Chapter 1

The Special Theory of relativity

1.1 Pre - relativistic physics

The starting point for our work are Newtons laws of motion. These can be stated as follows:

- **Free particles move with constant velocity.**
- **The vector force $\mathbf{F}$ is proportional to the rate of change of momentum i.e. $\mathbf{F} = \frac{d}{dt}(m\mathbf{v})$.**
- **To every action there is an equal and opposite reaction.**

The first of these laws singles out inertial frames as the non - accelerating ones.

Consider now a frame $\mathcal{O}$ [i.e. a set of spatial coordinates (x, y, z) and a time coordinate t] and another frame $\bar{\mathcal{O}}$ with coordinates $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ which moves in the x direction with uniform speed v relative to the frame $\mathcal{O}$.

Common sense suggests that the two sets of coordinates are related by

$$\begin{aligned} \bar{x} &= x - vt \, , \\ \bar{y} &= y \, , \\ \bar{z} &= z \, , \\ \bar{t} &= t \, . \end{aligned} \tag{1.1}$$

These are the Galelian transformations .

Figure 1.1: I do not know what I may appear to the world; but to myself I seem to have been only like a boy playing on the sea - shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me. *Isaac Newton*

If the particle has a velocity $\mathbf{u}$ with components (u_1, u_2, u_3) in $\mathcal{O}$, its velocity in $\bar{\mathcal{O}}$ is:

$$\begin{aligned} \bar{u}_1 &= u_1 - v \ , \\ \bar{u}_2 &= u_2 \ , \\ \bar{u}_3 &= u_3 \ , \end{aligned} \tag{1.2}$$

or

$$\bar{\mathbf{u}} = \mathbf{u} - \mathbf{v} \ , \tag{1.3}$$

where

$$u_1 = \frac{dx_1}{dt} \ , \quad u_2 = \frac{dx_2}{dt} \ , \quad u_3 = \frac{dx_3}{dt} \ . \tag{1.4}$$

More generally if the coordinate axes and the origins of $\mathcal{O}$ and $\bar{\mathcal{O}}$ differ then:

$$\bar{\mathbf{X}} = \mathcal{R}\mathbf{X} - \mathbf{v}t - \mathbf{d} \ , \tag{1.5}$$

where $\mathbf{X}$ has components (x, y, z). Here $\mathcal{R}$ is a rotation matrix aligning $\mathcal{O}$ and $\bar{\mathcal{O}}$, $\mathbf{v}$ is the relative velocity of $\mathcal{O}$ with respect to $\bar{\mathcal{O}}$ and $\mathbf{d}$ is the displacement of the origin from $\bar{\mathcal{O}}$.

Since the transformation is linear [constant velocity in $\mathcal{O} \Rightarrow$ constant velocity in $\bar{\mathcal{O}}$], $\bar{\mathcal{O}}$ is inertial if $\mathcal{O}$ is.

Thus there are an infinite set of inertial frames, all moving uniformly with respect to each other.

All of Newtons laws apply in any inertial frame since

$$\bar{\mathbf{a}} = \frac{d\mathbf{u}}{dt} = \frac{d\bar{\mathbf{u}}}{dt} \ , \tag{1.6}$$

and $\mathbf{F}$ is invariant. Thus we have Newtonian [Galelian] Relativity.

The laws of mechanics do not allow measurement of absolute velocity, however one can measure absolute acceleration.

Newton explained inertial frames in terms of absolute space identified with the center of mass of the solar system or a frame of "fixed stars". However this is unsatisfactory because:

- **There is no unique identification [many inertial frames].**
- **Philosophically unappealing since absolute space affects everything but is affected by nothing.**

1.2 The equations of electromagnetism

Maxwell's equations for the electromagnetic field [in units with $\epsilon_0 = \mu_0 = c = 1$] are:

$$\nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = 4\pi \mathbf{j} \quad , \qquad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \ ,$$
$$\nabla \cdot \mathbf{E} = 4\pi\rho \quad , \qquad \nabla \cdot \mathbf{B} = 0 \ . \tag{1.7}$$

These equations show that one can have electromagnetic waves in a vacuum which travel with a speed $c = 3 \times 10^8 \ ms^{-1}$, Thus light is a form of electromagnetic radiation. An Ether was postulated as the medium which carries these waves [i.e. the speed c is the speed with respect to the ether] and naturally this was identified with the absolute space of Newton.

