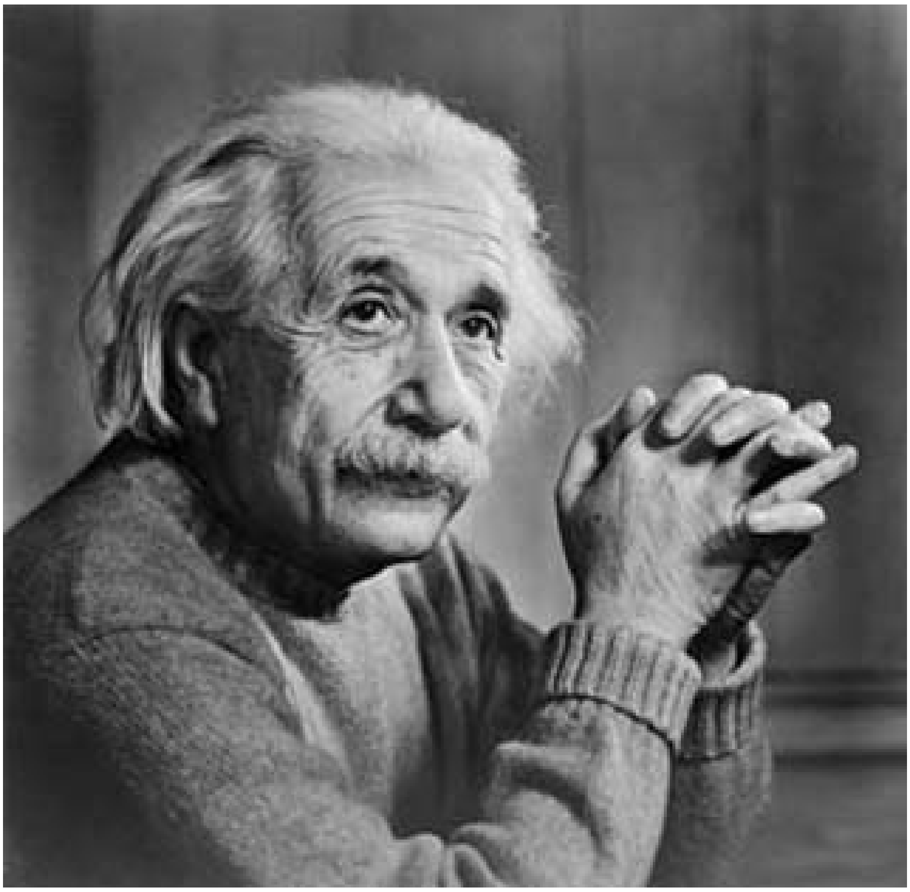
Summer of Science 2020

Final Report

Einstein's Theory of Relativity



by

Sanket J Hanamashetti

Roll No.-190070057, EE

Mentor-Shoaib Akhtar

Contents

List of Figures

Chapter 1

Failure of Classical Mechanics

This chapter follows the developments which led to the birth of Special Relativity in 1905. The Newtonian ideas of mechanics and electromagnetism and the apparent inconsistencies are discussed.

1.1 Frames of Reference

The study of Physics can be termed as *mathematical analysis* of nature which involves taking measurements of space, time, fields etc and searching for a relation amongst such quantities with the help of physical intuition. The observer (one who takes the measurements) , rather the observer's characteristics (like velocity, acceleration, etc [wrt[1] to another observer]), affects the outcome of the measurements. The observer is termed as a **frame of reference**. Formally, a frame of reference can be defined as a fixed observer with a set of axes, clocks and other instruments. There are two kinds of frames of reference which are discussed below.

Consider an isolated body (B_0) i.e., a body which is far away from other bodies, having no interaction with them. Such a body can naturally be defined as having zero acceleration ($a = 0$) for according to Newton, the accelaration of a body is given by $a = \frac{F}{m}$ and since the body is not interacting with any other body, the net force on it will be zero ($F = 0$).

Now, consider a frame of reference, in which the aforementioned isolated body B_0 is observed to be moving with a constant velocity v_0. Such a frame of reference is called an **Inertial Frame of Reference**. It is clear that the accelaration of B_0 as observed from that frame would be zero. Consider another frame of reference accodring to which B_0 is moving with a constant velocity u_0. Then, the former frame will see the latter frame to be moving with a constant velocity (between $u_0 - v_0$ and $u_0 + v_0$, depending on the directions), which is obvious from the basic Newtonian velocity addition formulae.

Consider another frame which observes B_0 to be moving with an acceleration $a (\neq 0)$. Such frames of reference are called **Non-Inertial Frames of Reference**.

1.2 Newtonian Ideas

According to Newton, length, mass and time are invariant. Any observer would report the same values for lengths, masses and time intervals according to him. The idea of a

[1] with respect to

'universal clock' ticking away was one of the causes for the downfall of the classical ideas.

Inertial frames of reference are a very special kinds of frames. It can be proved that, classically, the Newton's laws of motion are invariant (retain the same form) in different inertial frames of reference. Unless mentioned otherwise, whenever we say 'frame of reference' it means an inertial frame of reference.

Consider two inertial frames of reference, S and S' with origins O and O' respectively. According to S, let S' move with a constant velocity v_0 and let the position vector of O' at time t, according to S, be $R(t)$[2]. Let there be an object P, whose position vector wrt S be $r(t)$ and wrt S' be $r'(t)$ at time t as shown in Fig.1.1. According to classical

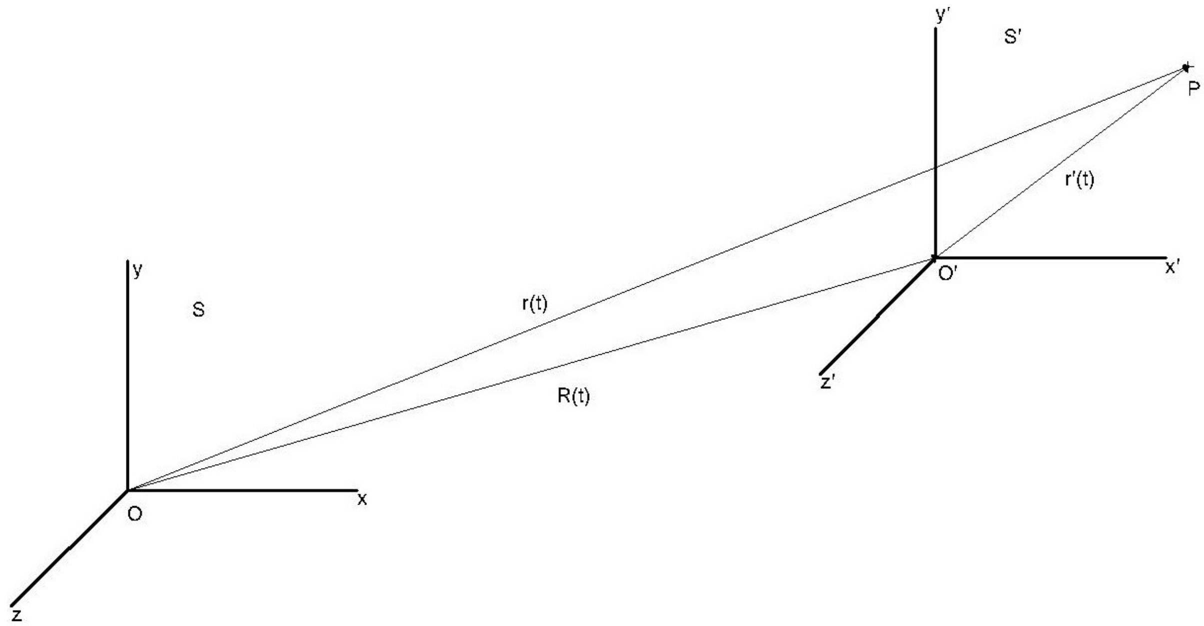


Figure 1.1: Arbitrary Inertial Frames of Reference

mechanics,

$$r(t) = R(t) + r'(t)$$

Differentiating the above equation once wrt t,

$$v(t) = v_0 + v'(t)$$

where $v(t)$ and $v'(t)$ are the velocities of the P wrt S and S' respectively. Differentiating the above equation wrt t,

$$a(t) = a'(t)$$

where $a(t)$ and $a'(t)$ are accelerations of P wrt S and S' respectively.
Hence, we proved that, as measured by two inertial frames, the *accelaration of a body is the same*[3]. Which meanse that a is universal (absolute, unlike velocity) in the Newtonian domain.

Thus, we can say that Newton's second law, $F = ma$, is universal and invariant amongst the inertial frames i.e., if it holds true in one inertial frame, it will be true for all inertial frames. If $F_{net} = 0$, a body would move unaccelerated wrt all inertial frames

[2]It should be noted that vectors do not depend on the frames of reference
[3]And hence, force is equal wrt all Inertial Frames

and would maintain its state of rest or motion unless acted upon by an external force which means that Newton's First Law too is universal and invariant amongst the inertial frames. Since, the value of forces is the same for all inertial frames, if the third law holds true in one inertial frame, it will be so in all inertial frames. Hence, Newton's laws of motion are universal for all inertial frames of reference.

The Newtonian prediction that the force on a body measured by all inertial frames is the same is a very bold statement. As we look further, we will see that it poses a few problems for the Classical ideas.

1.3 Galilean Transformations

Let's discuss about how two inertial frames report the space-time coordinates of the same event[4] in the Classical setup. Consider two inertial frames of reference S and S' with the x and x' axes coincident, y and y' axes parallel and z and z' axes too parallel. Let their origins O and O' coincide at time $t = t' = 0$. As seen by S, let S' be moving with a with a speed v in the positive x direction as shown in Fig.1.2.

Consider an object B_0 at rest in the S' frame at the point $P'(x', y', z')$ at time t'. Let

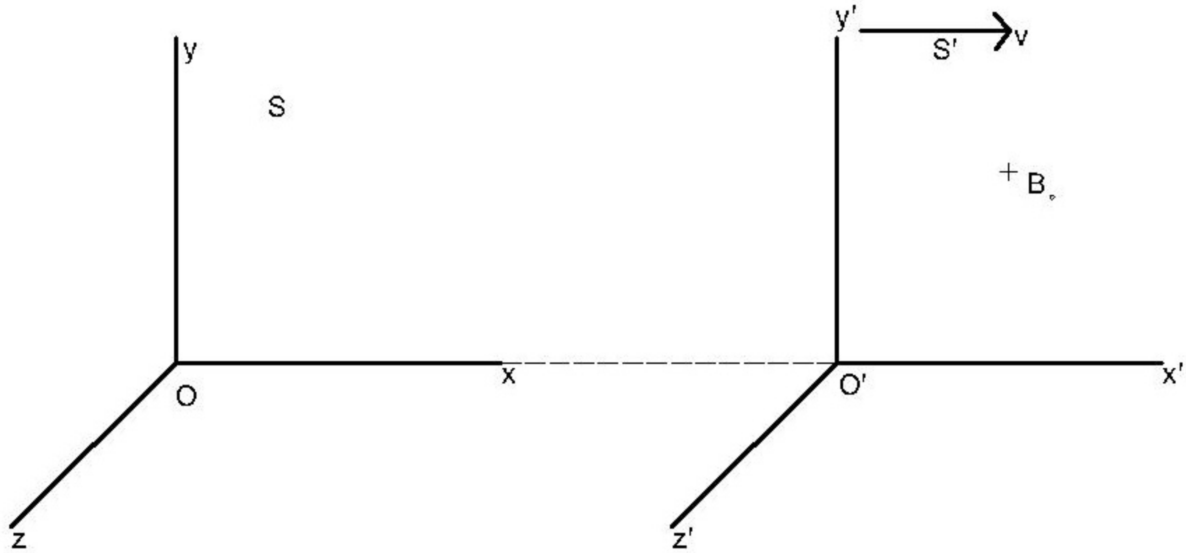


Figure 1.2: Standard Inertial Frames S & S'

the coordinates of the object wrt S be (x, y, z). The x-coordinate of O' is vt, clear from the way the frames are setup. Using the Classical relative coordinate idea, we at once see that the x-coordinate of B_0 will be

$$x = x' + vt$$

Since there is no motion of B_0 along the y' or z' axes and there is no motion of S' along y or z axes,

$$y = y'$$
$$z = z'$$

[4]An event is an occurence at a certain time which can be assigned space coordinates

Finally, as assumed in Classical Mechanics,

$$t = t'$$

We can write the transformation equation for the x coordinate in a more appealing manner,

$$x = x' + vt'$$

Thus, we have obtained the transformation laws governing the Classical domain. The inverse transformations are easily obtained-

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

$$t' = t$$

The above mentioned transformation equations are called *Galilean Transformations.*

We should make note of some obvious results. Consider two events A and B whose coordinates in S frame are (t_A, x_A, y_A, z_A) and (t_B, x_B, y_B, z_B) respectively. Their coordinates then transform into $(t_A, x_A - vt_A, y_A, z_A)$ and $(t_B, x_B - vt_B, y_B, z_B)$ in S' frame. It is clear that the time interval between these two events $(t_A - t_B)$is the same as measured by both S and S'. Suppose $t_A = t_B$ i.e., the two events are simultaneous. Then, the length intervals ($x_A - x_B$, etc)along the axes remain the same. Thus, in the Classical domain, length and time intervals remain indisputably same for all inertial observers. There is no length contraction or time dilation, unlike Special Relativity.

1.4 Newton, we have a problem!

One cannot go too far without problems while following the Classical ideas. Let's look at some of them

- Consider a charge (q) at rest in a combination of electric (**E**) and magnetic fields (**B**). The force on the charge is given by $\mathbf{F}= q\mathbf{E}$. Now, consider an inertial frame moving with a velocity **-v** wrt to the charge. Then, in that frame, the charge moves with a velocity **v**. Then, the force on the charge in that frame is given by $\mathbf{F}'= q(\mathbf{E}+\mathbf{v}\mathrm{x}\mathbf{B})$.
  According to Newton, $\mathbf{F} = \mathbf{F}'$. Clearly, $\mathbf{F}'-\mathbf{F}= q\mathbf{v}\mathrm{x}\mathbf{B} \neq 0$ for arbitrary **v**. This poses a problem for Classical Mechanics. It seems that the laws of electrodynamics don't reconcile with the Classical Laws.

- According to Maxwel's laws of electromagnetism, all electromagnetic waves (light) travel through vacuum at a speed given by $\frac{1}{\sqrt{\mu_0 \epsilon_0}}$ where μ_0 is the permeability of free space and ϵ_0 is the permittivity of free space. The speed through air can be taken as c.
  Consider three observers; one stationary on the ground, one in a car moving with constant speed v rightwards and another stationary on the ground. Let the car be between the two observers. Let the observer on the right switch on a torch leftwards. Then, the speed of the light emitted will be c according to the two

stationary observers. By the Galilean transformations, which govern this domain, the speed of light as measured by the observer in the car will be $v + c$.
But, Maxwell predicted that the speed of light is $\frac{1}{\sqrt{\mu_0 \epsilon_0}}$, not commenting on the reference frame. Does this mean that the fundamental constants(μ_0 and ϵ_0) vary according to the reference frame so as to keep the speed of light consistent with the Classical predictions?

The above two problems posed demonstrate the apparent disagreement between Mechanical and Electrodynamic laws under the Classical domain. The physicists at that time sought to resolve the problem by introducing a preferred medium called *Ether* or *Aether*. They argued that the speed of light as measured in a frame at rest wrt Ether will be given by $\frac{1}{\sqrt{\mu_0 \epsilon_0}}$ and it will be given by the Galilean transformations in all other frames. Thus, the value of the fundamental constants remained the same and the problem of the relative speed of light got solved too. Moreover, it comforted the Classical physicists of that time who thought that a medium was necessary for a wave to travel.
But that theory came crashing down when the most famous failed experiment in the history of humanity was conducted.

1.5 The Michelson-Morley Experiment

Michelson and Morley setup an ingenious interferometer to detect the motion of Earth through Ether. The setup is shown in Fig.1.3. There is a source S which emits light

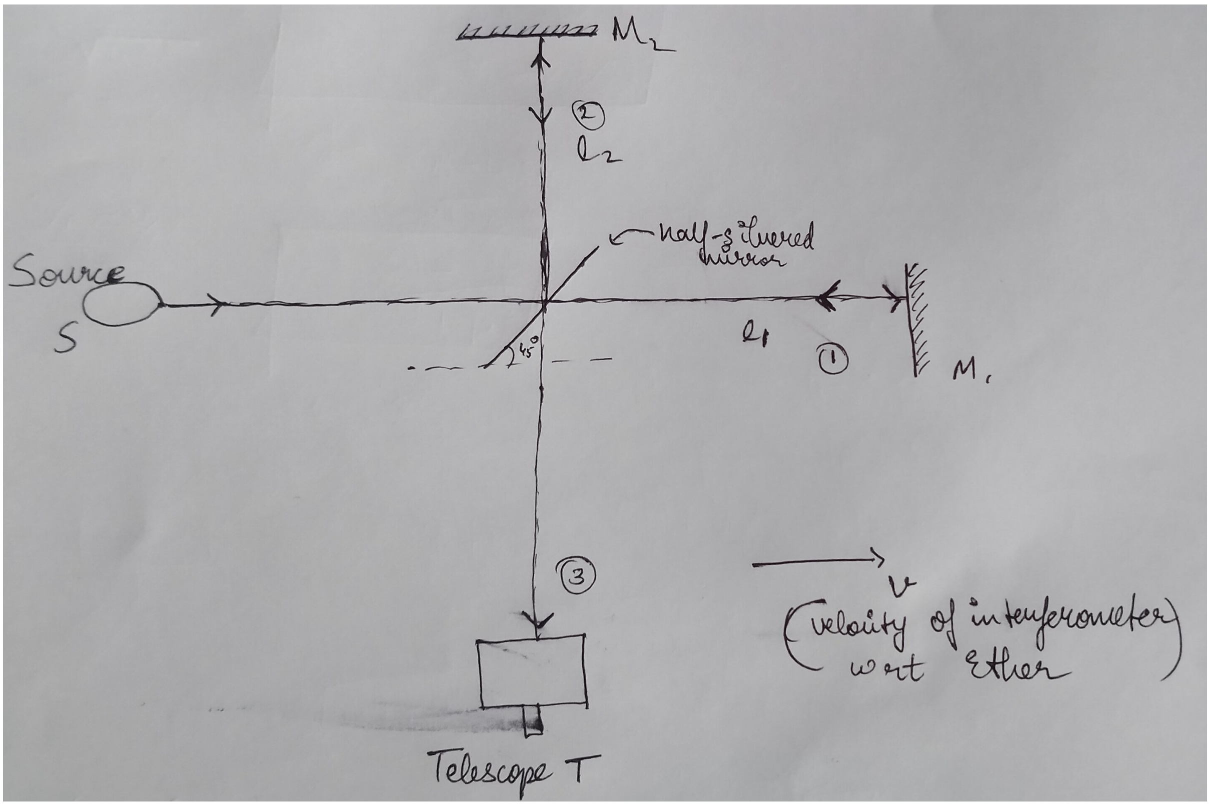


Figure 1.3: Experimental Setup used by Michelson and Morley

(monochromatic) which passes through the half-silvered mirror[5] inclined at an angle of

[5]It reflects half of the light incident and allows the other half to pass through

45° wrt the incident light. The reflected light hits M_2 and rebounds towards the telescope T passing through the half-silvered mirror. The other half hits M_1, rebounds and gets reflected towards T. The distance between M_1 and half-silvered mirror is l_1 (Arm 1)and that between M_2 and half-silvered mirror is l_2 (Arm 2). The whole setup is on Earth which we can approximate to be moving through Ether at a constant speed v rightwards at an instant.

