What does mass have to do with Energy anyways?

An intro to Special Relativity

Jujhaar Singh

Mentor: Nabeel Ahmed

Special Theory of Relativity

Jujhaar Singh

Mentor: Nabeel Ahmed

Contents

Introduction

This document contains a brief description of some of the key ideas of the Special Theory of Relativity. We follow the book "Special Relativity and Classical Field Theory" by Susskind and Friedman closely in its ideas and overall direction.

We shall first look at the effect on the coordinates on change of reference frame, which will be characterized by the Lorentz Transformation and find out how it arises and what it means physically. After taking a quick dive into some interesting thought experiments, we will define the quantities that we will be using in our new system and strive to write laws of motion for them.

1 Reference Frames and the Lorentz Transformation

The Special Theory of Relativity is mainly a theory about reference frames. We aim to find out that what remains the same between frames, and what changes, essentially, how physical quantities vary.

1.1 Reference Frames

In physics, we deal with the motion and behavior of things that we observe. Of course, if we are to talk about objects in space, we would need some way to express physical parameters of said objects. For this purpose, we define Frames of Reference.

> *Definition 1.1: Reference Frames*
>
> In space-time we say that a set of axis of x, y, z and along with that a reference for time is called a frame of reference.

Having defined frames of reference, we are ready to define a very special subcategory that is an integral part of the rest of our study.

> *Definition 1.2: Intertial Reference Frames*
>
> We define inertial frames of reference as those in which particles, on which no external force is acting, move in a straight line with constant speed. For a frame to be inertial, it must have zero acceleration.

Ones of the aims for both Einstein and Newton was to find a set of laws that can accurately predict the motion of particles, and are applicable, without modification, in all inertial reference frames. This led to Einstein figuring out that if the laws remain the same, then the speed of light should also be the same in all frames of reference and for all observers.

Units and Dimensions

In the units we generally use, the SI system, a velocity of magnitude 1 relates to an object moving one meter in one second. This also has the dimension of length over time. We shall call these "conventional or common units".

Since a lot of Relativistic physics deals with the speed of light, we would like to introduce a new set of units such that they have an inherent connection to the speed of light. In fact, this is a great way to reduce the complexity of many of the expressions we deal with. We define our new units in such a way that a velocity of magnitude 1 will be equal to the speed of light in conventional units. We will also

choose it in such a way that this quantity is dimensionless. We will refer to these units as "Relativistic units" for the rest of the document.

Converting to and fro

Later in the document, formulas may be represented in either set of units. To convert from one to the other, we simply use dimensional analysis and divide or multiply some power of c, the speed of light, to each term such that the eqaution becomes dimensionally correct.

1.2   Problem with Newtonian Frames of Reference

We consider a simple one dimensional system with the x axis positive on the right hand side and negative on the left hand side. For an inertial frame A, say we have another frame B moving to the right with a speed of v. We can write down a simple equation to describe its motion.

$$x = vt \tag{1.1}$$

Now, for the stationary frame, A, we can say that a light ray moving to the right would have the following equation to describe its motion.

$$x = ct \tag{1.2}$$

Looking at the same ray of light in the frame B, we get the following equation.

$$x = (c - v)t \tag{1.3}$$

This makes sense to us by the common knowledge that we have of Newtonian physics. Looking closely at the equations, we find that the speed of light seems to be different in both. But recall what we said towards the end of subsection 1.1 Reference Frames. If we examine that assumption, then we see that since the laws of physics remain the same in all inertial frames, the speed of light, which is a direct result of those laws, should also remain the same in all reference frames. Clearly, that is not followed in Newtonian reference frames as shown above.

