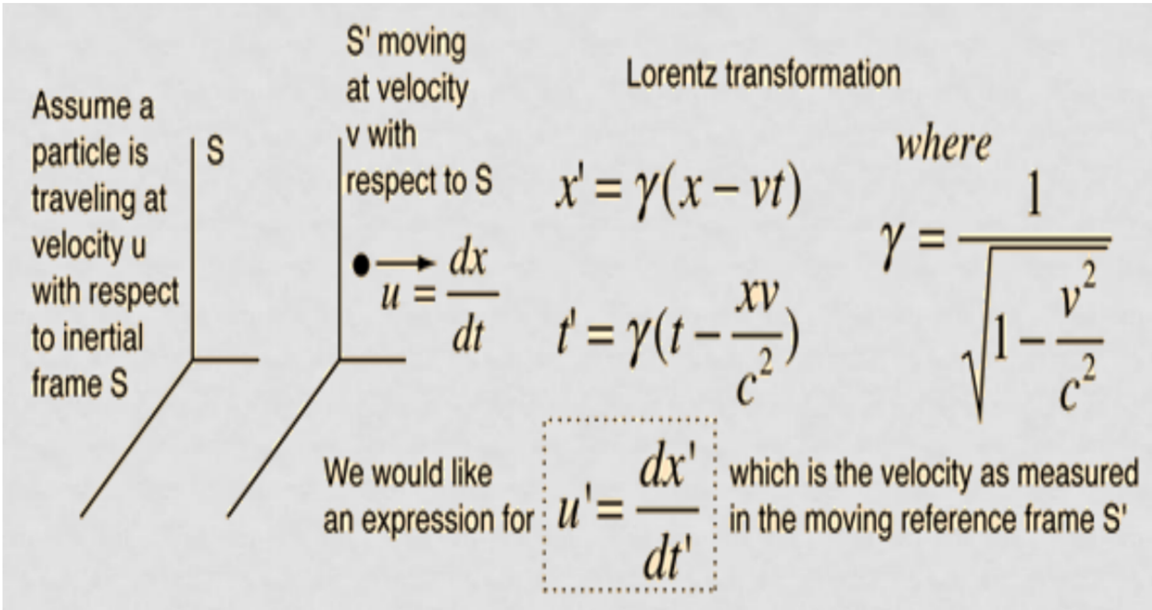
Introduction to Special Relativity



Lecture Notes for Bachelor of Science Course by Dr. E Bwayo

Contents

1 Introduction

At the turn of the twentieth century, the development of the Special Theory of Relativity brought into question many of the ideas in classical mechanics that had previously been regarded as fundamental. Maxwell's equations governing Electromagnetism had been formulated 30-40 years earlier and, although it was not appreciated until later, were to turn out to be closely interlinked with the new theory and supplied convincing evidence for its eventual acceptance. In fact it was Lorentz who laid the groundwork for relativity through his studies of electrodynamics, while Einstein contributed crucial concepts and placed the theory on a consistent and general footing. Beyond this, work throughout the twentieth century demonstrated that, even though its origins might have lain in electromagnetism and optics, Special Relativity can be applied to all types of interaction except large-scale gravitational phenomena. In modern physics, the theory serves as a benchmark for descriptions of the interactions between elementary particles, and relativistic features are now so well established that they form basic criteria to be built into any new theory.

Thus, if we consider only transformations in t and x, there must be constants $\alpha, \beta, \gamma, \delta$ such that $t' = \alpha t + \beta x$, $x' = \gamma t + \delta x$. Consider a point fixed in F (i.e., x fixed as t varies). Then $dx' = \gamma dt$, $dt' = \alpha dt$ and so

$$\frac{\gamma}{\alpha} = \frac{dx'}{dt'} = velocity \;\; of \;\; frame \;\; F \;\; with \;\; respect \;\; to \;\; F' = v(F, F') \tag{1.1}$$

One would expect $v(F, F') = -v(F', F)$ so that $\alpha = \delta$ (see below)

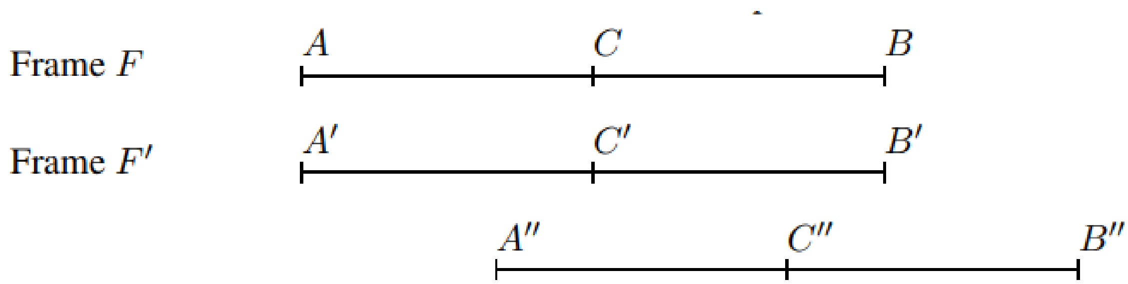
Practically, we can only consider relations between inertial frames such that our measuring apparatus (e.g., rulers and clocks) can actually be transferred from one to another. Such frames are said to be related. To go further we need two additional assumptions, that:

- the behaviour of apparatus transferred from F to F' is independent of the mode of acceleration.

- apparatus transferred from F to F' and then from F' to F" agrees with apparatus transferred directly from F to F"

With these assumptions and definitions, it is possible to state **The Principle of Special Relativity:** that all physical laws take equivalent forms in related inertial frames, so that we cannot distinguish between the frames.

Even in the 1900s, this was hardly new. Newton was aware of it, but he based his mechanics on the two fundamental premises (a) a rigid body has the same size in all frames, and (b) time is absolute. However, a very simple thought experiment shows why a revision of these ideas was needed. Consider two points A and B in an inertial frame F. Two events can be said to be simultaneous in F if light rays emitted from A and B at the time of the event meet at the mid-point C of AB.



Suppose a second frame F' moves with velocity v relative to frame F. The diagram shows that by the time the light rays meet at C, C' will have moved to $C'' \neq C$, so that events which are simultaneous in F cannot be simultaneous in F'. We conclude that simultaneity is not absolute but depends on the frame of reference under consideration.

Newtonian physics is a limiting case of the more correct Relativistic theory. Newtonian physics works just fine when the speeds you are dealing with are much less than the speed of light (which is about $3 \times 10^8 ms^{-1}$).

There are two main topics in relativity; General Relativity (which deals with gravity), and Special Relativity (which does not deal with gravity), Special Relativity is divided into two topics, kinematics and dynamics. Kinematics deals with lengths, times, speeds, etc and it is basically concerned with only the space and time coordinates, and not with masses, forces, energy, momentum, etc. Dynamics, does deal with masses, forces, energy, momentum. The theory is SPECIAL; it only applies to inertial reference frames, those for which the state of motion is not influenced by external forces.

The fundamental effects

The most striking effects of special relativity are:

1. the loss of simultaneity,
2. length contraction, and
3. Time dilation.

In this section we will discuss these three effects using some time-honored concrete examples (which are always nice to fall back on). In the following section we will derive the Galilean and Lorentz transformations using these three results. Relativity is used to solve problems involving large speeds. The theory of Relativity rests upon certain postulates.

