Chapter 2

Background History

Einstein's theory of relativity has a formidable reputation as being incredibly complicated and impossible to understand. It's not! The principle of relativity itself, the single, simple idea upon which Einstein's theory is based, has been around since the time of Galileo. As we shall see, when it is applied to objects that are moving extremely fast, the consequences seem strange to us because they are outside our everyday experience; but the results make sense and are all self-consistent when we think about them carefully. We can summarize the major corrections that we need to make to Newton's equations of motion as follows: Firstly, when an object is in motion, its momentum p is larger than expected, its length l shrinks in the direction of motion, and time t slows down, in each case by a factor

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}, \tag{2.1}$$

where v is the velocity of the object and c is the speed of light:

$$\begin{aligned} c &= 299792458 \mathrm{m/s} (\mathrm{exactly, by definition}) \\ &= 186,282 \mathrm{miles/second} \\ &= 30 \mathrm{cm/ns, in units we can grasp.} \end{aligned}$$

This seems confusing at first because we are used to assuming, for example, that the length l of any object should be constant. For everyday purposes, the correction is tiny — consider the International Space Station, moving at 8 km/s in orbit; its length is about 1 part per billion less than if it were at rest, and time on board moves more slowly by the same factor. But for many particles moving near the speed of light, the fact that time slows down (and hence lifetimes are longer) with velocity by the factor (2.1) has been well verified. The same applies to the increasing momentum – which demonstrates immediately the well-known principle that one can never push an object hard enough to accelerate it to the speed of light, since, as it goes faster and faster, you have to push harder and harder to obtain a given increase in velocity. Loosely speaking, it acts as though the *mass* increases with velocity.

Secondly, as we shall discover about halfway through the course, there emerges naturally what may well be the most famous equation in the world:

$$E = mc^2. \tag{2.2}$$

These formulae were not found experimentally, but theoretically, as we shall see.

Einstein's 1905 relativity paper, "On the Electrodynamics of Moving Bodies", was one of three he published that year, at age 26, during his spare time; he was at the time working as a patent clerk in Zurich. Another was a paper explaining Brownian motion in terms of kinetic theory (at a time when some people still doubted the existence of atoms), and the third proposed the existence of photons, thus laying the foundations for quantum theory and earning him the Nobel prize (relativity being too controversial then).

Einstein wrote two theories of relativity; the 1905 work is known as "special relativity" because it deals only with the special case of *uniform* (i.e. non-accelerating) motion. In 1915 he published his "general theory of relativity", dealing with gravity and acceleration. Strange things happen in accelerating frames; objects appear to start moving without anything pushing them... During this course we shall only deal with special relativity.

2.1 The Principle of Relativity

As we use our telescopes to look ever farther out into the universe, some relevant questions present themselves:

- Is space homogeneous? I.e., is it the same everywhere — are the laws of physics the same in distant galaxies as they are here on Earth?
- Is it isotropic — is it the same in all directions, or is there some defining "axis" or direction that is preferred in some way? Is, for example, the speed of light the same in all directions?
- Are the laws of physics constant in time?
- And finally, are the laws of physics independent of uniform relative motion?

By looking at light from the most distant visible galaxies, more than 10 billion light years away, we can recognise the spectra of hydrogen atoms. As far as we can tell, those hydrogen atoms are the same everywhere. And because the light was emitted so long ago, it seems clear that the laws of physics are indeed constant in time (with the possible exception of the gravitational constant G; its rate of change is very difficult to measure, but no variation has been seen so far).

The last of these questions lies at the core of relativity. If I perform an experiment on board a rocket that is moving uniformly through space (remember, we aren't dealing with acceleration or gravity here), will I get the same result as somebody doing the same experiment on another rocket moving at a different speed? "Common sense" suggests that there should be no difference. This "common sense" idea is known in physics as *the principle of relativity*, and it was first proposed by Galileo. Here is Newton's definitive statement of it as a corollary to his laws of motion:

> "The motions of bodies included in a given space are the same among themselves, whether that space is at rest or moving uniformly forward in a straight line."

Meaning: if a spaceship is drifting along at a uniform velocity, all experiments and phenomena inside the spaceship will be just the same as if the ship were not moving. *There is no "preferred" inertial (i.e. non-accelerating) frame of reference* which is "at rest" in the universe; and therefore, you cannot tell how "fast" a spaceship, or car, or whatever, is moving by doing experiments inside — you have to look outside to compare, in order to see how fast it is moving *relative to its surroundings*.

Galileo considered ordinary ships instead of spaceships: he pointed out that a rock dropped from the top of the mast will hit the same spot on deck whether the ship is stationary or moving along uniformly.

This is a simple and appealing idea which, of course, needed to be tested experimentally. Before we go any further, though, let us familiarise ourselves with the meaning of relativity in the everyday world of Newtonian mechanics.

2.2 Newtonian/Galilean Relativity

Consider two people, Tony (standing still) and Bill (walking past at velocity u). Tony has a "reference frame" S in which he measures the distance to a point on the pavement ahead of him, and calls it x. Bill, who walks past at time $t = 0$, has a "reference frame" S' in which he measures the (continually changing) distance to the same point, and calls it x'. Then, a *Galilean transformation* links the two frames:

$$\begin{aligned} x' &= x - ut \\ y' &= y \\ z' &= z \\ t' &= t. \end{aligned} \tag{2.3}$$

This is commonsense: we can see how the distance that Bill measures to the point decreases with time, until it goes negative when Bill actually walks past the point.

What about velocities? Suppose Tony is standing still to watch a bird fly past at speed v. He now calls the (changing)

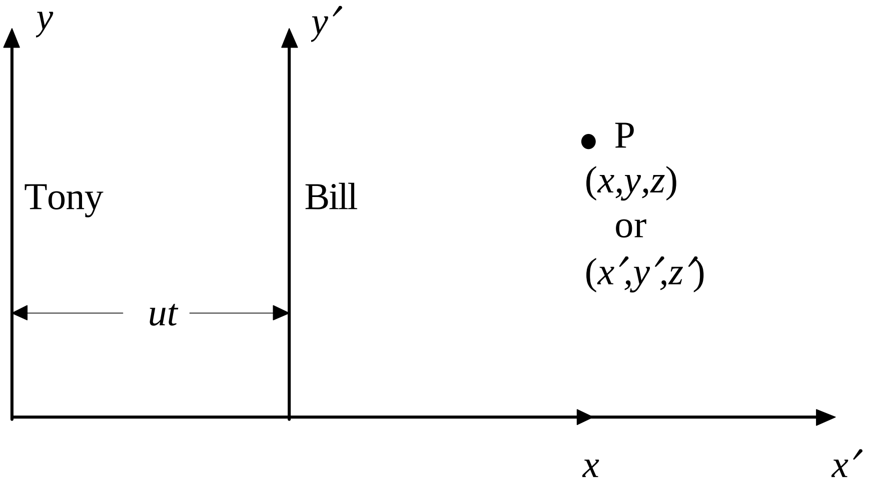

Figure 2.1: A pair of coordinate systems in relative motion.

distance to the bird x, and

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

If we differentiate 2.3, we see that, according to Bill, the bird is flying past with speed

$$\begin{aligned} v' &= \frac{dx'}{dt'} = \frac{dx'}{dt} \\ &= \frac{dx}{dt} - u \\ &= v - u. \end{aligned} \tag{2.4}$$

We see too that acceleration is the same in both frames,

$$a' = \frac{dv'}{dt'} = \frac{dv}{dt} = a, \tag{2.5}$$

so Newton's law, $F = ma$, will be the same in both frames of reference; likewise conservation of momentum holds true in both frames.

