|v| \leq 1$. Figure 2 shows the effect of the velocity transformation given different given values of u.

4 Special Theory of Relativity

Assuming the existence of ether, Section 3 shows that the Lorentz transformation can be used to explain the result of the Michelson-Morley experiment, prove that light always travels at speed c in any inertial reference frame. With the Lorentz transformation, Maxwell's equations also take exactly the same form in different inertial reference frames.

A. Einstein's special theory of relativity goes the other way around. It does not assume the existence of ether, or the Lorentz transformation. Instead, it postulates that

1. the same laws of electrodynamics and optics are valid for all inertial reference framess, and
2. light always propagates in empty space with a definite velocity c which is independent of the state of motion of the emitting body.

Einstein himself calls the first postulate the Principle of Relativity. The concept of the principle of relativity was not new. In fact G. Galilei formulated it more than two hundred years earlier, in the context of mechanics. Einstein extended the principle to newly discovered laws of electrodynamics and optics, and potentially other future laws of physics. These two postulates can be used to derive the Lorentz transformation, and they eliminated the need for the existence of ether.

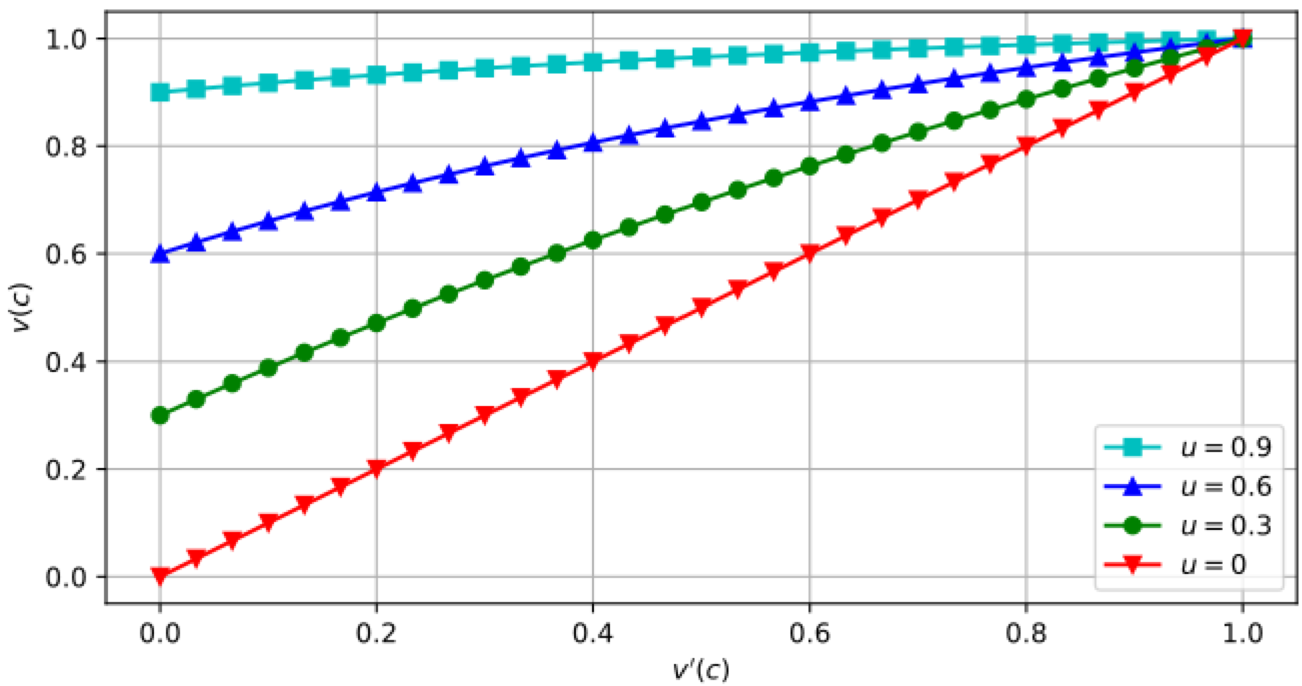

Figure 2: Velocity Transformation from S' to S.

4.1 Deriving the Lorentz Transformation

This section derives the Lorentz transformation based on the two postulates of the special theory of relativity, and the assumption that such transformation is linear. For simplicity, the transformation for y and z are skipped.