We will see that the principle assumption of Newtonian reference frames, that is, that the time in all frames moves at the same rate, is challenged by the theory of relativity. In fact, deleting that assumption is the only way to make the laws of physics consistent, as we will see later.
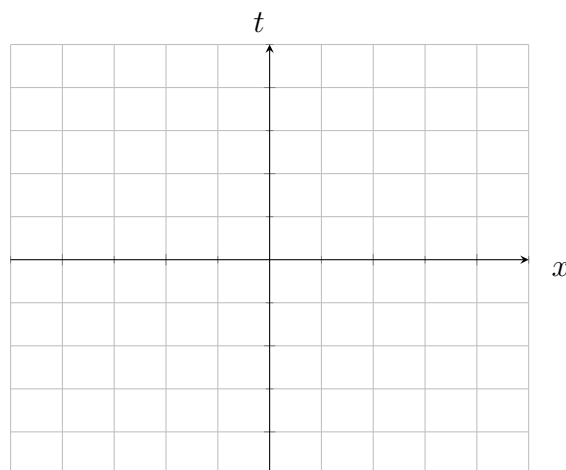
Figure 1: basic space-time plot

1.3   Transformations in Relativistic Reference Frames

1.3.1   Space-time

We will generally represent space-time with two axes, that is, one for time and one for distance, since we will consider a single dimension of space for simplicity. Space-time plots will generally look something like this basic space-time plot.

A more formal definition can also be given.

Definition 1.3: Space-time

The concepts of time and three-dimensional space regarded as fused in a four-dimensional continuum.

We will see that this same definition can be used in a simplified form for just the x axis. In this case, it would be the time dimension plus one dimensional space, giving us a two dimensional space representing both space and time.

Definition 1.4: Space-time events

We define a space-time event as one which has a time coordinate as well as space coordinates. In an n-dimensional space, a space-time event would look like $(t, x_1, x_2 \ldots x_n)$ where $x_1, x_2 \ldots x_n$ represent n space coordinates.

1.3.2   Synchronizing

To be able to talk about space-time events in different reference frames, we would need some sense of what "synchronous" events would mean. One simple definition

that can be used is given below.

> **Definition 1.5: Synchrony**
>
> Say we have 3 objects, A, B and C, moving through space in such a way that B is always at the midpoint of A and C. Let their paths be defined as $a(t)$, $b(t)$ and $c(t)$ respectively. Now, a light ray is sent from A to B at t_1 and another is sent from C to B at t_2. If these two reach B at the same time, then we say that the space-time events $(t_1, a(t_1))$ and $(t_2, c(t_2))$ synchronous.

The meaning of synchronous changes according to the observer and his frame of reference. With respect to space-time plots, two points that are synchronous should have the same time coordinate with respect to the axes of that frame to be considered synchronous. This of course only applies in the frame which we were considering, thus they may happen at different times in a different frame of reference.

1.3.3   The Lorentz Transformation

We know that any reference frame can be represented by the time and space axes. For yourself, your distance and time axes are perpendicular to each other. We take the time axis along the vertical and x(distance) along the horizontal.

If we would like to find how the axes looks for an observer who is moving relative to us, we can use the notion of Space-time events and their Synchrony to find the position of their axes.

Using the above discussed concepts of the synchrony of events, constancy of speed of light in all frames, and a few other concepts and manipulations, we arrive at the Lorentz Transformation. This transformation dictates how the space and time axes look for a person in a different inertial frame of reference from our resting frame.

$$t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \tag{1.4}$$

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} \tag{1.5}$$

$$y' = y \tag{1.6}$$

$$z' = z \tag{1.7}$$

This kind of a transformation is also referred to as a boost along the x direction.
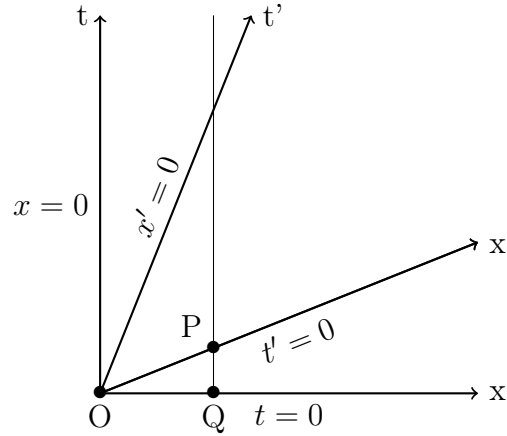
Figure 2: length contraction