2.2.1 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 (2.4):

$$\mathbf{v}' = \mathbf{v} - \mathbf{u}. \tag{2.6}$$

As above, $\mathbf{u}$ is the velocity of Bill relative to Tony. Let's call Eastwards the $\mathbf{i}$ direction and northwards the $\mathbf{j}$ direction. Then

$$\mathbf{u} = 25\mathbf{i}.$$

The velocity of the plane is, according to Tony,

$$\mathbf{v} = 200\mathbf{i}.$$

Therefore, the velocity of the plane relative to Bill is

$$\begin{aligned} \mathbf{v}' &= \mathbf{v} - \mathbf{u} \\ &= 200\mathbf{i} - 25\mathbf{i} = 175\mathbf{i}. \end{aligned}$$

As seen from the train, then, the plane is flying past eastwards at a speed of 175 m/s.

Tony calculates the velocity of the car to be

$$\mathbf{v} = 25\mathbf{j}.$$

Therefore, from Bill's frame of reference, the car is moving with velocity

$$\mathbf{v}' = 25\mathbf{j} - 25\mathbf{i}.$$

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

2.2.2 Example II

Tony is playing snooker. The white ball, which has mass m and moves with velocity

$$\mathbf{v} = 13\mathbf{i}\text{cm/s},$$

hits a stationary red ball, also of mass m, in an elastic collision. The white ball leaves the collision with velocity $\mathbf{v}_w = 11.1\mathbf{i} + 4.6\mathbf{j}$ (i.e. at 12 cm/s at an angle of 22.6° above the horizontal), and the red ball leaves at a velocity of $\mathbf{v}_r = 1.9\mathbf{i} - 4.6\mathbf{j}$ (which is 5 cm/s at an angle of 67.4° below the horizontal). You can check that momentum and energy are conserved. Suppose now that Bill is walking past with a velocity of $\mathbf{u} = 13\mathbf{i}$. 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 $0 - 13\mathbf{i} = -13\mathbf{i}$. After the collision, we obtain for the white ball

$$\begin{aligned} \mathbf{v}_w' &= (11.1\mathbf{i} + 4.6\mathbf{j}) - 13\mathbf{i} \\ &= -1.9\mathbf{i} + 4.6\mathbf{j}, \end{aligned}$$

whereas the red ball moves with velocity

$$\begin{aligned} \mathbf{v}_r' &= (1.9\mathbf{i} - 4.6\mathbf{j}) - 13\mathbf{i} \\ &= -11.1\mathbf{i} - 4.6\mathbf{j}. \end{aligned}$$

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.

2.2.3 Inverse Transformations

Naturally, if we have a set of coordinates in Bill's frame of reference and we want to know how they look from Tony's point of view, we just need to realise that, according to Bill, Tony is moving past with a velocity of $-u$, and so the (inverse) transformation is

$$\begin{aligned} x &= x' + ut' \\ y &= y' \\ z &= z' \\ t &= t'. \end{aligned} \tag{2.7}$$

2.2.4 Measuring Lengths

Tony is sitting on a train, which is moving at speed u. He has paced the corridors from one end to the other, and calculates that its length is 100 m. Bill, meanwhile, is standing by the tracks outside, and is curious to calculate for himself how long the train is. Several ways suggest themselves.

1. He can note where he is standing when the front of the train passes, then run towards the back end and note where he is standing when the back end passes, and subtract the two distances. This will obviously give the wrong

answer.

2. He can set up a row of cameras, which will take their pictures simultaneously. The separations between the cameras that see the front and the rear of the train gives the length of the train.
3. He can stand still, and time the train going past. Using his Doppler radar gun he measures the speed of the train, and from the speed and the time he calculates the length.
4. He can do some mixture of these, measuring the position of the front of the train at one time and the back of the train at another, and compensate for the train's velocity by calculating where the two ends would be at some particular moment in time, $t' = 0$.

Looking at this formally, suppose that, according to Tony, the front of the train is at x_2 and the rear is at x_1, so the length is $(x_2 - x_1)$. Bill measures x'_1 and x'_2 at times t'_1, t'_2. He calculates that at time $t' = 0$, the front of the train was at position

$$x'_2 - ut'_2,$$

and the rear was at position

$$x'_1 - ut'_1.$$

The length of the train is therefore given by the difference between these positions:

$$x'_2 - ut'_2 - x'_1 + ut'_1.$$

In using equation (2.3) to transform from Tony's to Bill's frame, we remember that Bill is moving past with a velocity of $-u$, so

$$
\begin{aligned}
x'_1 &= x_1 + ut_1 \\
x'_2 &= x_2 + ut_2 \\
t'_1 &= t_1 \\
t'_2 &= t_{2.}
\end{aligned}
$$

Therefore, the length will be

$$(x'_2 - x'_1) - u\,(t'_2 - t'_1) = (x_2 - x_1)\,, \tag{2.8}$$

which agrees with Tony's length measurement. Comparing with the options above, we have:

1. corresponds to forgetting that $t'_2 \neq t'_1$, just using $x'_2 - x'_1$, and getting the answer wrong.
2. corresponds to measuring the coordinates of the two ends at the same time; $t'_2 = t'_1$, so $x'_2 - x'_1 = (x_2 - x_1)$.
3. corresponds to measuring at the same place; $x'_2 = x'_1$, so the length is $u\,(t'_1 - t'_2)$.
4. corresponds to using (2.8) to compensate for the speed of the train.

This is just a matter of putting commonsense on a firm footing.

2.3 The Clash with Electromagnetism

Newton's laws reigned supreme in mechanics for more than 200 years. However, difficulties arose in the mid-19th century with studies of electromagnetism. All electrical and magnetic effects could be summarised nicely in *Maxwell's Laws*, which we won't go into here. The problem was that, unlike Newton's laws, these were *not* invariant under a Galilean transformation — the principle of relativity didn't seem to be valid for electricity or magnetism! Therefore, in a moving spaceship, it seemed that electromagnetic (including optical) phenomena would be different than they would be in a laboratory that was "at rest", and one should be able to determine the speed of the spaceship by doing optical or electrical experiments. Can you imagine, for example, if all of the magnets on the space shuttle were to get weaker in the first half of its orbit and stronger again in the second half, just because it was moving in different directions through space as it went around the Earth? Something was wrong!

In particular, the conflict became apparent where Maxwell's Laws predicted a *constant speed of light*, independent of the speed of the source. Sound is like this; it moves through the air at the same speed, regardless of the speed of the source. (The speed of the source changes the frequency, or pitch, of the sound, via the famous Doppler effect, but not its speed). As an example, suppose that Tony is standing still, on a calm day, and Bill is paragliding past at 30 m/s. Sound from the rear moves past Tony (in the same direction as Bill is going) at 330 m/s. The apparent speed of the sound wave, as measured by Bill, is

$$v' = v - u = 300\text{m/s}.$$

So, Bill can measure this speed, and deduce that he is moving at 30 m/s — relative to the air...

Suppose now Bill is on a spaceship, moving at $u = 2 \times 10^8$

m/s, and is overtaken by light moving at $c = 3 \times 10^8$ m/s. By measuring the speed of light going past, can Bill measure the speed of the spaceship? Newton's laws, using the Galilean transformation, would suggest that Bill would see the light going past at 1×10^8 m/s, from which he could deduce his speed; but Maxwell's laws, which predict that the light would pass Bill at 3×10^8 m/s regardless of his speed, clearly disagree.