1.1 Time Dilation

According to the Theory of Relativity, time dilation is a difference in the elapsed time measured by two events caused by their relative velocity difference or their different positions in a gravitational field. Einstein made one of the most significant contributions to physics by explaining the concept of space-time. However, it is critical to understand and distinguish between the general theory of relativity and the special theory of relativity. In this Physics article, we will learn time dilation formulas with some solved examples.

Definition

Time dilation is a difference in the elapsed time measured by two observers, due to slight differences between clocks on space and those on Earth. Time itself will also bend due to differences in either velocity or gravity. As a result of the nature of spacetime, a clock moving relative to an observer will be measured to tick slower than a clock at rest in the observer's own frame of reference. This effect is known as time dilation. It was first described by Albert Einstein's theory of special relativity.

There are two time dilation effects from the theories of relativity.

- Special Relativity: Moving clocks tick slower.
- General Relativity: Clocks in stronger gravitational fields tick slower

Special Relativity: Moving clocks tick slower.

In Figure 1 are the light clocks used to derive the time dilation formula in class. Focusing on the right triangle, we identified the lengths of the sides and applied the Pythagorean Theorem. This gives

$$\left(\frac{ct}{2}\right)^2 = \left(\frac{vt}{2}\right)^2 + \left(\frac{c\tau}{2}\right)^2 \tag{1.2}$$

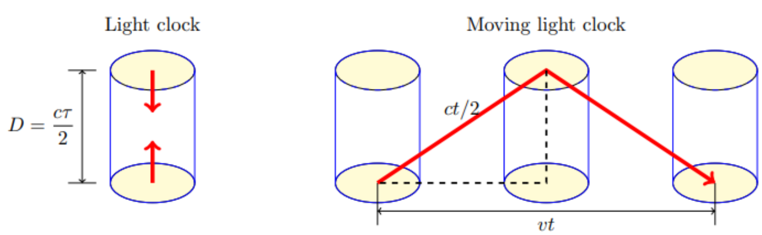


Figure 1: Stationary and moving clocks for time dilation computation.

We now solve for t in terms of τ. After cancelling all of the 2's

$$c^2t^2 = v^2t^2 + c^2\tau^2 \tag{1.3}$$

$$c^2t^2 - v^2t^2 = c^2\tau^2 \tag{1.4}$$

$$(c^2 - v^2)t^2 = c^2\tau^2 \tag{1.5}$$

$$\left(1 - \frac{v^2}{c^2}\right) t^2 = \tau^2 \tag{1.6}$$

$$t = \frac{\tau}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{1.7}$$

This is the time dilation formula. This result states that the time interval as measured by an observer in a stationary frame is longer than the one measured by an observer in a moving frame.

This expression can now be used to compare Alice and Bob's clock readings in Figure 2

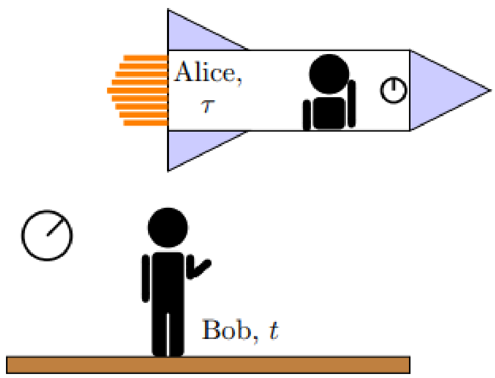


Figure 2: Alice's clock measures the proper time, τ. Bob's clock measures t.

Define

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{1.8}$$

Then, we can write $t = \gamma\tau$. Since $v < c$, then $v/c < 1$. So $\sqrt{1 - \frac{v^2}{c^2}} < 1$ and thus, $\gamma > 1$. This means that $t \geq \tau$. We conclude that moving clocks tick slower.

An example that I can give is that of a plane moving at 620 mph (277 m/s). What is γ? First compute

$$\frac{v}{c} = \frac{277}{3 \times 10^8} = 9.23 \times 10^{-7} \tag{1.9}$$

$$1 - \frac{v^2}{c^2} = 1 - (9.23 \times 10^{-7})^2 \tag{1.10}$$

Putting this in a calculator, you are bound to get an answer of 1 and eventually you will not get to find out by how much the clocks differ.

So, we need an approximation to the difference between the times. Assume that an hour passes on the Earth (t) and we want to know by how much the time (τ) slows down on the plane. Since $\tau < t$, this difference is $t - \tau$. Such an approximation is known (binomial approximation) and we have

$$t-\tau = t - t\sqrt{1-\frac{v^2}{c^2}} \tag{1.11}$$

$$= t - t\left(1-\frac{v^2}{2c^2}\right) \tag{1.12}$$

$$t-\tau = \frac{v^2}{2c^2}t \tag{1.13}$$

So, now the time difference in the problem is found as

$$t-\tau = \frac{v^2}{2c^2}t \approx 4.26\times 10^{-13}t \tag{1.14}$$

For $t = 1hr = 3600$ seconds, $t-\tau \approx 1.53$ nanoseconds.

1.2 Length contraction

One of the peculiar aspects of Einstein's theory of special relativity is that the length of objects moving at relativistic speeds undergoes a contraction along the dimension of motion. An observer at rest (relative to the moving object) would observe the moving object to be shorter in length. The amount of contraction of the object is dependent upon the object's speed relative to the observer.

An object has many lengths, just as it has many velocities, one for each reference frame you use. This contraction (more formally called Lorentz contraction or Lorentz–FitzGerald contraction), for standard objects, is negligible at everyday speeds, and can be ignored for all regular purposes. Only at greater speeds, or for electron motion, does it become significant.

Definition

According to the special theory of relativity, length contraction is the phenomenon that a moving object's length is measured to be shorter than its proper length, which is the length as measured in the object's own rest frame. This contraction is usually only noticeable at a substantial fraction of the speed of light.

Derivation of Length Contraction in the presence of Fields Consider a rod moving down ward with respect to a clock with speed v.
For observer at rest with respect to a rod, this rod is at rest with respect to him. Thus its length is equal to l_o. But the clock is moving with respect to him. Hence the time is given according to equation (22)

$$t = \frac{t_0}{\sqrt{1+\frac{2\phi}{c^2}-\frac{v^2}{c^2}}} \tag{1.15}$$

Thus the speed for him is given by

$$v = \frac{L_0}{t} \tag{1.16}$$

For observer at rest with respect to clock. The clock is at rest with respect to him. Thus the time is t_0. But

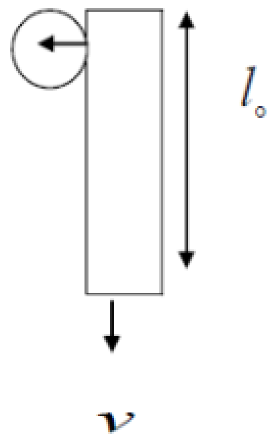


Figure 3: A rod moving down ward

the rod is moving with respect to him .Thus the rod length is l. The speed is given by

$$v = \frac{L}{t_0} \tag{1.17}$$

$$\frac{L}{t} = \frac{L_0}{t} \tag{1.18}$$

Thus the length in the presence of the field is given by

$$L = L_0\left(\frac{t_0}{t}\right) = L_0\sqrt{1 + \frac{2\phi}{c^2} - \frac{v^2}{c^2}} \tag{1.19}$$

$$L = L_0\sqrt{1 + \frac{2\phi}{c^2} - \frac{v^2}{c^2}} \tag{1.20}$$

This expression holds for any fields.