Attempts to measure the ether drift by Michelson and Morley in 1887 gave a null result, so effects like stellar aberration could not be due to the earth dragging the ether.

Another great problem was that Maxwell's equations did not appear to obey the principle of Galelian Relativity i.e. they were not invariant under the Galelian

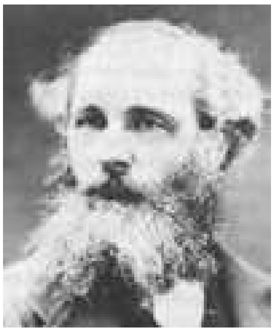

Figure 1.2: James Clerk Maxwell

transformations. This means that in a moving space ship the electric and optical phenomena should be different from those in a stationary ship!!! Thus one could use for example optical phenomena to determine the speed of the ship.

One of the consequences of Maxwell's equations is that if there is a disturbance in the field such that light is generated, these electromagnetic waves go out in all directions equally and at the same speed c. Another consequence of the equations is that if the source of the disturbance is moving, the light emitted goes through space at the same speed c. This is analogous to the case of sound being likewise independent of the motion of the source. This independence of the motion of the source in the case of light brings up an interesting problem.

Suppose we are riding in a car that is going at a speed u, and light from the rear is going past the car with speed c. According to the Galelian transformations, the apparent speed of the passing light, as we measure it in the car, should not be c, but $c - u$. This means that by measuring the speed of the light going past the car, one could determine the absolute speed of the car. This is crazy because we already know that the laws of mechanics do not permit the measurement of absolute velocity.

- **Something is wrong with the laws of Physics!!**

At first people thought that it was Maxwell's equations that were at fault and they modified them so that they were invariant under the Galelian transformations; but this led to the prediction of new electrical phenomena which could not be found experimentally.

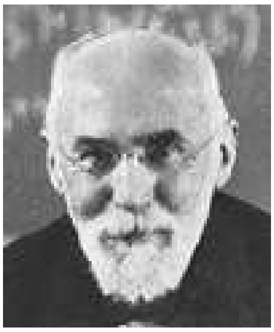

Figure 1.3: Hendrik Antoon Lorentz

Then in 1903 Lorentz made a remarkable and curious discovery. He found that when he applied the following transformations to Maxwell's equations

$$\begin{aligned} \bar{t} &= \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} , \\ \bar{x} &= \frac{x - vt}{\sqrt{1 - v^2/c^2}} , \\ \bar{y} &= y , \\ \bar{z} &= z , \end{aligned} \tag{1.8}$$

they remained invariant. These are the famous Lorentz transformations and we will discuss them in more detail in section 1.4.

1.3 The principle of Special Relativity

In 1905 Einstein generalized the Galelian relativity principle [applicable only to Newtons laws] to the whole of Physics by postulating that:

- **All inertial frames are equivalent for all experiments i.e. no experiment can measure absolute velocity.**
- **Maxwell's equations and the speed of light must be the same for all observers.**

Einstein's motivation was to avoid inconsistencies between Maxwell's equations and Galelian relativity. The Lorentz transformations must relate to actual space and time measurements.