2.4 The Invention of the Ether

Since, by the 1870s, Newton's Laws had stood the test of time for two centuries, and Maxwell's Laws, having a vintage of just 20 years or so, were young upstarts, the natural assumption was that Maxwell's Laws needed some modification. The first obvious conclusion was that, just as sound needed a medium to travel through — and the speed was constant relative to the medium — so light must need a medium too. (Remember that nobody had come across the idea of a wave without a medium until then). The Victorian scientists named this medium the *ether* (or æther, if you prefer). It had to have some strange properties:

- Invisibility, of course.
- It was massless.
- It filled all of space.
- High rigidity, so light could travel so quickly through it. (Something that springs back fast carries waves more quickly than something soft;sound travels faster through iron than air).

- It had no drag on objects moving through it: the Earth isn't slowed down in its orbit.
- Other curious properties had to be assumed to explain new experimental results, such as...

2.5 The Michelson-Morley Experiment

If the Earth is really a "spaceship" moving through the ether, the speed of light in the direction of Earth's motion should be lower than it is in a direction at right angles to this. By measuring these speeds, we should therefore be able to detect Earth's *absolute* velocity relative to the ether. The most famous experiment that tried to do this was the Michelson-Morley experiment, in 1887. Here is how it works (see diagram):

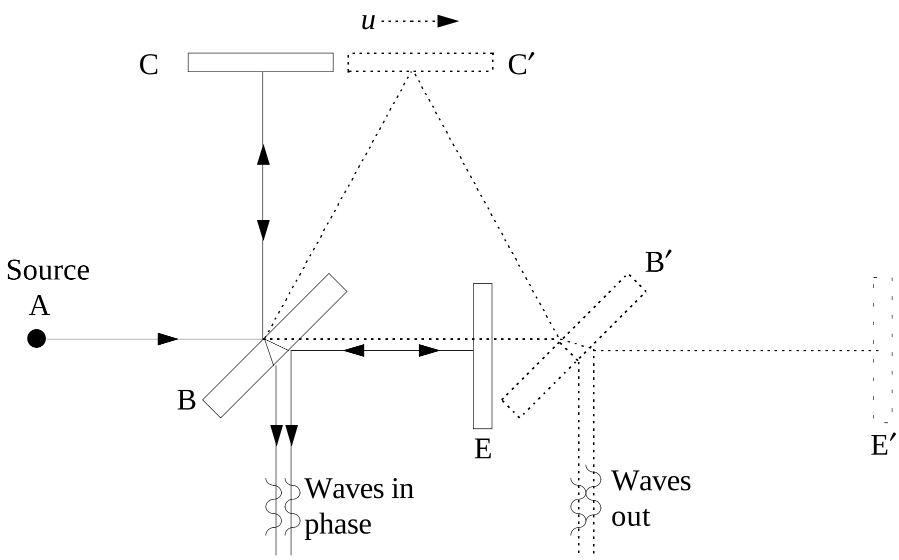

Figure 2.2: The Michelson-Morley experiment.

There is a light source at A. A glass plate at B is halfsilvered, so half of the light is reflected up to C (ignore the dashed lines for now), where it hits a mirror and comes back down. The other half of the light carries on through to E, where it also is reflected back. The two beams are recombined on the other side of B, where they make interference fringes (bright where crest meets crest, dark where crest meets trough).

Now, suppose the apparatus starts to move to the right at velocity u. Suppose the time now needed for the light to go from B to E is t_1. In this time, mirror E has travelled distance ut_1 to the right, so the total distance the light has to go is

$$L + ut_1 = ct_1,$$

since the light is travelling at speed c.

When it bounces back, B is moving in to meet it, so it has a shorter distance to go; if it takes time t_2, we have

$$L - ut_2 = ct_2.$$

Rearranging these equations,

$$\begin{aligned} t_1 &= \frac{L}{(c-u)} \\ t_2 &= \frac{L}{(c+u)} \end{aligned}$$

we find that the total time for the round trip is

$$\begin{aligned} t_1 + t_2 &= \frac{L\,(c+u) + L\,(c-u)}{(c-u)\,(c+u)} \\ &= \frac{2Lc}{(c^2 - u^2)} \end{aligned}$$

$$= \frac{2L/c}{(1 - u^2/c^2)}. \tag{2.9}$$

Now let's do the same calculation for the light that bounces off C. The ray follows the hypotenuse of a triangle, and so travels the same distance in each leg. If each leg takes time t_3, and is therefore a distance ct_3, we have

$$(ct_3)^2 = L^2 + (ut_3)^2 \, ,$$

or

$$\begin{aligned} c^2 t_3^2 - u^2 t_3^2 &= L^2 \\ t_3^2 \left(c^2 - u^2\right) &= L^2 \end{aligned}$$

so

$$t_3 = \frac{L}{\sqrt{(c^2 - u^2)}}.$$

Since the return trip is the same length, the total round trip takes a time of

$$2t_3 = \frac{2L/c}{\sqrt{(1 - u^2/c^2)}}. \tag{2.10}$$

So, the times taken to do the two round trips are *not the same.*

In fact, the lengths of the arms L cannot be made exactly the same. But that doesn't matter: what we have to do is to *rotate* the interferometer by 90°, and look for a *shift* in the interference fringes as we move through the ether in the direction of first one, and then the other arm.

The orbital speed of the Earth is about 30 km/s. Any motion through the ether should be at least that much at some time of the night or day at some time of the year — but nothing showed up! The velocity of the Earth through the ether could not be detected.

2.6 Frame Dragging and Stellar Aberration

When an aeroplane flies through the air, or a ship moves through the water, it drags a "boundary layer" of fluid along with it. Was it possible that the Earth in its orbit was somehow "dragging" some ether along? This idea — known as "frame dragging" — would explain why Michelson and Morley could not find any motion of the Earth relative to the ether. But this had already been disproved, by the phenomenon known as *stellar aberration*, discovered by Bradley in 1725.

Imagine a telescope, on a "still" Earth, pointed (for simplicity) to look at a star vertically above it. (See diagram). Now, suppose that the Earth is moving (and the telescope with it), at a speed v as shown. In order for light that gets into the top of the telescope to pass all the way down the moving tube to reach the bottom, the telescope has to be tilted by a small angle δ, where

$$\tan\delta = \frac{v}{c}.$$

(In fact, as we shall see, the angle is actually given by

$$\tan\delta = \frac{v/c}{\sqrt{1-v^2/c^2}},$$

but the correction factor is tiny). The angle δ oscillates with Earth's orbit, and, as measured, agrees with Earth's orbital speed of about 30 km/s. In this case, the ether cannot be being dragged along by the Earth after all, otherwise there would be no aberration.