1.3 Relativity of Simultaneity

Einstein began his rethinking of the nature of time by considering the concept of simultaneity and the methods used to synchronize clocks. He realized that the reckoning of simultaneity depended on the velocity of the observer. Einstein's rethinking resulted in the Special Theory of Relativity that states that the velocity dependence of simultaneity is a consequence of the relativity of space and time.

Let us consider the clock paradox from the point of view chosen here (Figure). If twin A remains at rest in the ether system for two years and twin B travels into space and returns with velocity v then B will be aged $T_2 = 2(1-v)^{1/2}$ m upon return. The reason for this is that in the ether theory clocks moving with respect to the ether are slow. Next consider a clock in B moving uniformly with velocity o with respect to the ether as shown in Figure 2b. Clock A is first at rest in the ether system and thus is fast as seen from the moving clock B (this statement does depend on a definition of simultaneity, e.g., on the convention about synchronization). At the time $s_l = T = (1-v^2)^{1/2}$ clock A starts to move uniformly towards clock B with velocity $w = 2v/a(2/a-a)^{-1}$ such as to reach clock B at time t = 2. Using (4.1) one easily finds that the proper time interval (as measured by A) $s_2 = (1-v^2)^{1/2}$ is spent by A on its way towards twin B.

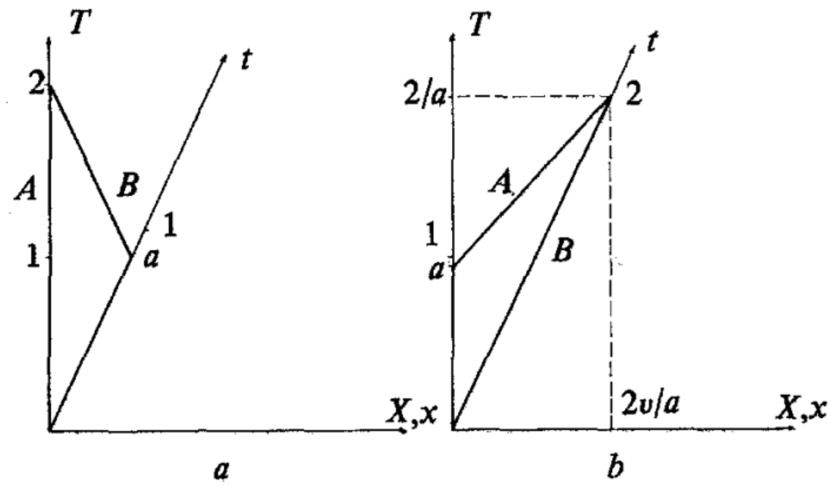


Figure 4: The twin paradox considered from the point of view of ether theory $[a = (1-v^2)^{l/2}]$.

Twin A spends the same (proper) time moving away from B as towards 2, so that the situation is analogous to Figure 2a. Owing to the very fast motion of A on the second part of her trip through space-time, A ages very little on.the second part of the trip and thus arrives at the meeting point with B after a proper time interval $s = 2(1-v^2)^{1/2} < 2$ has elapsed. We thus arrive (not surprisingly) at a result completely equivalent to the standard result of special relativity.

1.4 The postulates of Special Relativity

Two basic ideas are important in the structured formulation of Special Relativity, helping to explain where Newton went wrong and how new thinkers, such as Einstein, Minkowski and Lorentz put the theories to rights. First, we have the idea of simultaneity, implicit in the statement that two clocks at points A and B are said to be *synchronized* if they read the same time at the mid-point of AB. Secondly, there is the concept of an *inertial frame*, defined to be a frame in which particles acting under no forces move with constant velocity

Using ideas from projective geometry, it is fairly easy to prove that transformations between such frames must be linear. More formally: *The time and position coordinates (t, x, y, z) of a particle with respect to a frame of reference F are linearly related to those (t', x', y', z') in another frame F', the frames both being inertial.*

The first postulate holds that the speed of light has the same value in any inertial frame. i.e. the speed of light, $c = 299,792kms^{-1}$, is the same for all inertial observers, independent of their velocity of motion relative to the source of light. The speed of light has the same value in any inertial (that is, non-accelerating) reference frame.

This speed is much greater than the speed of everyday objects, so most of the consequences of this new theory are not noticeable. If we lived in a world similar to ours except for the fact that the speed of light was much less than the speed of light, then the consequences of relativity would be everywhere, and we wouldn't think twice about time dilations, length contractions, and so on.

Illustration

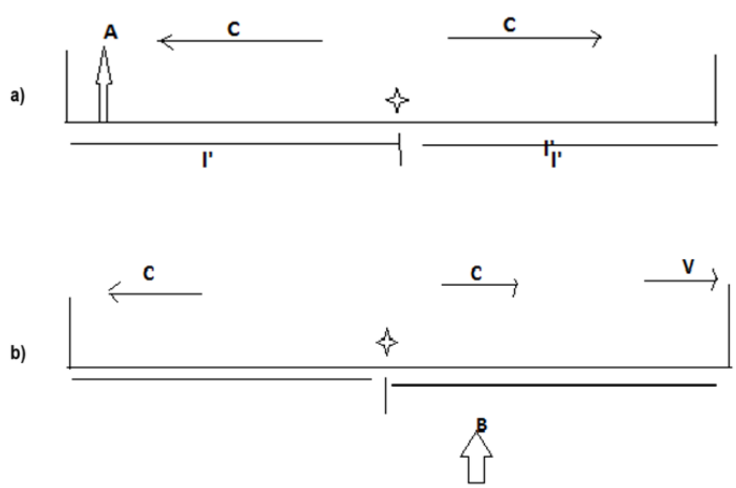
Consider a train moving along the ground at constant velocity (that is, it is not accelerating - inertial frame). Someone on the train shines a light from one point on the train to another. Let the speed of the light with respect to the train be c ($\frac{1}{4} of c$). This postulate holds that a person on the ground also sees the light move at speed c. The truth of this postulate cannot be demonstrated from first principles. No statement with any physical content in physics (that is, one that isn't purely mathematical can be proven.

The second postulate states that, the laws of physics are the same in all inertial frames, or all inertial frames are 'equivalent'. This basically maintains that one inertial frame is no better than any another. There is no preferred reference frame. That is, it makes no sense to say that something is moving; it only makes sense to say that one thing is moving with respect to another.

Suppose two frames S and S_0, then S should see things in S0 in exactly the same way that S_0 sees things in S (because we could simply switch the labels of S and S_0). It also says that empty space is homogeneous, isotropic, that is, that all directions look the same (because we could pick any axis to be, say, the x-axis of a coordinate system).

The principle of relativity essentially states that, the laws of physics take the same mathematical form in all frames of reference moving with constant velocity with respect to one another. Explicitly recognized in this statement is the empirical fact that the laws of nature, almost without exception, can be expressed in the form of mathematical equations. Why this should be so is a profound issue that is not fully understood, but it is nevertheless the case that doing so offers the most concise way of summarizing the observed behavior of a physical system under reproducible experimental conditions.

Consider the following setup. In A's reference frame, a light source is placed midway between two receivers, '0 meters from each other. The light source emits a flash. From A's point of view, the two receivers receive the light at the same time t, seconds after the flash). Another observer, B, is travelling by to the left at speed v.



In B's frame, the receivers (and the whole train) are moving to the right at speed v, and the light is travelling in both directions at speed c (with respect to B, and not with respect to the light source, in B's frame). So the relative speed of the light and the left receiver is $c+v$, and the relative speed of the light and the right receiver is $c–v$ (as viewed by B). Let be the distance from the source to the receivers, as measured by B. Then the light hits the left receiver at t_l and the right receiver at t_r, where

$$t_l = \frac{l}{c+v} \tag{1.21}$$

and

$$t_r = \frac{l}{c-v} \tag{1.22}$$

These are not equal if $v \neq 0$.