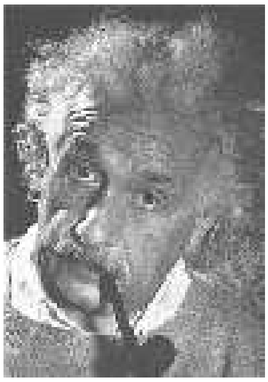

Figure 1.4: In the light of knowledge attained, the happy achievement seems almost a matter of course, and any intelligent student can grasp it without too much trouble. But the years of anxious searching in the dark, with their intense longing, their alternations of confidence and exhaustion, ant the final emergence into the light - only those who have themselves experienced it can understand that. *Albert Einstein.*

Special Relativity abolishes the idea of absolute space [for example the ether] and absolute time, but it leaves unexplained the origin of inertial frames. Mach's principle says that the inertial frames are determined be the rest of the matter in the universe [i.e. those that are non - accelerating with respect to the rest of the universe]. Einstein tried later to incorporate this idea into General Relativity.

1.4 Spacetime diagrams and the Lorentz transformations

Probably the easiest way to understand what the physical consequences of the above postulates are, is through the use of simple spacetime diagrams. In Figure 1.5 we illustrate some of the basic concepts using a two dimensional slice of spacetime. A single point, of fixed x and t is called an <u>event</u>. A particle or observer moving through spacetime maps out a curve $x = x(ct)$, and so represents the position of the particle at different times. This curve is called the particles <u>world - line</u> . The gradient of the world - line is related to the particle's velocity,

$$\frac{d(ct)}{dx} = \frac{c}{v} \ , \tag{1.9}$$

so light rays [$v = c$] move on 45^o lines on this diagram.

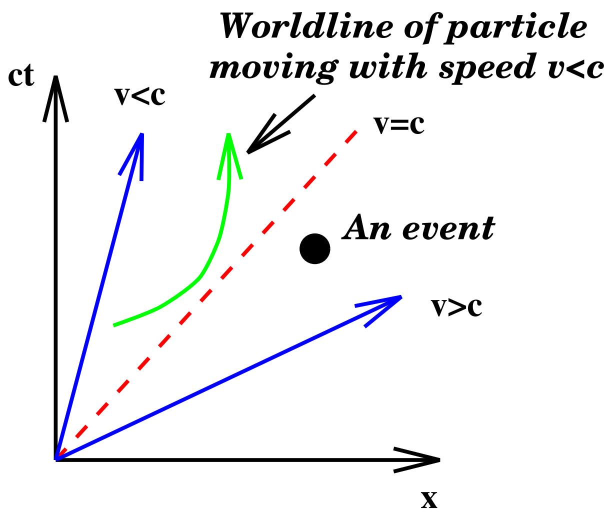

Figure 1.5: A simple spacetime diagram

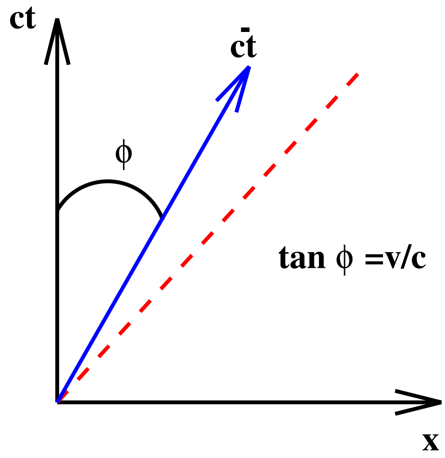

Figure 1.6: The $c\bar{t}$ axis of a frame whose velocity is v relative to $\mathcal{O}$.

Suppose an observer $\mathcal{O}$ uses coordinates ct and x as in Figure 1.5, and that another observer, with coordinates $c\bar{t}$ and $\bar{x}$, is moving with velocity v in the x direction relative to $\mathcal{O}$. It is clear from the above discussion that the $c\bar{t}$ axis corresponds to the world-line of $\bar{\mathcal{O}}$ in the spacetime diagram of $\mathcal{O}$ [see Figure 1.6]. We will now use Einstein's postulates to determine where the $\bar{x}$ goes in this diagram. Consider the following three events in the spacetime diagram of $\bar{\mathcal{O}}$ shown in Figure 1.7 [A, B and C] defined as follows. A light beam is emitted from the point A in $\bar{\mathcal{O}}$ $(c\bar{t} = 0, \bar{x} = -a)$ [event A]. It is then reflected at $(c\bar{t} = a, \bar{x} = 0$ [event B]. Finally it is received at $(c\bar{t} = 0, c\bar{x} = a)$ [event C]. How do these three