First, let the linear transformation for x' be

$$x' = ax + bt. \tag{13}$$

An event at $(x'_1, t'_1) = (0, t'_1)$ in S' will be viewed as $(x_1, t_1) = (ut_1, t_1)$ in S. After all, that is definition of the speed of S' relative to to S. Plugging this information in (13), it follows that

$$0 = aut_1 + bt_1. \tag{14}$$

Combining (13) and (14) derives

$$x' = ax - aut = a(x - ut). \tag{15}$$

Since S is moving at speed $-u$ in S', by the first postulate of the special relativity, replacing u with $-u$ and renaming the variables get the translation from S' to S. That is

$$x = a(x' + ut'). \tag{16}$$

Both (15) and (16) are true, multiplying the corresponding sides of them gives

$$xx' = a^2(xx' + uxt' - utx' - u^2tt').$$

Next according to the second postulate of the special theory of relativity, $v = v' = \frac{dx'}{dt'} = \frac{dx}{dt} = c$ is true. That is

$$\frac{dx'}{dt'} = \frac{a(dx - udt)}{dt'} \tag{17}$$

$$\frac{dx'}{dt'} = c \tag{18}$$

$$\frac{dx}{dt} = \frac{a(dx' + udt')}{dt} \tag{19}$$

$$\frac{dx}{dt} = c. \tag{20}$$

Combining (18), (19) and (20) derives

$$\frac{dx}{dt} = \frac{a(c + u)dt'}{dt} = c. \tag{21}$$

Combining (17), (18) and (20) derives

$$\frac{dx'}{dt'} = \frac{a(c - u)dt}{dt'} = c. \tag{22}$$

Multiplying the corresponding sides of (21) and (22) gives

$$a^2(c^2 - u^2) = c^2$$
$$a = \frac{1}{\sqrt{1 - u^2/c^2}}.$$

So finally we have

$$x' = \frac{x - ut}{\sqrt{1 - u^2/c^2}} \tag{23}$$
$$x = \frac{x' + ut'}{\sqrt{1 - u^2/c^2}}. \tag{24}$$

Combining (23) and (24) gives

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

These complete the Lorentz transformation.

4.2 Time Dilation

Besides length contraction, the Lorentz transformation also predicts a phenomenon called Time Dilation: a clock that is moving relative to the observer will be measured to tick more slowly than a clock at rest in the observer's reference frame. This phenomenon can be easily derived from the Lorentz transformation.

Suppose at location x in S, one event happened at t_1, and then another event happened at t_2. In S', what is the time lapse between the two events? We have

$$t_1' = \gamma(t_1 - ux)$$
$$t_2' = \gamma(t_2 - ux).$$

Subtracting the first equation from the second equation gives

$$\Delta t' = (t_2' - t_1') = \gamma(t_2 - t_1) = \gamma \Delta t.$$

This is called time *dilation* because $\gamma \geq 1$. Unlike length contraction which is only observed in the direction of motion, time dilation is independent of the direction of motion.

For observers in S and S', everything feels normal. When observers in S observe S' however, objects are shorter along the x'-axis and everything seems to happen more slowly. People are walking slower, heartbeats are slower, cancers develop more slowly, and people take longer to get hungry as well. Observers in S' notice exactly the same oddity when observing S.

If someone rides in a very fast spaceship, say at $\sqrt{3}/2$ the speed of light, for a year and then returns to the earth. His friends will notice the traveler is a lot younger than he should be. A very fast spaceship allows one to advance to the "future" of the earth, and treats his cancer with technology inconceivable if he were to stay on the earth, but does not live a longer life. Even though people on earth think he has traveled for 2 ($\gamma = 2$ when $u = \sqrt{3}c/2$) years, in S', he has traveled for 1 year and had 365 dinners only!

5 Paradoxes

Incomplete or inappropriate application of length contraction and time dilation can often lead to paradoxes.

5.1 Twin Paradox

Suppose A and B are twins. A stays on the earth, while B takes a spaceship travel. When B returns to the earth, according to the special theory of relativity, A is older. However, from B's perspective, A is traveling, and therefore when they meet again A should be older. This is the famous twin paradox.

The key to resolving the twin paradox is to notice A and B are not symmetric. For B to return to the earth, there has to acceleration that he can feel, and that A cannot. So A is older.

5.2 Ladder Paradox

The ladder paradox involves a ladder, parallel to the ground, traveling horizontally at relativistic speed (near the speed of light) and therefore undergoing a length contraction. The ladder is imagined passing through the open front and rear doors of a garage or barn which is shorter than the ladder's proper length, so if the ladder was not moving it would not be able to fit inside. To a stationary observer, due to the contraction, the moving ladder is able to fit entirely inside the building as it passes through. On the other hand, from the point of view of an observer moving with the ladder, the ladder will not be contracted, and it is the building which will be contracted to an even smaller length. Therefore, the ladder will not be able to fit inside the building as it passes through. This poses an apparent discrepancy between the realities of both observers. This apparent paradox results from the mistaken assumption of absolute simultaneity. The paradox is resolved when it is considered that in relativity, simultaneity is relative to each observer, making the answer to whether the ladder fits inside the garage also relative to each of them.