The moral of this exercise is that it makes no sense whatsoever to say that one event happens at the same time as another (unless they take place at the same location). Simultaneity depends on the frame in which the observations are made.
Example

Two clocks are positioned at the ends of a train of length L (as measured in its own frame). They are synchronized in the train frame. If the train travels past you at speed v, it turns out that you will observe the rear clock showing a higher reading than the front clock, calculate the time difference in the readings

Solution:

let's put a light source on the train, but now let's position it so that the light hits the clocks at the ends of the train at the same time in your frame, the relative speeds of the photons and the clocks are $c+v$ and $c-v$ (as viewed in your frame). We therefore need to divide the train into lengths in this ratio. The desired lengths are easily seen to be $\frac{L(c-v)}{2c}$ and $\frac{L(c+v)}{2c}$ to the left and right respectively.

Consider now the situation in the frame of the train (see Fig.). The light must travel an extra distance of

$$\frac{L(c+v)}{2c} - \frac{L(c-v)}{2c} = \frac{Lv}{c} \tag{1.23}$$

to reach the rear clock. The extra time is therefore, is $\frac{Lv}{c^2}$.

Hence, the rear clock reads $\frac{Lv}{c^2}$. more when it is hit by the backward photon, compared to what the front clock reads when it is hit by the forward photon. Now, let the instant you look at the clocks be the instant the photons hit them (that's why we chose the hitting to be simultaneous in your frame). Then you observe the rear clock reading more than the front clock by an amount, (difference in readings) $= \frac{Lv}{c^2}$.

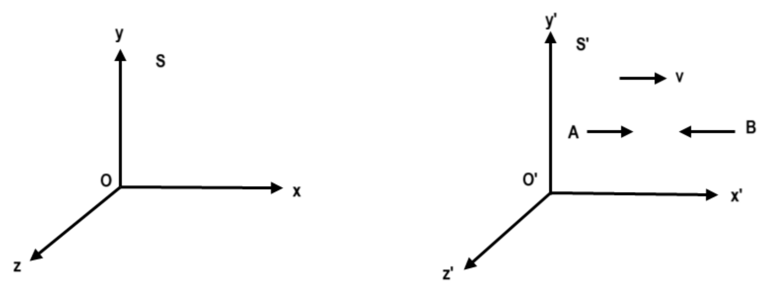
The fact that the rear clock is ahead of the front clock in your frame means that the light hits the rear clock after it hits the front clock in the train frame. Note that our result does not say that you see the rear clock ticking at a faster rate than the front clock. They run at the same rate. (Both have the same time-dilation factor relative to you; the back clock is simply a fixed time ahead of the front clock, as seen by you.

1.5 Variation of mass with velocity

There is variation of mass with velocity in relativity that is mass varies with the velocity when the velocity is comparable with the velocity of the light.

Let there are two inertial frames of references S and S'. S is the stationary frame of reference and S' is the moving frame of reference. At time t=t'=0 that is in the start, they are at the same position that is Observers O and O' coincides. After that S' frame starts moving with a uniform velocity v along x axis.

Suppose there are two particles moving in opposite direction in frame S'. velocity of particle A will be u' and of B will be –u' according to the observer O'. Let us study the velocities and mass of these particles from frame S



Velocity of A is u_1 and B is u_2 from frame S and these are given by relativistic addition of velocity relation respectively

$$u_1 = \frac{(u'+v)}{1+u'v/c^2} \tag{1.24}$$

$$u_2 = \frac{(-u'+v)}{1-u'v/c^2} \tag{1.25}$$

Let m_1 and m_2 are the mass of A and B from frame S respectively. As the particles are moving to each other, at certain instant they will collide and momentarily came to rest. But even when they came to rest, they travel with the velocity of the frame S' that is with v

According to the law of conservation of momentum:

Momentum before collision = momentum after collision

$$m_1u_1 + m_2u_2 = (m_1 + m_2)v = m_1v + m_2v \tag{1.26}$$

or

$$m_1(u_1 - v) = m_2(u_2 - v) \tag{1.27}$$

Put equations (1.3) and (1.5) in above equations, we get

$$m_1\left(\frac{(u'+v)}{(1+u'v/c^2)} - v\right) = m_2\left(v - \frac{(-u'+v)}{1-u'v/c^2}\right) \tag{1.28}$$

Then take LCM of terms in the bracket and solve, we get

$$m_1\left(\frac{1}{(1+u'v/c^2)}\right)=m_2\left(\frac{1}{(1-u'v/c^2)}\right) \tag{1.29}$$

or

$$\frac{m_1}{m_2}=\frac{(1+u'v/c^2)}{(1-u'v/c^2)} \tag{1.30}$$

Now square equation (1.4), then divide both sides by c_2 and subtract both sides by 1, we get

$$1-\frac{u_1^2}{c^2}=1-\left[\frac{(u'+v)/c}{(1+u'v/c^2)}\right]^2 \tag{1.31}$$

By taking LCM on RHS and solving, we get

$$1-\frac{u_1^2}{c^2}=\frac{(1+u'^2v^2/c^2-u'^2/c^2-v^2/c^2)}{(1+u'v/c^2)^2} \tag{1.32}$$

Similarly by squaring equation (1.5), then dividing both sides by c_2 and subtracting both sides by 1, we get

$$1-\frac{u_2^2}{c^2}=\frac{(1+u'^2v^2/c^4-u'^2/c^2-v^2/c^2)}{(1-u'v/c^2)^2} \tag{1.33}$$

On dividing equation (1.12) by (1.13), we get

$$\frac{(1-u_2^2/c^2)}{(1-u_1^2/c^2)}=\frac{(1+u'v/c^2)^2}{(1-u'v/c^2)^2} \tag{1.34}$$

Take square root on both sides

$$\frac{(1-u_2^2/c^2)^{1/2}}{(1-u_1^2/c^2)^{1/2}}=\frac{(1+u'v/c^2)}{(1-u'v/c^2)} \tag{1.35}$$

Now compare equations (1.14) and (1.15), we get

$$\frac{m_1}{m_2}=\frac{(1-u_2^2/c^2)^{1/2}}{(1-u_1^2/c^2)^{1/2}} \tag{1.36}$$

This is more of a complicated result. To make this result simple, let us assume that the particle B is in the state of rest from frame S that is it has zero velocity before collision. Thus $u_2=0$. And $m_2=m_0$. Where m_0 is the rest mass of the particle, Therefore equation (1.16) becomes

$$\frac{m_1}{m_0}=\frac{1}{(1-u_1^2/c^2)^{1/2}} \tag{1.37}$$

Also assume $u_1=v$ and $m_1=m$. Therefore above equation becomes

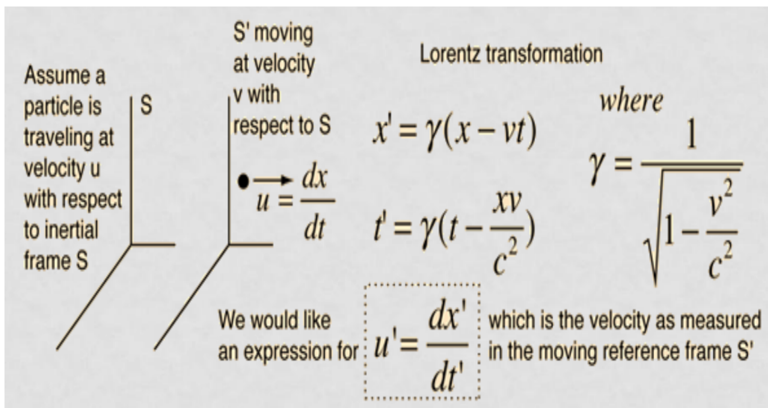
$$\frac{m}{m_0}=\frac{1}{(1-v^2/c^2)^{1/2}} \tag{1.38}$$