1.4   Length Contraction and Time Dilation

Length Contraction

Say I am stationary and standing at the origin. I have a meter stick in my hand, which extends from the origin, O, to the point Q in my frame of reference. Now suppose a person A, who is moving towards the positive x direction at a speed v, moves past me. At the point that he crosses me, if he were to measure my meter stick with his set of axes, will the meter stick have a length of $1m$? (assume the time that he crosses me is $t = 0$ in both our clocks)

In fact, it will not, for him the stick will appear shorter. This is the phenomena of length contraction and it is represented in Figure 2(t, x are my axes and t', x' represent his). For person A, points O and Q are not simultaneous since they do not lie on the x' axis. For him, the stick is along OP when he crosses me. That is, the length of the stick for him is the length along the axis x' from the point O to the point P.

Something looks a little suspicious here. Doesn't OP look longer than OQ? Then how come the length he sees is shorter? This is because the ticks along the x' axis have a different spacing than the ticks along the x axis. It is a simple consequence of the Lorentz Transformation and you can easily verify this.

Thus, the length of the meter stick measured by person A will be given by Equation 1.8 [1].

$$\sqrt{1 - v^2} \tag{1.8}$$

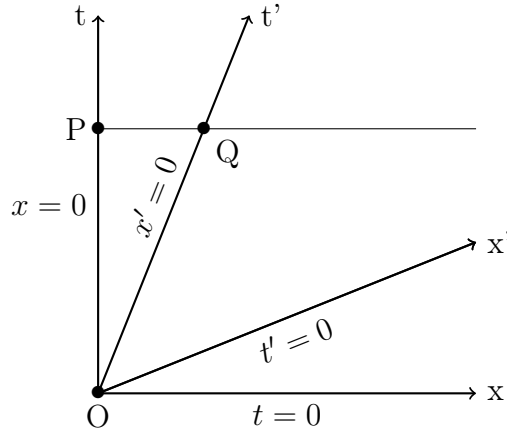[1]note that we are working in relativistic units

Figure 3: time dilation

Time dilation

Time dilation works in a similar way. Say I am now moving with a clock in my hand while you are standing still. I started the clock at $x = 0$ and at point Q I see that the clock reads time $t' = 1$. The question now is, what is the time on your clock at this point? This situation is represented in Figure 3

Notice that the surface PQ is what you call synchronous. Using the Lorentz Transformation once again, you can obtain the expression in Equation 1.9 [2] for time passed in the stationary reference frame.

$$\frac{1}{\sqrt{1 - v^2}} \tag{1.9}$$

1.5   Minkowski Space

The space-time diagrams we draw exist in what is called Minkowski Space. This is similar to our notion of cartesian space which contains the x, y, and z dimensions. In Minkowski Space, instead of having just the three spatial dimensions, we also have time represented along another direction. For obvious reasons we cannot really represent, or even imagine, a Minkowski space in 3 space dimensions and 1 time dimension since we run out of direction to orient our axes in. Thus we generally use diagram draw for 2 spatial dimensions and 1 time dimension.

---

[2]again, we are working in relativistic units

Invariants

The concept of invariants is very widely used in physics, but you may not have pondered on it too much. This is because in our box of Newtonian mechanics, we have a very intuitive invariant. Time.

In Relativistic terms however, time is clearly not an invariant as we have seen above. From frame to frame we have different measurements of time and all observers do not agree on it. It can be found with some algebra that the invariant in Relativistic cases is something that we call "proper time". It's symbol is τ.

$$\tau^2 = t^2 - x^2 - y^2 - z^2 \tag{1.10}$$

This quantity, just like time in the Newtonian world, is constant over all frames of reference, for all observers and through any Lorentz transformation. Thus it is an invariant.