In this thought experiment, S is the garage, and S' is the moving ladder. Let's assume S' is moving at speed $u = \sqrt{3}/2$ such that $\gamma = 2$, and the proper lengths of the garage and ladder are L_g and L'_l, respectively. Let's further assume $L_g = L'_l = 1$ (yes, this ladder and this garage are very very long). We'll call the observer stationary in S the garage man, and the observer stationary in S' the ladder man.

5.2.1 Garage's View

First, let's examine the garage man's point of view. When the left end of the ladder meets the left end of the garage, the garage man makes, at the same time t, the following measurements:

1. the left end of the ladder: x_{ll},
2. the left end of the garage: x_{lg},
3. the right end of the ladder: x_{rl}, and
4. the right end of the garage: x_{rg}.

The following relationships exist among these measurements:

$$\begin{aligned} x_{ll} &= x_{lg} \\ x_{rl} &= x_{ll} + L_l = x_{ll} + L'_l/\gamma = x_{ll} + \frac{1}{2} \\ x_{rg} &= x_{lg} + L_g = x_{ll} + c. \end{aligned}$$

Therefore

$$x_{ll} = x_{lg} < x_{rl} < x_{rg}. \tag{25}$$

If the garage man waits an infinitesimal amount of time, he can close the left and right doors of the garage at the same time and the garage will fully contain the ladder briefly.

5.2.2 Ladder's View

Now let's examine the ladder man's view, which sees a garage moving rapidly to the left at speed $u = -\sqrt{3}/2$. When the right end of the garage meets the the right end of the ladder, the ladder man makes, at the same time t', the follow measurements:

1. the left end of the ladder: x'_{ll},
2. the left end of the garage: x'_{lg},
3. the right end of the ladder: x'_{rl}, and
4. the right end of the garage: x'_{rg}.

The following relationships exist among these measurements:

$$\begin{aligned} x'_{rl} &= x'_{rg} \\ x'_{lg} &= x'_{rg} - L'_g = x'_{rg} - L_g/\gamma = x'_{rl} - \frac{1}{2} \\ x'_{ll} &= x'_{rl} - L'_l = x'_{rl} - 1. \end{aligned}$$

Therefore

$$x'_{ll} < x'_{lg} < x'_{rg} = x'_{rl} \tag{26}$$

and the garage cannot contain the ladder.

5.2.3 Resolution

The apparent paradox is the statement "the garage can contain the ladder" cannot be both true and false at the same time. The resolution lies in recognizing that the relativistic view does not recognize the concept of simultaneity across different inertial reference frames. First, let's examine (25) a bit more closely. How does the the garage man's measurement actions look to the ladder man? To find that out, we need to transform variables $(x_{ll}, t_{ll}), (x_{lg}, t_{lg}), (x_{rl}, t_{rl})$, and (x_{rg}, t_{rg}) to their S' counterparts using the Lorentz transformation as below:

$$\begin{aligned} x'_{ll} &= \gamma(x_{ll} - ut_{ll}) \\ x'_{lg} &= \gamma(x_{lg} - ut_{lg}) \\ x'_{rl} &= \gamma(x_{rl} - ut_{rl}) \\ x'_{rg} &= \gamma(x_{rg} - ut_{rg}) \\ t'_{ll} &= \gamma(t_{ll} - ux_{ll}) \\ t'_{lg} &= \gamma(t_{lg} - ux_{lg}) \\ t'_{rl} &= \gamma(t_{rl} - ux_{rl}) \\ t'_{rg} &= \gamma(t_{rg} - ux_{rg}). \end{aligned}$$

Because we have $x_{ll} = x_{lg}$ and $t_{ll} = t_{lg} = t_{rl} = t_{rg}$, we conclude that

$$\begin{aligned} t'_{lg} &= t'_{ll} \\ t'_{rl} &= t'_{ll} + \gamma u \cdot (x_{ll} - x_{rl}) = t'_{ll} - \gamma u L_l = t'_{ll} - \gamma u L'_l/\gamma = t'_{ll} - \sqrt{3}/2 \\ t'_{rg} &= t'_{ll} + \gamma u \cdot (x_{ll} - x_{rg}) = t'_{ll} + \gamma u \cdot (x_{lg} - x_{rg}) = t'_{ll} - \gamma u L_g = t'_{ll} - \sqrt{3} \end{aligned}$$

and as a result

$$t'_{rg} < t'_{rl} < t'_{lg} = t'_{ll}.$$

Now the ladder man figures out why the garage man was "wrong": he was measuring the right end of the garage first, and then after a $\sqrt{3}/2$ seconds, the right end of the ladder, and after another $\sqrt{3}/2$ seconds, the left end of the garage and the ladder. Of course, that will make the garage "artificially" longer and the garage man will believe the garage is long than the ladder.