This equation represents the equation of the variation of mass with the velocity

1.6 Relativistic Velocity addition

In relativistic physics, a velocity-addition formula is an equation that relates the velocities of objects in different reference frames. It is 3-dimensional in nature. It also relates different frames, that is, the formulas applies to successive Lorentz transformation. Accompanying velocity addition is a kinematic effect called the Thomas precession. Successive non-collinear Lorentz boosts affects a rotation to the coordinate system

No two objects can have a relative velocity greater than c! But what if I observe a spacecraft travelling at 0.8c and it fires a projectile which it observes to be moving at 0.7c with respect to it!? Velocities must transform according to the Lorentz transformation, and that leads to a very nonintuitive result called Einstein velocity addition.



Just taking the differentials of these quantities leads to the velocity transformation. Taking the differentials of the Lorentz transformation expressions for x' and t' above gives

$$\frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma\left(dt - \frac{vdx}{c^2}\right)} = \frac{\frac{dx}{dt} - v}{1 - \frac{v\frac{dx}{dt}}{c^2}} \tag{1.39}$$

Putting this in the notation introduced in the illustration above:

$$u' = \frac{u - v}{1 - \frac{uv}{c^2}} \tag{1.40}$$

The reverse transformation is obtained by just solving for u in the above expression. Doing that gives

$$u = \frac{u' + v}{1 + \frac{u'v}{c^2}} \tag{1.41}$$

The formulas for boosts in the standard configuration follow most straight forwardly from taking differentials of the inverse Lorentz boost in standard configuration,

$$dx = \gamma(dx' + vdt'), \qquad dy = dy', \qquad dz = dz', \qquad dt = \gamma\left(dt' + \frac{V}{c^2}dx'\right) \tag{1.42}$$

Divide the first three equations by the fourth,

$$\frac{dx}{dt} = \frac{\gamma(dx' + vdt')}{\gamma\left(dt' + \frac{V}{c^2}dx'\right)}, \qquad \frac{dy}{dt} = \frac{dy'}{\gamma\left(dt' + \frac{V}{c^2}dx'\right)}, \qquad \frac{dz}{dt} = \frac{dz'}{\gamma\left(dt' + \frac{V}{c^2}dx'\right)} \tag{1.43}$$

or

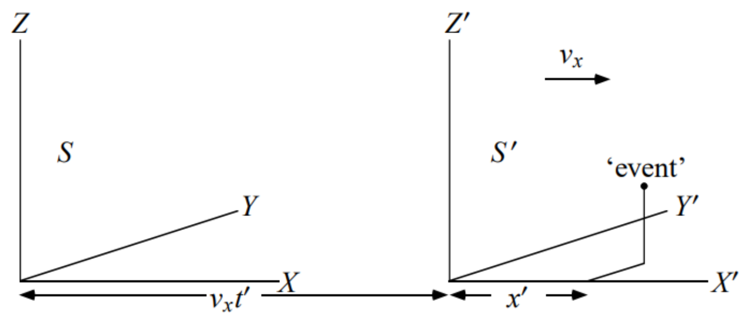
$$\frac{dx}{dt}=\frac{(dx'+vdt')}{dt'\left(1+\frac{V}{c^2}dx'\right)},\qquad \frac{dy}{dt}=\frac{dy'}{\gamma dt'\left(1+\frac{V}{c^2}dx'\right)},\qquad \frac{dz}{dt}=\frac{dz'}{\gamma dt'\left(1+\frac{V}{c^2}dx'\right)} \tag{1.44}$$

which is

$$v_x=\frac{v'_x+V}{1+\frac{V}{c^2}v'_x},\qquad v_y=\frac{\sqrt{1-\frac{V^2}{c^2}}v'_y}{1+\frac{V}{c^2}v'_x},\qquad v_z=\frac{\sqrt{1-\frac{V^2}{c^2}}v'_Z}{1+\frac{V}{c^2}v'_x} \tag{1.45}$$

1.7 Events

Colloquially, an event is something that occurs at a localized region in space over a localized interval in time, or, in an idealized limit, at a point in space at an instant in time. Thus, the motion of a particle through space could be thought of as a continuous series of events, while the collision of two particles would be an isolated event, and so on. However, it is useful to release this term 'event' from being associated with something happening.



After all, the the coordinate network spread throughout space, and the clocks ticking away the hours will still be labelling points in space, along with 'the time' at each point in space, irrespective of whether or not anything actually takes place at a particular locality and at a particular time. The idea then is to use the term 'event' simply as another name for a point in space and time, this point specified by the spatial coordinates of the point in space, and the reading of a clock at that point.

An event will have different coordinates in different reference frames. It is then important and useful to be able to relate the coordinates of events in one reference frame to the coordinates of the same event in some other reference frame. In Newtonian physics, this relation is provided by the Galilean transformation equations, and in special relativity by the Lorentz transformation. It is the latter transformation law, and special relativity in particular that we will be concerning ourselves with from now on.

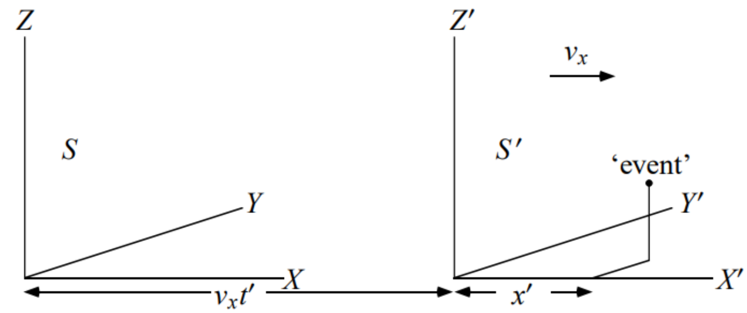
2 The Galilean Transformation

2.1 The Galilean System of Co-ordinates

The fundamental law of the mechanics of Galilei-Newton, the law of inertia, states that, a body removed sufficiently far from other bodies continues in a state of rest or of uniform motion in a straight line. This law not only says something about the motion of the bodies, but it also indicates the reference-bodies or systems of coordinates, permissible in mechanics, which can be used in mechanical description.

The visible fixed stars are bodies for which the law of inertia certainly holds to a high degree of approximation. Now if we use a system of co-ordinates which is rigidly attached to the earth, then, relative to this system, every fixed star describes a circle of immense radius in the course of an astronomical day, a result which is opposed to the statement of the law of inertia. A system of co-ordinates of which the state of motion is such that the law of inertia holds relative to it is called a **"Galileian system of co-ordinates"**. The laws of the mechanics of Galilein-Newton can be regarded as valid only for a Galileian system of co-ordinates

To derive these transformation equations, consider an inertial frame of reference S and a second reference frame S_0 moving with a velocity v_x relative to S.