Even though right now it may not mean much to you, τ^2 does hold physical and experimental significance. Say we are sitting at the origin and someone, person A, crosses us. When t seconds pass in our frame of reference person A is at x_0 in our frame. Say at this point, t' seconds have passed in person A's frame. Since the proper time of the event at the location of person A must be same in both the frame, we get the following.

$$\tau'^2 = t'^2 - x'^2 \text{ but } x' = 0 \text{ for person A} \tag{1.11}$$
$$\implies \tau'^2 = t'^2 \tag{1.12}$$
$$\text{also, } \tau^2 = t^2 - x^2 \tag{1.13}$$
$$\implies t'^2 = t^2 - x^2 \tag{1.14}$$

We also define the space-time interval as the following. Of course this quantity too is an invariant.

$$s^2 = x^2 + y^2 + z^2 - t^2 = -\tau^2 \tag{1.15}$$

The Light Cone and Space-time Separations

Imagine a two dimensional plane on which you represent x and y. Now add a third line perpendicular to the plane in order to represent time. This now forms a Minkowski Space of two space dimensions and one time dimension.

Now imagine how you would represent a light "ring" emitting from the point $x = 0$ and $y = 0$ at time $t = 0$. This is the analog of a point source of light in two dimensions. Notice that it would form a cone in the Minkowski Space. This is illustrated in Figure 4 Light Cone.
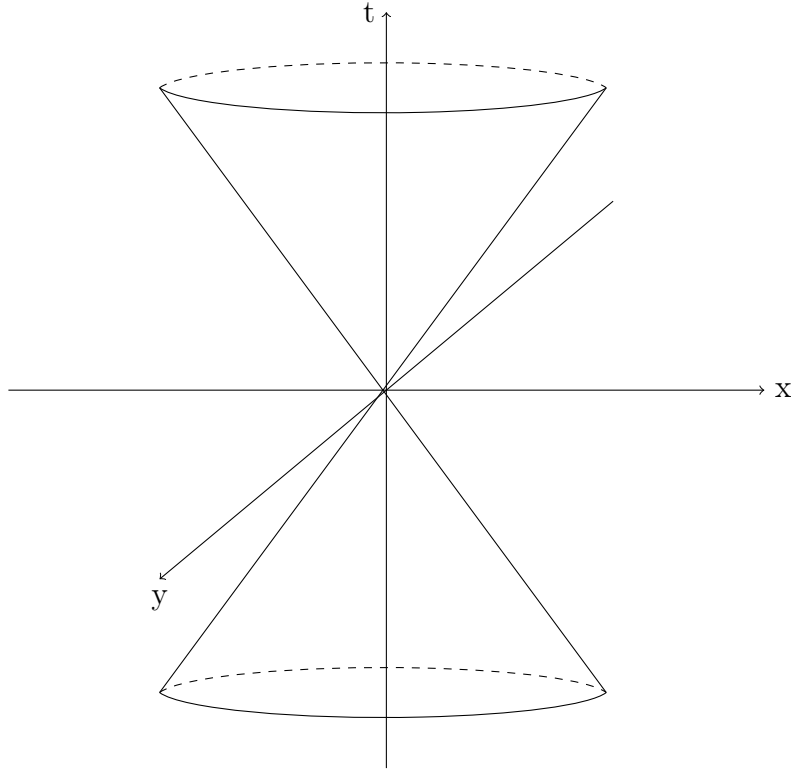
Figure 4: Light Cone

The equation of this cone is given by $x^2 + y^2 + z^2 = c^2t^2$ as you might have guessed. But in relativistic coordinates of course this simplifies to $x^2 + y^2 + z^2 = t^2$.

Now we define a "Time-like" event.

Definition 1.6: Time-like event

When an event lies within the light cone, ie $x^2 + y^2 + z^2 < t^2$, then we say that the separation is "time-like" with respect to the origin. The time separation is larger than the space separation from the origin for these events.