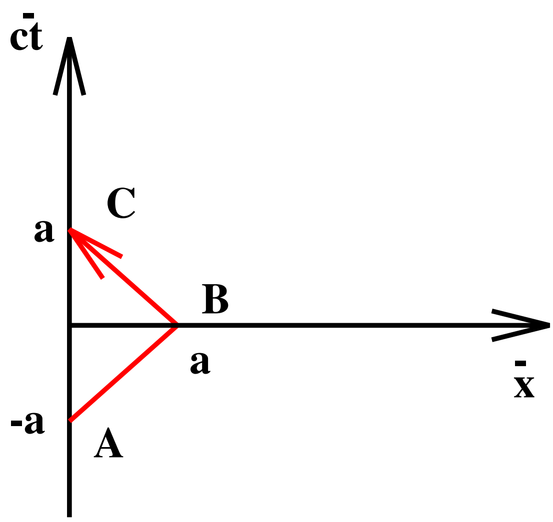

Figure 1.7: Light reflection in $\bar{\mathcal{O}}$

events look in the spacetime diagram of $\mathcal{O}$? We already know where the $c\bar{t}$ axis lies [see Figure 1.6]. Since this line defines $\bar{x} = 0$, we can locate events A and C [at $c\bar{t} = -a$ and $c\bar{t} = a$]. The second of Einstein's postulate states that light travels with speed c in all frames. We can therefore draw the same light beam as before, emitted from A and traveling on a 45^o line in the spacetime diagram of $\mathcal{O}$. The reflected beam must arrive at C, so it is a 45^o line with negative gradient which passes through C. The intersection of these two lines defines the event of reflection B in $\mathcal{O}$. It follows therefore, that the $\bar{x}$ axis is the line which passes through this point and the origin [see Figure 1.8].

One of the most startling results which follows from this geometrical construction is that events simultaneous to $\bar{\mathcal{O}}$ are not simultaneous to $\mathcal{O}$!

Let us now derive the Lorentz transformations using the geometrical arguments above and the principle of Special Relativity discussed in the last section. Assuming that we orient our axes so that $\bar{\mathcal{O}}$ moves with speed v along the positive x axis relative to $\mathcal{O}$, the most general linear transformations we can write down are

$$\begin{aligned} c\bar{t} &= \alpha ct + \beta x \ , \\ \bar{x} &= \delta ct + \gamma x \ , \\ \bar{y} &= y \ , \end{aligned}$$

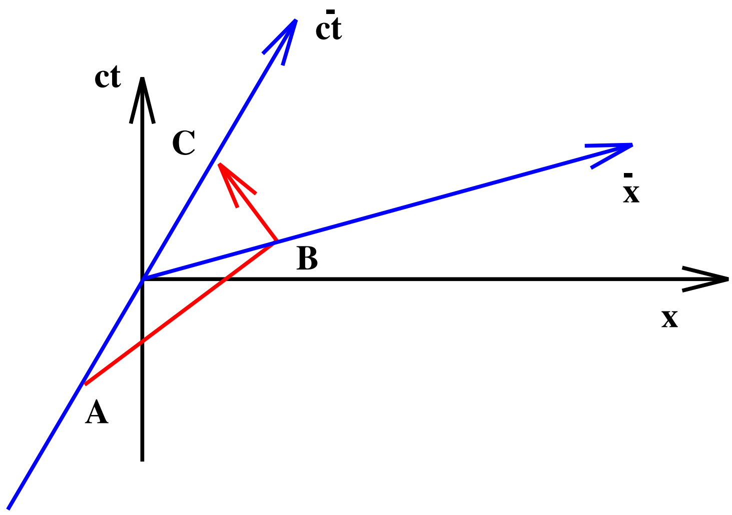

Figure 1.8: Light reflection in $\bar{\mathcal{O}}$ as measured by $\mathcal{O}$.