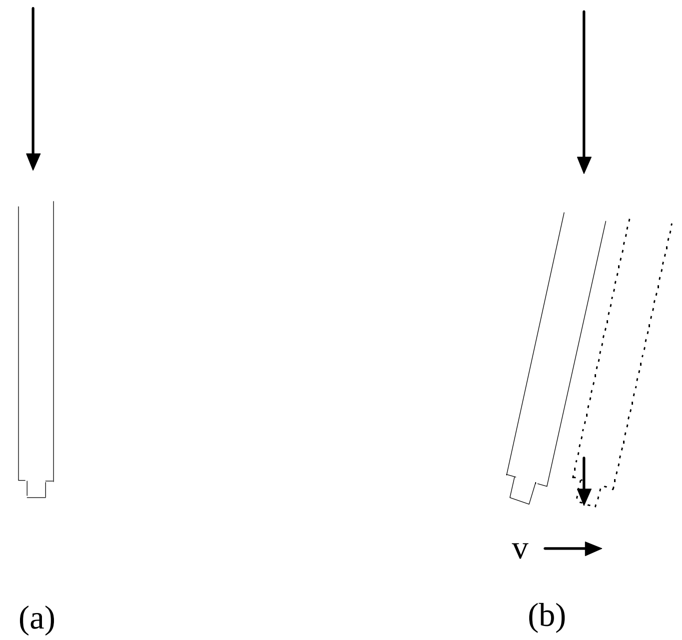

Figure 2.3: The aberration of starlight. (a) Stationary telescope. (b) Moving telescope.

2.7 The Lorentz Transformation

Notice that the difference in travel times for the two arms of the Michelson interferometer is a factor of

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

Lorentz suggested that, because all of the electrons in all of the materials making up the interferometer (and everything else) should have to interact with the ether, moving through the ether might make materials contract by just this amount in the direction of motion, but not in transverse directions. In that case, the experiment would give a null result! (Fitzgerald had also noticed that this "fix" would work, but could not suggest what might cause it). In developing this idea further, Lorentz found that clocks that were moving through the ether should run slowly too, by the same amount. He noticed that if, instead of the Galilean transformations 2.3, he made the substitutions

$$\begin{aligned} x' &= \frac{x - ut}{\sqrt{1 - u^2/c^2}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t - ux/c^2}{\sqrt{1 - u^2/c^2}}. \end{aligned} \tag{2.11}$$

into the Maxwell equations, they became invariant! These equations are known as a *Lorentz transformation*.

2.8 Poincaré and Einstein

To everyone else, naturally, Lorentz's solution looked like an artificial fudge, invented just to solve this problem. People continued to try to discover an "ether wind", and every time an explanation had to be made up as to why nature was "conspiring" to thwart these measurements. In the end, Poincaré pointed out that *a complete conspiracy is itself a law of nature*; he proposed that there *is* such a law of nature, and that it is not possible to discover an ether wind by *any* experiment; that is, there is no way to determine an absolute velocity. Poincaré's *principle of relativity* (1904) states that:

> The laws of physics should be the same in all reference frames which move in uniform motion with respect to one another.

To this, Einstein added his *second postulate*:

> The velocity of light in empty space is the same in all reference frames, and is independent of the motion of the emitting body.

The second postulate as stated is not very general, and it implies that there is something special about the behaviour of light. This is not the case at all. We should restate it as:

> There is a finite speed that is the same relative to all frames of reference.

The fact that light *in vacuo* happens to travel at that speed is a consequence of the laws of electromagnetism – and of the first postulate. Let us re-emphasize: there is nothing particularly special about light that makes it somehow magically change the properties of the universe. It is the fact that there is a finite speed that is the same for all observers that runs counter to our instincts, and that has such interesting consequences.

We will from now on assume that these postulates are true, and see what experimental results we can predict from them. In fact, as we stated earlier, relativity has passed every experimental test that has ever been proposed for it. It is now so deeply ingrained in our thinking that, unlike other "laws" or "theories" that apply to particular branches of physics, relativity is used as a sort of check — any theory that we produce has to be consistent with relativity or else it is at best an approximation.

Chapter 3

The Lorentz Transformations

3.1 Transverse Coordinates

Let us take two rulers that are exactly the same, and give one to a friend who agrees to mount it on his spaceship and fly past us at high speed in the x direction. His ruler will be mounted in the transverse (y) direction. As he flies by, we hold pieces of chalk at the 0 and 1 m marks on our ruler, and hold it out so they make marks on his ruler. When he comes back, we look to see where those marks are. What do we find? They must, of course, be one metre apart, because if they were different then we would have a way of knowing which of us was "really" moving. Thus, for transverse coordinates, the transformations between his frame and ours must be

$$y' = y,$$
$$z' = z.$$

3.2 Time Dilation

Imagine a simple clock, a "light clock", consisting of light bouncing between two mirrors separated by a distance l (see diagram). Each time the light hits one of the mirrors, the clock gives out a "tick". Let us make a pair of these, and give one to our friend to take abord his spaceship, while we keep the other on Earth. The clock on the spaceship is mounted perpendicular to the direction of motion, just like the ruler that we used last time.

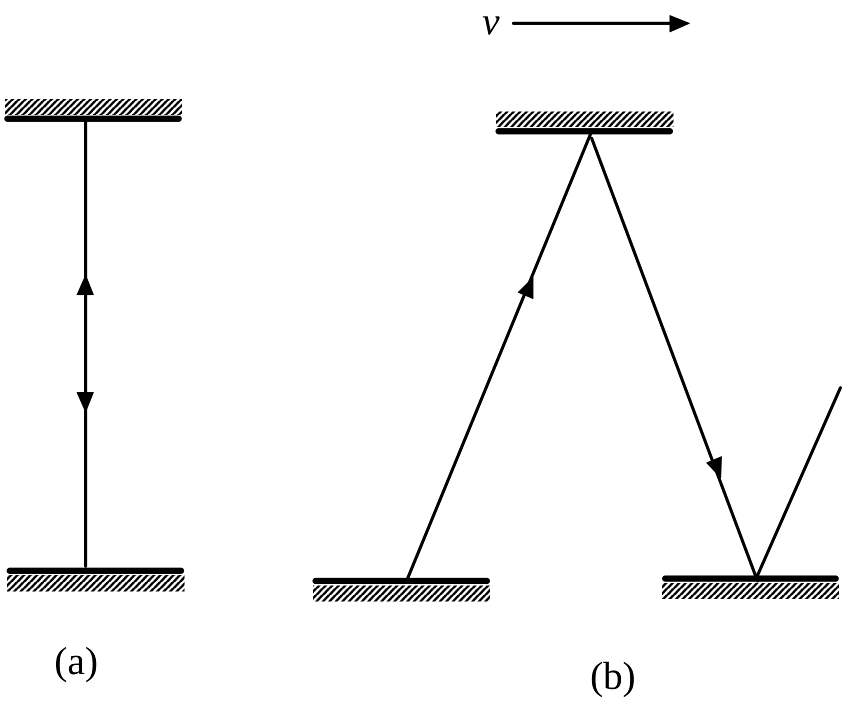

Figure 3.1: A "light clock", as seen in its rest frame (a) and from a frame (b) in which it is moving with velocity v.

As our friend flies past, we watch the light bouncing between the mirrors. But to us, instead of just going up and down, the light makes a zigzag motion, which means that it has to go further. Between "ticks", therefore, whereas the light in our clock covers a distance l in time

$$t = l/c,$$

the light in the clock on board the spaceship covers a distance

$$\sqrt{l^2 + v^2 t'^2},$$

which it must do in time

$$t' = \frac{\sqrt{l^2 + v^2 t'^2}}{c}.$$

Substituting for l, we obtain

$$\begin{aligned} t' &= \frac{\sqrt{c^2 t^2 + v^2 t'^2}}{c} \\ &\Rightarrow t'^2 = t^2 + \left(v^2/c^2\right) t'^2 \\ &\Rightarrow t^2 = \left(1 - v^2/c^2\right) t'^2 \end{aligned}$$

and therefore

$$t = t' \sqrt{1 - v^2/c^2}. \tag{3.1}$$

So, if it takes time t for our clock to make a "tick", it takes time $t' = t/\sqrt{1 - v^2/c^2}$ — which is always greater than t — for the clock in the spaceship to make a tick. In other words, it looks to us as though the moving clock runs slowly by a factor $\sqrt{1 - v^2/c^2}$. Is this just an illusion — something to do with the type of clock? No: suppose instead we had a pair of another type of clock (mechanical, quartz, a "biological clock", whatever), that we agree keeps time with our simple "light clock". We keep one on Earth, and check that it keeps time. Our friend takes the other one along; but if he notices any discrepancy between the clocks, then we have a way to tell who is "really" moving — and that is not allowed!