The speed with which light goes to and fro in Arm 2 is $\sqrt{c^2+v^2}$ wrt Earth where $c = \frac{1}{\sqrt{\mu_0\epsilon_0}}$, which is evident on applying Galilean transformation between Ether and Earth. The speed of light when it travels towards M_1 is $c+v$ and when it gets reflected by M_2, the speed of light is $c-v$. The time taken by light to go to and fro in Arm 1 is

$$t_1 = \frac{l_1}{c-v} + \frac{l_1}{c+v} = \frac{2l_1c}{c^2-v^2} = \frac{2l_1}{c}[\frac{1}{1-\frac{v^2}{c^2}}]$$

and time taken by light to go to and fro in Arm 2 is

$$t_2 = \frac{2l_2}{\sqrt{c^2-v^2}} = \frac{2l_2}{c}[1-\frac{v^2}{c^2}]^{\frac{-1}{2}}$$

If $v << c$, by Binomial Expansion,

$$t_2 = \frac{2l_2}{c}[1+\frac{v^2}{2c^2}]$$

and

$$t_1 = \frac{2l_1}{c}[1+\frac{v^2}{c^2}]$$

Let $l_1 = l_2 = l$, the time difference for the two rays to reach T is

$$\Delta t = t_1 - t_2 = \frac{lv^2}{c^3}$$

Even if there were no difference in speed of light, this setup will produce a fringe pattern, when viewed through T, because different rays of the beam get reflected differently by the mirrors and they interfere at T. Due to the time delay, the fringe pattern shifts a bit. This absolute shift cannot be detected as we don't know the original fringe pattern. But we can compare the fringe pattern produced after rotating the apparatus by 90° (Arms 1 and 2 exchange places). The time difference after rotation will be

$$\Delta t' = -\Delta t = -\frac{lv^2}{c^3}$$

The fringe shift between the two arrangements will be given by

$$\Delta N = \frac{\Delta t - \Delta t'}{T} = \frac{\frac{2lv^2}{c^3}}{\frac{\lambda}{c}} = \frac{2lv^2}{\lambda c}$$

where T is the time period of light. If the Sun were assumed to be stationary in Ether frame, then $v = 10^{-4}c$ (orbital speed). The other experimental parameters were $l = 11m$ and $\lambda = 5.5\times10^{-7}m$. The expected fringe shift was $\Delta N = 0.4$. The fringes were expected to be shifted by about half a fringe width.

But no fringe shift was ever observed between the two arrangements! The experiment was repeated many times in different seasons at day and night, all in vain. This suggested that $v = 0$ at all times wrt Earth. Other theories such as Ether drag theory, Emission theory etc were proposed but they too bit the dust as they contradicted experimental findings.

1.6 Birth of Special Theory of Relativity

One way to explain the results of the experiments by Michelson and Morley was to say that the speed of light wrt all inertial observers is constant (c). That's what Einstein did. The whole Special Theory of Relativity is based on just two postulates:

1. The laws of Physics are the same for all Inertial Frames of Reference. No preferred Inertial Frame exists. (**The Principle of Relativity**)

2. The speed of light in free space has the same constant value (c) in all Inertial Frames. (**The Principle of Constancy of Speed of Light**)

The Special Theory of Relativity (SR) has successfully withstood the test of time till date. It successfully explained the results of those experiments which the other theories failed to and also predicted a few phenomena (E.g. Time Dilation, Transverse Doppler Effect). It plays an important role in the working of Global Positioning System (GPS) and in reconciling the ideas of Mechanics and Electrodynamics. In the next chapter, we will look at the relativistic equivalent of Galilan Transformations; the *Lorentz Transformations.*

Chapter 2

Basics of Special Relativity

In this chapter we explore the basic principles underlying Special Relativity and its consequences which are unimaginable for the Clasical mind.

2.1 Simultaneity

We saw in Section 1.3 that the time intervals and length intervals (for 'simultaneous' events) between two events are the same for all inertial observers, according to Classical formalisms. But, the dawn of relativity shattered the concepts of a 'Universal' clock. The fact that the simultaneity of two events is subject to the reference frame can be shown without any reference to mathematical formulae. Let's consider a thought experiment performed while keeping the postulates of relativity in mind.

Consider Fig.2.1 wherein, there are two observers; one on a stationary (wrt Earth) platform (O) and the other one in a train cabin (O') moving with a constant speed of v towards right wrt O. The observer O' is stationary wrt the train and is at the dead centre of the train cabin. At the moment when the two observers are momentarily facing each other, lightning strikes at the two ends of the cabin simultaneously wrt O (Fig.2.1(a)). Let's look at the situation from the platform's perspective. The red light ray moves, from

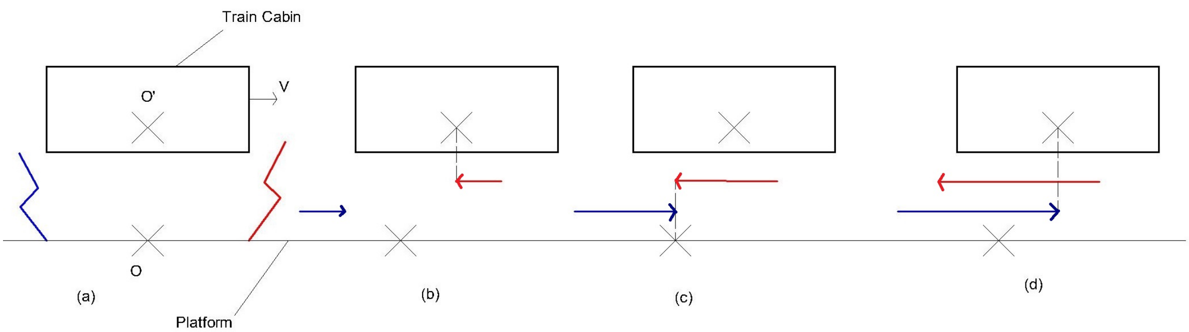


Figure 2.1: Relativity of Simultaneity

point where the lightning struck, leftwards and O' moves rightwards in the same time. At a point of time, the red ray reaches O' but the blue ray, which has moved the same distance (since c is a constant), is still far behind (2.1(b)). Later, the two rays meet where O is standing and hence, O percieves that the two lightnings struck at the same time (Fig.2.1(c)). Then, the blue light ray meets O' at a later point of time (Fig.2.1(d)).

Hence, the lightnings struck simultaneously wrt O but O' percieves that the red lightning struck earlier than the blue one.
Now let's jump into the train frame. O moves with a speed v leftwards wrt O'. Since O reports the two events as simultaneous, it means that the rays of light must reach O at the same time. But O is moving leftwards, towards the point where blue lightning strikes. This means that, according to O', the red lightning must strike earlier than the blue one so that the rays reach O at the same time. If this situation were to happen in real life, the time difference observed by O' would be very low as the train's speed is not much as compared to c.

We should note the crucial role played by the fact that the speed of light wrt both observers is the same. The above mentioned thought experiment demonstrates that, under the relativistic domain, simultaneity is a matter subject to the observer. No observer is wrong; both are correct. If c were infinite, there would be no such disagreement on simultaneity of events.

This means that two seemingly caliberated clocks, in two different inertial frames, will give different readings for the same occurence as the simultaneity of the arrangement of the clock hands (or digits) in a particular manner and the occurence is debatable. Hence, the concept of absolute universal time is banished. This calls for a replacement of the Galilean Transformations.

2.2 Lorentz Transformations

The way the spacetime coordinates are transformed under Special Relativity is dictated by the Lorentz Transformations. Here, we won't implicitly assume the universality of time for we have explored the relativity of simultaneity in the previous section. We can derive the Lorentz Transformations from scratch using the two basic postultes of Special Relativity.
We start by assuming homogenous space and time (measurements do not depend on when and where they were made). This necessitates a linear relation between the coordinates of the two frames S and S' (Fig.1.2) else, the condition of homogeneity of space and time will be violated. Consider that when O and O' coincide, $t = t' = 0$. Let the relations be

$$x' = a_{11}x + a_{12}y + a_{13}z + a_{14}t$$

$$y' = a_{21}x + a_{22}y + a_{23}z + a_{24}t$$

$$z' = a_{31}x + a_{32}y + a_{33}z + a_{34}t$$

$$t' = a_{41}x + a_{42}y + a_{43}z + a_{44}t$$

If the dependence were $x' = a_{11}x^2$, then the length measurements $(x'_1 - x'_2)$ made in S' would depend on the locations in S frame. Hence, linear dependence is a reasonable assumption. Now let's eliminate some coefficients. If an object is on the x-axis in S frame, then clearly it should be on the x′-axis of S' frame ($y = z = 0 \Rightarrow y' = z' = 0$). This requires $a_{21} = a_{24} = a_{31} = a_{34} = 0$. If an object is in the x-y plane in S then the object should be in the x′-y′ plane in S' ($z = 0 \Rightarrow z' = 0$). Similarly, $y = 0 \Rightarrow y' = 0$. This requires $a_{23} = a_{32} = 0$. The reduced equations are

$$x' = a_{11}x + a_{12}y + a_{13}z + a_{14}t$$

$$y' = a_{22}y$$

$$z' = a_{33}z$$

$$t' = a_{41}x + a_{42}y + a_{43}z + a_{44}t$$

Now consider a rod of unit length (wrt S) lying along the y-axis in the S frame. Its length, as measured by S', will be a_{22} units. Consider another rod of unit length (wrt S') lying along the y'-axis in the tS' frame. Its length, as measured by S' will be $\frac{1}{a_{22}}$ units. According to the first postulate, all physical laws should be equivalent for all inertial observers. This means that $a_{22} = \frac{1}{a_{22}}$, else the frames will be physically unequivalent. Hence, $a_{22} = a_{33} = 1$ (proceeding similarly for a_{33}). Consider an object at $x' = 0$ (y$'$-z$'$ plane) to be at rest wrt S'. According to S, $x = vt$ regardless of y and z.

$$\Rightarrow x' = a_{11}(x - vt)$$

$$\Rightarrow a_{14} = -a_{11}v$$

Furthermore, t' should not depend on y and z as points symmetricaly opposite to x-axis would give different time readings as seen by S', contradicting the isotropy of space. Therefore, $a_{42} = a_{43} = 0$. So far,

$$x' = a_{11}x - a_{11}vt$$

$$y' = a_{22}y$$

$$z' = a_{33}z$$

$$t' = a_{41}x + a_{44}t$$

Consider that when O and O' coincide, a spherical electromagnetic wave originates from the origins. According to the second postulate, it spreads with the speed c in all the directions in both frames. Its equation in the two frames will be given as

$$x^2 + y^2 + z^2 = (ct)^2 \tag{2.1}$$

$$x'^2 + y'^2 + z'^2 = (ct')^2 \tag{2.2}$$

Substitute (t', x', y', z') in terms of (t, x, y, z) in Eq.2.2. We get

$$x^2(a_{11}^2 - c^2a_{41}^2) + y^2 + z^2 = t^2(a_{44}^2c^2 - a_{11}^2v^2) + 2(va_{11}^2 + c^2a_{41}a_{44})xt \tag{2.3}$$

Eq.2.1 and Eq.2.3 represent the same thing and hence, should be the same. Comparing the coefficients,

$$a_{11}^2 - c^2a_{41}^2 = 1$$

$$a_{44}^2c^2 - a_{11}^2v^2 = c^2$$

$$va_{11}^2 + c^2a_{41}a_{44} = 0$$

Upon solving these equations, we get

$$a_{11} = a_{44} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$a_{41} = \frac{v}{c^2}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

By convention, $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and $\beta = \frac{v}{c}$ ($\gamma \geq 1$ and $\beta < 1$). Finally, we can write the Lorentz Transformations as

$$x' = \gamma(x - vt)$$
$$y' = y$$
$$z' = z$$
$$t' = \gamma(t - \beta\frac{x}{c})$$

and the Inverse Transformations as

$$x = \gamma(x' + vt)$$
$$y = y'$$
$$z = z'$$
$$t = \gamma(t' + \beta\frac{x'}{c})$$

While writing the inverse transformations, we have just interchanged the primed and unprimed coordinates and replaced v with $-v$ which should do the trick, keeping in mind the first postulate. These transformations reduce to the Galilean Transformations for $v << c$ because in that case, $\beta \approx 0 \Rightarrow \gamma \approx 1$. This explains why Galilean Transformations seem to work for speeds to which we are usually accustomed. We must keep in mind the assumptions we made while deriving these equations, namely:

- x-axis and x$'$-axis are coincident while y-axis$\|$y$'$-axis and z-axis$\|$z$'$-axis
- $t = t' = 0$ when O and O' coincide
- S and S' experience relative motion only in x and x$'$ axes

2.3 Consequences of Lorentz Transformations

2.3.1 Length Contraction

Suppose a rod lies on the x-axis, at rest wrt S frame, with one end at $x_1 = 0$ and the other at $x_2 = l$. Then, the length of the rod as measured in S will be l. Note that,in general, the measurement is made by noting the coordinates of the rod's ends at the same time (say t) and then finding the distance between them. Let's jump into S' frame. The coordinates of the ends of the rods in the two frames are related by $x_1 = \gamma(x'_1 - vt'_1)$ and $x_2 = \gamma(x'_2 - vt'_2)$. This implies

$$l = x_2 - x_1 = \gamma(x'_2 - x'_1 - v(t'_2 - t'_1)) \tag{2.4}$$

If $t'_2 = t'_1$, $x'_2 - x'_1$ will be the length of the rod as measured by S' (say l'). Then, Eq.2.4 suggests

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

That is, the length of the rod, as measured in a frame moving wrt the rod, is shortened by a factor of γ. This is a consequence of the relativity of simultaneity. The observers do not agree on the simultaneity of the coordinate measurement of the rod-ends. We must observe that the length of the rod is maximum in the frame which is stationary wrt the rod. Thus, we define the *proper length* as the length of a body measured in a frame in which the body is at rest and that frame is called *proper frame*.

2.3.2 Time Dilation

Suppose there is a clock at O', at rest wrt S'. This clock moves rightwards in S frame with a speed v. Let this clock tick t' units in S' frame ($t'_2 = t'_1 + t'$ and $x'_1 = x'_2 = 0$). Let's look at the situation from the S frame. The equations connecting the time coordinates are

$$t_1 = \gamma(t'_1 - \beta\frac{x'_1}{c}) = \gamma t'_1$$

$$t_2 = \gamma(t'_2 - \beta\frac{x'_2}{c}) = \gamma t'_2$$

$$\Rightarrow t = t_2 - t_1 = \gamma(t'_2 - t'_1) = \gamma t'$$

The above equation shows us, mathematically, that the time interval measured in a general frame is γ times larger than the time interval measured in the clock's frame. In other words, *moving clocks run slow*. The time taken by the moving clock to tick off a second is larger than a second in S frame. This time dilation is reciprocal i.e., both observers say that the other's clock to run slower.

Just like proper length, *proper time* between two events is defined as the time ticked off by a clock at rest wrt the body. It is usually denoted by the Greek letter τ. Note that the proper time is measured by a single clock but time interval in any other frame is measured by two different *synchronized* clocks.

2.3.3 Synchronization of Clocks

Let's discuss about the ways in which time and space coordinates are measured in a frame of reference. The space coordinates' measurement are easy to understand. Three standard axes (x, y, z) are chosen and lengths are marked off on the axes. Now any point in space can be given unique coordinates. For measuring the time coordinates, we can consider that every point in space has its own clock so that if any event occurs, a clock is right at the position to measure the time coordinate. All the clocks are synchronized. If one were to stop all clocks at once (not really possible), all the clocks would be stuck at the same value. This is true for all inertial frames.

Consider the standard inertial frames of reference, S ad S'. According to S, all clocks of S' appear to run at the same slower rate but the reading of each clock of S' differs from the other clocks'. The same is true when S' observes S's clocks. In other words, each observer disagrees on the synchronization of the other's clocks. We must note that this is not due to the finite time taken by light to reach the obserever's eyes but is an inherent property of the Lorentz Transformation $t = \gamma(t' + \beta\frac{x'}{c})$. For a constant t, t' decreases as x' increases and a similar relation is shown by $t' = \gamma(t - \beta\frac{x}{c})$.

This apparent asynchronization of other's clocks allows both observers to rightly say that the other's clock rates are slower without any contradiction or paradox. We must note that the observer who measures a moving clock, does so using two 'synchronized' clocks and compares it with the reading of a single moving clock. But the observer in the moving clock's frame would say that the two clocks are not synchronised and hence do not show the correct time difference. Both observers are correct in their own sense.

An observer should not be seen as a 'watcher' but as a 'measurer'. Whenever someone says that an observer 'sees', it is meant that the observer 'measures'.

2.4 Relativistic Velocity Addition

It is obvious that the Classical relative velocity addition formula is not followed in the real world as objects would exceed the speed of light in some frame if they followed it. Let's derive the relativistic velocity addition formula. Consider Fig.2.2. There are the

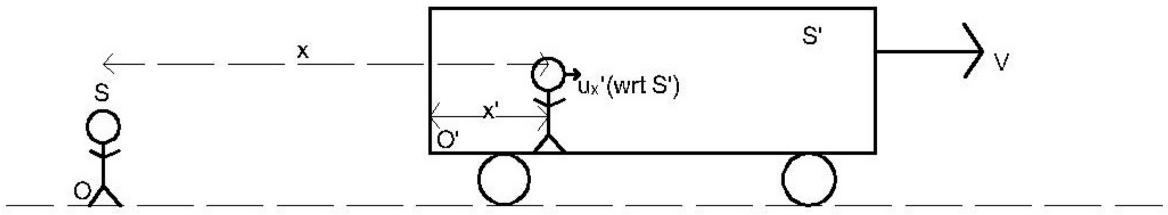


Figure 2.2: Relativistic Velocity Addition

usual frames of reference, S and S'. S is a stationary observer on the ground and S' is a vehicle moving with speed v rightwards. The origin O is where S stands and O' is the bottome left corner of the vehicle as indicated. Within S', there is a person moving with a velocity u'_x rightwards, wrt S'. We wish to find the velocity of the person wrt S. Let's start with the relation between the spacetime coordinates

$$x = \gamma(x' + vt')$$

$$t = \gamma(t' + \beta\frac{x'}{c})$$

Taking differentials on both sides,

$$dx = \gamma(dx' + vdt') \tag{2.5}$$

$$dt = \gamma(dt' + \beta\frac{dx'}{c}) \tag{2.6}$$

Dividing Eq.2.5 by Eq.2.6, we get

$$\frac{dx}{dt} = \frac{dx' + vdt'}{dt' + \beta\frac{dx'}{c}}$$

Dividing the numerator and denominator of RHS by dt',

$$u_x = \frac{dx}{dt} = \frac{\frac{dx'}{dt'} + v}{1 + \beta\frac{dx'}{dt'}c} = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}} \tag{2.7}$$

Hence, the velocity of the person wrt S is given by $\frac{u'_x + v}{1 + \frac{u'_x v}{c^2}}$. Note that if $u_x << c$ and $v << c$, this formula reduces to $u'_x + v$ which is the classical velocity addition formula. If $u'_x = c$, $u_x = \frac{c+v}{1+\frac{v}{c}} = c$ i.e., speed of light is invariant. And $u'_x < c$, $v < c \Rightarrow u_x < c$.

Let's discuss about the transverse velocities. Do they remain the same? Consider an object moving in the y' direction in S' with a speed u'_y. The relation between y and y' will be

$$y = y'$$

Differentiating both sides,

$$dy = dy' \tag{2.8}$$

Dividing Eq.2.8 by Eq.2.6, we get

$$u_y = \frac{dy}{dt} = \frac{dy'}{\gamma(dt' + \beta\frac{dx'}{c})}$$

Dividing the numerator and deominator of the RHS by dt', we get

$$u_y = \frac{u'_y}{\gamma(1 + \frac{u'_x v}{c^2})}$$

Proceeding on similar lines, we get

$$u_z = \frac{u'_z}{\gamma(1 + \frac{u'_x v}{c^2})}$$

An interesting thing to note is that the transverse velocities depend on the velocity in the x′ direction too. The inverse relations are easily obtained by switching the primed and unprimed quantities and replacing v with $-v$.

2.5 Relative Acceleration

Suppose the body about which we have been discussing is accelerating wrt S' with an acceleration a'_x. What will be the acceleration of the body wrt S? Let's find out. Differentiating Eq.2.7, we get

$$du_x = \frac{du'_x}{\gamma^2(1 + \frac{u'_x v}{c^2})^2} \tag{2.9}$$

Dividing Eq.2.9 with Unlike Newtonian Mechanics, acceleration is not invariant under relativistic considerations, even in inertial frames! As $\frac{u'_x}{c}; \frac{v}{c} \to 0$, $\gamma \to 1$ and $a_x \to a'_x, a_y \to a'_y$ and $a_z \to a'_z$. This shows that for low velocities Relativistic Ideas reduce to the Classical Ideas. This explains why the Classical Ideas seem appropriate for are everyday exploits.