$$\bar{z} = z \; , \tag{1.10}$$

where α, β, γ and δ depend only on the velocity v. Looking at Figure 1.8, we see that the $c\bar{t}$ and $\bar{x}$ axes have the following equations:

$$\begin{aligned} \frac{v}{c}(ct) - x &= 0 \; , \\ \frac{v}{c}x - ct &= 0 \; . \end{aligned} \tag{1.11}$$

Together with (1.10), these straight line equations imply

$$\frac{\delta}{\gamma} = -\frac{v}{c} \; , \qquad \frac{\beta}{\alpha} = -\frac{v}{c} \; . \tag{1.12}$$

which simplify the first two transformation equations giving

$$\begin{aligned} c\bar{t} &= \alpha\left(ct - \frac{v}{c}x\right) \; , \\ \bar{x} &= \gamma\,(x - vt) \; . \end{aligned} \tag{1.13}$$

For the speed of light to be the same in both $\mathcal{O}$ and $\bar{\mathcal{O}}$ we require that

$$\frac{\bar{x}}{c\bar{t}} = \frac{\gamma\,(x - vt)}{\alpha\left(ct - \frac{v}{c}x\right)} = 1 \ . \tag{1.14}$$

Dividing top and bottom by c gives

$$\begin{aligned} \frac{\gamma\,(c-v)}{\alpha\,(c-v)} &= 1 \\ \Rightarrow \quad \gamma &= \alpha \ . \end{aligned} \tag{1.15}$$

We are therefore left with the following transformation law for $\bar{x}$ and $\bar{t}$:

$$\begin{aligned} c\bar{t} &= \gamma\left(ct - \frac{v}{c}x\right) \ , \\ \bar{x} &= \gamma\,(x - vt) \ . \end{aligned} \tag{1.16}$$

The principle of Special Relativity implies that if $x \to \bar{x}$, $y \to \bar{y}$, $z \to \bar{z}$ and $t \to \bar{t}$, $v \to -v$. This gives the inverse transformations

$$\begin{aligned} c\bar{t} &= \gamma\left(ct + \frac{v}{c}x\right) \ , \\ \bar{x} &= \gamma\,(x + vt) \ . \end{aligned} \tag{1.17}$$

Substituting for $\bar{x}$ and $\bar{t}$ from (1.16) in (1.17) gives, after some straightforward algebra,

$$\gamma = \pm\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \ . \tag{1.18}$$

We must choose the positive sign so that when $v = 0$ we get an identity rather than an inversion of the coordinates. The complete Lorentz transformations are therefore,

$$\begin{aligned} \bar{t} &= \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \ , \\ \bar{x} &= \frac{x - vt}{\sqrt{1 - v^2/c^2}} \ , \\ \bar{y} &= y \ , \\ \bar{z} &= z \ . \end{aligned} \tag{1.19}$$

1.5 The spacetime interval

Consider the effect of the Lorentz transformations on the spacetime interval

$$ds^2 = dx^2 - c^2dt^2 \ . \tag{1.20}$$

Substituting for x and t from the above Lorentz transformations one obtains

$$\begin{aligned} d\bar{s} &= d\bar{x}^2 - c^2d\bar{t}^2 = \gamma^2\left[dx^2 + v^2dt^2 - 2vdxdt - c^2dt^2 - \tfrac{v^2}{c^2}dx^2 + 2vdxdt\right] \\ &= dx^2 - c^2dt^2 \\ &= ds^2 \ . \end{aligned} \tag{1.21}$$

Generalizing to four dimensions we see that the spacetime interval

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \tag{1.22}$$

is invariant under the Lorentz transformations.