Let us suppose that the clocks in S and S' are set such that when the origins of the two reference frames O and O' coincide, all the clocks in both frames of reference read zero i.e. $t = t' = 0$. According to "common sense", if the clocks in S and S_0 are synchronized at $t = t' = 0$, then they will always read the same, i.e. $t = t'$ always. Suppose now that an event of some kind, e.g. an explosion, occurs at a point (x', y', z', t_0) according to S'. Then, by examining the Fig. above, according to S, it occurs at the point

$$x = x' + v_x t' \tag{2.1}$$

$$y = y' \tag{2.2}$$

$$z = z' \tag{2.3}$$

$$t = t' \tag{2.4}$$

These equations together are known as the Galilean Transformation,
Implication

The equations tell us how the coordinates of an event in one inertial frame S are related to the coordinates of the same event as measured in another frame S_0 moving with a constant velocity relative to S. Now suppose that in inertial frame S, a particle is acted on by no forces and hence is moving along the straight line path given by:

$$r = r_o + ut \tag{2.5}$$

where u is the velocity of the particle as measured in S. Then in S_0, a frame of reference moving with a velocity relative to S, the particle will be following a path

$$r' = r_o + (u - v)t' \tag{2.6}$$

where we have simply substituted for the components of r using above. And since the particle is being acted on by no forces, S' is also an inertial frame, and since v is arbitrary, there is in general an infinite number of such frames. Incidentally, if we take the derivative of Eq. (3.3) with respect to t, and use the fact that $t = t_0$, we obtain

$$u' = u - v \tag{2.7}$$

which is the familiar addition law for relative velocities.

2.2 Galilean Principle of Relativity

Galileo, who is regarded to be the first true natural scientist, made a simple observation which turned out to have far reaching consequences for modern physics.

Galileo noticed that it was difficult to tell whether a ship was anchored or coasting at sea by means of mechanical experiments carried out on board of this ship.

Suppose a boat is moving with uniform velocity along a canal and we are looking at it. and the crow's nest drops a heavy weight onto the deck. It will hit the captain below because it passes to a frame of reference moving with the boat. The frame at rest with respect to the canal is an inertial frame of reference.

Galileo stated that, **the laws of dynamics are the same in all frames of reference which are in uniform motion with respect to an inertial frame of reference**. Now if we drop something from rest in frame S it will fall vertically down. Therefore, if we drop something from rest in frame it will fall vertically down. The boxed statement is Galilean Relativity follows in Isaac Newton's account of dynamics because in frame S ,

$$m\frac{d^2x}{dt^2} = F(x,t) \tag{2.8}$$

But to transform $\bar{S}$ we set,

$$\bar{x} = x - ut \tag{2.9}$$

And hence,

$$m\frac{d^2\bar{x}}{dt^2} = F(\bar{x}, ut) \tag{2.10}$$

Note that Galileo assumed that the passengers in the boat would use the same coordinate t. In principle one might have thought that one would also have to change the time coordinate to a new coordinate $\bar{t}$ for this equivalence to workout but both Galileo and Newton agreed that is an absolute coordinate that is, it takes the same value in all inertial frames of reference

$$\bar{t} = t \tag{2.11}$$

Equations 2.9, 2.11 constitute a Galilean Transformation

$$\bar{x} = x - ut \qquad\qquad \bar{t} = t \tag{2.12}$$

Relativity (both the Special and General theories), quantum mechanics, and thermodynamics are the three major theories on which modern physics is based. What is unique about these three theories, as distinct from say the theory of electromagnetism, is their generality. Embodied in these theories are general principles which all more specialized or more specific theories are required to satisfy.

What the principle of relativity essentially states is the following:

The laws of physics take the same mathematical form in all frames of reference moving with constant velocity with respect to one another.

Using these transformation equations, the mathematical statement of any physical law according to one observer can be translated into the law as written down by another observer. The principle of relativity then requires that the transformed equations have exactly the same form in all frames of reference moving with constant velocity with respect to one another, in other words that the physical laws are the same in all such frames of reference.

This statement contains concepts such as 'mathematical form' and 'frame of reference' and 'Galilean transformation' which we have not developed, so perhaps it is best at this stage to illustrate its content by a couple of examples. In doing so it is best to make use of an equivalent statement of the principle, that is:

Given two observers A and B moving at a constant velocity with respect to one another, it is not possible by any experiment whatsoever to determine which of the observers is 'at rest' or which is 'in motion'
Example

Suppose Tony is standing by the railway tracks, watching a train go past to the east at 25 m/s. At the same time, a plane is flying overhead (again eastwards) at 200 m/s. Meanwhile, a car drives away to the north at 25 m/s. What does the scene look like to Bill, who is sitting on the train?

We have to remember here that velocity is a vector. In order to transform from Tony's frame of reference to Bill's, we will have to use a vector version of (4):

$$v' = v - u \tag{2.13}$$

As above, u is the velocity of Bill relative to Tony. Let's call Eastwards the i direction and northwards the j direction, then,

$$u = 25i \tag{2.14}$$

The velocity of the plane is, according to Tony, $v = 200i$, the velocity of the plane relative to Bill is

$$v' = v - u = 200i - 25i = 175i \tag{2.15}$$

As seen from the train, then, the plane is flying past east wards at a speed of 175 m/s. Tony calculates the velocity of the car to be $v = 25j$. Therefore, from Bill's frame of reference, the car is moving with velocity

$$v' = 25j - 25i \tag{2.16}$$

Thus, the car is moving in a northwesterly direction relative to the train.
Example II

Tony is playing snooker. The white ball, which has mass m and moves with velocity $v = 13icm/s$, hits a stationary red ball, also of mass m, in an elastic collision. The white ball leaves the collision with velocity

$$v_w = 11.1i + 4.6j \tag{2.17}$$

(i.e. at 12 cm/s at an angle of 22.6^o above the horizontal), and the red ball leaves at a velocity of

$$v_r = 1.9i - 4.6j \tag{2.18}$$

(which is 5 cm/s at an angle of 67.4^o below the horizontal). You can check that momentum and energy are conserved. Suppose now that Bill is walking past with a velocity of

$$u = 13i \tag{2.19}$$

What does the collision look like to him? We know straight away that, since he is moving with the same speed as the white ball had initially, it is at rest in his reference frame; this of course agrees with equation (2.6. What about the red ball before the collision? In his frame, it is no longer at rest; instead, it is moving 'backwards', with velocity

$$0 - 13i = -13i \tag{2.20}$$

After the collision, we obtain for the white ball

$$v_w' = (11.1i + 4.6j) - 13i = -1.9i + 4.6j \tag{2.21}$$

where as the red ball moves with velocity

$$v_r' = (1.9i - 4.6j) - 13i = -11i - 4.6j \tag{2.22}$$

From Bill's point of view, then, the collision is essentially a mirror image of the collision as seen from Tony's reference frame. For Bill, it is the red ball that moves in and hits the stationary white ball. As expected, momentum and energy are conserved in the two frames.