Furthermore, while it appears to us that his clocks run slowly, it appears to him that our clocks also run slowly, by exactly the same reasoning! Each sees the other clock as running more slowly.

A time span as measured by a clock in its own rest frame is called the *proper time*. This is not meant to imply that there is anything wrong with the measurement made from another frame, of course.

3.2.1 Warning

You should be *very careful* when you use formula 3.1 — it is very easy to become confused and to use it the wrong way around, thus contracting instead of dilating time!

3.2.2 The Lifetime of the Muon

How can we test this, without having any extremely fast spaceships handy? One early test was provided by looking at the lifetime of a particle called the mu-meson, or muon. These are created in cosmic rays high in the atmosphere, and they decay spontaneously after an average of about 2.2×10^{-6} s; thus, even travelling close to the speed of light, they should not be able to travel more than about 600 m. But because they are moving so close to the speed of light, their lifetime (as measured on the Earth) is "dilated" according to equation (3.1), and they live long enough to reach the surface of the Earth, some 10 km below... Muons have been created in particle accelerators, and their lifetimes measured as a function of their speed; the values are always seen to agree with the formula.

3.2.3 The Twins Paradox

The most famous apparent "paradox" in relativity concerns two twins, Peter and Paul. When they are old enough to drive spaceships, Paul flies away at very high speed. Peter, who stays on Earth, watches Paul's clocks slowing down; he walks and talks and eats and drinks more slowly, his heart beats more slowly, and he grows older more slowly. Just as the muons lived longer because they were moving, so Paul lasts longer too. When, in the end, he gets tired of travelling around and comes back to Earth to settle down, he finds that he is now younger than Peter! Of course, according to Paul, it is Peter who has moved away and come back again — so shouldn't Peter be the younger one?

In fact, the two reference frames are not equivalent. Paul's reference frame has been accelerating, and he knows this; he is pushed towards the rear of the spaceship as it speeds up, and towards its front as it slows down. There is an absolute difference between the frames, and it is clear that Paul is the one who has been moving; he really does end up younger than Peter.

3.3 Lorentz Contraction

Consider one of the apparently long-lived muons coming down through the atmosphere. In its rest frame, it is created at some time $t = 0$, at (let us say) the origin of coordinates, $x = 0$. At some later time t (but at the same spatial coordinate $x = 0$), the surface of the Earth moves up rather quickly to meet it. If, say, $t = 10^{-6}$ s, and $v = 0.995c$, the length of atmosphere that has moved past it is $vt = 300$ m. But this is far less than the 10 km distance separating the events in Earth's frame! It seems that not just time, but also length is changed by relative motion.

To quantify this, let the distance that the muon travels in Earth's frame be $x' = vt'$; note that, from the muon's point of view, it is the Earth that is moving, and so we use primes to denote Earth's reference frame. The length of atmosphere moving past the muon in its rest frame is $x = vt$. Thus,

$$\begin{aligned} x &= vt = vt'\sqrt{1 - v^2/c^2} \\ &= x'\sqrt{1 - v^2/c^2}. \end{aligned} \tag{3.2}$$

Therefore, the length of atmosphere x as measured by the muon is contracted by a factor $\sqrt{1 - v^2/c^2}$ relative to the length measured on Earth. Likewise, lengths in the muon's rest frame are contracted as seen by Earth-bound observers; but, because the muon is a point-like particle, such lengths are difficult for us to measure directly.

3.3.1 Alternative Approach

We can work this out again from scratch by mounting our spaceship clock in the direction of motion (see diagram). The light leaving the rear mirror is now "chasing" the front mirror, and has to move further to meet it than it would if the clock were stationary. The distance between the mirrors is l for our stationary clock on Earth; let's call it l' for the moving clock. If the time (as seen from Earth) for the light to travel from the rear to the front mirror is t'_1, then

$$ct'_1 = l' + vt'_1.$$

For the return journey, where the rear mirror moves up to meet the light, the distance is shorter:

$$ct'_2 = l' - vt'_2.$$

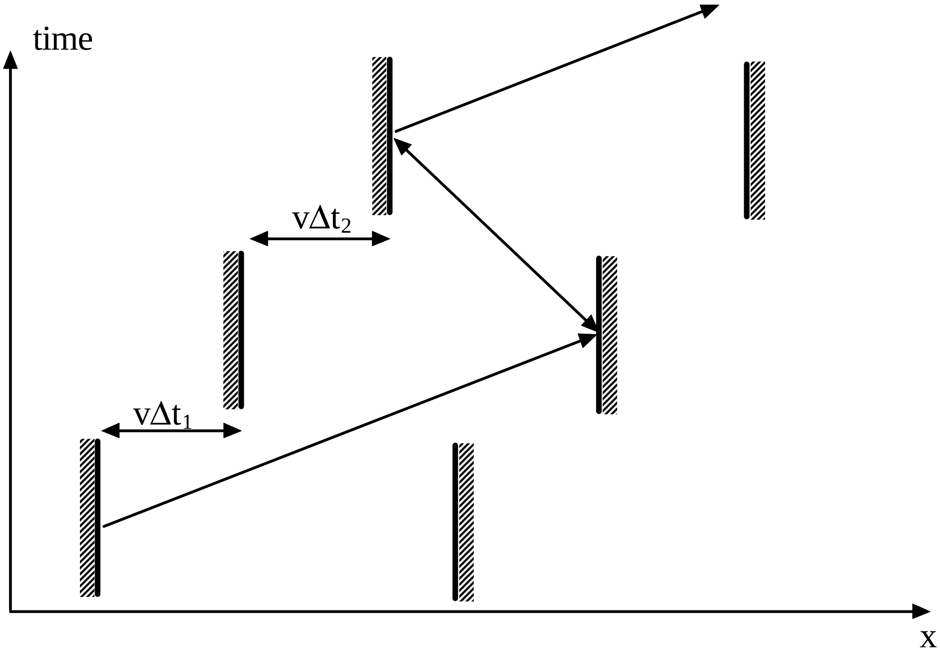

Figure 3.2: A light clock moving parallel to its axis. Note the vertical axis here represents time.

Therefore,

$$t'_1 \;=\; \frac{l'}{c-v};$$
$$t'_2 \;=\; \frac{l'}{c+v}$$

and the total time for the light to travel in both directions is

$$\begin{aligned} t' \;&=\; t'_1 + t'_2 = \frac{l'\,(c+v) + l'\,(c-v)}{(c-v)\,(c+v)} \\ &=\; \frac{2l'c}{c^2 - v^2}. \end{aligned}$$

Now, for the stationary clock the total "there-and-back" time is

$$t = \frac{2l}{c}.$$

But we already know that the moving clock is running slowly;

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

Therefore, the total time

$$\begin{aligned} \frac{2l'c}{c^2 - v^2} &= t' = \frac{t}{\sqrt{1 - v^2/c^2}} \\ &= \frac{2l}{c\sqrt{1 - v^2/c^2}}, \end{aligned}$$

and so

$$l' = l\sqrt{1 - v^2/c^2}.$$

This shows us that lengths are actually contracted in the direction of motion.