For a time-like event, we can always find a frame such that the two events occur at the same spatial location in that frame but at different times. We also say that there is no frame in which these two events occur simultaneously.

Similarly we define "Space-like" events.

> **Definition 1.7: Space-like event**
>
> When an event lies outside the light cone, ie $x^2 + y^2 + z^2 > t^2$, then we say that the separation is "Space-like" with respect to the origin. The space separation is larger than the time separation from the origin for these events.

For a space-like event, we can always find a reference frame such that the two events occur simultaneously but at different spatial locations in that frame. We cannot, however, find a reference frame in which both the events occur at the same location in space.

Finally we also define "light-like" intervals.

> **Definition 1.8: Light-like Events**
>
> When an event lies on the light cone, ie $x^2 + y^2 + z^2 = t^2$, then we say that the separation is "light-like" with respect to the origin. The time separation is equal to the space separation from the origin for these events.

If two events are light-like it means that there is a possibility of a light ray going from one and reaching the other.

2   4-Vectors and Velocities

2.1   4-Vectors

> **Definition 2.1: 4-Vector**
>
> If a vector with 4 coordinates transforms according to the rules of the Lorentz transform then we call it a 4-Vector.

A 4-Vector can be used to define an event in a space-time diagram. It consists of one time coordinate and three space coordinates. The first coordinate is time and the next three are space coordinates, say x, y, and z.

Some notation

We may use the following notation to represent a 4-Vector.

$$(X^0, X^1, X^2, X^3) \tag{2.1}$$

The above representation implies the following.

$$t = X^0 \tag{2.2}$$
$$x = X^1 \tag{2.3}$$
$$y = X^2 \tag{2.4}$$
$$z = X^3 \tag{2.5}$$

We will also use X^i to talk about only the space components and X^μ to talk about all four components together.

$$X^i \implies (X^1, X^2, X^3) \tag{2.6}$$
$$X^\mu \implies (X^0, X^1, X^2, X^3) \tag{2.7}$$

2.2   4-Velocity

This is just the counterpart to velocity for 4-Vectors. One major difference from normal velocity is that here we consider variation over proper time instead of time. This is convenient because proper time is an invariant too. Thus we define 4-Velocity as the following.

$$U^0 = \frac{\mathrm{d}X^0}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau} \tag{2.8}$$

$$U^1 = \frac{\mathrm{d}X^1}{\mathrm{d}\tau} = \frac{\mathrm{d}x}{\mathrm{d}\tau} \tag{2.9}$$

$$U^2 = \frac{\mathrm{d}X^2}{\mathrm{d}\tau} = \frac{\mathrm{d}y}{\mathrm{d}\tau} \tag{2.10}$$

$$U^3 = \frac{\mathrm{d}X^3}{\mathrm{d}\tau} = \frac{\mathrm{d}z}{\mathrm{d}\tau} \tag{2.11}$$

2.3   More on 4-Velocity

There is an interesting relation between normal velocity and 4-Velocity. We will denote the 4-velocity using U^i and the normal velocity using V^i.

$$U^0 = \frac{1}{\sqrt{1 - v^2}} \tag{2.12}$$

$$U^i = \frac{V^i}{\sqrt{1 - v^2}} \tag{2.13}$$

Another property of 4-velocity is that its components are connected by an invariant. This is analogous to how proper time is an invariant for position represented

by 4-vectors. This invariant quantity is equal to 1 and further implies that only any three components of the 4-velocity are independent and the fourth is a dependent.

$$(U^0)^2 - (U^1)^2 - (U^2)^2 - (U^3)^2 = 1 \qquad (2.14)$$

3   Relativistic Laws of Motion

3.1   Principle of Least action, Lagrangian and Hamiltonian

This principle states that a process will happen in such a way that a certain defined quantity called action is minimized over the process. The action is the integral of Lagrangian over time during the process. Based on the system we define the Lagrangian.