2.6 Aberration of Light

With the tools which have been developed till here, a few phenomena can be explained. We will proceed with the standard set of frames S and S' as shown in Fig. We wish to find the angle which the light ray makes wrt the x-axis in the S frame. By doing so, we will uncover some interesting stuff. Let's split the light's velocity into its components. The x′ component is $u'_x = c \cdot cos\theta'$ whilst the y′ component is $u'_y = c \cdot sin\theta'$. The corresponding

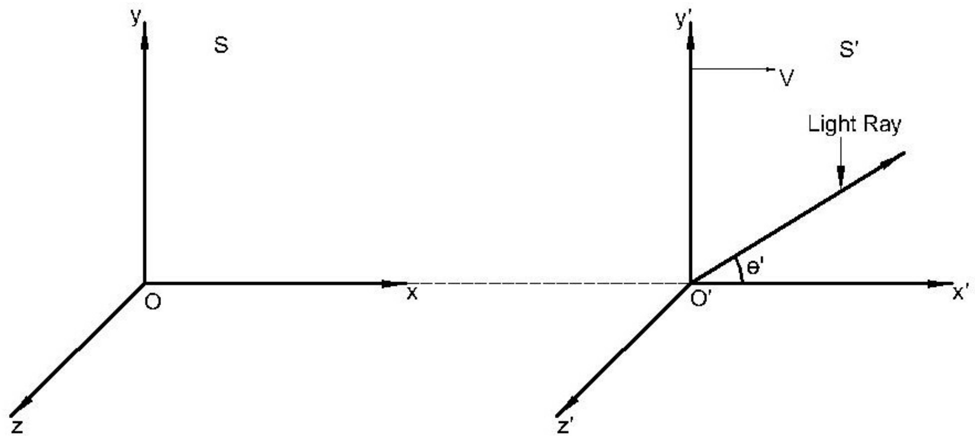


Figure 2.3: Aberration of Light

x and y components in the S frame will be given by the relativistic velocity addition formula. Applying those formulae, we get

$$u_x = c\frac{v + c \cdot cos\theta'}{c + v \cdot cos\theta'}$$

$$u_y = \frac{c^2 \cdot sin\theta'}{\gamma(c + v \cdot cos\theta')}$$

Note that we could have found out u_y using the Principle of Constancy of the Speed of Light ($u_x^2 + u_y^2 = c^2$). Let the light ray in the S frame make an angle θ wrt x-axis. Since the speed of light is c, $u_x = c \cdot cos\theta$ and $u_y = c \cdot sin\theta$. Thisi means $tan\theta = \frac{u_y}{u_x}$ i.e.,

$$tan\theta = \frac{c \cdot sin\theta'}{\gamma(v + c \cdot cos\theta')}$$

The above equation reveals that the angle made by the light ray with x-axis is not the same as percieved by S' (unless $\theta' = 0$). This phenomenon is called *aberration*. It is often observed while viewing stars and is called *stellar aberration*. The angle which the light beam from the star makes wrt the normal to the Earth's surface as seen in the Sun's frame is not the same as perceved by us when observing the star. This was earlier explained using Classical Ideas wherein the factor of γ was missing from the denominator and hence worked for speeds much less than that of light. The formula we derived perfectly explains the phenomenon of aberration.

Later on, we will derive Doppler Effect using the concept of four-momentum without the hassle of writing complex wave equations, in a very elegant manner.

Chapter 3

Relativistic Dynamics

Till now we dealt with basic Lorentz Transformations, relative velocities and related phenomena. In this chapter, we will discuss about relativistic acceleration, forces and momenta. We shall uncover concepts like rest mass and relativistic mass. We shall do all this staying in inertial frames of reference.

3.1 Conservation Laws

After the introduction of relativistic ideas, many Classical Ideas should change. The concept of simultaneous action and reaction must be abandoned as simultaneity is a relative concept except when the events occur at the same point in space. Action at a distance is not acceptable as no information can travel faster than light. Fields are introduced.

In Classical Mechanics, conservation laws (energy, mass, momentum) play a very important role in simplifying quite a few problems and also make some sense to the rational mind when reasoned with our everyday experience. But do those laws hold in relativistic realm? Not if we define momentum as $\mathbf{p}= m_0\mathbf{u}$. We must redefine momentum and a few other things if we want to continue dealing with such conservation laws. While doing so, we must do away with one of the very fundamental ideas; the invariance of mass.
First, we define the *rest mass* of a body as the mass of a body as measured in a frame in which the body is at rest. It's usually denoted by m_0 and is an invariant quantity. Then, we define the mass of a body as measured in a frame in which it is moving to be

$$m = \gamma m_0$$

where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and v is the speed of the body wrt the frame. We define the momentum of a body as

$$\mathbf{p} = m\mathbf{v} = \gamma m_0 \mathbf{v} \tag{3.1}$$

Hence, by defining momentum as $\gamma m_0\mathbf{v}$, we have allowed the law of conservation of momentum to exist in the relativistic realm. But there is an important point to note. We can't replace m in every classical formula by γm_0 and claim that we have developed a relativistic version of that quantity. Moreover, we could have defined momentum as $\mathbf{p} = m_0\frac{d\mathbf{r}}{d\tau}$ where τ is the proper time. This is just a different way of writing Eq.3.1.

3.2 Force and Kinetic Energy

Force on an object is given by

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}\left(\frac{m_0\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}\right)$$

The law of conservation of momentum is immediate from the above equation. Now, the Kinetic Energy of an object can be defined as the work done upon an object to raise its velocity from rest. Hence, Kinetic Energy of a body is given by

$$K = \int_{u=0}^{u=v} \mathbf{F}\cdot d\mathbf{l}$$

For simplification, let the motion be along the x-axis only. Then,

$$K = \int_{u=0}^{u=v} \mathbf{F}dx = \int_{u=0}^{u=v} \frac{d}{dt}\left(\frac{m_0 u}{\sqrt{1-\frac{u^2}{c^2}}}\right) dx$$

$$\Rightarrow K = \int_{u=0}^{u=v} d\left(\frac{m_0 u}{\sqrt{1-\frac{u^2}{c^2}}}\right)\frac{dx}{dt} = \int_{u=0}^{u=v} d\left(mu\right)u$$

$$\Rightarrow K = \int_{u=0}^{u=v} u^2 dm + mudu$$

Now consider $m = \frac{m_0}{\sqrt{1-\frac{u^2}{c^2}}}$. In other words,

$$m_0^2c^2 + m^2u^2 = m^2c^2$$

Differentiating throughout, we get

$$u^2dm + mudu = c^2dm$$

Substituting this in the integral, we get

$$K = \int_{u=0}^{u=v} c^2dm = \int_{m=m_0}^{m=m_v} c^2dm$$

$$\Rightarrow K = m_vc^2 - m_0c^2$$

where m_v is just γm_0 and can be replaced by m. For any body with a nonzero rest mass to attain the speed of light, $K \to \infty$ for $m \to \infty$ as $v \to c$. Hence, speed of light is unattainable for bodies with non zero rest mass.
The quantity mc^2 is called the *total energy* (E) of the body and m_0c^2 is called the *rest energy* of the particle.

$$E = K + m_0c^2$$

Consider $p = \gamma m_0 v$. Squaring both sides and cross-multiplying the denominator, we get

$$p^2c^2 - p^2v^2 = m_0^2c^2v^2$$

$$\Rightarrow p^2c^2 = v^2\left(p^2 + m_0^2c^2\right)$$

Substituting $p = mv$ in the RHS, we get

$$m^2c^2 = p^2 + m_0^2c^2$$

Multiplying the whole equation by c^2, we get

$$m^2c^4 = p^2c^2 + m_0^2c^4$$

The RHS is nothing but E^2. Taking the square root of the equation, we get

$$E = c\sqrt{p^2 + m_0^2c^2}$$

Hence, we got the following relations

$$E = c\sqrt{p^2 + m_0^2c^2} = K + m_0c^2 = mc^2$$

3.3 Acceleration and its direction

We know that $\mathbf{F} = m\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{dm}{dt}$ and $m = \frac{E}{c^2}$. This implies

$$\frac{dm}{dt} = \frac{1}{c^2}\frac{dE}{dt} = \frac{1}{c^2}\frac{d}{dt}\left(K + m_0c^2\right) = \frac{1}{c^2}\frac{dK}{dt}$$

But $\frac{dK}{dt} = \mathbf{F}\cdot\frac{d\mathbf{l}}{dt} = \mathbf{F}\cdot\mathbf{v}$, which means

$$\frac{dm}{dt} = \frac{1}{c^2}\mathbf{F}\cdot\mathbf{v}$$

Substituting this in the original equation, we get

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{\mathbf{F}\cdot\mathbf{v}}{c^2}$$

$$\Rightarrow \mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{\mathbf{F}}{m} - \frac{\mathbf{v}}{mc^2}\left(\mathbf{F}\cdot\mathbf{v}\right)$$

A very interesting thing is observed from the above eqaution; acceleraion, in general, is *NOT* in the direction of force. Acceleration is in the direction of force only if

- either $\mathbf{v} \parallel \mathbf{F}$
- or $\mathbf{v} \perp \mathbf{F}$ (so that $\mathbf{F}\cdot\mathbf{v} = 0$)

3.4 Rest Mass and Energy

Consider a perfectly inelastic collison of two bodies (say A and B), both of rest mass m_0, with speed v and moving in oppposite directions. After the collision, the two bodies come to a complete halt and combine into a single body of rest mass M_0. A classical mind would expect that $M_0 = 2m_0$. To find if this is true relativistically, consider a frame which is moving with the same speed v rightwards. In this frame, one of the bodies (say A) is staionary initially, while the other one moves with a speed $v'_B = \frac{2v}{1+\frac{v^2}{c^2}}$ leftwards (calculated using the relativistic velocity addition formula). After the collision, the combined body moves with a speed v leftwards. Let's apply conservation of momentum in the new frame. Initial momentum is

$$p_i = -m_0 \frac{1}{\sqrt{1-\frac{v_B'^2}{c^2}}} v'_B = \frac{-2m_0 v}{1-\frac{v^2}{c^2}}$$

and final momentum is

$$p_f = -M_0 \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} v$$

Equating the initial and final momenta, we get

$$\frac{-2m_0 v}{1-\frac{v^2}{c^2}} = -M_0 \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} v$$

$$\Rightarrow M_0 = \frac{2m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

Surprisingly, $M_0 \neq 2m_0$. This is valid in all frames as rest mass is an invariant quantity. Note that $M_0 > 2m_0$. How should we account for this 'extra' mass? Consider the lost Kinetic Energy in the frame where the combined mass is at rest, to make calculations easy. The kinetic energy lost is equal to the intial kinetic energy which is

$$\Delta K = 2\left(mc^2 - m_0c^2\right) = 2m_0c^2 \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)$$

and the extra mass is

$$\Delta m = 2m_0 \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)$$

Notice that $\Delta mc^2 = \Delta K$. In inelastic collisions, K is converted to internal energies (heat, stress, potential, etc). This internal energy results in the increased rest mass of the body. Consider the intial total energy

$$E_i = 2m_c^2 = \frac{2m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}}$$

and the final total energy

$$E_f = M_0c^2 = \frac{2m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}}$$

Albeit K is not conserved, total energy (K+rest-mass energy) is conserved. Since $E = mc^2$, Conservation of Total Energy $\equiv$ Conservation of Mass. K can be regarded as the external energy and m_0c^2 as the internal energy.
The mass which we associate with various internal energies ($\frac{E}{c^2}$) has all the properties like inertia, weight, contribution to centre of mass, etc.

3.5 Momentum and Force Transformations

Just like the space-time coordinates and velocities transform according to the Lorentz Transformations, momentum and force too obey it. Here we consider the standard frames S and S'. The transformations for momentum are

$$p_x = \gamma\left(p'_x + \frac{E'v}{c^2}\right)$$

$$p_y = p'_y$$

$$p_z = p'_z$$

$$E = \gamma\left(E' + vp'_x\right)$$

Note that energy plays the same role for momentum as time plays for space coordinates. Energy is regarded as the time component of the four-momentum, which we shall discuss later. The transformation laws for force are

$$F_x = F'_x + \frac{u'_y v}{c^2 + u'_y v}F'_y + \frac{u'_x v}{c^2 + u'_x v}F'_z$$

$$F_y = \frac{F'_y}{\gamma\left(1 + \frac{u'_x v}{c^2}\right)}$$

$$F_z = \frac{F'_z}{\gamma\left(1 + \frac{u'_x v}{c^2}\right)}$$

All the symbols have the usual meaning. The inverse transformations can be obtained by interchanging the primed and unprimed quantities and replacing v with $-v$.

Chapter 4

Relativistic Electromagnetism

In this chapter we will discuss about Electric and Magnetic fields and their transformation properties under the relativistic regime.

4.1 Current Carrying Wires

In a current carrying wire, there are no net charges. The electrons move randomly within the wire, with a net movement opposite to the field direction (drift velocity, say v). It can be thought of as electrons moving with a uniform spacing between them amongst the positive charged ions, fixed in a lattice, separated by the same uniform spacing (Fig.4.1(a)). At any given time, if we were to inspect a particular length of the wire, we would find equal number of positive and negative charges and hence, find the wire to be neutral. This implies that there will be only magnetic field around the wire but no electric field. The above discussion is valid in the wire's frame.

Let's observe it from a frame in which the moving electrons are at rest (Fig.4.1(b)). We

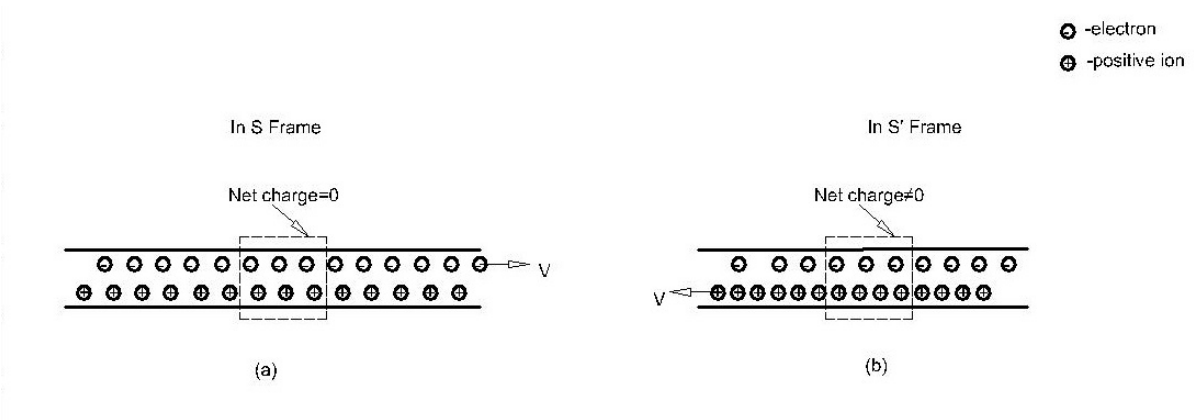


Figure 4.1: Charge of a Current Carrying Wire

assume *charge invariance* i.e., charge on a body is independent of the frame from which we observe it. Since, we changed frames, the electrons appear at rest with a uniform spacing between them greater than the value which was previously observed (as we are observing the proper length now). But, the positive ions seem to move in the opposite direction and hence, by length contraction, the separation between them reduces from the previous value. This means, if we choose a particular length of wire, we may find a net positive charge as the electrons are more widely separated than the positive charges. This implies that we would observe electric field in addition to magnetic field. Hence, the electromagnetic fields are frame dependent and interconvertible.

An interesting thing to discuss is that we assumed charge invariance and yet, we observed that the wire was neutral earlier and has a net positive charge after changing frames. How do we explain this contradiction? It seems perplexing at first but has a very simple explanation. For current to flow in a wire, it should form a closed loop. We considered only one part of the loop. In the other part of the loop, the electrons will be moving in the opposite direction and hence, the length contraction for them would increase further, leading to a net negative charge in that part of the wire. Thus, the wire, as a whole, remains neutral in all frames but parts of wire may develop charges in some frames.

4.2 Transformation Laws for Electromagnetic Fields

From the above example, it is clear that electric and magnetic fields transfrom when frames are changed. Let's find the transformation laws. Consider the standard frames S and S' and a charge q instantaneously at rest wrt S'. Let the electric and magnetic fields in S be $\mathbf{E}$ and $\mathbf{B}$ and in S' be $\mathbf{E}'$ and $\mathbf{B}'$ respectively. For the particle we have considered, the force transformation equations are given by

$$F_x = F'_x \tag{4.1}$$

$$F_y = \frac{F'_y}{\gamma} \tag{4.2}$$

$$F_z = \frac{F'_z}{\gamma} \tag{4.3}$$

In the frame S, the instantaneuos velocity of the charge is $\mathbf{v}$ in x-direction. Then

$$F_x = q(E_x + (\mathbf{v} \times \mathbf{B})_x)$$

Since $v_y = v_z = 0$,

$$F_x = qE_x$$

In the S' frame, $v' = 0$

$$\Rightarrow F'_x = qE'_x$$

Proceeding in a similar manner, we get

$$F_y = q(E_y - v_x B_z)$$

$$F'_y = qE'_y$$

$$F_z = q(E_z + v_x B_y)$$

$$F'_z = qE'_z$$

Substituting the repective values in Eqs.4.1, 4.2 and 4.3, we get

$$E'_x = E_x$$

$$E'_y = \gamma(E_y - v_x B_z)$$
$$E'_z = \gamma(E_z + v_x B_y)$$

Thus, we have obtained the transformation laws for electric fields.
For getting the transformations laws for magnetic fields, let the charge move in the y′ direction in S' with a speed u'_y. By the relativistic velocity addition formula, we get the speed of the charge in y-direction in S is given by $u_y = \frac{u'_y}{\gamma}$. Proceeding similarly as before, we obtain

$$B'_x = B_x$$
$$B'_y = \gamma(B_y + \frac{v}{c^2}E_z)$$
$$B'_z = \gamma(B_z - \frac{v}{c^2}E_y)$$

For getting the last direction, we will have to let the charge move in the z′-direction. Note that we will have to use the previously derived results for electric fields for obtain the above results for magnetic fields. The inverse transformations can be obtained by the usual procedure.

Since, the direction of axes is arbitrary, we can rewrite the transformations in another form

$$E'_\parallel = E_\parallel$$
$$E'_\perp = \gamma(E_\perp + (\mathbf{v} \times \mathbf{E})_\perp)$$
$$B'_\parallel = B_\parallel$$
$$B'_\perp = \gamma(B_\perp - \frac{(\mathbf{v} \times \mathbf{E})_\perp}{c^2})$$

where $\parallel$ denotes the component of field in the direction of relative velocity of the frames while $\perp$ denotes the component of field in the perpendicular direction to the relative velocity. The trasformation laws for **E** and **B** are dependent on both **E** and **B**, emphasizing the fact that they are not separate entities.

4.3 Fields due to a moving charge

Consider the standard frames S and S' and charge (q) stationary in the S' frame. The charge moves rightward in the S frames and hence would produce both electric and magnetic fields whilst in the S' frame, it produces just an electric field. Hence, using the above transformations, we can derive the values for electric and magnetic fields in S frame by calculating just the electric field in the S' frame.
In the S' frame, the electric field is given by $\mathbf{E}' = \frac{1}{4\pi\epsilon_0}\frac{q}{r'^3}\mathbf{r}'$ where $r' = \sqrt{x'^2 + y'^2 + z'^2}$ and ofcourse, $\mathbf{B}' = 0$. Let's break **E'** into its components.

$$E'_x = \frac{1}{4\pi\epsilon_0}\frac{qx'}{r'^3}$$

$$E'_y = \frac{1}{4\pi\epsilon_0}\frac{qy'}{r'^3}$$

$$E'_z = \frac{1}{4\pi\epsilon_0}\frac{qz'}{r'^3}$$

Using the transformation equations for the fields and substituting the S coordinates in place of S' coordinates, we get

$$E_x = \frac{1}{4\pi\epsilon_0}\frac{\gamma q\left(x - vt\right)}{\left(\gamma^2\left(x-ut\right)^2 + y^2 + z^2\right)^{\frac{3}{2}}}$$

$$E_y = \frac{1}{4\pi\epsilon_0}\frac{\gamma q y}{\left(\gamma^2\left(x-ut\right)^2 + y^2 + z^2\right)^{\frac{3}{2}}}$$

$$E_z = \frac{1}{4\pi\epsilon_0}\frac{\gamma q z}{\left(\gamma^2\left(x-ut\right)^2 + y^2 + z^2\right)^{\frac{3}{2}}}$$

$$B_x = 0$$

$$B_y = \frac{-u}{c^2}E_z$$

$$B_z = \frac{u}{c^2}E_y$$

An interestig thing to note is that though $\mathbf{E}'$ is spherically symmetric in the S' frame, $\mathbf{E}$ is not so in the S frame. This is evident if we write $\mathbf{E}$ as

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\frac{\gamma q\mathbf{r}}{\left(\gamma^2\left(x-ut\right)^2 + y^2 + z^2\right)^{\frac{3}{2}}} - \frac{1}{4\pi\epsilon_0}\frac{\gamma qut}{\left(\gamma^2\left(x-ut\right)^2 + y^2 + z^2\right)^{\frac{3}{2}}}\hat{\mathbf{x}}$$

Clearly, the strength of the field is not the same at say, $(0, 0, a)$ and $(a, 0, 0)$, for arbitrary $a \in \mathbb{R}$.