3.4 Derivation of Lorentz Transformations

We have seen that lengths and times are both modified when bodies are in motion. We now derive the Lorentz transformations; this involves just a little algebra; the procedure is entirely based on Einstein's two postulates.

Let us start with a fairly general set of linear transformations:

$$\begin{aligned} x' &= ax - bt \\ t' &= dx + et. \end{aligned}$$

The distances and times correspond to measurements in *reference frames* S, S'.We have here assumed

- common origins ($x' = 0$ at $x = 0$) at time $t = t' = 0$.

- linear transformations; there are no terms in, e.g., x^2. To see that this is reasonable, consider the point $x' = 0$, which is the origin of the S' frame. It is moving with velocity

$$v = dx/dt = b/a = \text{constant};$$

if there were x^2 terms present, the speed would depend on position, and the velocity would not be uniform.

- Now, the origin of S (i.e., $x = 0$) moves with velocity $-v$ in the S' frame, so for $x = 0$,

$$\begin{aligned} x' &= -bt,\ t' = et \\ &\Rightarrow \frac{dx'}{dt'} = -\frac{b}{e} = -v. \end{aligned}$$

  Therefore, $b = ev$; but since we have $b = av$, we know that $e = a$.

So far, then,

$$\begin{aligned} x' &= ax - avt \\ t' &= dx + at. \end{aligned}$$

- Now we require light to travel with velocity c in all frames, in other words $x = ct$ is equivalent to $x' = ct'$.

$$\begin{aligned} ct' &= a.ct - avt \\ t' &= d.ct + at. \end{aligned}$$

Substitute for t':

$$dc^2t + act = act - avt.$$

Therefore,

$$dc = -av/c,$$

giving

$$\begin{aligned} x' &= a\,(x - vt) \\ t' &= a\left(t - vx/c^2\right). \end{aligned} \tag{3.3}$$

The constant a we can guess from our previous look at time dilation and space contraction... but here we derive it by requiring symmetry between the observers. Suppose that we are sitting in the S' frame, looking at the S frame. We certainly expect the same relation to hold true the other way around, if we swap

$$\begin{aligned} x &\longleftrightarrow x' \\ t &\longleftrightarrow t', \end{aligned}$$

but we have to remember that the velocity is reversed:

$$v \to -v.$$

So,

$$\begin{aligned} x &= a\,(x' + vt') \\ t &= a\left(t' + vx'/c^2\right). \end{aligned} \tag{3.4}$$

As both 3.3 and 3.4 must hold simultaneously,

$$\begin{aligned} x' &= a\left\{a\,(x' + vt') - va\left(t' + vx'/c^2\right)\right\} \\ &= a^2 x'\left(1 - v^2/c^2\right) \end{aligned}$$

so

$$a = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

- We normally give a the symbol γ :

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}},$$

- ...and define $\beta = v/c$, so
$$\gamma = \frac{1}{\sqrt{1-\beta^2}}.$$

The complete Lorentz transformations are therefore
$$\begin{aligned} x' &= \gamma\left(x - \beta.ct\right) \\ y' &= y \\ z' &= z \\ ct' &= \gamma\left(ct - \beta x\right). \end{aligned} \tag{3.5}$$
The inverse transformations are then
$$\begin{aligned} x &= \gamma\left(x' + \beta.ct'\right) \\ ct &= \gamma\left(ct' + \beta x'\right), \end{aligned}$$
with the y and z coordinates remaining unaffected as before. Using ct instead of just t gives all of the variables the dimensions of distance, and displays the implicit symmetry between the first and last transformation equations.

Notice, as usual, the limiting speed of light: if one frame moves faster than light with respect to another, γ becomes imaginary, and one or other frame would then have to have imaginary coordinates x', t'.

3.4.1 Matrix Formulation

It is clear that (3.5) can be written as
$$\begin{pmatrix} x' \\ y' \\ z' \\ ct' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ ct \end{pmatrix}.$$

3.5 Help! When Do I Use Which Formula?

A little thought will tell you when you can use the simple Lorentz contraction / time dilation formulae, and when you must use the full Lorentz transformations.

We saw earlier that, if Bill wants to calculate the length of Tony's train as it passes, he can either measure both ends simultaneously or else make separate measurements but allow for the distance the train has moved in the time between measurements. However, when we specify *events*, the distance between them depends upon the frame of reference even in Galilean relativity. For example, a 100 m long train is moving at 50 m/s. At $t = 0$, the driver at the front spills his coffee; one second later, the guard at the back of the train drops his sandwich. From Tony's point of view, these events took place 100 m apart; but in Bill's frame of reference, they are only separated by 50 m. The Galilean transformations take care of the distance the guard travels forwards during the 1 s between the events. The same principle applies in Einstein's relativity, of course.

Basically, then, if you are concerned with objects — rulers, rockets, galaxies — you can calculate the lengths in the various frames of reference just by using the Lorentz contraction. Likewise, if you are concerned with the time that has passed on a clock (atomic, biological, whatever) that is at rest in one particular frame of reference, you can safely use the time dilation formula. But, if you are considering separate events, each with their own space and time coordinates, you must use the full Lorentz transformations if you want to get the right answer.

3.6 The Doppler Shift

Consider a source of light of a given frequency ν, at rest. Every $t = 1/\nu$ seconds it emits a new wavefront.

Now let the source move slowly towards us (so that there are no relativistic effects), at speed u. During the time t between one emission and the next, the source "catches up" a distance ut with its previous wavefront, so the distance λ between wavefronts is no longer c/ν but instead

$$\begin{aligned} \lambda &= \frac{c}{\nu} - ut = \frac{c}{\nu} - \frac{u}{\nu} \\ &= \frac{c}{\nu}\left(1 - \beta\right). \end{aligned}$$

Its apparent frequency is therefore

$$\nu' = \frac{c}{\lambda} = \frac{\nu}{(1-\beta)}.$$

This is the classical Doppler shift (as it applies to sound, for example). But when the source is moving towards us, its "clock" runs more slowly by a factor $\gamma = 1/\sqrt{1-\beta^2}$, so the rate at which it emits pulses is no longer ν but ν/γ. Therefore, the frequency actually shifts to

$$\begin{aligned} \nu' &= \frac{\nu}{\gamma\left(1-\beta\right)} \\ &= \nu\sqrt{\frac{1+\beta}{1-\beta}}. \end{aligned} \tag{3.6}$$

This is the equation for the *relativistic Doppler shift.* A classic example is the redshift seen due to the expansion of the universe. Hubble was the first to observe that the further away a galaxy lies from us, the faster it is moving away. As it moves away, the frequency of the light it emits drops, and the wavelength therefore increases towards the red end of the spectrum. (The opposite effect, the increase in frequency as sources approach, is called "blueshift"). From the distances and speeds of the galaxies, and taking into account the reduction in gravitational attraction as objects move apart, the age of the universe has been calculated to be about 14$\pm$ 2 billion years.

Note that (3.6) gives the measured frequency in a moving frame, starting with the source frequency. Often, of course, the source will be in the moving (primed) frame, and the shift is then inverted; just as with the Lorentz contraction and time dilation, it is important to think and to make sure that the shift is in the right direction. Later, we shall derive an expression for the Doppler shift for light emitted at a general angle θ to the direction of motion of the source.