The most general transformation between $\mathcal{O}$ and $\bar{\mathcal{O}}$ will be more complicated but it must be linear. We can write it as:

$$\bar{x}^\alpha = \sum_{\beta=0}^{3} A_{\alpha\beta}x^\beta + b^\alpha \ , \tag{1.23}$$

where $x^\alpha \rightarrow (ct, x, y, z)$. This linear transformation is called the generalized Lorentz transformations. It contains ten parameters: four correspond to an origin shift (b^α), three correspond to a Lorentz boost [which depends on $\mathbf{v}$] and three to the rotation which aligns the axes of $\mathcal{O}$ and $\bar{\mathcal{O}}$. The last six are contained in the 4×4 matrix $A_{\alpha\beta}$ [six because $A_{\alpha\beta}$ is symmetric i.e. $\frac{1}{2}(4 \times 3)$].

- **Note that these parameters for a group galled the Poincaré group.**

Later we will show that the Poincaré transformations preserve Maxwells equations as well as light paths.

1.6 Minkowski spacetime

Since ds is frame independent we can interpret this result geometrically by treating points x^α in the four dimensional spacetime, with ds representing the four dimensional distance [interval] between neighboring points.

Figure 1.9: The views of space and time that I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself and time by itself are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality. *Hermann Minkowski*

The expression for ds looks Euclidian if the time component is taken to be ict, $i = \sqrt{-1}$:

$$ds^2 = dx^2 + dy^2 + dz^2 + c^2dt^2 \ . \tag{1.24}$$

The line element allows one to classify spacetime intervals into three different classes: Consider a curve Γ [world-line of a particle, say] and an interval on that curve. If:

$$ds^2 < 0 \Rightarrow c^2dt^2 > dx^2 + dy^2 + dz^2 \ , \tag{1.25}$$

the curve is timelike in that interval;

$$ds^2 > 0 \Rightarrow c^2dt^2 < dx^2 + dy^2 + dz^2 \ , \tag{1.26}$$

the curve is spacelike in that interval; and

$$ds^2 = 0 \Rightarrow c^2dt^2 = dx^2 + dy^2 + dz^2 \ , \tag{1.27}$$

the curve is null in that interval.

Note that the sign of ds^2 is a matter of convention, invariance between frames is what is important.

1.6.1 The null cone

Consider a null ray [for example a light ray propagating in the x direction: then $\frac{dx}{dt} = \pm c$, so the null ray represents a particle moving at the speed of light. If we

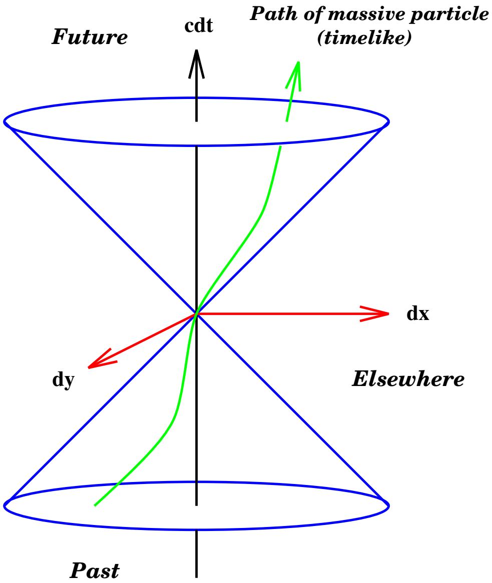

Figure 1.10: The lightcone.

consider 3D Minkowski spacetime then the null rays lie on the surface of a cone [see Figure 1.9]:

Before we look at some of the wonderful consequences of the Einstein postulates, let us return briefly to the Lorentz transformations. If we define $\tanh\Phi = \frac{v}{c}$ we can write them in the following way:

$$\begin{aligned} \bar{x} &= x\cosh\Phi - ct\sinh\Phi \ , \\ c\bar{t} &= -x\sinh\Phi + ct\cosh\Phi \ . \end{aligned} \tag{1.28}$$