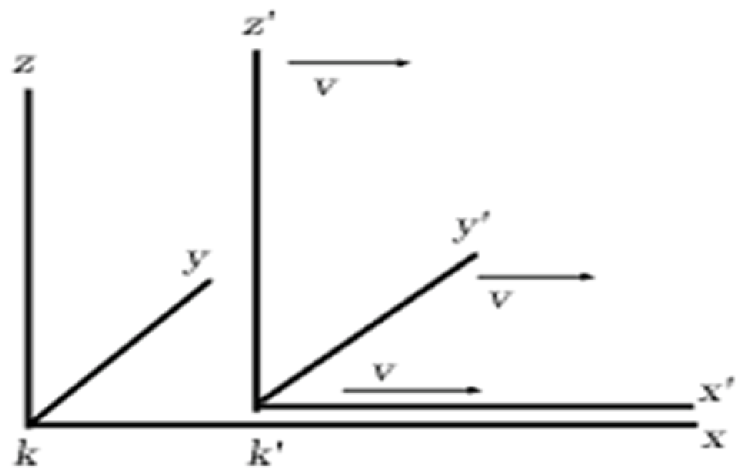
2.3 The Principle of Relativity (in the restricted sense)

Let us imagine a raven flying through the air in such a manner that its motion, as observed from the embankment, is uniform and in a straight line. Consider a mass m is moving uniformly in a straight line with respect to a co-ordinate system K, then it will also be moving uniformly and in a straight line relative to a second co-ordinate system k' provided that the latter is executing a uniform translatory motion with respect to K, it follows that: If K is a Galileian co-ordinate system, then every other co- ordinate system k' is a Galileian one, when, in relation to K, it is in a condition of uniform motion of translation. Relative to k' the mechanical laws of Galilei-Newton hold good exactly as they do with respect to K.

If, relative to K, k' is a uniformly moving co-ordinate system devoid of rotation, then natural phenomena run their course with respect to k' according to exactly the same general laws as with

respect to K. This statement is called the **Principle of Relativity** (in the restricted sense).

3 The Lorentz Transformation



We can imagine the train travelling with the velocity v to be continued across the whole of space, so that every event, no matter how far off it may be, could also be localized with respect to the second framework. Without committing any fundamental error, we can disregard the fact that in reality these frameworks would continually interfere with each other, owing to the impenetrability of solid bodies. In every such framework we imagine three surfaces perpendicular to each other marked out, and designated as "co-ordinate planes" ("co-ordinate system").

A co-ordinate system K then corresponds to the embankment, and a co-ordinate system k' to the train. An event, wherever it may have taken place, would be fixed in space with respect to K by the three perpendiculars x, y, z on the co-ordinate planes, and with regard to time by a time value t. Relative to k', the *same event* would be fixed in respect of space and time by corresponding values x', y', z', which of course are not identical with x, y, z, t.

The relations must be so chosen that the law of the transmission of light in vacuo is satisfied for one and the same ray of light (and of course for every ray) with respect to K and k'. For the relative orientation in space of the co- ordinate systems indicated in the diagram, this problem is solved by means of the equations:

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3.1}$$

$$y' = y \tag{3.2}$$

$$z' = z \tag{3.3}$$

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3.4}$$

This system of equations is known as the **"Lorentz transformation"**.

If in place of the law of transmission of light we had taken as our basis the implied assumptions of the older mechanics as to the absolute character of times and lengths, then instead of the above we should have obtained the following equations:

$$\bar{x} = x - vt \tag{3.5}$$
$$\bar{t} = t \tag{3.6}$$
$$\bar{y} = y \tag{3.7}$$
$$\bar{z} = z \tag{3.8}$$

This system of equations is often termed the "Galilei transformation." The Galilei transformation can be obtained from the Lorentz transformation by substituting an infinitely large value for the velocity of light c in the latter transformation.

Aided by the following illustration, we can readily see that, in accordance with the Lorentz transformation, the law of the transmission of light in vacuo is satisfied both for the reference-body K and for the reference-body k'. Alight-signal is sent along the positive x- axis, and this light-stimulus advances in accordance with the equation

$$x = ct \tag{3.9}$$

i.e. with the velocity c. According to the equations of the Lorentz transformation, this simple relation between x and t involves a relation between x' and t'.

3.1 Deriving Lorentz transformation

In deriving this transformation, we will eventually make use of the constancy of the speed of light, but first we will derive the general form that the transformation law must take purely from kinematic/symmetry considerations. Doing so is based on two further assumptions which seem to be entirely reasonable:

1. Homogeneity: The intrinsic properties of empty space are the same everywhere and for all time. In other words, the properties of the rulers and clocks do not depend on their positions in (empty) space, nor do they vary over time.

2. Spatial Isotropy: The intrinsic properties of space is the same in all directions. In other words, the properties of the rulers and clocks do not depend on their orientations in empty space. There is a third, much more subtle condition:

3. No Memory: The extrinsic properties of the rulers and clocks may be functions of their current states of motion, but not of their states of motion at any other time

The starting point is to consider two inertial frames S and S' where S' is moving with a velocity v_x relative to S.

Let us suppose that when the two origins coincide, the times on the clocks in each frame of reference are set to read zero, that is $t = t' = 0$. Now consider an event that occurs at the point (x, y, z, t) as measured in S . The same event occurs at (x', y', z', t') in S'. What we are after is a set of equations that relate these two sets of coordinates.

We are going to assume a number of things about the form of these equations, all of which can be fully justified, but which we will introduce more or less on the basis that they seem intuitively reasonable.

First, because the relative motion of the two reference frames is in the X direction, it is reasonable to expect that all distances measured at right angles to the X direction will be the same in both S and S', i.e.

$$y = y' \qquad and \qquad z = z' \tag{3.10}$$

We now assume that (x, t) and (x', t') are related by the linear transformations

$$x' = Ax + Bt \tag{3.11}$$

$$t' = Cx + Dt \tag{3.12}$$

Why linear? Assuming that space and time is homogeneous tells us that a linear relation is the only possibility. What it amounts to saying is that it should not matter where in space we choose our origin of the spatial coordinates to be, not should it matter when we choose the origin of time, i.e. the time that we choose to set as t = 0. Now consider the origin O' of S'. This point is at x' = 0 which, if substituted into Eq. (4.3) gives

$$Ax + Bt = 0 \tag{3.13}$$

where x and t are the coordinates of O' as measured in S , i.e. at time t the origin O' has the X coordinate x, where x and t are related by $Ax + Bt = 0$. This can be written

$$\frac{x}{t} = -\frac{B}{A} \tag{3.14}$$

but x/t is just the velocity of the origin O' as measured in S . This origin will be moving at the same speed as the whole reference frame, so then we have

$$-\frac{B}{A} = v_x \tag{3.15}$$

which gives $B = -v_x A$ which can be substituted into Eq. (4.3) to give

$$x' = A(x - v_x t) \tag{3.16}$$

If we now solve Eq. (4.3) and Eq. (4.4) for x and t we get

$$x = \frac{Dx' + v_x At'}{AD - BC} \tag{3.17}$$

$$t = \frac{At' - Cx'}{AD - BC} \tag{3.18}$$

If we now consider the origin O of the reference frame S , that is, the point x = 0, and apply the same argument as just used above, and noting that O will be moving with a velocity $-v_x$ with respect to S', we get

$$-\frac{v_x A}{D} = -v_x \tag{3.19}$$

which then gives

$$A = D \tag{3.20}$$

and hence the transformations Eq. (4.9) and Eq. (4.10) from S' to S will be, after substituting for D and B:

$$x = \frac{(x' + v_x t')}{A + v_x C} \tag{3.21}$$

$$t = \frac{(t' - (C/A)x')}{A + v_x C} \tag{3.22}$$

which we can compare with the original transformation from S to S'

$$x' = A(x - v_x t) \tag{3.23}$$

$$t' = A(t + (C/A)x) \tag{3.24}$$

At this point we will introduce a notation closer to the conventional notation i.e. we will now write