Chapter 5

Spacetime Diagrams

In this chapter we will geometrically visualize the concepts of SR using the Spacetime Diagrams, also known as the Minkowski Diagrams. The invariance of speed of light, relativity of simultaneity, time dilation and length contraction will be easier to observe using such diagrams.

5.1 Natural Units

Before discussing about the spacetime diagrams, we must discuss about a new system of units, a more 'natural' one. If we use the SI units, the value of the speed of light is 299792458 ms^{-1}, a ridiculously high value! Also, as we will see further, since we are considering time and space in a single coordinate system, it would be better if they had the same units.

In the new system, one unit of time is defined as the time taken by light to travel $1m$ and is expressed in *metres*. If a light ray travels a distance of say $14.32m$, then the time taken by the light ray would be $14.32m$ in this new system. By this definition, the speed of light would be 1 in the new system. The speed of massive bodies will always be less than one. Let t denote time in natural system and t' denote time in SI. Then the relationship between the two is $t = 299792458t'$. Measurement of length is unaffected.

This has several consequences. Speed is a unitless quantity in this system as time and space have the same units of measurement. The Lorentz Transformations become symmetric, as $c = 1$ in this system, as shown below

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

$$y' = y$$

$$z' = z$$

$$t' = \gamma(t - vx)$$

We must keep in mind that v and t are measured in the natural system here and $\gamma = \frac{1}{\sqrt{1-v^2}}$. The inverse transformations are obtained in the usual manner.

5.2 Constructing the Diagrams

Till now we were discussing SR using the standard space-coordinate axes and dealing with time separately. Now, we introduce a time axes in addition to the usual space-coordinate axes. To keep things simple, we will consider just one space coordinate (usually x). In Minkowski diagrams (another name for spacetime diagrams), the time axis is drawn where we usually place our y-axis and the x-axis retains its purpose. Time axis is given by $x = 0$ and the space axis is given by $t = 0$. It is shown in Fig. 5.1.

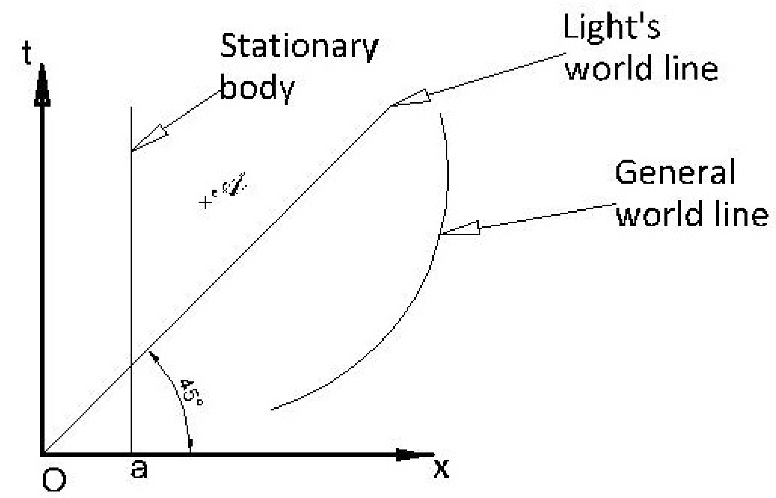


Figure 5.1: The Spacetime Diagram (2D)

In this diagram, every point can be uniquely represented by two coordinates (we have assumed y and z coordinates to be constant), one time and another space. Each point is called as an event. We will denote an event by a cursive, capital letter (like $\mathcal{A}$). The path traced by an object in the spacetime diagram is called its *word line* (given by a relation $x = x(t)$). The tangent to the world line tells us about the particle's velocity. The slope of a tangent is given by

$$\frac{dt}{dx} = \frac{1}{v}$$

A stationary body will have a vertical world line with an equation $x = a$ where a is the constant space coordinate of the particle. The slope of the world line of light is always 1 (as $v = c = 1$) i.e., it is equally inclined to the axes. Coordinates of an event is denoted in the form (t, x, y, z). General coordinates will be denoted by (x^0, x^1, x^2, x^3)[1]. We will denote a general `space-time` coordinate by $x^{greek\ letter}$ (like $x^a lpha$, x^β, etc) and a general `space` coordinate by $x^{latin\ letter}$ (like x^i, x^j, etc).

5.3 Representing the Axes of another Frame

Consider an object moving with a velocity v in x-direction wrt our frame. We will represent the frame of reference of the object (S') in our spacetime diagram (S). Note that frame now includes both time and space axes. First, let's draw the time axis of the object. The time axis of S' will be represented as $x' = 0$. By the Lorentz Transformation written above, $x' = 0 \Rightarrow x = vt$. This can be physically interpreted as the world line of the object in S frame as the object resides at $x' = 0$ in the S' frame.

[1]They are not powers but indices

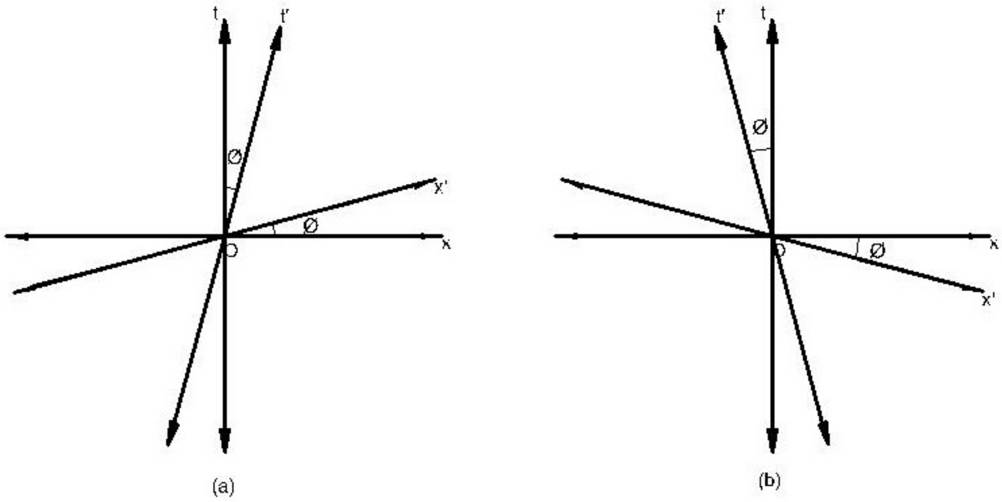


Figure 5.2: Axes of another frame in a Minkowski Diagram

The space axis of S' frame is given by $t' = 0$. By the above Lorentz Transformations, $t' = 0 \Rightarrow t = vx$. We could have found this axis by finding the locus of those points where light rays sent from points, equidistant from O, on t' axis would meet. The x$'$ and t$'$ axes are inclined at the same angle (say ϕ) with x and t axes respectively, given by $\phi = tan^{-1}|v|$. The situation for positive and negative v is shown in Fig.5.2(a) and (b) respectively.

5.4 Interval

Just like distance in Euclidean Space ($\Delta x^2 + \Delta y^2 + \Delta z^2$), there is a quantity called *interval* in the Minkowski Space. Consider two events, with coordinates (t, x, y, x) and $(t + \Delta t, x + \Delta x, y + \Delta y, z + \Delta z)$. The interval between the two events is defined as

$$\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$$

where $\Delta s^2 = (\Delta s)^2$. Δs^2 can be negative too. It won't matter as we won't take its square root. Two events are said to be

- *spacelike* separated if $\Delta s^2 > 0$
- *timelike* separated if $\Delta s^2 < 0$
- *lightlike* or *null* separated if $\Delta s^2 = 0$

Just like distance between two points is indepedent of the coordinate axes chosen in Euclidean Space, the interval Δs^2 between two events is independent of the frame of reference. For our frames S and S', the intervals between two events are related as

$$\Delta s^2 = \Delta s'^2$$

The above relation will help us to calibrate the t' and x' axes drawn in the S frame.

5.5 Calibration of the Axes

Consider the hyperbolae $-t^2 + x^2 = a^2$ and $-t^2 + x^2 = -b^2$. These hyperbolae represent the locus of the points which are at intervals of a^2 and $-b^2$ respectively, from the origin. Since we know that interval is an invariant quantity, we can write the equations as $-t'^2 + x'^2 = a^2$ and $-t'^2 + x'^2 = -b^2$. The points where these intersect the t$'$ and x$'$ axes can be found out by putting $x' = 0$ and $t' = 0$ in appropriate equations. This will yield the points $t' = \pm b$ and $x' = \pm a$. Hence, the axes can be calibrated by checking the points of intersection of the hyperbolae with the axes. These 'calibration curves' have the world lines of light as their asymptotes. The event $\mathcal{B}$ looks further from the origin

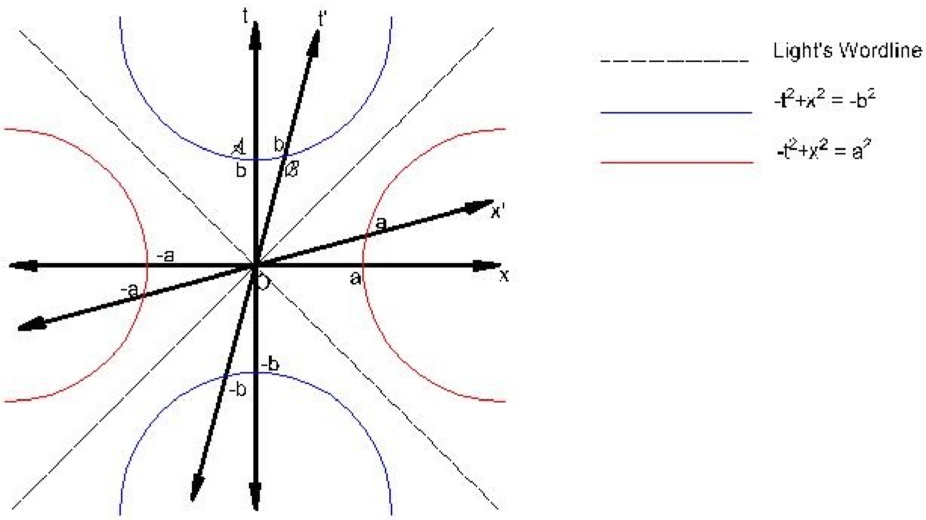


Figure 5.3: Calibration of Axes

than $\mathcal{A}$, though they denote the same coordinates in their respective frames. This is due to the fact that we are accustomed to measure the distance between two points as $\Delta x^2 + \Delta y^2 + \Delta z^2$ in Euclidean Space. But we must taken into account the $-\Delta t^2$ in the interval too as we are dealing in Minkoswki Space.

5.6 Future, Past and Elsewhere

Consider an event $\mathcal{A}$. Draw the path of light rays intersecting at $\mathcal{A}$. The cone thus formed is called a *light cone*. It separates spacelike, timelike and lightlike events wrt $\mathcal{A}$. The events within the cone are timelike separated from $\mathcal{A}$, the events on the cone are lightlike separated from $\mathcal{A}$ while the events outside the cone are spacelike separated from $\mathcal{A}$.

The events which are timelike separated and forward in time wrt $\mathcal{A}$ constitute the *future* wrt $\mathcal{A}$. Similarly, the events which are timelike separated from $\mathcal{A}$ and backward in time wrt $\mathcal{A}$ constitute the *past* wrt $\mathcal{A}$. The event $\mathcal{A}$ denotes the *present* or *NOW* wrt $\mathcal{A}$. All other events constitute elsewhere. The future is those set of events which a physical object can reach and the past are those set of events where a physical object could have been. Elsewhere denotes unreachable points. This contradicts the Classical Notion (Refer Fig.5.4(b)). All observers would agree about the future, past, present and elsewhere of an event as interval is an invariant quantity.

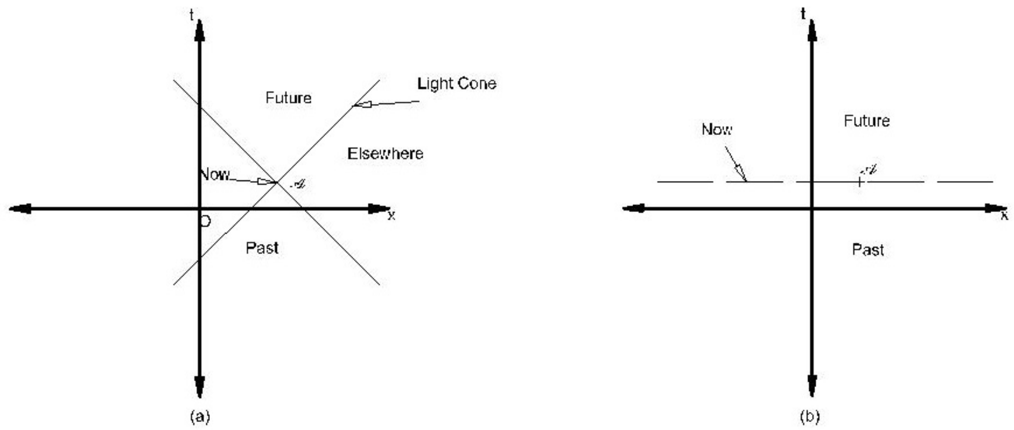


Figure 5.4: Future, Past and Now in Relativistic and Classical Domains

5.7 Relativity of Simultaneity

We previously introduced the idea of relativity of simultaneity using a thought experiment. Now we will visualise it using a Spacetime Diagram. In a particular frame, two events are said to be simultaneous if they occur at the same time i.e., have the same time coordinate. This requires the two events to lie on a line parallel to the space axis of that frame so that they only differ in the space coordinate. Consider Fig.5.5. Here, $\mathcal{A}$ and

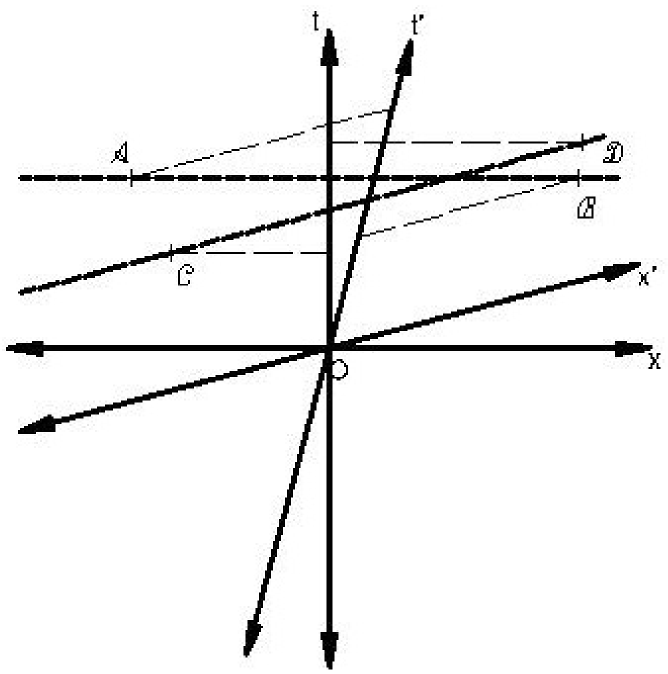


Figure 5.5: Simultaneity visualised using Spacetime Diagram

$\mathcal{B}$ lie on a line parallel to the x-axis and hence, are simultaneous wrt S. But, wrt S', $\mathcal{B}$ occurs before $\mathcal{A}$. Similarly, events $\mathcal{C}$ and $\mathcal{D}$ lie on a line which is parallel to the x$'$-axis and hence, are simultaneous wrt S'. But in the S frame, $\mathcal{C}$ occurs before $\mathcal{D}$. Hence, the Spacetime Diagram brings out the relativity of simultaneity naturally. As we shall see, simultaneity plays a very important role in time dilation and length contraction.

5.8 Time dilation

While discussing time dilation previously, we observed that the effect is reciprocal i.e., both observers see each other's clocks running slow. But how's that possible? After all, we must be able to tell which clock is slower. We will reason out why we can't do so using the Spacetime Diagrams.

Consider Fig5.6. Let there be a clock in S' at its spatial origin. We, sitting in the S frame, start observing the clock when $t = t' = 0$. Then, after some time (say 1s in our frame), we observe it once more. We observe the clock at $\mathcal{A}$. In our frame, it is simultaneous with $\mathcal{B}$, and hence the time we measure is the value at $\mathcal{B}$. But, when we look at the clock at $\mathcal{A}$, it is displaying a lower value, as evident by the calibration curve drawn through it. Hence, we conclude that time runs slower in S'. Notice that we have

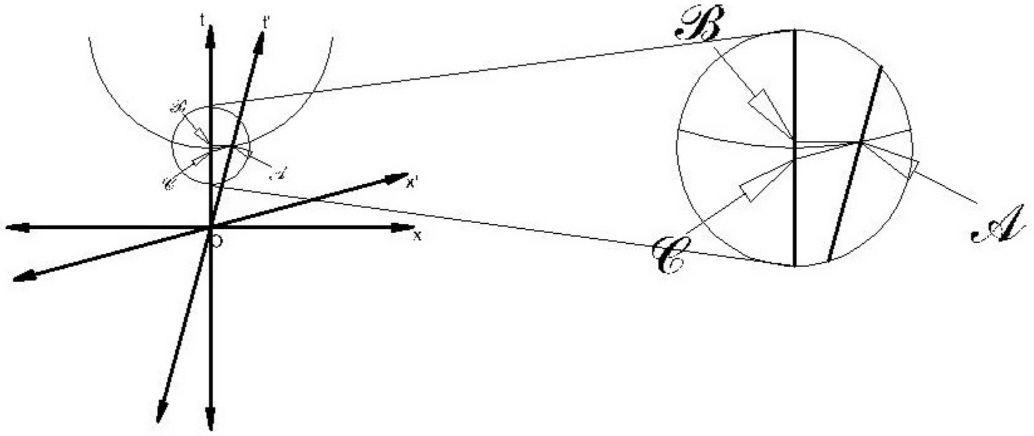


Figure 5.6: Time Dilation visualised using Spacetime Diagram

used our two clocks (one at our spatial Origin and one at the space coordinate of $\mathcal{A}$) to compare with a single moving clock (at the spatial origin of S'). The observer S' doesn't agree with our conclusion as wrt S', $\mathcal{B}$ is not simultaneous with $\mathcal{A}$ and hence, it would be wrong to compare say that both have same time coordinates. Instead, S' argues that $\mathcal{C}$ is simultaneous with $\mathcal{A}$ and hence, the time at those two events should be compared. S' observes that the clock of S at $\mathcal{C}$ reads a time which is less than what S' reads on its own clock, which is evident due to the calibration curve. Here, yet again, we observe that S' has used two clocks (one at its spatial origin and another at the spatial coordinates of $\mathcal{C}$) to compare with single moving clock (one at the the spatial origin of S).

We can define proper time in terms of the interval between two events. Proper Time (τ) between two events is defined as the time ticked off by a single clock which passes through both the events. In the clock's frame, the interval between the two events will be given by

$$??\Delta s^2 = -\Delta\tau^2 \tag{5.1}$$

Since, the interval between two events is invariant, the proper time between two events is simply given by $\sqrt{-\Delta s^2}$. As noted in the second chapter, $\Delta\tau = \frac{\Delta t}{\gamma}$, were Δt is the time measured in any general frame. This relation can also be obtained geometrically from the Spacetime Diagram.

5.9 Length Contraction

Consider Fig5.7. It depicts the situation wherein a rod lies at rest in the S' frame. The two lines parallel to t$'$-axis depict the worldlines of the ends of the rod. The length of any object is measured by noting the spatial coordinates of the body *simultaneously*. Since different observers disagree on simultaneity of events, it is only obvious from the above discussion that they would disagree on the lengths of objects too.