In a simple Newtonian system where no forces act on a particle, the Lagrangian is simply the kinetic energy of the particle.

The Hamiltonian of a system is generally representative of the total energy of a system and a simple relation exists between the Hamiltonian and the Lagrangian given by the following.

$$H = \sum_i \dot{X}^i P^i - \mathcal{L} \qquad (3.1)$$

Here P^i represent the components of the particle's momentum.

3.2   Relativistic Action

In relativity, we define the action as the following.

$$Action = -m \int_a^b d\tau = -m \int_a^b \sqrt{1 - v^2} dt \qquad (3.2)$$

This leads us to defining the Lagrangian for relativistic systems.

$$\mathcal{L} = -m\sqrt{1 - v^2} = -m\sqrt{1 - (\dot{X}^i)^2} \qquad (3.3)$$

If we rewrite this expression in terms of conventional units we obtain the following.

$$\mathcal{L} = -mc^2\sqrt{1 - \frac{v^2}{c^2}} \qquad (3.4)$$

3.3   Relativistic Momentum

In Lagrangian mechanics we define the moment as the derivative of the Lagrangian $\mathcal{L}$ with respect to $\dot{X}$. So now we try to extend this notion to try and define relativistic momentum.

$$P^i = \frac{\partial \mathcal{L}}{\partial \dot{X}^i} \tag{3.5}$$

This gives us the following.

$$P^i = \frac{mV^i}{\sqrt{1 - \frac{v^2}{c^2}}} = mU^i = m\frac{\mathrm{d}X^i}{\mathrm{d}d\tau} \tag{3.6}$$

3.4   Relativistic Energy

To define energy, we take help of (3.1) and try to extend its definition beyond what it is meant for. This gives us the following.

$$H = \frac{m}{\sqrt{1 - v^2}} \tag{3.7}$$

Interestingly this is the "zeroth" component of our relativistic momentum, given by (3.8).

$$P^0 = mU^0 = H = \frac{m}{\sqrt{1 - v^2}} \tag{3.8}$$

This P^0 with our previously discovered P^i gives us a 4-vector! The implications of this are that the energy and the momentum of a particle get mixed up during a lorentz transform. This is why we may find the particle having a momentum in one frame but not the other or another particle having energy in one frame but not the other.

Another important take away is that our previous notion of conservation of momentum now turns into the conservation of the 4-momentum instead.

This energy written in conventional units gives us (3.9)

$$H = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3.9}$$

Slow Particles and mc^2

Now imagine the case of a particle for which v is far less than the speed of light. This gives us the following relation using approximations.

$$H = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} = mc^2\left(1 + \frac{v^2}{2c^2}\right) = \frac{1}{2}mv^2 + mc^2 \tag{3.10}$$

This tells us that when the particle is at rest, ie $v = 0$, the energy of the particle will be given by the famous Einstein equation (3.11).

$$E = mc^2 \tag{3.11}$$

This quantity represents the energy of the body in a frame where it is at rest. This equation also famously represents mass-energy equality and is thus an equation of utmost importance.

Relation between energy, momentum and mass

Now recall the relation (2.14). This relation defines the invariant quantity for 4-velocity. As it turns out, we can use it to arrive at a relation between the above mentioned quantities. We simply multiply the whole equation with m^2 and then substitute various quantities based on the definitions given by us. We finally obtain:

$$E = \sqrt{P^2 + m^2} \tag{3.12}$$

Rewriting this in conventional units gives us a less famous but equally important relation.

$$E = \sqrt{P^2c^2 + m^2c^4} \tag{3.13}$$

This holds significance because it gives us a complete description of the energy term. It is valid in all cases including for particles with no mass, no momentum, etc.

4   Conclusion

With this we get to the end of this paper. While it has been short, I certainly hope that it was of value to the reader. One major objective was to explain what the whole fuss about the mystical relation between energy and mass is. I hope that this paper has been adequate to convey the relevance of the same.