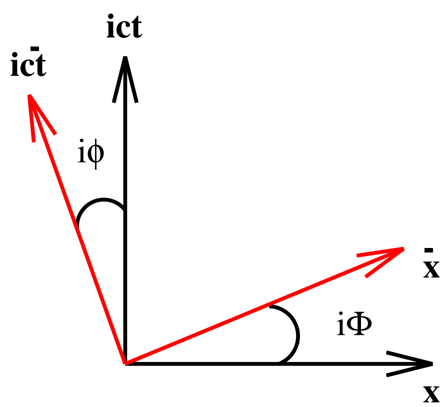

Figure 1.11: ict space

Formally this corresponds to a rotation through an angle $i\Phi$ in (ict, x) space [see Figure 1.11].

1.7 Consequences of the Einstein postulates

1.7.1 Time dilation

Suppose we have someone in a spaceship $\bar{\mathcal{O}}$ moving at some velocity v relative to an observer $\mathcal{O}$. The guy in the spaceship turns his stereo on and starts listening to some rap music!! With a clock he measures the time interval between two beats $\Delta\bar{t} = \bar{t}_B - \bar{t}_A$. what time interval does the observer $\mathcal{O}$ measure? To calculate this it is easier to use the inverse Lorentz transformations. By the Relativity principle we have:

$$t = \gamma\left(\bar{t} + v\bar{x}/c^2\right) \ , \tag{1.29}$$

so

$$t_B - t_A = \gamma\left[\bar{t}_B - \bar{t}_A + \tfrac{v}{c^2}\left(\bar{x}_B - \bar{x}_A\right)\right] \ . \tag{1.30}$$

But the stereo is stationary relative to $\bar{\mathcal{O}}$ so $\bar{x}_A = \bar{x}_B$, so we end up with:

$$\begin{aligned} \Delta t &= t_B - t_A = \gamma\Delta\bar{t} \\ &= \Delta t = \frac{\Delta\bar{t}}{\sqrt{1-\frac{v^2}{c^2}}} \ . \end{aligned} \tag{1.31}$$

Since $\Delta t > \Delta\bar{t}$, time in the spaceship slows down! That is, when the clock in the spaceship records 1 second elapsed, as seen be the man in the ship, it shows $1/\sqrt{1-v^2/c^2}$ seconds to the man outside.

The slowing of clocks in a moving system is a real effect and it applies equally to all kinds of time for example biological, chemical reaction rates, even to the rate of growth of a cancer in a cancer patient! How is this so you may ask? If the rate of growth of the cancer was the same for a stationary patient as for a moving one, it would be possible to use the rate of cancer development to determine the speed of the ship!

A very interesting example of the slowing down of time is that of the cosmic ray muons. These are particles that disintegrate spontaneously after an average

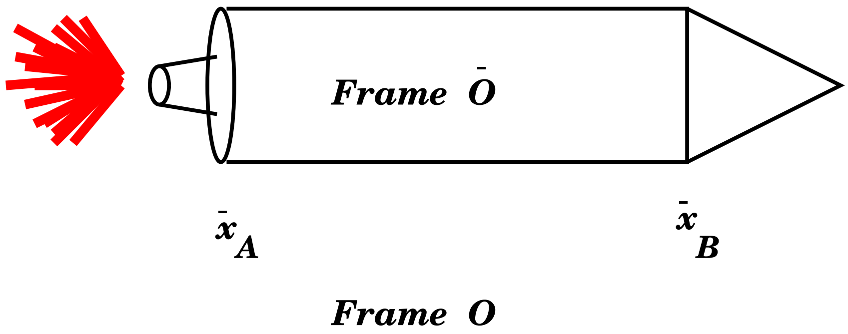

Figure 1.12: Length contraction

lifetime of about 2.2×10^{-6} seconds. It is clear that in its short lifetime a muon cannot, even at the speed of light, travel more than 600 m. But although the muons are created at the top of the atmosphere, some 10 km up, we can detect them down here on earth. How can that be!!? From the muon point of view (i.e. their frame of reference) they only live about 2 μs. However from our point of view they live considerably longer, indeed long enough to reach the surface of the earth (by a factor of $1/\sqrt{1 - v^2/c^2}$).