$$A = \gamma \qquad and \qquad C/A = K \tag{3.25}$$

so that the sets of equations above become

$$x = \frac{(x' + v_x t')}{\gamma(1 + v_x K)} \tag{3.26}$$

$$t = \frac{(t' - Kx')}{\gamma(1 + v_x K)} \tag{3.27}$$

and

$$x' = \gamma(x - v_x t) \tag{3.28}$$

$$t' = \gamma(t + Kx) \tag{3.29}$$

We now want to make use of some of the symmetry properties listed above to learn more about γ and K. In doing this, it should be understood that the quantities γ and K are not constants. While it is true that they do not depend on x or t, they still potentially depend on v_x. However, the assumed isotropy of space means that γ cannot depend on the sign of v_x. If we write $\gamma = \gamma(v_x)$ and $\gamma' = \gamma(-v_x)$, (with a similar meaning for K and K'), this means that

$$\gamma = \gamma' \tag{3.30}$$

A symmetry property we have already used is that if S' is moving with a velocity vx relative to S, then S must be moving with velocity $-v_x$ relative to S'. We now make use of this fact to reverse the transformation equations Eq. (4.17) to express x and t in terms of x' and t'. We do this by making the substitutions $v_x \leftarrow -v_x$, $x \approx x'$, and $t \approx t'$ which leads to

$$
\begin{aligned}
x &= \gamma(x' + v_x t') && (3.31)\\
t &= \gamma(t' + K'x') && (3.32)
\end{aligned}
$$

By comparison with Eq. (4.16) we have

$$\gamma = \frac{1}{\gamma(1 + v_x K)} \tag{3.33}$$

$$\frac{-K}{\gamma(1 + v_x K)} = \gamma K' \tag{3.34}$$

which tells us that

$$\gamma^2 = \frac{1}{1 + v_x K} \qquad \textit{and} \qquad K = -K' \tag{3.35}$$

The second of these two equations tells us that we can write K as

$$K = -v_x/V^2 \tag{3.36}$$

where V^2 will not depend on the sign of v_x though it could still depend on v_x. We are motivated to write K in this way because by doing so the quantity V will have the units of velocity, which will prove to be convenient later. There is nothing physical implied by doing this, it is merely a mathematical convenience. Thus we have

$$\gamma = \frac{1}{\sqrt{1 - (v_x/V)^2}} \tag{3.37}$$

The transformation laws now take the form

$$
\begin{aligned}
x' &= \frac{x - v_x t}{\sqrt{1 - (v_x/V)^2}} && (3.38)\\
t' &= \frac{t - (v_x/V^2)x}{\sqrt{1 - (v_x/V)^2}} && (3.39)
\end{aligned}
$$

To determine the dependence of V on v_x, we will suppose there is a further reference frame S'' moving with a velocity $\bar{v_x}$ relative to S'. The same argument as used above can be applied once again to give

$$
\begin{aligned}
x'' &= \frac{x' - \bar{v_x} t'}{\sqrt{1 - (\bar{v_x}/\bar{V})^2}} && (3.40)\\
t'' &= \frac{t' - (\bar{v}/\bar{V}_x^2)x'}{\sqrt{1 - (\bar{v_x}/\bar{V})^2}} && (3.41)
\end{aligned}
$$

where we have introduced a new parameter $\bar{V}$. If we now substitute for x', y', z' and t' in terms of $x, y, z,$ and t from Eq. (4.24) and rearrange the terms we get

$$x'' = \frac{1 + \overline{v_x} v_x/V^2}{\sqrt{[1 - v_x^2/V^2][1 - \overline{v_x^2/\overline{V}^2}]}} \left[x - \frac{v_x + \overline{v_x}}{1 + v_x \overline{v_x}/V^2} t \right] \tag{3.42}$$

$$t'' = \frac{1+\overline{v_x}v_x/\overline{V}^2}{\sqrt{[1-v_x^2/V^2][1-\overline{v_x^2}/\overline{V^2}]}}\left[t - \frac{v_x/V^2+\overline{v_x}/\overline{V}^2}{1+\overline{v_x}v_x/V^2}x\right] \tag{3.43}$$

This is now the transformation law relating the coordinates of events in S to their coordinates in S''.

The Galilean transformation

$$t = \bar{t} \qquad \bar{t} = t \tag{3.44}$$
$$x = \bar{x}v\bar{t} \qquad \bar{x} = x - vt \tag{3.45}$$
$$y = \bar{y} \qquad \bar{y} = y \tag{3.46}$$
$$z = \bar{z} \qquad \bar{z} = z \tag{3.47}$$

Is inconsistent with the second postulate of Special Relativity. The new transformation should have the form;

$$t = f(\bar{t}, \bar{x}, v) \qquad \bar{t} = f(t, x - v) \tag{3.48}$$
$$\bar{x} = g(\bar{t}, \bar{x}, v) \qquad \bar{x} = g(t, x - vt \tag{3.49}$$
$$y = \bar{y} \qquad \bar{y} = y \tag{3.50}$$
$$z = \bar{z} \qquad \bar{z} = z \tag{3.51}$$

just because of the symmetry between the two frames. Indeed, the only difference between the frames S and $\bar{S}$ is the direction of relative motion: If $\bar{S}$ moves with speed v relative to S then s moves with speed –v relative to $\bar{S}$.

Here comes the change in the sign of v in the equations of direct and inverse transformations (2.7). y and z coordinates are invariant because lengths normal to the direction of motion are unchanged. Now we need to find functions f and g. (2.6) in a linear transformation. Assume that (2.7) is linear as well (if our derivation fails through we will come back and try something less restrictive). Then

$$x = \gamma(v)\bar{x} + \delta(v)\bar{t} + \eta(v) \tag{3.52}$$

Clearly we should have

$$\bar{x} = -v\bar{t} \tag{3.53}$$

for any $\bar{t}$, if $x = 0$, thus

$$(-v\gamma + \delta)t + \eta(v) = 0 \tag{3.54}$$

for any $\bar{t}$, $\eta = 0$, and $\delta = v\gamma$, thus;

$$\bar{x} = \gamma(x - vt) \quad (3.55)$$
$$x = (\bar{x} + v\bar{t}) \quad (3.56)$$

In principle, the symmetry of direct and inverse transformation is preserved both if

$$\gamma(-v) = \gamma(v) \quad (3.57)$$

and

$$\gamma(-v) = -\gamma(v) \quad (3.58)$$

However, it is clear that for $\bar{x} \to +\infty$, we should have $x \to +\infty$ as well. This condition selects $\gamma(-v) = \gamma(v) > 0$. Now to the main condition that allows us to fully determine the transformation however, suppose that a light signal is fired at time $t = \bar{t} = 0$ in the positive direction of the x axis. (Note that we can always ensure that the standard time-keeping clocks located at the origins of S and S' show the same time when the origins coincide). Eventually, the signal location will be $\bar{x} = c\bar{t}$ and , Substitute these into eq. (2.10, 2.9) to find

$$ct = \gamma c\bar{t}(1 + v/c) \quad (3.59)$$
$$c\bar{t} = \gamma ct(1 - v/c) \quad (3.60)$$

Now we substitute from eqn 2.12 into 2.11 and derive

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (3.61)$$

We immediately recognize the Lorentz factor. In order to find function γ of Eq.2.7 we simply substitute x from Eq.2.9 into Eq.2.10 and then express t as a function of $\bar{t}$ and $\bar{x}$

$$\bar{x} = \gamma(\gamma\bar{x} + v\gamma\bar{t}) - vt \quad (3.62)$$