3.7 Synchronisation of Clocks

Einstein, at age 14, is reported to have wondered what it would "look like" to ride along with a beam of light. He realised that time would appear to be "frozen"; if the light was emitted from a clock which said 12:00 noon, the image of that light would still say 12:00 noon when it arrived at the observer, however far away he was... This raises the question of what the time "really" is, and how we should measure it in a consistent way.

Events can be measured by a set of synchronised clocks and rigid rods with which the reference frame is provided. There are two ways in which we can synchronise distant clocks:

1. We can synchronise them when they are right next to each other and then separate them *very* slowly so that the time dilation correction is negligible.

2. We can send a beam of light from one clock to another. Suppose we send a light beam out in this way from clock A, which reads time t_1, to clock B and back (see diagram). Clock B reads time t_2 when the light pulse arrives, and A reads t_3 when the pulse returns to its starting point. The clocks are synchronised if t_2 is equal to the average of t_1 and t_3, in other words

$$t_2 = \frac{1}{2}(t_1 + t_3)\,.$$

This just allows for the travel time of the light between the clocks.

3.8 Summary

We have seen how (a) lengths contract along the direction of motion, and (b) moving clocks slow down. This is true for all observers in relative motion: just as A sees B's clocks slow down, so B sees A's clocks slow down. The apparent discrepancy is resolved because to measure the length of a moving object, you have to measure simultaneously the positions of the two ends, which are spatially separated; however, the observers disagree about the timing of spatially-separated

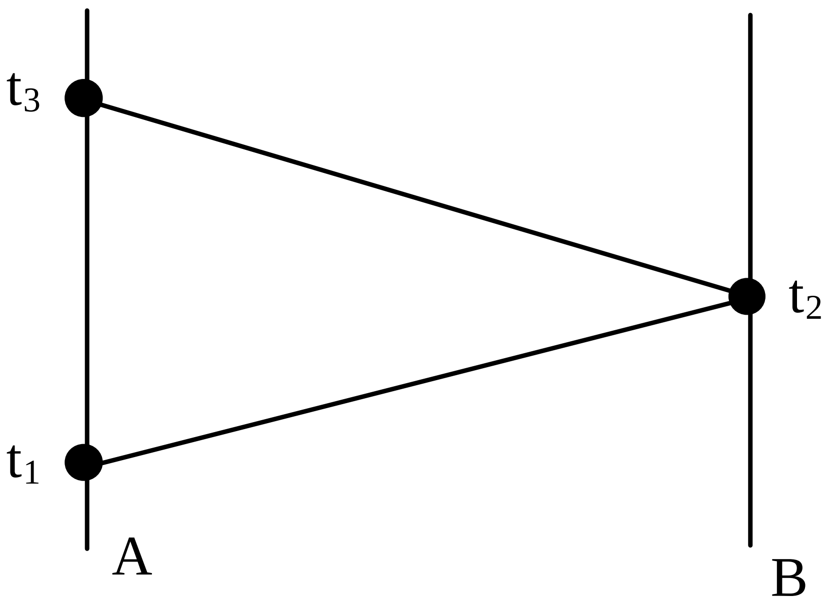

Figure 3.3: How to synchronise remote clocks.

events, and the disagreement in time measurements exactly compensates the disagreement in length measurements between the two frames.

Chapter 4

Spacetime

4.1 Spacetime Events and World Lines

We often represent space and time on a single *spacetime* diagram, with time on the vertical axis (see diagram). An *event* is something that happens at a given point in space and time, and is represented by a point $E\,(x,t)$ on a spacetime diagram.

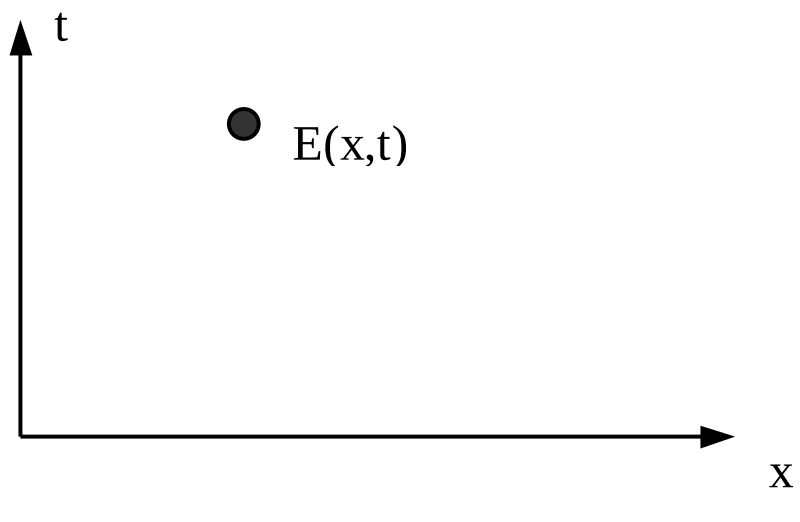

Figure 4.1: Spacetime diagram of an event.

The path traced out by an object in a spacetime diagram is called its *world line*.

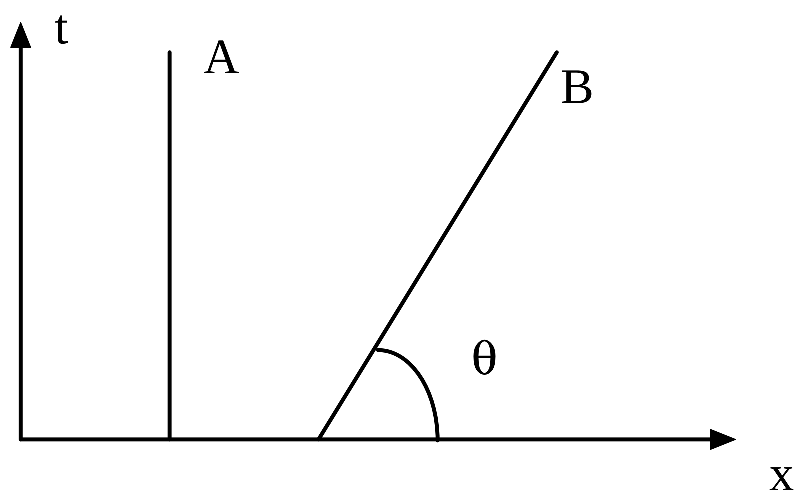

Figure 4.2: Spacetime diagram showing world lines of stationary and moving objects.

- A vertical line (A in diagram) represents a stationary object.
- An object moving with velocity v is represented by a line at angle θ to the horizontal (B in diagram), where
$$v = \cot\theta.$$
- A light ray has $\cot\theta = c$; we usually measure x in units of "ct" (e.g. "light seconds"), so $\theta = 45°$ for photons. Note that nothing can have a world line with $\theta < 45°$.