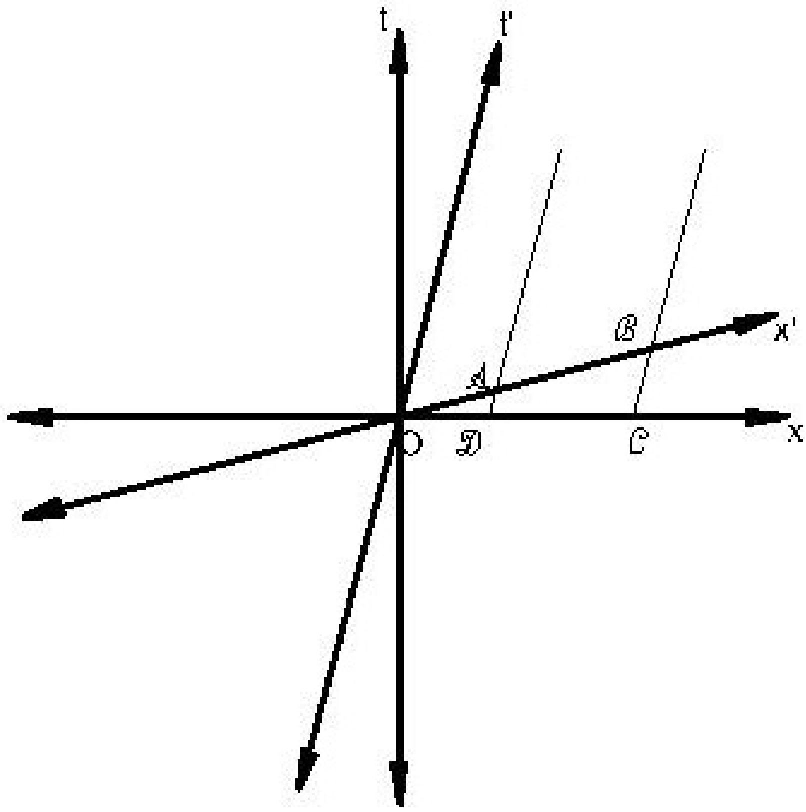


Figure 5.7: Length Contraction visualised using Spacetime Diagram

For S, the length of the rod can be measured by measuring the spatial distance between $\mathcal{C}$ and $\mathcal{D}$ as those two events are simultaneous for S. But for S', those two events are anything but simultaneous. Instead, it observes the events $\mathcal{A}$ and $\mathcal{B}$ as simultaneous and hence, claims that the length of the rod is the distance measured by S' between $\mathcal{A}$ and $\mathcal{B}$. Though we shouldn't do this without calibration curves, we can roughly see that the length measured by S' is more than what S measures. Hence, the length of the rod is contracted for S. A similar situation can be analysed by keeping the rod stationary in the S frame.

We can define proper length of a body using intervals. Proper length of the body can be defined as $\sqrt{\Delta s^2}$ between the two points where the worldlines of the ends of the rod intersect a line parallalel to the spatial axis of the frame in which the rod is at rest. In the above figure, two such points are $\mathcal{A}$ and $\mathcal{B}$.

Chapter 6

Four Vectors

In this chpter will introduce the concept of four vectors (four velocity, four momentum, etc) which are four-dimensional vectors. The notion of dot-product is modified to suit the Minkowski Space.

6.1 Vectors

As we had remarked earlier, that we will represent a point in the format (time, space). We will follow that convention for vectors too. Consider a displacement vector $\Delta\vec{x}$. A vector is independent of the coordinate system we use to analyse it. It points in the same direction. But its components may change on changing the coordinate system. Consider our usual frame S. The components of $\Delta\vec{x}$ in the S frame can be written as $(\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3)$. The components are more generally represented as Δx^α (index can be any greek letter). The space coordinates can be represented by Δx^i (index can be any latin letter).

On changing our frame to S', the vector, $\Delta\vec{x}$, remains the same but the components change. Similar to the way we proceeded for S, we can represent $\Delta\vec{x}$ as $(\Delta x^{0'}, \Delta x^{1'}, \Delta x^{2'}, \Delta x^{3'})$ and the components can be generally referred to as $\Delta x^{\alpha'}$. Similarly, the space components are represented as $\Delta x^{i'}$.

6.2 Einstein Summation Convention

We can relate the components of the displacement vector using the Lorentx Transformations. The relation for the time coordinate is

$$\Delta x^{0'} = \gamma\left(\Delta x^0 - v\Delta x^1\right)$$

The above relation can be compactly written as

$$\Delta x^{0'} = \sum_{\beta=0}^{3} \Lambda^{0'}_{\beta} \Delta x^\beta$$

where $\Lambda^{0'}_0 = \gamma$, $\Lambda^{0'}_1 = -\gamma v$ and $\Lambda^{0'}_2 = \Lambda^{0'}_3 = 0$. In general, for all components, the Lorentz Transformations can be written as

$$\Delta x^{\beta'} = \sum_{\alpha=0}^{3} \Lambda^{\beta'}_{\alpha} \Delta x^\alpha$$

We will be dealing with such summations regularly and it will be very tedious to write the summation symbol ($\sum$) over and over again. As a soothing, smart and elegant solution, we introduce the *Einstein Summation Convention.* In this convention, we omit writing any summation signs. Then how will we know which indices to sum you ask? Well, whichever indices occur in both the superscript and the subscript (of two different elements) are the indices over which the summation is to be performed. The range of summation is clear from the context.
The aforementioned summations can be succinctly re-written as

$$\Delta x^{0'} = \Lambda^{0'}_{\beta} \Delta x^{\beta}$$

$$\Delta x^{\beta'} = \Lambda^{\beta'}_{\alpha} \Delta x^{\alpha}$$

In the above equation, α is called the *dummy index*, which iterates over all possible values and can be replaced by any other greek letter. On the other hand, β' is called the *free index*, which can take four values $(0, 1, 2, 3)$, thus forming four separate equations. The equation can also be written as

$$\Delta x^{\beta'} = \Lambda^{\beta'}_{0} \Delta x^{0} + \Lambda^{\beta'}_{i} \Delta x^{i}$$

for latin indices take values from the set $\{1, 2, 3\}$.

Coming back to the Lorentz Transformations, we can represent the coefficients $(\Lambda^{\beta'}_{\alpha})$ in the form of a 4×4 matrix. We must note that these transformations apply to any displacement vector in the Minkowski Space.
E.g. If the vector $\vec{A}$ has the components (A^0, A^1, A^2, A^3) $(\equiv A^{\alpha})$ in the S frame, then its components in the S' frame are given by

$$A^{\beta'} = \Lambda^{\beta'}_{\alpha} A^{\alpha}$$

Note that $\Lambda^{\beta'}_{\alpha}$ depends only one the relative velocity v. The inverse transformations $(\Lambda^{\beta'}_{\alpha})$ are obtained by replacing v with $-v$ in $\Lambda^{\beta'}_{\mu}$. That's the only difference between the two. It can be shown (by repeated application of transformations) that

$$\Lambda^{\beta'}_{\alpha} \Lambda^{\alpha}_{\mu} = \delta^{\beta'}_{\mu}$$

Where δ is our beloved *Kronecker Delta.* The above relation can be visualised in terms of matrices too.

6.3 Magnitude and Scalar Product of Vectors

Consider a vector $\vec{A}$ with components A^{α} in the S frame. The *magnitude* of $\vec{A}$ is given by

$$\vec{A}^2 = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2$$

It can be written in a more elegant manner as

$$\eta_{\alpha\beta} A^{\alpha} A^{\beta}$$

where

$$\eta_{00} = -1$$

and

$$\eta_{ij} = \delta^i_j$$

and rest all terms are zero. It seems a bit weird to the mind accustomed to the Euclidean Space. Proceeding along the same lines, the *scalar product* of two vectors is defined as

$$\vec{A} \cdot \vec{B} = \eta_{\alpha\beta} A^\alpha B^\beta = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3$$

The scalar product is a commutative operation. Note that the magnitude of $\vec{A}$ is also given by $\vec{A} \cdot \vec{A}$. By definition, the magnitude of a vector is frame independent as we defined it analogous to the interval and it transforms under Lorentz Transformations. We can prove that the scalar product too is an invariant. Consider two vectors $\vec{A}$ and $\vec{B}$. Then, the magnitude of $\vec{A} + \vec{B}$ is given by $(\vec{A} + \vec{B}) \cdot (\vec{A} + \vec{B})$. Expanding the expression, we get

$$(\vec{A} + \vec{B})^2 = \vec{A}^2 + 2\vec{A} \cdot \vec{B} + \vec{B}^2 \tag{6.1}$$

We know that magnitude of a vector is frame invariant. This means that $(\vec{A} + \vec{B})^2$, $\vec{A}^2$ and $\vec{B}^2$ are frame invariant. To satisfy these properties of magnitude, Eq.6.1 forces $\vec{A} \cdot \vec{B}$ to be a frame invariant quantity.

6.4 Basis Vectors

Every coordinate system, in Minkowski Space, has four basis vectors wrt which the components are defined. For S, we denote the standard basis by $\vec{e}_\alpha$ and those of S' are denoted by $\vec{e}_{\beta'}$. With respect to S, the components of $\vec{e}_\alpha$ are given as

$$\vec{e}_0 = (1, 0, 0, 0)$$

$$\vec{e}_1 = (0, 1, 0, 0)$$

$$\vec{e}_2 = (0, 0, 1, 0)$$

$$\vec{e}_3 = (0, 0, 0, 1)$$

For a vector $\vec{A}$, the following relation is obvious

$$\vec{A} = A^\alpha \vec{e}_\alpha = A^{\beta'} \vec{e}_{\beta'}$$

The basis vectors of S' have identical components in the S' frame. What is the relation between the basis vectors of the two different frames? We explore that below.
We start with the above relation i.e.,

$$A^\alpha \vec{e}_\alpha = A^{\beta'} \vec{e}_{\beta'}$$

Applying the Lorentz Transformations to $A^{\beta'}$ in the RHS, we get

$$A^\alpha \vec{e}_\alpha = \Lambda^{\beta'}_{\mu} A^\mu \vec{e}_{\beta'}$$

Since μ is a dummy index, we can replace it with any other dummy index. Here, we replace it with α. The equation becomes

$$A^\alpha \vec{e}_\alpha = \Lambda^{\beta'}_{\alpha} A^\alpha \vec{e}_{\beta'}$$

The RHS is a double sum. The order of summation doesn't matter for finite summations. We can re-write the equation as

$$A^\alpha \vec{e}_\alpha = A^\alpha \Lambda^{\beta'}_\alpha \vec{e}_{\beta'}$$

Since this should be true for any general $\vec{A}$, we deduce, by transferring all variables to one side and taking A^α out as a factor, that

$$\vec{e}_\alpha = \Lambda^{\beta'}_\alpha \vec{e}_{\beta'}$$

Hence, we have derived the relation between basis vectors of two different coordinate systems. Comparing this with $A^\alpha = \Lambda^\alpha_{\beta'} A^{\beta'}$, we observe a difference in the transformation coefficients. They are opposite (in the order of the indices). This property is necessary for the invariance of vectors.

6.5 Four Velocity

Before discussing the four velocity of an object, we must discuss about *Momentarily Co-moving Reference Frame*(MCRF). It's obvious from the name that an MCRF is a frame moving with the same velocity as the body at a particular moment. An MCRF remains the same for an object moving with uniform velocity. But, for an accelerating object, MCRF changes at every instant. For a particular body at a certain instant, there can be infinitely many MCRFs; each differing from the other in the orientation of the axes. But all move with the same velocity at a given instant.

We define the *four velocity* of an object as the tangent to the object's worldline such that it has a unit time-component in the object's MCRF and is represented as $\vec{U}$. Let's analyse this definition. The four velocity, as observed from a body's MCRF, will be a vertical vector with a unit time-component as the object is stationary in its MCRF and hence its world line can't have any components in the spatial direction. In other words, the four velocity of a body is just the standard basis vector in the temporal direction of its MCRF ($\vec{U} \equiv (1, 0, 0, 0)$ in an MCRF).

Since we know that magnitude of a vector is frame invariant, $\vec{U}^2 = -1$ in all Lorentz frames. We can also define $\vec{U}$ in terms of τ. Consider an infinitesimal displacement $d\vec{s} \equiv (dt, dx, dy, dz)$ of a body in the Minkowski Space. Its magnitude is just the interval between two infinitesimally separated events and accroding to Eq.**??**,

$$d\vec{s}^2 = -d\tau^2$$

In the MCRF of the body, there are no spatial components of $d\vec{s}$. Hence, the time component of $d\vec{s}$ must be nothing but $d\tau$. Hence, in the MCRF,

$$d\vec{s} \equiv (d\tau, 0, 0, 0)$$

$$\Rightarrow \frac{d\vec{s}}{d\tau} \equiv (1, 0, 0, 0)$$

The RHS looks somewhat familiar. Ofcourse, they are the components of the four velocity of the body in its MCRF. Hence,

$$\vec{U} = \frac{d\vec{s}}{d\tau} = \left(\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}\right)$$

We know that $\Delta\tau = \frac{\Delta t}{\gamma}$ from our previous discussions. Letting, $\Delta\tau \to 0$, we get $\frac{dt}{d\tau} = \gamma$. Also, consider the ratio

$$\frac{U^1}{U^0} = \frac{\frac{dx}{d\tau}}{\frac{dt}{d\tau}} = \frac{dx}{dt} = v_x$$

where v_x is the x-component of the usual three velocity. We can write similar expressions for v_y and v_z. Considering these relations, which are valid in all inertial frames, $\vec{U}$ can be written in terms of the components of the three velocity of the body in the following manner

$$\vec{U} \equiv (\gamma, \gamma v_x, \gamma v_y, \gamma v_z)$$

6.6 Four Momentum

The *four momentum* of a body, analogous to the Classical definition, is defined as

$$\vec{p} = m_0\vec{U}$$

where m_0 is the rest mass of the body. Here we need not distinguish between rest mass and general mass as that effect has been implicitly taken care of by the $d\tau$ factor of $\vec{U}$ in the denominator. But for the sake of continuity we will continue to refer m_0 as rest mass. The components of the four momentum in any frame are

$$\vec{p} \equiv \left(E, p^1, p^2, p^3\right)$$

where E is the energy of the body and p^1, p^2 and p^3 are the components of the usual momentum three vector. Hence, $p^0 = E$.

For an isolated system, it can be shown that the four momentum of the system is conserved. Although the ideas of 'before' and 'after' are subject to arguments for different observers, the total four momentum of an isolated system retains its components, though different, for each inertial frame.

Consider $\vec{p} \cdot \vec{p}$. It is given by

$$\vec{p} \cdot \vec{p} = -E^2 + \sum_{i=1}^{i=3} (p^i)^2$$

Considering the fact that $\vec{p} = m_0\vec{U}$ and $\vec{U} \cdot \vec{U} = -1$, we get

$$\vec{p} \cdot \vec{p} = -m_0^2$$

Equating the two equations, we get

$$E^2 = m_0^2 + \sum_{i=1}^{i=3} (p^i)^2$$

which is a familiar relation from the previous chapters.

Let there be an observer with a velocity $\vec{U}_{obs}$. In an MCRF of the observer, $\vec{U}_{obs}$ has only one non zero component which is 1 in the temporal direction. Let there be a body with a momentum $\vec{p}$. In an MCRF of the observer, consider the dot product of $\vec{p}$ and $\vec{U}_{obs}$

$$\vec{p} \cdot \vec{U}_{obs} = -E \times 1 + p^1 \times 0 + p^2 \times 0 + p^3 \times 0 = -E$$

Hence the energy of any object with a given momentum in the MCRF of an arbitrary observer is merely

$$-\vec{p}\cdot\vec{U}_{obs} \tag{6.2}$$

. Note that we will get the result irrespective of the components of the vectors in a frame as the scalar product of two vectors is frame invariant.

6.7 Four Acceleration

We define the *four acceleration* ($\vec{a}$) of a body as

$$\vec{a} = \frac{d}{d\tau}\vec{U}$$

Consider $\vec{U}\cdot\vec{U}$. Differentiating it wrt τ, we get

$$\frac{d}{d\tau}\left(\vec{U}\cdot\vec{U}\right) = 2\vec{U}\cdot\frac{d}{dt}\vec{U}$$

But we know that $\vec{U}\cdot\vec{U} = -1$. Hence, $\vec{U}\cdot\frac{d}{d\tau}\vec{U} = 0$. Hence, $\vec{a}$ has the following properties

$$\vec{a} = \frac{d}{d\tau}\vec{U} \text{ and } \vec{a}\cdot\vec{U} = 0$$

In an MCRF, the time component of $\vec{a}$ is zero because $U^0 = 1$, a constant.

6.8 Photons

An interesting thing about photons is that they have no four velocity as there is no MCRF for a photon and hence, we can't define a unit vector in the temporal direction in its MCRF albeit we can find a vector tangent to its worldline. The four momentum of a photon is well defined for it contains the energy information of a photon. Also, $\vec{p}\cdot\vec{p} = -m_0^2 = 0$ as the rest mass of a photon is zero. This forces the spatial momentum to be equal to the energy of the photon. It can be shown that if a body has zero rest mass, then it must move with the speed of light in every frame.

6.9 Relativistic Doppler Effect

As promised in an earlier chapter, we will derive the relativistic Doppler Effect in a very elegant manner. Let a beam of light, of momentum $\vec{p}$, be emmited by a body (say B) in the positive x direction. Let the frequency of the light beam be ν wrt the body.
Then, the components of $\vec{p}$ should be $(h\nu, h\nu, 0, 0)$ since, $\vec{p}\cdot\vec{p} = -m_0^2 = 0$ as the rest mass of a photon is zero and the light beam moves only in the x direction.
Let there be an observer moving with a velocity $\vec{U}$ wrt B ($\vec{U} \equiv (\gamma, \gamma v_x, \gamma v_y, \gamma v_z)$ wrt B).
Let the frequency of the light beam be ν' wrt the observer. Then, from Eq.6.2

$$E' = h\nu' = -\vec{p}\cdot\vec{U} = -(-h\nu\times\gamma + h\nu\times\gamma v_x + 0 + 0)$$

$$\Rightarrow \nu' = \gamma\nu(1 - v_x)$$

The above equation predicts a very interesting thing which didn't exist in the Classical realm. Even if we put $v_x = 0$ in the above equation i.e., the motion of the observer is perpendicular to the light beam, we still observe the Doppler Effect (since $\nu' = \gamma\nu \neq \nu$). This was indeed confirmed by experiments.

Chapter 7

Tensors

In this chapter we will introduce tensors. We will discuss about the different types of tensors and their properties. We have already seen the metric tensor ($\eta_{\alpha\beta}$) in the previous chapter.

7.1 Definition of Tensors

A *tensor* of type $\binom{0}{N}$ is a function of N vectors into Real numbers which is linear in each of its N arguments. Let's define a function $\mathbf{g}$ as

$$\mathbf{g}(\vec{A},\vec{B}) = \eta_{\alpha\beta}A^{\alpha}B^{\alpha}$$

It takes two vectors as arguments and gives their scalar product (a real number). It is trivial to show that it is linear in both of its arguments. Thus, $\mathbf{g}$ can be regarded as a tensor of type $\binom{0}{2}$. It is called the *metric tensor*. Another example of a tensor would be $f(x) = x + 2$, where $x \in \mathbb{R}$. It is a $\binom{0}{0}$ tensor. It takes 0 vectors as arguments and produces a real number.
Tensors are frame invariant i.e., they give the same output for the same vectors irrespective of their components in a frame.

7.2 Components of a Tensor

The *components* of a $\binom{0}{N}$ tensor in a frame are defined as the outputs produced by a tensor when operated upon the basis vectors of the frame. The components of a tensor are frame dependent for the basis vectors are different for every frame.
E.g. Components of the metric tensor are given by $\mathbf{g}(\vec{e}_\alpha,\vec{e}_\beta) = \vec{e}_\alpha \cdot \vec{e}_\beta = \eta_{\alpha\beta}$. It has 16 components as α and β can take four values each. They can be arranged in a 4×4 matrix.

7.3 One forms

A tensor of type $\binom{0}{1}$ is called a *one form*. Due to some other reasons, they are also known as *covectors* or *covariant vectors*. So, a one form takes a single vector as an argument and produces a real number as an output. We shall denote a general one form as $\tilde{p}$ and as we know, $\tilde{p}(\vec{A}) \in \mathbb{R}$.