Let's not stop here, and actually calculate how the ladder man sees the garage man's location measurements as well, as below:

$$\begin{aligned} x'_{lg} &= x'_{ll} \\ x'_{rl} &= x'_{ll} - \gamma(x_{rl} - x_{ll}) = x'_{ll} - \gamma L_l = x'_{ll} - \gamma L'_l/\gamma = x'_{ll} - 1 \\ x'_{rg} &= x'_{ll} - \gamma(x_{rg} - x_{ll}) = x'_{ll} - \gamma(x_{rg} - x_{lg}) = x'_{ll} - 2. \end{aligned}$$

Now the ladder is more relieved, because he sees that the garage man's "wrong" measurements show that the garage's length is $2c$ and the ladder's c. This explains completely why the garage man made the "wrong" claim.

Similarly, the garage man can also observe the ladder man's measurements and figure out why the ladder man is "wrong".

6 Minkowski Spacetime

We can represent both space and time in one vector (x, y, z, t) called spacetime.

6.1 Light Cone, Past, Future and Causality

If we focus on space with only one dimension, spacetime can be represented in a Cartesian coordinate system as in Figure 3. By convention, x is described by the horizontal axis and t the vertical axis. The two dashed lines are the traces of an object moving at the speed of c or $-c$. If we consider two dimensional space (like x and y), these two lines become a cone. Therefore these two lines are often called the Light Cone. The light cone separates the entire space into four regions: top, bottom, left and right. According to relativistic mechanics to be discussed below, no information or signal travels faster than light. If some signal is sent from O at a certain speed, its trace will stay in the top region. That is to say, this region is all we can affect from now at O, hence it is called the Future of O. Similarly, if some signal were to reach us now, it has to be sent from the bottom region and remain in the bottom region. Thus the bottom region is what can affect O from the past, and is called the Past of O. The left and the right regions of the light cone cannot affect O or be affected by O.

The future of O and O can potentially have a causal relationship, so can O and the past of O. One question one may raise is whether a causal relationship is invariant in different inertial frames of reference. That is, if two points (x_1, t_1) and (x_2, t_2) have a causal relationship in S, are they guaranteed to have a causal relationship in S'? If they don't in S, can they have one in S'? This question can be phrased in mathematics like below:

$$\left|\frac{x_2 - x_1}{t_2 - t_1}\right| < 1 \Longleftrightarrow \left|\frac{x'_2 - x'_1}{t'_2 - t'_1}\right| < 1$$

First, let's start with some rewriting:

$$\begin{aligned} \left|\frac{x'_2 - x'_1}{t'_2 - t'_1}\right| &= \left|\frac{\gamma(x_2 - ut_2) - \gamma(x_1 - ut_1)}{\gamma(t_2 - ux_2) - \gamma(t_1 - ux_1)}\right| = \left|\frac{(x_2 - x_1) - u(t_2 - t_1)}{(t_2 - t_1) - u(x_2 - x_1)}\right| \\ &= \left|\frac{\frac{x_2 - x_1}{t_2 - t_1} - u}{1 - u\frac{x_2 - x_1}{t_2 - t_1}}\right| \end{aligned}$$

Let

$$v = \frac{x_2 - x_1}{t_2 - t_1},$$

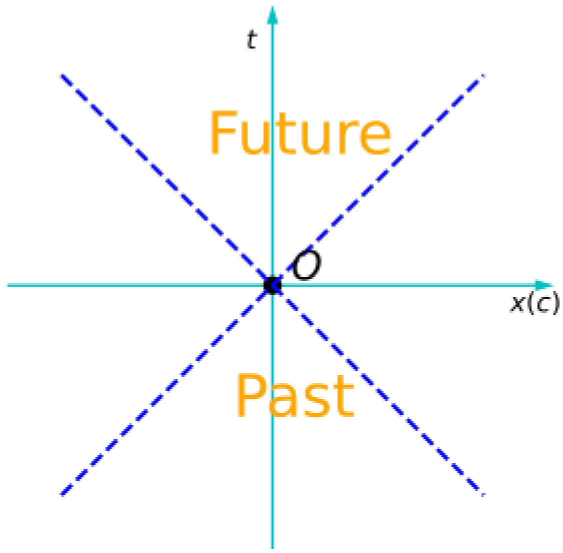

Figure 3: Light cone, past and future

and the problem becomes

$$|v| < 1 \Longleftrightarrow |\frac{v-u}{1-uv}| < 1 \text{ if } |u| < 1$$

which is also equivalent to

$$v^2 < 1 \Longleftrightarrow (\frac{v-u}{1-uv})^2 < 1 \text{ if } u^2 < 1.$$

Proof:

$$\begin{aligned}
&v^2 < 1 \\
\Longleftrightarrow &v^2(1-u^2) < (1-u^2) \text{ because } u^2 < 1 \\
\Longleftrightarrow &v^2 - v^2u^2 < 1 - u^2 \\
\Longleftrightarrow &v^2 + u^2 < 1 + v^2u^2 \\
\Longleftrightarrow &v^2 + u^2 - 2uv < 1 + v^2u^2 - 2uv \\
\Longleftrightarrow &(v-u)^2 < (1-uv)^2 \\
\Longleftrightarrow &(\frac{v-u}{1-uv})^2 < 1
\end{aligned}$$

Now we just proved that a causality relationship does not change across different frame of reference. Therefore the idea of riding on a super fast spaceship hoping to change our past to affect our present is hopeless, at least according to the special theory of relativity.

Also in case it was clear, every point (x, t) in spacetime has its own light cone, its past and its future.

6.2 Lorentz Invariance

H. Minkowski noticed that the Lorentz transformation keeps the following scalar unchanged, when the unit of space and velocity is in c.

$$x^2 + y^2 + z^2 - t^2 \quad (27)$$

This can be easily shown mathematically, at least in our simplified case where the motion of S' relative to S is only in the x direction, as shown below.

$$\begin{aligned}
x'^2 + y'^2 + z'^2 - t^2 &= \gamma^2(x-ut)^2 + y^2 + z^2 - \gamma^2(t-ux)^2 \\
&= \gamma^2(x^2 + u^2t^2 - 2utx - t^2 - u^2x^2 + 2utx) + y^2 + z^2 \\
&= \gamma^2[x^2(1-u^2) + t^2(u^2-1)] + y^2 + z^2 \\
&= x^2 + y^2 + z^2 - t^2
\end{aligned}$$

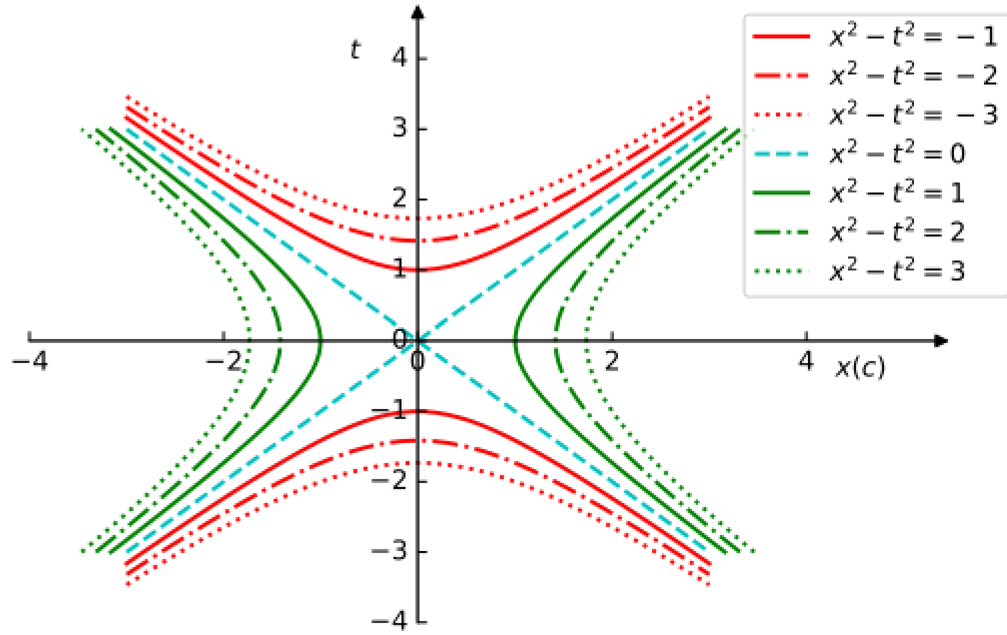

Figure 4: Minkowski Spacetime

In 3-D Euclidean space, the square of the length of a vector (x, y, z) is $x^2 + y^2 + z^2$. Rotation does not change this quantity. By analogy, $x^2 + y^2 + z^2 - t^2$ is defined as the interval length of (x, y, z, t) in the 4-D Minkowski space. This quantity can be negative, and it is kept unchanged under the Lorentz transformation.