1.7.2 Length contraction

Now suppose the observer $\mathcal{O}$ wants to measure the length of the spaceship. He can only do this by making an instantaneous measurement of the spatial coordinates of the end of the ship i.e. x_A and x_B. The Lorentz transformations give you:

$$\begin{aligned} \bar{x}_1 &= \gamma\,(x_1 - vt_1) \\ \bar{x}_2 &= \gamma\,(x_2 - vt_2)\ , \end{aligned} \tag{1.32}$$

so the length is

$$\begin{aligned} L &= x_2 - x_1 \\ &= \gamma^{-1}\,(\bar{x}_2 - \bar{x}_1) \\ &= \gamma^{-1}\bar{L}\ , \end{aligned} \tag{1.33}$$

since $t_1 = t_2$ [instantaneous measurement by $\mathcal{O}$]. Writing this result in terms of v and c we have:

$$L = \sqrt{1 - v^2/c^2}\,\bar{L}\ . \qquad (1.34)$$

Note that this is not a physical effect on the rod but an effect of spacetime itself.

1.7.3 The twin paradox

To continue our discussion of the Lorentz transformations and relativistic effects, we consider the famous "twin paradox" of Peter and Paul. When they are old enough to drive a spaceship, Paul flies away from earth at very high speed. Because Peter who is left on the ground, sees Paul going so fast, all of Pauls' clocks appear to go slower , from Peter's point of view. Of course, Paul notices nothing unusual. After a while he returns and finds that he is younger that Peter! Now some people might say "heh, heh, heh, from the point of view of Paul, can't we say that Peter was moving and should therefore appear to age more slowly? By symmetry, the only possible result is that they are both the same age when they meet". But in order for them to come back together and make a comparison, Paul must turn around which involves decelerating and accelerating, and during that period he is not in an inertial frame. This breaks the apparent symmetry and so resolves the paradox.

1.8 Velocity composition law

Suppose a particle has a velocity $\mathbf{u} = (u_1, u_2, u_3)$ in $\mathcal{O}$, then at time t its spacetime coordinates are $\mathbf{x} = (ct, u_1 t, u_2 t, u_3 t)$. In a frame $\bar{\mathcal{O}}$ moving at a speed v relative to $\bar{\mathcal{O}}$, its spacetime coordinates will be:

$$\bar{\mathbf{x}} = \left[\gamma c\left(t - \tfrac{vu_1}{c^2}\right)\ , \gamma\left(u_1 t - vt\right)\ , u_2 t\ , u_3 t\right] \qquad (1.35)$$

The velocity in $\bar{\mathcal{O}}$ simply comes from dividing the space parts by the time part i.e.

$$\bar{u}_1 \quad = \quad \frac{d\bar{x}_1}{d\bar{t}} = \frac{u_1 - v}{1 - \frac{vu_1}{c^2}}\ ,$$

$$
\begin{aligned}
\bar{u}_2 &= = \frac{d\bar{x}_2}{d\bar{t}} = \frac{u_2}{\gamma\left(1 - \frac{vu_1}{c^2}\right)} \,, \\
\bar{u}_3 &= \frac{d\bar{x}_3}{d\bar{t}} = \frac{u_3}{\gamma\left(1 - \frac{vu_1}{c^2}\right)} \,. \qquad (1.36)
\end{aligned}
$$

If $u_1 = u$ and $u_2 = u_3 = 0$ then

$$\bar{u} = \frac{u - v}{1 - \frac{uv}{c^2}} \,. \qquad (1.37)$$

This looks like the Galelian relative velocity if $|u|\,, |v| \ll c$. If $u = c$, $\bar{u} = c$, so we confirm the universality of the the speed of light.