If $\tilde{s} = \tilde{p} + \tilde{q}$ and $\tilde{r} = a\tilde{p}$ ($a \in \mathbb{R}$), then we define $\tilde{s}(\vec{A}) = \tilde{p}(\vec{A}) + \tilde{q}(\vec{A})$ and $\tilde{r}(\vec{A}) = a\tilde{p}(\vec{A})$. With these definitions for addition and scalar multiplication, the space of all oneforms satisfies the conditions for a vector space and is called the *dual vector space.*

Consider a one form $\tilde{p}$. Its components are given by $p_\alpha = \tilde{p}(\vec{e_\alpha})$ which are four in number. Using the properties of tensors, we can show that

$$\tilde{p}(\vec{A}) = \tilde{p}(A^\alpha \vec{e}_\alpha) = A^\alpha \tilde{p}(\vec{e_\alpha}) = A^\alpha p_\alpha$$

The expression $A^\alpha p_\alpha$ is called the *contraction* of vector $\vec{A}$ with the one form $\tilde{p}$ and is fundamental to tensor analysis. Note the convention, lower index for one form components and upper index for vector components. The transformation law for one forms in Minkowski Space can be derived in the following manner

$$p_{\beta'} = \tilde{p}(\vec{e}_{\beta'}) = \tilde{p}(\Lambda^\alpha_{\beta'} \vec{e}_\alpha) = \Lambda^\alpha_{\beta'} p_\alpha$$

The transformation of one form components is just like that of the basis vectors (opposite to that of vector components) and this guarantees the invariance of tensors, just like vectors.

7.4 Basis One Forms

Just like the standard basis vectors, there exist standard basis one forms. They are denoted as $\tilde{\omega}^\alpha$ and are related to $\tilde{p}$ as $\tilde{p} = p_\alpha \tilde{\omega}^\alpha$ (analogous to $\vec{A} = A^\alpha \vec{e}_\alpha$). Let's find the components of the standard basis one forms. Using the properties of tensors, it can be shown that

$$\tilde{p}(\vec{A}) = p_\alpha \tilde{\omega}^\alpha (A^\beta \vec{e}_\beta) = p_\alpha A^\beta \tilde{\omega}^\alpha (\vec{e}_\beta)$$

But we know that $\tilde{p}(\vec{A}) = p_\alpha A^\alpha$

$$\Rightarrow \tilde{\omega}^\alpha (\vec{e}_\beta) = \delta^\alpha_\beta$$

δ^α_β gives the β^{th} component of the α^{th} basis one form. The above equation defines the one form basis in terms of the basis vectors and hence, they are known as dual bases. The relation is not between individual pairs like $\tilde{\omega}^0$ and $\vec{e}_0$. They can be listed as

$$\tilde{\omega}^0 = (1, 0, 0, 0)$$

$$\tilde{\omega}^1 = (0, 1, 0, 0)$$

$$\tilde{\omega}^2 = (0, 0, 1, 0)$$

$$\tilde{\omega}^3 = (0, 0, 0, 1)$$

There are many other basis one forms but these are the standard ones and we shall use these. It can be easily proven that

$$\tilde{\omega}^{\alpha'} = \Lambda^{\alpha'}_\beta \tilde{\omega}^\beta$$

Oneforms are sometimes visualised as surfaces in space and the number of surfaces pierced by a vector represents the value produced by the oneform when it acts on that vector.

Also, let's define *normal oneforms* to a surface. The oneforms that produce a value of zero when they act on vectors tangential to a surface are known as normal oneforms. If the surface is closed, the normal oneforms are called *outward* if they produce a positive value for vectors pointing out of the surface.

7.5 Gradient as a oneform

Consider a function $\phi(t, x, y, z)$. Since we can write the spacetime coordinates uniquely in terms of the proper time, we can represent our function by $\phi(\tau)$ (parameterisation). Let's differentiate ϕ wrt τ, which we can express as

$$\frac{d\phi}{d\tau} = \frac{d\phi}{dt}\frac{dt}{d\tau} + \frac{d\phi}{dx}\frac{dx}{d\tau} + \frac{d\phi}{dy}\frac{dy}{d\tau} + \frac{d\phi}{dz}\frac{dz}{d\tau}$$

$$\Rightarrow = \frac{d\phi}{dt}U^0 + \frac{d\phi}{dt}U^1 + \frac{d\phi}{dt}U^2 + \frac{d\phi}{dt}U^3$$

Now, $(\frac{d\phi}{dt}, \frac{d\phi}{dx}, \frac{d\phi}{dy}, \frac{d\phi}{dz})$ is the gradient of ϕ. $\frac{d\phi}{d\tau}$ is a real number formed by some combination of the components of the gradient and the four velocity. Hence, gradient can be regarded as a oneform as it clearly is linear on vectors and produces a real number when it acts on them. It is represented as $\tilde{d\phi}$ and has the obvious components.

We shall use a different convention for denoting derivatives of a function. It is shown below

$$\phi_{,x} = \frac{d\phi}{dx}$$

$$\phi_{,\alpha} = \frac{d\phi}{dx^\alpha}$$

An interesting thing to note is that $x^\alpha_{,\beta} = \delta^\alpha_\beta$ (obviously) and $\tilde{\omega}^(\alpha\vec{e}_\alpha) = \delta^\alpha_\beta$. This means that the basis oneforms can be represented as

$$\tilde{\omega}^\alpha = d\tilde{x}^\alpha$$

7.6 $\binom{0}{2}$ Tensors

We have already defined such tensors (they take two vectors and produce a real number). An apt example will be the metric tensor $\mathbf{g}$. A $\binom{0}{2}$ tensor can also be formed from two $\binom{0}{1}$ tensors by their *outer product*. Consider two oneforms $\tilde{p}$ and $\tilde{q}$. Then, $\mathbf{f} = \tilde{p} \otimes \tilde{q}$ is a $\binom{0}{2}$ tensor formed from by the outer product of two oneforms. Note that

$$\mathbf{f}(\vec{A}, \vec{B}) = \tilde{p} \otimes \tilde{q}(\vec{A}, \vec{B}) = \tilde{p}(\vec{A}) \cdot \tilde{q}(\vec{B})$$

The outer product is thus, not commutative. The components of $\mathbf{f}$ are given by

$$f_{\alpha\beta} = \mathbf{f}(\vec{e_\alpha}, \vec{\beta})$$

$\mathbf{f}$ can be expressed in terms of the basis $\binom{0}{2}$ tensors as given below

$$\mathbf{f} = f_{\alpha\beta}\tilde{\omega}^{\alpha\beta}$$

One can show that

$$\tilde{\omega}^{\alpha\beta} = \tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta$$

We can arrange the components $f_{\alpha\beta}$ into a 4×4 matrix. If the matrix is symmetric (anti-symmetric), $\mathbf{f}$ is symmetric (anti-symmetric) too. We can produce symmetric or anti-symmetric $\binom{0}{2}$ tensors from an ordinary $\binom{0}{2}$ tensor just like we do so with ordinary matrices.

7.7 The Metric Tensor

The metric tensor plays a special role in defining the relation between oneforms and vectors. Consider the entity $\mathbf{g}(\vec{V},)$. It's the metric tensor with just one argument, craving for another vector. If provided with a vector it will produce a real number, just like a oneform. Hence, one is justified in writing the below expression

$$\mathbf{g}(\vec{V},) = \tilde{V}() = \mathbf{g}(,\vec{V})$$

It is obvious that $\tilde{V}(\vec{A}) = \vec{V}\cdot\vec{A}$. The components of $\tilde{V}$ are

$$V_\alpha = \tilde{V}(\vec{e_\alpha}) = \mathbf{g}(V^\beta\vec{e}_\beta,\vec{e_\alpha}) = \eta_{\alpha\beta}V^\beta$$

Hence, $\tilde{V} \equiv (V_0, V_1, V_2, V_3) \equiv (-V^0, V^1, V^2, V^3)$, the components of $\tilde{V}$ and $\vec{V}$ differ in just the sign of the time component. The inverse transformation is given by

$$V^\alpha = \eta^{\alpha\beta}V_\beta$$

where $\eta^{\alpha\beta}$ is identical to $\eta_{\alpha\beta}$ in terms of its components.

In the usual 3D space with which we are well acquainted, there is no distinction between oneforms and vectors. But, Minkowski space forces us to distinguish between them (may be due to the 'weird' scalar product). Such distinction is also observed in other fields of physics like Quantum Mechanics (Hilbert Space). Hence, in Minkowski space, the gradient vector differs from the gradient oneform (in the time component).

The magnitude of a oneform is defined as the same as that of the associated vector i.e.,

$$\tilde{p}^2 = \vec{p}^2 = \eta_{\alpha\beta}p^\alpha p^\beta$$

It can be shown that

$$\tilde{p}^2 = \eta^\alpha p_\alpha p_\beta$$

Ofcourse, it is frame invariant (as magnitude of a vector is invariant).
Normal vectors are the ones associated with normal oneforms. Normal vectors, or for that matter normal oneforms, are called unit normal if their magnitude is ± 1. An outward normal vector is the vector associated with a outward normal oneform.

Note that one can regard vectors too as tensors, of type $\binom{1}{0}$, because when provided with a oneform they can produce a real number and are linear functions of oneforms. Hence, to place oneforms and vectors at the same pedestal, we shall denote the contraction of a vector and oneform by

$$\left\langle \tilde{p}, \vec{V} \right\rangle$$

7.8 $\binom{M}{N}$ Tensors

Before discussing $\binom{M}{N}$ tensors, consider $\binom{M}{0}$ tensors. They take M oneforms as their arguments and produce as real number. E.g. $\vec{V}\otimes\vec{W}\ldots\otimes\vec{X}$ is a $\binom{M}{0}$ type tensor where $\vec{V},\vec{W}\ldots\vec{X}$ are M vectors.

Now, $\binom{M}{N}$ tensors are just functions, linear in each argument, which take M oneforms and N vectors as inputs and produce Real numbers as output. E.g. $\vec{V}\otimes\vec{W}\ldots\otimes\vec{X}\otimes\tilde{p}\otimes$

$\tilde{q} \ldots \otimes \tilde{r}$ is a $\binom{M}{N}$ tensor where $\vec{V}, \vec{W} \ldots \vec{X}$ are M vectors and $\tilde{p}, \tilde{q} \ldots \tilde{r}$ are N oneforms. The components of a $\binom{2}{2}$, for instance, are represented as

$$T^{\alpha\beta}_{\mu\nu} = \mathbf{T}(\tilde{\omega}^{\alpha}, \tilde{\omega}^{\beta}, \vec{e}_{\mu}, \vec{e}_{\nu})$$

The metric tensor converts a $\binom{M}{N}$ into a $\binom{M-1}{N+1}$ in the following manner,

$$T^{\alpha}_{\beta} = \eta_{\mu\beta} T^{\alpha\mu}$$

where it converts a $\binom{2}{0}$ tensor into a $\binom{1}{1}$ tensor. The inverse converts a $\binom{M}{N}$ tensor into a $\binom{M+1}{N-1}$ tensor in the following manner,

$$T^{\alpha\beta} = \eta^{\beta\mu} T^{\alpha}_{\mu}$$

The above operations are called *index lowering* and *index raising* respectively.

Chapter 8

Fluids in SR

In this chapter we will explore the concept of fluids from a relativistic point of view and employ tensors in our analysis.

8.1 Dust

Before discussing dust, let's define continuum. A *continuum* is a collection of particles so innumerable that we have to deal with average quantities, like number density, energy density, etc, to describe it. An *element* is a homogeneous collection of particles in which the average quantities, like average velocity, KE and interparticular spacing, are the same everywhere. A continuum can be constituted of many elements. A *fluid* is a continuum wherein the antislipping forces between adjacent elements is very weak.

Dust is a collection of particles such that all of them are at rest in some Lorentz Frame. This means that all the constituent particles move with the same velocity. The *number density* (n) of dust is defined as the number of particles per unit volume in the dust's MCRF. Infact, any property's density, like energy density and momentum density, can be regarded as that property per unit volume. The number density may differ from point to point. Let the total number of particles be N in a rectangular volume V in the dust's MCRF. Then,

$$n = \frac{N}{V}$$

If in some frame, say S', the rectangular volume of dust moves parallel to one of its sides with a velocity v, the volume appears to be $V' = \frac{V}{\gamma}$ due to Lorentz Contraction of the side parallel to the velocity. The number of particles does not change while changing the frame. Hence, the number density in this frame becomes

$$n' = \gamma n$$

Flux of some quantity can be defined as the amount of that quantity crossing a unit area per unit time. Let, in the frame S', all the particles be moving with a velocity v in the x$'$ direction. Let R be a surface perpendicular to the x$'$ axis with an area of ΔA. In time Δt, $(\gamma n) \times (v\Delta t \times \Delta A)$ particles cross the surface R. Hence, the flux of the particles across R will be the number of particles crossing R per unit area per unit time, which is

$$(Flux)^{x'} = \frac{(\gamma n) \times (v\Delta t \times \Delta A)}{\Delta A \times \Delta t} = nv\gamma$$

If the particles' velocity had a y′ component too, the number of particles crossing R in time Δt would be $(\gamma n) \times (v_x \Delta t \times \Delta A)$ (It depends only only on the x′ component of the velocity). Then, the flux would be

$$Flux = \frac{(\gamma n) \times (v_x \Delta t \times \Delta A)}{\Delta A \times \Delta t} = n v_x \gamma$$

Note that $(Flux)^{x'} \neq Flux$ as γ in $Flux$ is $\frac{1}{\sqrt{1-(v_x^2+v_y^2)}}$.

8.2 The Number Flux Vector

Define $\vec{N} = n\vec{U}$ where $\vec{U}$ is the four velocity of the dust particles. $\vec{N}$ is known as the *Number Flux Vector*. As we know that $\vec{U} \equiv (\gamma, \gamma v_x, \gamma v_y, \gamma v_z)$, we can write the components of $\vec{N}$ as

$$\vec{N} \equiv (\gamma n, n v_x \gamma, n v_y \gamma, n v_z \gamma)$$

We immediately see that N^0 is the number density of dust in the frame in which we are observing it, N^1 is the flux of the particles through a surface of constant x and similar are the interpretations for N^2 and N^3. The thing to note is that $\vec{N}$ is frame independent. Although its components change, we need just one vector $\vec{N}$ to describe the quantities like number density and fluxes of particles across surfaces in all Lorentz frames.

We can visualise number density of particles as the timelike flux of particles i.e., the number of particles crossing a surface of constant time which is nothing but the number of particles in a spatial volume at a particular instant. Hence, the components of $\vec{N}$ describe the fluxes across various surfaces. Usually a surface is defined in the form $\phi(t,x,y,z) = constant$. $d\tilde{\phi}$ defines the normal oneform to the surface at each point and hence, can define the surface uniquely. The unit normal oneform to the surface is defined as $\tilde{n} = \frac{d\tilde{\phi}}{|d\tilde{\phi}|}$ where $|d\tilde{\phi}|$ is $\sqrt{\eta^{\alpha\beta} d\tilde{\phi}_\alpha d\tilde{\phi}_\beta}$. The surface element is defined as $\tilde{n} dx^\alpha dx^\beta dx^\mu$ where α, β and μ are all distinct. It can be easily verified that the flux across a surface of constant ϕ is nothing but

$$\left\langle \tilde{n}, \vec{N} \right\rangle$$

8.3 The Stress Energy Tensor

Let all dust particles have rest mass m. In an MCRF of dust, the energy is just m (rest mass energy) as they are not moving. Hence, the energy density (ρ) is given by

$$\rho = nm$$

In some other frame say S', the number density is γn and the energy of each particle is γm. Hence, the energy density in S is

$$\rho' = \gamma^2 nm = \gamma^2 \rho$$

This transformation involves two factors of γ and hence, can't be represented in a frame invariant manner using a vector. We will need a $\binom{2}{0}$ tensor for this job, called the *stress energy tensor*.

The stress energy tensor is denoted by $\mathbf{T}$. The component $T^{\alpha\beta} = \mathbf{T}(d\tilde{x}^\alpha, d\tilde{x}^\beta)$ denotes the flux of the α^{th} component of momentum across a surface of constant x^β. Hence

- T^{00} is the flux of p^0 across a surface of constant t which is nothing but energy density.
- T^{0i} is the flux of energy across a surface of constant x^i.
- T^{i0} is the i^{th} momentum density.
- T^{ij} is the flux of i^{th} momentum across a surface of constant x^j.

The mathematical form for the tensor is

$$\mathbf{T} = \rho\vec{U} \otimes \vec{U}$$

We immediately realise from the above equation that $T^{\alpha\beta} = T^{\beta\alpha}$ i.e., $\mathbf{T}$ is a symmetric tensor. All the fluxes are in the increasing direction of the constant component of the surface.

Before moving further, we must note something about the momentum flux. Momentum flux through a surface is nothing but the pressure on the surface as the flux of , say, i^{th} momentum across a surface of constant x^j is the rate of change of i^{th} momentum at the surface per unit area which is

$$\frac{\frac{dp}{dt}}{A} = \frac{F}{A} = P$$

8.4 General Fluids

Now that we have gained insights from the special case of dust, we will deal with general fluids. The particles of a general fluid may have random individual velocities. There may also be some interparticle forces giving rise to potential energy apart from the rest mass energy. We denote energies other than rest mass energy per particle by Π. Hence,

$$\rho = n\Pi + \rho_0$$

where ρ_0 is the rest mass energy density. We must note that the scalar quantities are defined in the MCRF of each element. We will deal in MCRF of each element.

Let's discuss about the inerpretation of $\mathbf{T}$ in General Fluids. According to the first law of Thermodynamics,

$$\Delta Q = \Delta E + P\Delta V$$

where ΔE is the total eneergy of the system (here, an element), ΔQ is the heat absorbed by the system and $P\Delta V$ is the work done by the system, all of which are macroscopic quantities. We can show that this equation, through some manipulations, transforms into

$$\Delta q = T\Delta S$$

where Δq is the heat absorbed per particle, T is temperature and ΔS is specific entropy (entropy per particle). The above equation relates the properties of each particle.

T^{0i} gives the energy flux, which is related to heat conduction across a surface. T^{i0} gives the momentum density associated with the heat flow. As discussed before, T^{ij} represents the force between two elements.

8.5 Conservation Laws

Consider a cubic element at a particular instant with one corner at the origin and each side of length l. The energy content of the cube is l^3T^{00}. The energy entering the cube from the surface $x = 0$ is $l^2T^{01}(x = 0)$ and that from the surface $x = l$ is $-l^2T^{01}(x = l)$. Similar are the expressions for influxes through other surfaces of the cube. The energy influx must be equal to the rate of change of energy within the cube, according to Conservation of Energy, in other words

$$\frac{\partial}{\partial t}(l^3T^{00}) = l^2\left[T^{01}(x=0) - T^{01}(x=l) + T^{02}(y=0) - T^{02}(y=l+T^{03}(z=0) - T^{03}(z=l))\right]$$

Divide the equation by l^3 throughout. Applying the limit $l \to 0$, we get

$$\frac{\partial}{\partial t}(T^{00}) = -\frac{\partial}{\partial x}(T^{01}) - \frac{\partial}{\partial y}(T^{02}) - \frac{\partial}{\partial z}(T^{03})$$

$$\Rightarrow T^{0\alpha}_{,\alpha} = 0$$

The above equation is a mathematical expression for energy conservation in terms of the stress energy tensor. A similar expression for momentum conservation is $T^{\alpha\beta}_{,\beta} = 0$.

Similarly, the conservation of number of particleswithin an element (the change in number of particles in an element can only be through the flux of particles through the boundaries)can be mathematically expression as $N^{\alpha}_{,\alpha} = \left(nU^{\alpha}_{,\alpha}\right) = 0$ which can be derived in a similar manner. We just need to replace the stress energy tensor by the number flux vector in the derivation and get the expression.

8.6 Perfect Fluids

Perfect Fluids are defined to be those kinds of fluids which do not conduct heat and have no viscosity (no inter-elemental forces parallel to the boundary surface) in their MCRF. This implies that in the MCRF, $T^{0i} = T^{i0} = 0$ as there is no heat conduction. We know that T^{ij} represents force on the surface with constant x^j in the x^i direction. Since the forces parallel to a surface are zero for a perfect fluid, $T^{ij} = 0$ unless $i = j$. These forces must be zero in all MCRFs and the only way that's possible is if $T^{ii} = P$ for all values of i. These implications render $T^{\alpha\beta}$ to be a diagonal matrix with just two unique values (ρ and P) for the matrix elements.