Figure (4) shows various lines formed by all the points with $x^2 - t^2$, skipping y and z for now, equal to different constants. For example, the two solid red lines are the hyperbolas described by $x^2 - t^2 = -1$, the two green dotted lines are the hyperbolas described by $x^2 - t^2 = 3$, and the two cyan dashed straight lines are described by $x^2 - t^2 = 0$.

6.3 Minkowski Diagram

The Lorentz transformation is central to the special relativity. It makes sense to have more than one frame of references described in the same Cartesian coordinate system.

First, we'd like to plot the x' and t' axes of S' on the Cartesian coordinate system of S. Pick the unit vectors along the x' and t' axes of S': $(x'_1, t'_1) = (0, 1)$ and (x'_2, t'_2)=$(1, 0)$ and see how they should appear in S. Simple application of the Lorentz transformation tells us that

$$(x'_1, t'_1) = (0, 1) \Rightarrow (x_1, t_1) = (\gamma u, \gamma)$$
$$(x'_2, t'_2) = (1, 0) \Rightarrow (x_2, t_2) = (\gamma, \gamma u).$$

Since $\frac{t_2}{x_2} = \frac{x_1}{t_1} = u$, the angle between the t axis and the t' axis is the same as that between the x axis and the x' axis between. The magnitude of the angle is is defined by $\tan\theta = u$. Note that (x_1, t_1) and (x_2, t_2) are not unit vectors. In fact, their length, in Euclidean sense, not Minkowski sense, is

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

Figure 5 shows a pair of green (x', t') axes and another pair of red (x', t') axes, besides the original (x, t) axes. One simple way to visualize these (x', t') axes is to think about rotating the (x, t) axes and shrink the angle between them symmetrically.

If we use L_u to denote the matrix representation of the Lorentz transformation from S to S', we have

$$L_u = \gamma \begin{pmatrix} 1 & -u \\ -u & 1 \end{pmatrix}$$

and the matrix representation of the Lorentz transformation from S' to S is

$$L_{-u} = \gamma \begin{pmatrix} 1 & u \\ u & 1 \end{pmatrix}.$$

What will the coordinates of (x'_0, t'_0) be on the transformed (back to S) (x', t') axes?

$$\begin{aligned}(x_0, t_0)^T &= L_{-u}(x'_0, t'_0)^T \\ &= L_{-u}(x'_0(1, 0)^T + t'_0(0, 1)^T) \\ &= x'_0(L_{-u}(1, 0)^T) + t'_0(L_{-u}(0, 1)^T)\end{aligned}$$

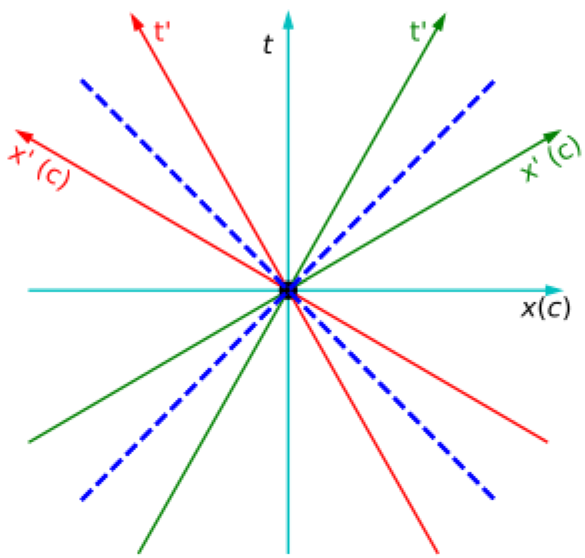

Figure 5: Three frames of reference

Since $L_{-u}(1,0)^T$ and $L_{-u}(0,1)^T$ are exactly the transformed version of (x',t') axes, (x'_0,t'_0)'s coordinates will be $x'_t\beta$ and $t'_0\beta$ regarding the transformed (x',t') axes. The factor β is to compensate for the fact that $L_{-u}(1,0)^T = (\gamma, \gamma u)^T$ and $L_{-u}(0,1)^T = (\gamma u, \gamma)^T$ are not unit vectors.

A lot of videos and articles online use Minkowski Diagrams to visually show the length contraction and time dilation effects. They are almost all wrong as they ignored the β scaling factor.

6.4 Loedel Diagram

Minkowski Diagrams have two problems that can confuse people. First, though the x axis and the t axis are perpendicular, the x' axis and the t' axis form an acute angle. Second, we can not directly compare lengths, again in Euclidean sense, of different vectors visually because the coordinates of (x',t') contain a scaling factor β.