It can be shown that the stress energy tensor for a perfect fluid is

$$\mathbf{T} = (\rho + P)\vec{U} \otimes \vec{U} + P\mathbf{g}^{-1}$$

where $\mathbf{g}^{-1}$ is the inverse of the metric tensor ($\eta^{\alpha\beta}$). The pressure arises due to random motion of the particles. The conservation laws can be shown to hold true for perfect fluids too. It can also be shown that the flow of particle conserving perfect fluid is adiabatic i.e., $\frac{dS}{d\tau} = 0$.

8.7 Gauss' Law in Minkowski Space

We know the Gauss' law in Euclidean Space. It has the following form in the Minkowski Space

$$\int W^{\alpha}_{,\alpha} d^4x = \oint W^{\alpha} n_{\alpha} d^3S$$

where W^α are components of a vector $\vec{W}$, $\tilde{n}$ is the unit normal oneform to the surface. Sense of normal is outward.d^3S denotes the three volume of the $3-D$ hypersurface bounding the four dimensional volume of integration (d^4x).

Chapter 9

An Overview of General Relativity

In this chapter we will discuss a brief overview of General Relativity without delving deep into the math but we will discuss the physical implications of the theory.

9.1 The Need for GR

We discussed about Special Relativity in the previous sections. We made a distinction between accelerating and non-accelerating frames in the Special Relativity regime. But laws of physics must treat every frame of reference on an equal footing, be it accelerating, uniformly moving or spinning. Also, there was a need to incorporate gravity. Einstein was occupied with this problem after publishing Special Relativity. To incorporate all frames on an equal footing and gravity, General Relativity was born.

9.2 Principle of Equivalence

The Principle of equivalence was put forth by Einstein in 1911, changing the way we look at frames of reference and gravity. Gravitational force, unlike other fundamental forces, affects every particle regardless of its charge and other properties and causes them to accelerate at the same rate (given they are put under similar conditions). Since $m = E$, it is reasonable to expect that even photons are affected by Gravitational force. It produces the same acceleration for all particles, regardless of their mass. But the Classical explanation uses the action-at-a-distance concept which has no room in the relativistic construct as no signal or information can travel faster than light. Hence, there was a need to come up with an alternative explanation for the Gravitational force and the principle of equivalence was the first step towards it.

Consider two frames of reference, say S and S'. Let there be a uniform gravitational field in S and let there be no gravitational field in S' which is uniformly accelerating wrt an inertial frame. The *Principle of Equivalence* states that the above two frames, S and S', are physically equivalent and indistinguishable by experiments conducted in one particular frame. Consider a spacecraft, far away from any source of gravitaional force. Within that spacecraft is an astronaut with a ball in his hand. The spacecraft starts accelerating wrt an inertial frame. The direction of acceleration is in the direction of the astronaut's stance as seen from the inertial frame. The astronaut lets go of the ball in his hand and observes that it falls to the floor of the spacecraft, just like it would happen on Earth. He concludes that the spacecraft is now in a gravitational field. Moreover, he experiences a normal force from the floor just like he would experience while standing on Earth. This scenario justifies the Principle of Equivalence beautifully.

Now consider a container, much smaller than Earth, in a free fall in Earth's gravity. Inside the container is a person with a few belongings. Since, the container is much smaller than Earth, we can assume that the Earth's gravitational field is uniform for the container. The person within the windowless container will observe all his belongings floating freely and would feel no force pushing him downwards towards the container's floor. These observations lead him to conclude that the he is in a gravity free region. Hence in regions with uniform gravity, we cannot tell if we can't tell if we are in a gravity free region or not.

9.3 Tidal Forces

The situation is a bit different in non-uniform gravitational fields. In such fields, we have to deal with *Tidal Forces*. Tidal forces is the named used to refer to the difference in gravitational field between nearby points. Tidal forces due to the moon on Earth cause the tides and hence the name. Now consider our planet Earth. Let a ball be dropped from a very great height. We shall ignore all effects except gravity. If we observe the ball from a Lorentz frame falling freely under Earth's gravity. The ball, according to the frame, appears to move with a uniform velocity. But what if there was a substantial distance between the origin of the Lorentz frame and the ball? The acceleration of both would be different as the gravitational field is non-uniform. This means that the ball would appear to accelerate if the Lorentz frame were at a distance from it. Hence, in the presence of non-uniform gravity, Lorentz frames can describe just 'local' events but not 'global' ones. Consider a global Lorentz frame with its origin falling freely under the Earth's gravity. Clearly, not all points of the frame fall freely under earth's gravity i.e., not all points are falling radially inwards. Hence, 'global' Lorentz Frames cannot be used to describe events in the presence of gravity. In non-uniform Gravitational Fields, wordlines that begin parallel do not necessarily remain parallel. This gives us a preface to curvature of spacetime, as we shall see later. We can say that the laws of physics in a freely falling reference frame are the same as that in SR if the tidal forces are negligible.

9.4 Rotating Frames of Reference

We have never considered accelerating frames of reference in SR (MCRFs are not accelerating), leave alone rotating. Let's have a look at them. Consider a window-less cylindrical container, free from any kind of influences, floating in deep space. Let there be two identical clocks, one at the centre of the floor and one at the circular edge of the floor. We are looking at the events from distance outside the container (Assuming the walls are transparent for us). Let the floor start rotating at a uniform angular velocity, about its central axis, wrt us. Then, the clock at the centre will appear to be at rest while that at the edge will appear to be moving along a circle wrt us. The latter clock will experience an outward force which we ascribe to the centrifugal force. But in the container's frame, the clocks regard the container to be at 'rest'. They ascribe the outward force which they experience to a (radial) gravitational field which is allowed by the Principle of Equivalence.

Now, according to us the outer clock is moving but the inner one is stationary (we may as well say that the inner clock is in sync with ours). From our study of SR, we conclude that the outer clock ticks at a slower rate than the inner one. It is fine for us but how will we explain this in the container's frame? We say that clocks are affected by gravitational fields too and that the rates at which they tick depends on where they are in the gravitational field.

Now that we have discussed about the temporal coordinates in gravitational fields, let's have a look at the spatial coordinates. Let there be a rod, much smaller than the radius of the container. When the rod is kept in the container, parallel (approximately) to the edge, we observe the rod's length to be a bit smaller than it really is (Lorentz Contraction). This means that an observer within the container will obtain a larger value for the circumference than us. Now, if the rod is re-oriented along the radius of the container, we percieve the rod at its original length as the translation of the rod is perpendicular to its length. This means that an observer in the container will measure the same value for diameter as we do. Now, on dividing the circumference by itsi diameter, we shall obtain $\pi = 3.142\ldots$ but the observer in the container will obtain a value higher than that. If the container were at rest, both would have obtained the same value for π. Hence, the ideas of Euclidean geometry won't apply here if we attribute the same length to a rod in all positions and orientations in a frame (container).

9.5 Deflection of Light

Equiped with our knowledge of the Equivalence Principle, we can infer that light deflects in the presence of massive bodies or gravitational fields. Consider a cuboidal cabin with holes on two opposite facing sides. Let the holes be parallel to the cabin's floor. Let there be a light pulse emitter which is in a position such that a pulse emitted by it would pass through both holes. Inititally, both the emitter and the cabin are at rest in a uniform gravitational field (wrt a frame far from any gravitational influence) acting downward towards the floor of the cabin. Now, let the emitter emit a pulse and as soon as the pulse enters the cabin through the first hole, it starts falling freely in the uniform gravitational field.

By the Principle of Equivalence, the freely falling cabin can be considered as a Lorentz frame in which SR holds. According to SR, the velocity of light is constant in a Lorentz frame. Hence according to the cabin, the light pulse enters the first hole, travels across the cabin parallel to its floor and leaves the cabin through the second hole. Now let's look from another frame which is still stationary in the gravitational field. According to this frame, the light enters the cabin through the first hole and leaves it from the second hole with a finite time interval between both the events. Between those events, the cabin falls freely under gravity. Hence, the path of light is not a straight line in this frame but instead, it is bent towards the source of gravity. Hence, we can say that light bends in the presence of gravity. It seems that the law of constancy of velocity of light doesn't hold in GR; except in freely falling reference frames.

The light doesn't interact with the cabin at all and thus, the cabin must be merely regarded as a frame of reference and nothing else. The above phenomenon must happen in the absence of a cabin too. But what about non-uniform gravitational fields? We have seen that tidal forces exist in non-uniform gravitaional fields and also that Euclidean geometry is no good in describing events in such fields. Hence, we will have to deal with the situation using 'local' freely falling cabins (frames of reference) which are infinitesimal both in spatial and temporal extents. In this way, we can apply the Principle of Equivalence to each local cabin and find the local bend in light's path. Later we can patch together all the inifinitesimal frames in an appropriate manner to get the global picture. We must note that the patching of frames must not be done in an Euclidean manner but must incorporate curvature of space and time.

9.6 Curvature of Spacetime

We have seen that Euclidean Spaces don't comply with the idea of GR and we need new Spaces with different characteristics. Enter Curved Spaces. We shall see that the idea of gravity and its effects will be beautifully embodied within curved spaces and something called 'geodesics'. Suppose there are two bodies moving in two parallel straight lines in a Euclidean Space. We are currently dealing in 2D space. It is obvious that the paths of the two bodies will never meet each other unless the forces on them are non-zero. If they meet each other at some point in the future, in the Newtonian domain, it will be attributed to the gravitational attraction between them. In this process we shall have to consider force equations, accelerations, etc. But suppose that instead of a flat Euclidean plane, the bodies were on the surface of a sphere. Now, let the bodies start off on straight parallel lines. But wait, what are 'straight' lines on the curved surface of a sphere? *Geodesics* are lines which are the straightest possible ones on that surface. On a sphere, geodesics are the great circles. Any body which wants to move on a straight line on a sphere will move along one of the great circles. It is noteworthy that every great circle meets another one at two diametrically opposite points.

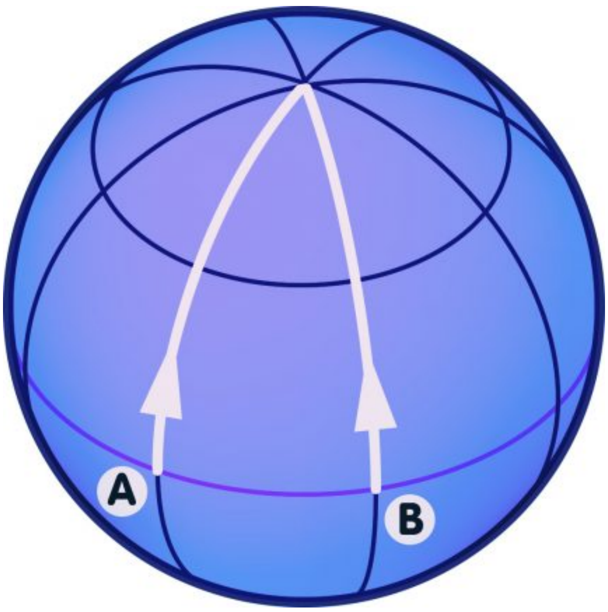


Figure 9.1: Straight paths of two bodies on a sphere Courtesy: [14]

Coming back to the two bodies, if they start on straight and parallel paths, they will meet at one of the points where their great circles intersect; even if no other force acts on them! We can see a connection between geodesics on curved surfaces and the paths of two bodies which are mutually attracted through their gravitational fields. Einstein saw this and used this relation beautifully. He didn't characterise gravity as a 'force' like we do in Classical Mechanics. Instead he made it a consequence of the geometry of spacetime. According to GR, the spacetime is Euclidean locally but not globally. We can say that it is made of infinitesimal Euclidean frames sticthed in a particular manner which determines the curvature of the spacetime, the metric of the spacetime and other things. The source of the curvature of spacetime is attributed to matter and energy (stress-energy tensor). The spacetime where a massive body resides is curved due to it and has geodesics defined on it. Any other particle around it will follow the geodesics and note that the path which leads directly towards a massive body is a geodesic for that curved spacetime too. In this way, Einstein made gravity a part of the geometry of the spacetime in which we live. In a way, matter follows geodesics and also creates them.

In a nutshell, in a Euclidean Space, two straight and parallel lines do not meet at any point in the Space. But, in a curved space, two lines which start off parallel may meet each other at a point in the space. Einstein exploited this fact and said that massive bodies (or components of the stress-energy tensor) warp and curve the space around them and hence, change the definition of 'straight' and 'parallel' lines. In this way, Einstein beautifully embodied gravity into a geometrical theory. John Archibal Wheeler, a theoretical physicist of the 20th century, elegantly summarised GR in the following manner, *"Spacetime tells matter how to move; matter tells spacetime how to curve"*.

9.7 Experimental Confirmations of GR

The predictions of GR have been experimentally confirmed many times. A few of them are listed here.

- **Mercury's Perhelion Shift**- We know that all planets revolve around the Sun in elliptical orbits, with the Sun at one of the focii. Their orbits have perihelion (closest point to Sun) and aphelion (furthest point to Sun). The planets of our Solar System revolve in nearly circular orbits, except Mercury whose orbit is highly eccentric and elliptical. If we include the gravitational influence of other planets on Mercury, there must be a shift of the axes of its orbit in the Sun's frame. This was predicted by Classical Mechanics too but was off by some 43 seconds per century. On the application of GR to this case, the discrepancy was settled to an acceptable extent. Thus, this proved as a successful test for GR. The flower-petal-like orbital precession of Mercury is shown in Fig. 9.2

- **Effect of Gravitation on Path of Light**- We saw that GR predicted the bending of light under gravitational influence. This too was experimentally confirmed in 1919 during a Solar eclipse. The Royal Astrnomical Society sent out two expeditions to Africa to observe a Solar eclipse and photograph the event. They compared the photos taken during the eclipse with other photos taken at night where the same stars were visible in the night sky. It was observed that the stars near the sun appeared to be shifted in their position which was attributed to the fact that light rays bend as they pass near the Sun.
  Another effect which has been observed is *Gravitational Lensing*. It is the term given to the phenomenon when light rays from far away sources pass through regions with high masses and undergo optical effects similar to lenses. They usually result in multiple images of the same object and sometimes form beautifull rings (Fig. 9.3).

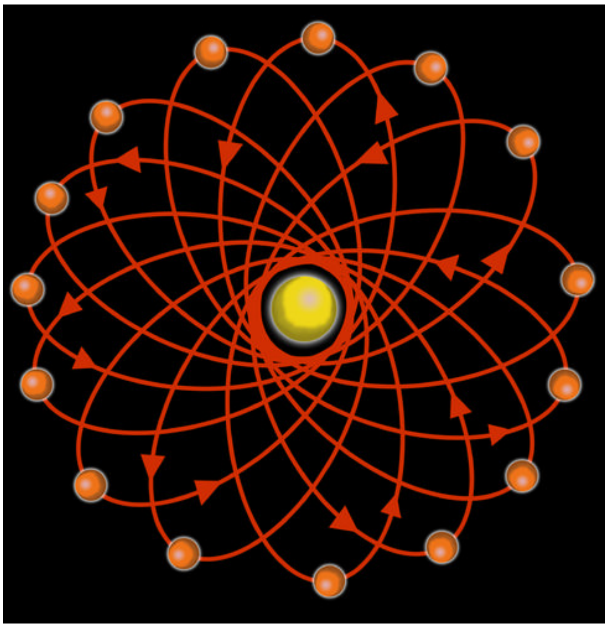


Figure 9.2: An exaggerated orbital precession of mercury Courtesy: [15]

- **Black Holes**- Recently (on April 10, 2019) the first ever image of the shadow of a black hole was released by the Event Horizon Telescope (EHT) confirming the existence of black holes which were predicted by General Relativity. The black hole in the image (Fig. 9.4) is the one at the centre of the galaxy Messier 87 (M87) called M87*. The EHT also aims to take an image of the black hole (Saggitarius A* or Sgr A*) at the center of our galaxy (Milky Way Galaxy).

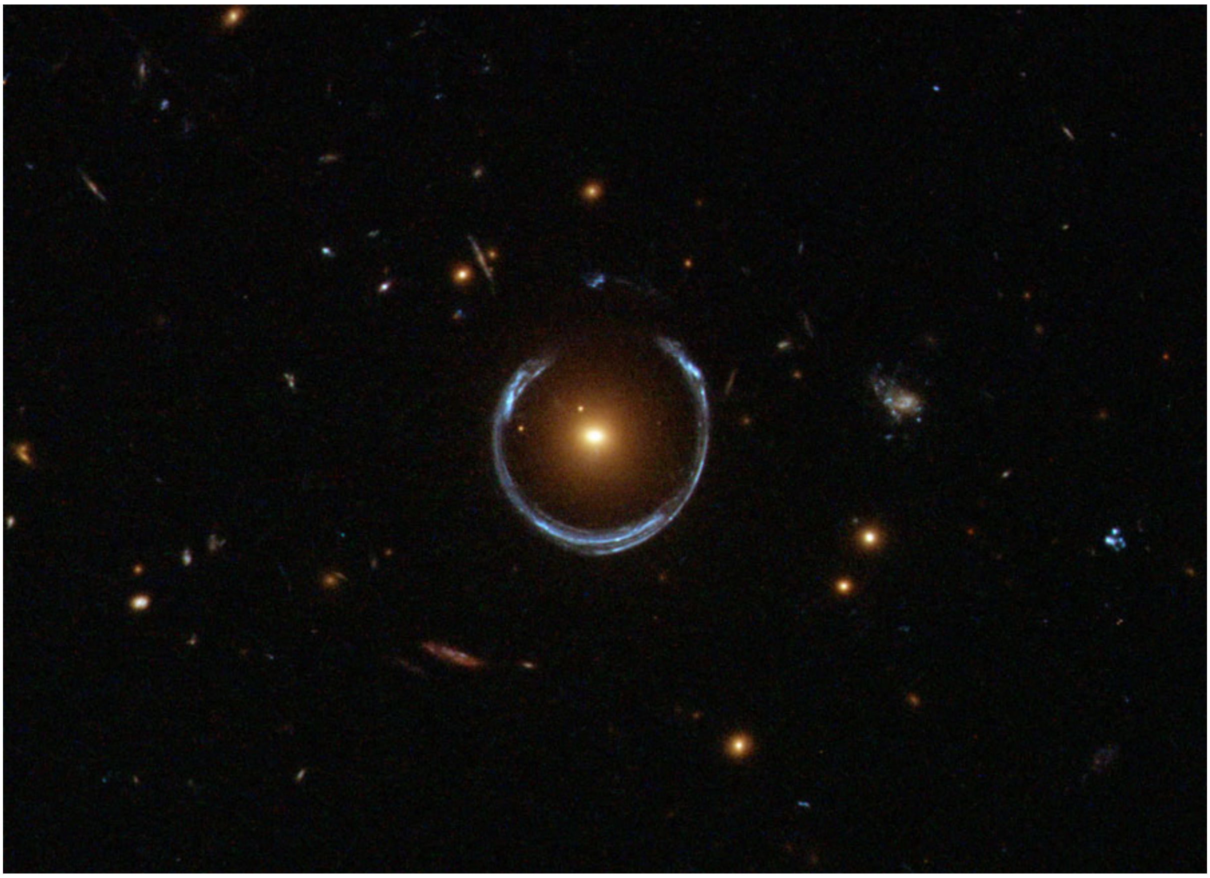


Figure 9.3: An Einstein 'Horseshoe' Courtesy: [16]

Figure 9.4: Blackhole M87* Courtesy: [17]

Chapter 10

Paradoxes and Problems

In this chapter we will explore two interesting 'paradoxes' and explain them based on our understanding of SR. We shall also solve two interesting problems.

10.1 Twin Paradox

This paradox is perhaps the most famous one; known to every Relativity enthusiast or a general Sci-Fi fan. ***Statement***: There exist identical twin brothers; Karan and Arjun. Both of them are *Vyomnauts*[1] in ISRO. They decide to conduct an experiment to test Special Relativity. Karan boards a spaceship which will travel at a speed of 0.6 towards an exoplanet which is at a distance of about 6 lightyears from Earth and then return back at the same speed towards Earth. All accelerations and decelerations are assumed to be instantaneous. Earth is assumed to be non accelerating for this problem. For Arjun, Karan's journey must take $2 \times \frac{6}{0.6} = 10$ years. According to Arjun, Karan must experience time dilation, as he is moving relative to him, and hence must be younger than Arjun when he returns to Earth. But Karan argues that while he is aboard the spaceship, Arjun appears to be in motion wrt him and hence, Arjun must experience time dilation wrt him. This renders Arjun younger than Karan at the end of the journey. This seems to be a paradox. Who's correct? How to resolve this apparent paradox?