Now if a median reference frame S_0 is introduce that is moving at speed $u/2$ relative to S. The axes of S and the axes of S', when transformed to S_0, will be symmetric and have the same scaling factors. So visual comparisons between S and S' can be done directly. Such diagrams are called Loedel Diagram, as depicted in Figure 6. The (x_0,t_0) axes are omitted as they are not often used. It is also easy to show that t' and x are perpendicular, so are t and x'.

6.4.1 Time Dilation Revisited

Loedel diagrams can be used to visually show the time dilation effect, the phenomenon where two events occurring at the same location but different times in S are observed in S' to have a greater time gap. That is

$$\begin{cases} x_1 = x_2 \\ t_2 > t_1 \end{cases} \Rightarrow t'_2 - t'_1 > t_2 - t_1.$$

If we pick, for simplicity, $x_1 = x_2 = 0$ and $t_1 = 0$, (x_1,t_1) is the origin, (x_2,t_2) is on the t axis, (x'_1,t'_1) is also the origin, and (x'_2,t'_2) won't be on the t' axis, but t'_2 will be larger than t_2, as shown in the left plot of Figure 7. Remember that t'_2 is simply the second event's coordinate in the t' axis. The right plot in Figure 7 demonstrate the same time dilation phenomenon when observed in S.

6.4.2 Length Contraction Revisited

Loedel diagrams can also be used to visually demonstrate the length contraction effect, the phenomenon where objects in S appear to be shorted in S' along the x-axis. Keeping in mind that an object's length has to be measured at the same time in the frame of reference moving relative to the object, we can translate the length contraction effect in math as below:

$$\begin{cases} x_2 > x_1 \\ t'_1 = t'_2 \end{cases} \Rightarrow x'_2 - x'_1 < x_2 - x_1.$$

For simplicity, we can pick $t'_1 = t'_2 = 0$ and $x'_1 = 0$, so the first event will be at the origin, and the second event is on the t' axis. The second event's x coordinate $x_2$will be larger than its x' coordinate x'_2 as demonstrated by Figure 8.