Solution: The main cause of the paradox is that we assume both the frames (Karan's and Arjun's) to be similar in all respects (except the velocity ofcourse). But this assumption breaks down if we look at the situation closely. Karan's frame experiences an acceleration when his spaceship turns around for the return journey whereas Arjun's frame does not and we will show that this very difference turns the tables. Since we assumed the accelerations to be instantaneous, we could also say that Karan boards a different frame of reference the moment his spacecraft turns around. This is easily visualised in the Spacetime Diagram (Fig.10.1).

The blue coloured line represents the worldline of Karan's spaceship or in other words, the time axis of his frame(s). We note that at the event $\mathcal{A}$, his spaceship accelerates instantaneously and turns around. This changes his time axis. Values (in years) have been marked off on the time axes of both observers. The justification for the markings is as follows. We already have determined that in Arjun's frame, Karan takes 5 years to reach the exoplanet and another 5 years to return and hence, the markings on Arjun's

[1] The Indian word for 'Astronauts' equivalent to the Russian 'Cosmonauts'

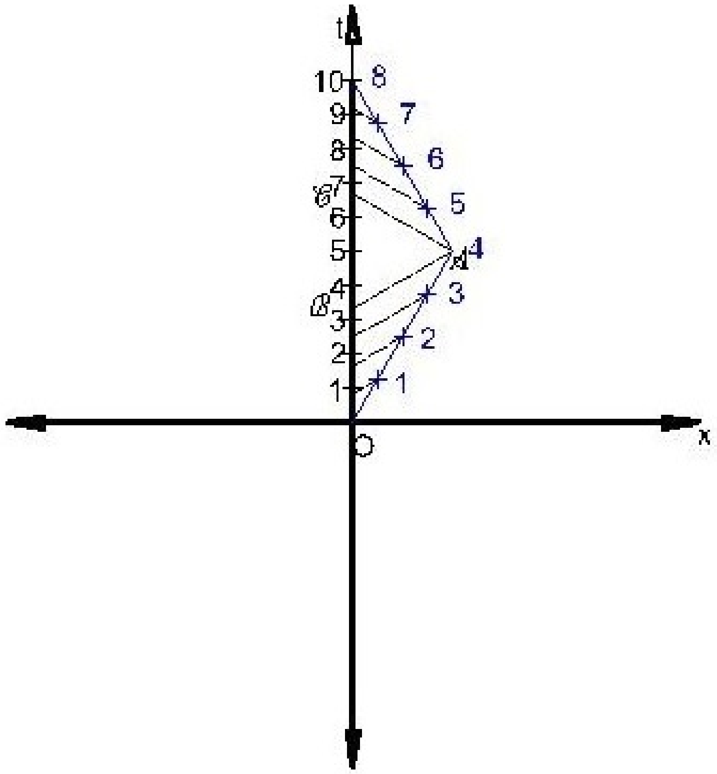


Figure 10.1: The Twin Paradox in a Spacetime Diagram

time axis are justified. Now, wrt Arjun, $\gamma = 1.25$ for Karan. Hence, according to Arjun, it takes Karan $\frac{5}{1.25} = 4$ years, in Karan's frame, to reach the exoplanet and hence, the markings on the receding time axis are justified. Similarly, on the return journey, $\gamma = 1.25$ and Arjun determines that Karan experiences 4 more years.
Let's explore Karan's point of view. The dashed lines in Fig.10.1 represent lines of simultaneity in Karan's frame. When Karan first reaches the exoplanet, just before turning, he observes the clocks of Arjun to be showing a value of something between 3 and 4 years ($\mathcal{B}$) and this confirms that according to Karan, Arjun's clocks tick slower. But just after Karan turns around, he sees the clocks of Arjun to be displaying a value somewhere between 6 and 7 years ($\mathcal{C}$). For the rest of his journey, his clocks tick off years while Arjun's tick off around 3 years, confirming that Arjun's clocks run slower than him. The whole difference is made when he instantaneously turns around; while Karan's clock register negligible change, Arjun's clock change drastically by around 3 years.

The above discussion justifies that by the end of the journey, Arjun would have aged by 10 years while Karan would hav eaged by only 8 years. All the difference was made just by the fact that Karan accelerated while Arjun did not.

10.2 The Pole and The Barn Paradox

This is one of the lesser known paradoxes of SR and its solution is pretty straightforward. It involves length contraction as opposed to time dilation in the previous paradox.
Statement: Consider a person holding a long pole (a pole vaulter, maybe?), at a distance from a barn. The barn has two doors, one at the front and one in the back (as shown in Fig.10.2). We are stationary on the ground, observing the sequence of events unfold

before us. The rest length of the barn is l_0 ad that of the pole is $2l_0$. The person, along with the pole, starts running at a speed of $\frac{\sqrt{3}}{2}$ towards the barn. In our frame, $\gamma = 2$ for the pole and hence, its length is exactly the same as that of the barn. This means that if we momentarily shut and open the barn doors (both front and back doors simultaneously) at the appropriate moment, the pole can be contained within the barn (albeit for a very tiny moment). This is possible in our frame ofcourse. But, problem arises when we shift our frames to that of the pole. In the pole's frame, the length of the barn appears to be $\frac{l_0}{2}$ and hence, it can't be contained within the barn, even for the tiniest moment.

Moreover, if we keep the front door closed in our frame and close the back door when the pole fully enters the barn we can entrap the pole within the barn. But in the pole's frame, this won't work as, apparently, the pole's back end will never be within the barn if the front door is closed. How's it possible that the pole can be contained within the barn in one frame but not in another? How can we resolve this?

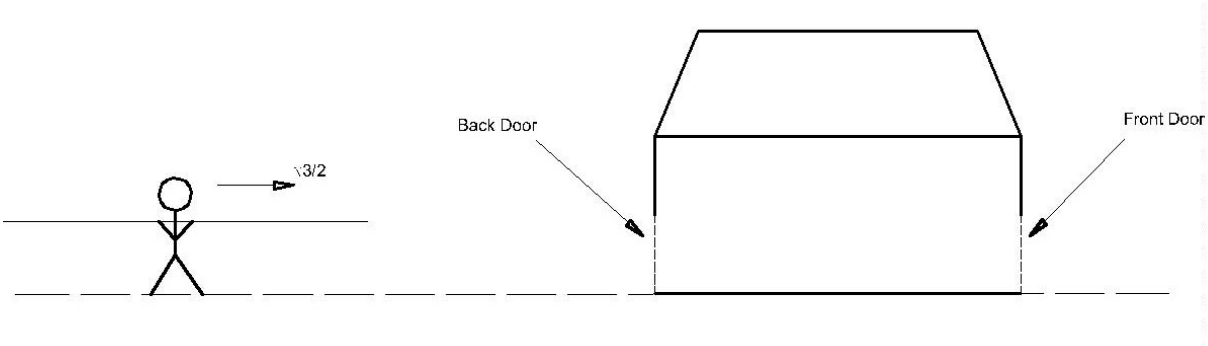


Figure 10.2: The Pole and The Barn Paradox

Solution: The solution to the first part of the problem hinges on the concept of Relativity of Simultaneity. Events which are simultaneous in one frame may not be so in any other frame. Hence, when we say that we *simultaneously* close the barn doors in our frame, the closure of the doors is not simultaneous in the pole's frame. Instead, in the pole's frame, the front door closes and opens first and then when the back end of the pole is beyond the back door, the back door closes and opens. We can prove it. Let the pole be the S' frame and let the ground be S frame. Let the coordinates of the back and front doors be x_1 and x_2 in the S frame and x'_1 and x'_2 in the S' frame. Let t_1 and t_2 denote the closure and opening of back and the front doors in the S frame and t'_1 and t'_2 in the S' frame.It is given that the closure of the doors is simultaneous in S frame i.e., $\Delta t = t_2 - t_1 = 0$. From the Lorentz Transformations, we know that

$$t' = \gamma(t - vx)$$

This means that

$$\Delta t' = \gamma(\Delta t - v\Delta x) = -\gamma v \Delta x$$
$$\Rightarrow t'_1 - t'_2 = \gamma v \Delta x$$

Since $\Delta x = x_2 - x_1 > 0$, the above equation indicates that $t'_2 < t'_1$ i.e., the front door opens and closes before the back door does so, as expected. The Spacetime Diagram (Fig.10.3) helps visualize the situation.In this way, the pole is allowed to pass the barn without any problems.

As for the second part, we must invoke the fact that no information can travel faster than the speed of light. Hence, in the pole's frame, if the front door remains closed

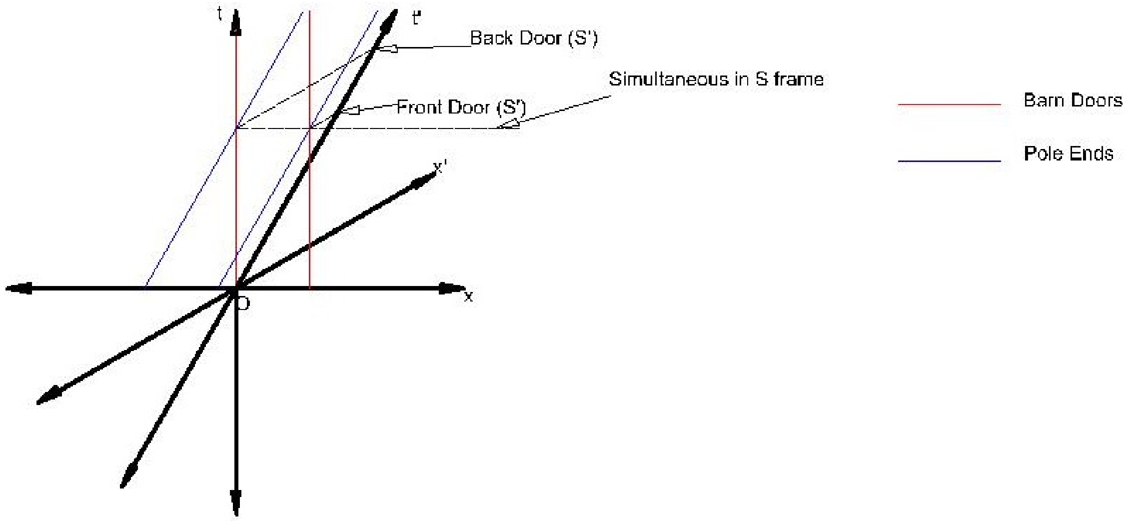


Figure 10.3: The Pole and The Barn Paradox in a Spacetime Diagram

and the front end of the pole hits it, it decelerates quickly to a halt but the rest of the pole continues to move because the information of deceleration due to the closed front door takes finite time to reach the other end of the pole. In this manner, the back end continues moving despite the front end being stationary and this cause the rod to lose its 'rigidity' and break into pieces. The pole was destined to break in our frame too because on closing both the doors, the pole would come to an abrupt halt within the barn and become longer than the barn and hence, break into pieces.

10.3 Invariance of Wave Equation of Light

In this section will prove that the wave equation of light is invariant under a Lorentz Transformation but not so under a Galilean transformation. To do so, consider the following wave equation

$$\frac{-\partial^2\phi}{\partial t^2} + \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0 \tag{10.1}$$

The above equation is written in Natural Units. In SI, the term $\frac{-\partial^2\phi}{\partial t^2}$ changes to $\frac{-1}{c^2}\frac{\partial^2\phi}{\partial x^2}$. Consider the same old frames S and S'. We are given that the above equation is valid in the S frame. We are supposed to find its mathematical form in S' under both the transformations. In doing so, we shall use the following fact from multi-variable calculus

$$\frac{\partial\phi}{\partial x^\alpha} = \sum_{\beta'=0}^{\beta'=3} \frac{\partial\phi}{\partial x^{\beta'}} \frac{\partial x^{\beta'}}{\partial x^\alpha} \tag{10.2}$$

We can draw a similar expression for second order partial derivatives, which can be written as

$$\frac{\partial^2\phi}{\partial^2 x^\alpha} = \sum_{\mu'=0}^{\mu'=3} \frac{\partial}{\partial x^{\mu'}} \left(\sum_{\beta'=0}^{\beta'=3} \frac{\partial\phi}{\partial x^{\beta'}} \frac{\partial x^{\beta'}}{\partial x^\alpha} \right) \frac{\partial x^{\mu'}}{\partial x^\alpha}$$

10.3.1 Galilean Transformation

To recall, the Galilean Transformation can be written as

$$t' = t$$

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

Then, we can arrange the derivatives of various coordinates in the form of 4×4 matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where the element in the β'^{th} row and α^{th} column represents the derivative of $x^{\beta'}$ wrt x^{α} (the row and column indices start from 0). Hence, from Eq.10.2 and the above matrix,

$$\frac{\partial \phi}{\partial t} = -v\frac{\partial \phi}{\partial x'} + \frac{\partial \phi}{\partial t'}$$

$$\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial x'}$$

$$\frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial y'}$$

$$\frac{\partial \phi}{\partial z} = \frac{\partial \phi}{\partial z'}$$

Differentiating the above equations once again, we get

$$\frac{\partial^2 \phi}{\partial^2 t} = \frac{\partial^2 \phi}{\partial t'^2} - 2v\frac{\partial^2 \phi}{\partial x' \partial t'} + v^2\frac{\partial^2 \phi}{\partial x'^2}$$

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \phi}{\partial x'^2}$$

$$\frac{\partial^2 \phi}{\partial^2 y} = \frac{\partial^2 \phi}{\partial y'^2}$$

$$\frac{\partial^2 \phi}{\partial z^2} = \frac{\partial^2 \phi}{\partial^2 z'}$$

Substituting the above values in Eq.10.1, we get

$$-\frac{\partial^2 \phi}{\partial t'^2} + 2v\frac{\partial^2 \phi}{\partial x' \partial t'} + \left(1 - v^2\right)\frac{\partial^2 \phi}{\partial x'^2} + \frac{\partial^2 \phi}{\partial y'^2} + \frac{\partial^2 \phi}{\partial z'^2} = 0$$

The above equation is not in the same form as Eq.10.1 and hence, clearly the wave equation of electromagnetic radiation (light) is *not* invariant under Galilean Transformation.

10.3.2 Lorentz Transformation

Recall that the Lorentz Transformation equations are

$$t' = \gamma(t - vx)$$
$$x' = \gamma(x - vt)$$
$$y' = y$$
$$z' = z$$

The matrix of derivatives (similar to the previous subsection) can be written as

$$\begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence, from Eq.10.2 and the above matrix, we can write

$$\frac{\partial\phi}{\partial t} = \gamma\frac{\partial\phi}{\partial t'} - v\gamma\frac{\partial\phi}{\partial x'}$$
$$\frac{\partial\phi}{\partial x} = -v\gamma\frac{\partial\phi}{\partial t'} + \gamma\frac{\partial\phi}{\partial x'}$$
$$\frac{\partial\phi}{\partial y} = \frac{\partial\phi}{\partial y'}$$
$$\frac{\partial\phi}{\partial z} = \frac{\partial\phi}{\partial z'}$$

Applying the above rules to get the second derivatives, we have

$$\frac{\partial^2\phi}{\partial t^2} = \gamma^2\frac{\partial^2\phi}{\partial t'^2} - 2v\gamma^2\frac{\partial^2\phi}{\partial x'\partial t'} + v^2\gamma^2\frac{\partial^2\phi}{\partial x'^2}$$
$$\frac{\partial^2\phi}{\partial x^2} = v^2\gamma^2\frac{\partial^2\phi}{\partial t'^2} - 2v\gamma^2\frac{\partial^2\phi}{\partial x'\partial t'} + \gamma^2\frac{\partial^2\phi}{\partial x'^2}$$
$$\frac{\partial^2\phi}{\partial y^2} = \frac{\partial^2\phi}{\partial y'^2}$$
$$\frac{\partial^2\phi}{\partial z^2} = \frac{\partial^2\phi}{\partial z'^2}$$

Substituting the above expressions in Eq.10.1, we get

$$-\left(\gamma^2\frac{\partial^2\phi}{\partial t'^2} - 2v\gamma^2\frac{\partial^2\phi}{\partial x'\partial t'} + v^2\gamma^2\frac{\partial^2\phi}{\partial x'^2}\right) + v^2\gamma^2\frac{\partial^2\phi}{\partial t'^2} - 2v\gamma^2\frac{\partial^2\phi}{\partial x'\partial t'} + \gamma^2\frac{\partial^2\phi}{\partial x'^2} + \frac{\partial^2\phi}{\partial y'^2} + \frac{\partial^2\phi}{\partial z'^2}$$

$$\Rightarrow -\gamma^2\left(1 - v^2\right)\frac{\partial^\phi}{\partial t'^2} + \gamma^2\left(1 - v^2\right)\frac{\partial^\phi}{\partial x'^2} + \frac{\partial^2\phi}{\partial y'^2} + \frac{\partial^2\phi}{\partial z'^2}$$

Since $\gamma^2 = \frac{1}{1-v^2}$, the above expression gets simplified to

$$\frac{-\partial^2\phi}{\partial t^2} + \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0$$

Lo and behold! The above expression is similar to Eq.10.1. This means that the wave equation of light is invariant under Lorentz Transformation.

10.4 Massless Particles and Their Speeds

Inspired by a question posed in AFCIGR ([6]), I asked the following question on stack-exchange ([20])
Question- Do all massless particles travel with the speed of light?

Solution- Short answer- yes, they do. We will prove it in the following lines. Consider a particle (rest mass = $m \neq 0$) moving with a speed v in the x-direction of our frame. The four velocity of such a particle would be $\vec{U} = (\gamma, \gamma v, 0, 0)$ and hence, its four momentum is

$$\vec{p} = m\vec{U} = (\gamma m, \gamma m v, 0, 0)$$

The total energy of the particle will be $E = \gamma m$, which we assume is finite and non zero. If we take the ratio of the magnitude of the spatial momentum and total energy, we get

$$\frac{p_x}{E} = v \tag{10.3}$$

Eq.10.3 is valid for massless particles too as we can apply the limit $m \to 0$ without any problem.

In general, we know that

$$E^2 = m^2 + \sum_{i=1}^{i=3} \left(p^i\right)^2$$

In our case, we have a massless particle ($m = 0$) and the only spatial momentum is in the x direction. So, we get

$$E^2 = p_x^2$$

$$\Rightarrow \frac{p_x}{E} = 1 \tag{10.4}$$

We chose positive sign for the ratio as we are dealing with magnitudes. From Eq.10.3 and Eq.10.4, we get

$$v = 1$$

Hence, we have proved that all massless particles travel with the speed of light.

Bibliography

[1] The image on the cover page is a courtesy of *Wikimedia Commons*

[2] Robert Resnick, *Introduction to Special Relativity*, John Wiley & Sons Inc., 1968

[3] Prof. Shiv Prasad, IITB, *Special Relativity*, *NPTEL Video Lectures*, 2017

[4] Bruce Knuteson, *8.20 Introduction to Special Relativity*, January IAP 2005, *Massachusetts Institute of Technology: MIT OpenCourseWare*

[5] Albert Einstein, *Relativity: The Special and The General Theory*, General Press, 2014

[6] Bernard F. Schutz, *A First Course in General Relativity*, Second Edition, Cambridge University Press, 2009

[7] Resources used for Twin Paradox include *TED-Ed*, *Fermilab*, *minutephysics*

[8] Resources used for Tensor Algebra include *The Science Asylum*, *Dan Fleisch*

[9] Markus Pössel, AEI, "The elevator, the rocket, and gravity: the equivalence principle" in: *Einstein Online* **Band 01** (2005), 01-1009

[10] Markus Pössel, "The equivalence principle and the deflection of light" in: *Einstein Online* **Band 04** (2010), 02-1020

[11] Markus Pössel, "Gravity: from weightlessness to curvature" in: *Einstein Online* **Band 01** (2005), 03-1011

[12] Steven S. Shapiro, Irwin I. Shapiro, "Gravitational deflection of light" in: *Einstein Online* **Band 04** (2010), 03-1003

[13] Markus Pössel, "The equivalence principle and the deflection of light" in: *Einstein Online* **Band 04** (2010), 02-1020

[14] *Einstein Online General Relativity Elementary Tour*

[15] *Mercury Orbit Precession*

[16] Einstein 'Horseshoe'- ESA/Hubble NASA (*APOD* 21 Dec, 2011)

[17] *Event Horizon Telescope*

[18] *Veritasium*

[19] Resources used for The Pole And The Barn Paradox include *Fermilab*, *Up and Atom*, *The Science Asylum*

[20] *AlmostClueless* (User on Stackexchange), *Physics Stackexchange*

[21] *Wikipedia*