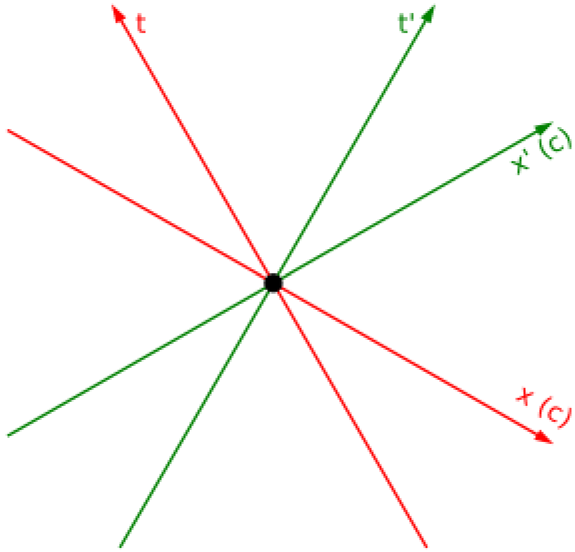

Figure 6: Loedel Diagram

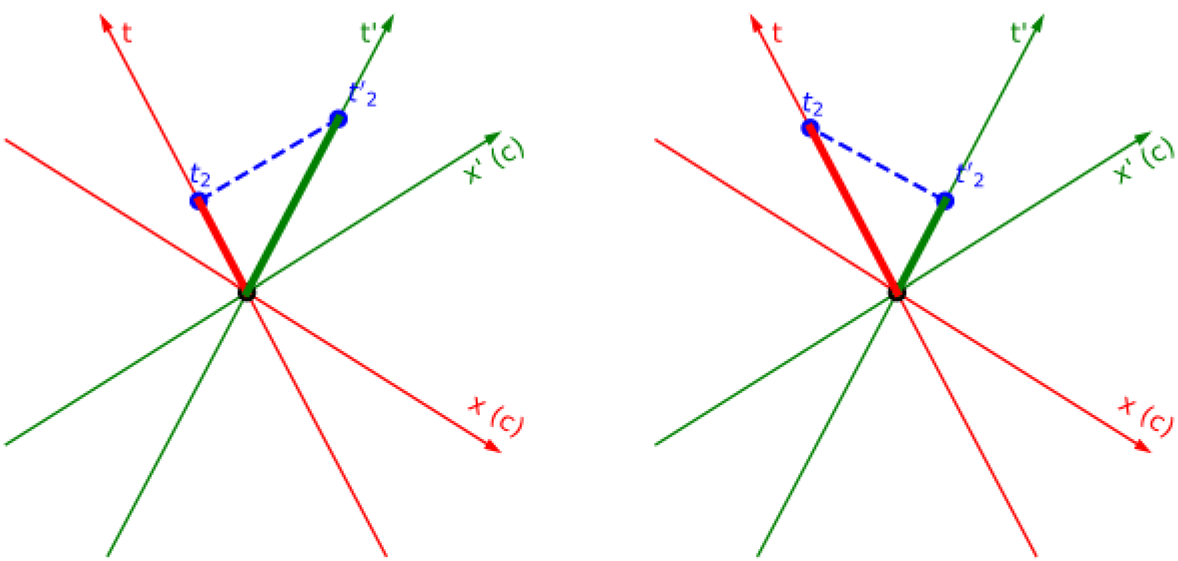

Figure 7: Time Dilation Demonstrated with a Loedel Diagram

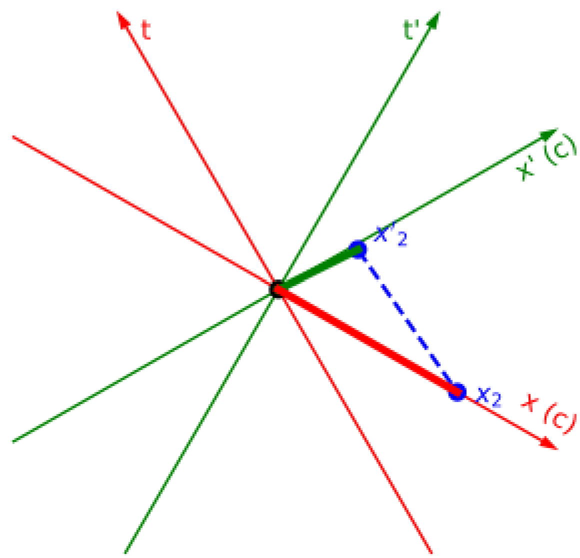

Figure 8: Length Contraction Demonstrated with a Loedel Diagram

7 Relativistic Mechanics

7.1 Mass, Momentum and Dynamics

In Newtonian mechanics, the second Newtonian law is

$$F = \frac{dP}{dt} = \frac{d(mv)}{dt} = m\frac{dv}{dt}.$$

The last step of the derivation is valid because m is constant. In relativistic mechanics, m is no longer constant. Instead it if a function of m_0 which is the object's mass when it is stationary and v:

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}}. \tag{28}$$

The definition of mass in (28) means it is no longer a constant, but instead increases with velocity. However in everyday situations, v is so small compared with c, m can be approximated by a constant, and classical Newtonian mechanics works very well.

The definition of P and F do not require any further adaptations; therefore

$$P = mv = \frac{m_0}{\sqrt{1 - v^2/c^2}}v$$

and

$$F = \frac{d(\frac{m_0}{\sqrt{1-v^2/c^2}}v)}{dt}.$$

If action and reaction are still equal, the conservation of momentum still applies. One important conclusion of relativistic mechanics is that an object cannot be accelerated to the speed of light, simply because as the object's velocity increases, its mass keeps increasing, and the effect of the same amount of force on the mass decreases. When the object's speed approaches that of light, its mass approaches infinity. The object's momentum however, can continue to increase forever given a constant force.

The derivation of (28) involves the application of the special relativity and the conversation of momentum and energy in different inertial frames of references, respectively. We'll skip the details here.

7.2 Equivalence of Mass and Energy and $E = MC^2$

Consider the motion of molecules in a tank of gas. When the gas is heated, the speed of molecules increases and the mass also increases. When v is relatively small, we can apply Taylor expansion to (28) and get

$$\begin{aligned} m &= \frac{m_0}{\sqrt{1 - v^2/c^2}} \\ &= m_0(1 - v^2/c^2)^{-1/2} \\ &= m_0(1 + \frac{1}{2}v^2/c^2 + \frac{3}{8}v^4/c^4 + \cdots) \\ &\simeq m_0 + \frac{1}{2}m_0 v^2/c^2 \\ &= m_0 + \Delta(K.E.)/c^2. \end{aligned}$$

Multiplying both sides of the above equation by c^2, we have

$$mc^2 = m_0c^2 + \frac{1}{2}m_0v^2 + \cdots .$$

The term on the left is the total energy of a body, if we agree with Einstein, and the second term on the right is the ordinary kinetic energy. Einstein interpreted the large constant term m_0c^2 to be an intrinsic energy known as the "rest energy."