4.2 Intervals

Consider the quantity
$$S^2 = c^2t^2 - \left(x^2 + y^2 + z^2\right). \tag{4.1}$$
The last three terms are the distance (squared) from the origin to the point at which an event occurs; the first term is

the distance that light can travel in the available time. Let us evaluate the same quantity in another frame of reference, moving with velocity v and having the same origin when $t = t' = 0$.

$$\begin{aligned} S^2 &= c^2\gamma^2\left(t' + \frac{vx'}{c^2}\right)^2 - \gamma^2\left(x' + vt'\right)^2 - y'^2 - z'^2 \\ &= \gamma^2\left\{c^2t'^2\left(1 - \frac{v^2}{c^2}\right) + 2vt'x' - 2vt'x' - x'^2\left(1 - \frac{v^2}{c^2}\right)\right\} - y \\ &= c^2t'^2 - \left(x'^2 + y'^2 + z'^2\right) = S'^2. \end{aligned}$$

Therefore, the quantity S^2, which is known as the *interval*, is the same in *all* inertial reference frames; it is a *Lorentz invariant.* Remember this — it's important, and very useful! The behaviour is very similar to the way in which distance, in three-dimensional space, is invariant when we rotate our coordinate axes; the interval stays unchanged when we move from x, t to x', t' axes via a Lorentz transformation.

Spacetime can be divided into three parts, depending upon the sign of the interval:

1. $S^2 > 0$: Timelike (with respect to the origin). This class of events contains $x = 0$, $t \neq 0$, which corresponds to changes in time of a clock at the origin.

2. $S^2 < 0$: Spacelike (with respect to the origin). This class contains $t = 0, x \neq 0$ events, which are simultaneous with but spatially separated from the origin.

3. $S^2 = 0$: Lightlike (with respect to the origin). Rays of light from the origin can pass through these events. Consider a spherical wavefront of light spreading out from the

origin; at time t it has radius $r = ct$, so it is defined by the equation

$$x^2 + y^2 + z^2 = r^2 = c^2 t^2,$$

which of course defines an interval $S^2 = 0$. Naturally, in any other frame, light also travels at speed c, so its wavefront must be determined by the same equation, but using primed coordinates.

We can also define intervals between two events, instead of relating it to the origin. In this case,

$$S^2 = c^2 \left(t_2 - t_1\right)^2 - \left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2 + \left(z_2 - z_1\right)^2.$$

4.3 Simultaneity

Let's look more closely at the Lorentz transformations 3.5. The first three are just the same as the Galilean transformations, except that we have the length contraction factor γ in the direction of motion x. The fourth equation looks similar to the Galilean transformation $t' = t$ with the time dilation factor γ — but we now have an additional, unexpected term $\gamma v x / c^2$. What does it mean?

If, in the spaceship, there are two events that are separated in space — suppose they occur at positions x_1, x_2 — but *at the same time* t_0, then according to our observer on Earth, they occur at times

$$\begin{aligned} t'_1 &= \gamma \left(t_0 - \beta x_1 / c\right), \\ t'_2 &= \gamma \left(t_0 - \beta x_2 / c\right). \end{aligned}$$

In that case, the events as seen from the Earth are *no longer simultaneous*, but are separated by a time

$$t'_2 - t'_1 = \gamma\,(x_1 - x_2)\,\beta/c.$$

This is known as *failure of simultaneity at a distance*, and it lies at the heart of most of the problems and "paradoxes" of relativity. The whole idea of simultaneity just breaks down: events (separated in space) that are simultaneous in one reference frame are not in any other.

As an example, suppose the spaceship pilot is standing in the middle of his spaceship, and he emits a flash of light which reaches both ends at the same time. As seen by the man on Earth, though, the rear of the spaceship is moving up to meet the light, whereas the front of the spaceship is moving away from it; so the backwards-going pulse of light reaches the rear of the spaceship before the forward-going pulse reaches the front (see diagram).

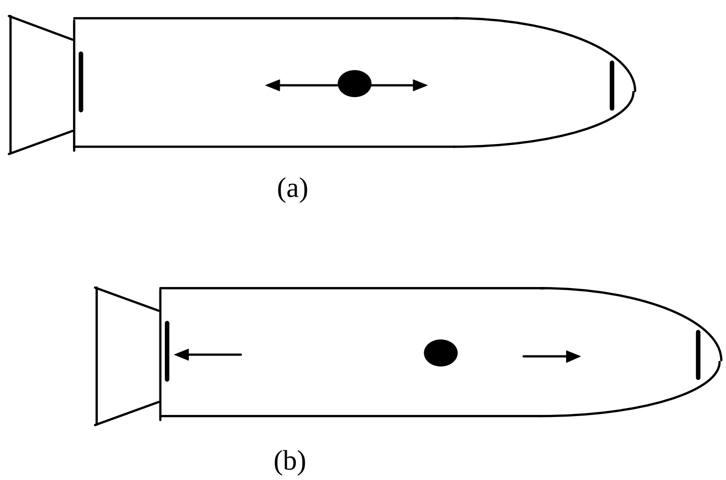

Figure 4.3: An example of the failure of simultaneity at a distance.

Note that from the point of view of a second rocket which is overtaking the first, the first one is moving backwards, in which case the light will hit the front mirror first.

In order to appear simultaneous in all Lorentz frames, a pair of events must coincide in both space and time — in which case they are really just one event.

4.4 The Principle of Causality

Easily stated — this is simply that

causes always occur before their resultant effects.

If A caused B, then A happened before B. That's all. It prevents nasty paradoxes: for example, you can't go backwards in time and prevent your own birth, because if you did, you would never have been born, and so you could not have travelled back... and so on. We don't have any proof of it; it's just an observation of the way things seem to be. It keeps life simple.

But in relativity, it has an interesting consequence.

Look again at the rocket in the previous section. In its frame of reference, light hits each mirror at the same time. In Earth's frame, light hits the back mirror first. In the frame of an overtaking rocket, light hits the front mirror first. Suppose that, just as light hits the back mirror, a guard standing beside it drops his sandwich; and just as light hits the front mirror, the pilot, standing beside it, spills his coffee. Is it possible that the guard dropping his sandwich caused the pilot to spill his coffee? Although it might seem like it from the Earth, it clearly cannot be, since from the overtaking rocket the pilot seemed to spill his coffee first. The order of the sequence depends upon the frame of reference, and so there is no cause-and-effect relationship.

On the other hand, if one event does cause another, it will occur first in *all* reference frames.This can happen if and only if the interval between them is timelike or lightlike; in other words, for one event to influence another, there has to have been enough time for light to propagate between them. Therefore, *it is clear that nothing — no information — can travel faster than the speed of light.* In fact,

- If events are physically related, their order is determined absolutely.
- If they are *not* physically related, the order of their occurrence depends upon the reference frame.

(N.b. "physically related" here means that information has had time to travel between the events, not necessarily that one caused another).

4.4.1 Example: Gunfight on a Train

Another example. Imagine a train that moves at 0.6 c. A man on the ground outside sees man A, at the rear of the carriage, start shooting at B, who is standing about 10 m ahead of him. After 12.5 ns, he sees B start shooting back.

But the passengers all claim that B shot first, and that A retaliated after 10 ns! Who did, in fact, shoot first?

The answer depends upon the frame of reference. Since the light could not travel the 10 m between the protagonists in the 10 ns or so required, the events are spacelike — there is no cause-and-effect relationship, and *the sequence is relative.*

4.4.2 Example: A Lighthouse

Suppose two satellites are positioned 6×10^8 m apart, and midway between them is a lighthouse, which rotates once every two seconds. The satellites are programmed to start broadcasting when the beam of the lighthouse sweeps past them. The "spot" of light passes one satellite, which duly begins transmitting; one second later, the spot of light has travelled halfway around the circle of radius 3×10^8 m, and it reaches the second satellite, which also begins broadcasting. The light spot has travelled 10^9 m in just one second — a speed of $10c$; and the satellites also began transmitting within a second of each other, despite being separated by 6×10^8 m — an apparent transmission of information at $2c$. Is this a problem? No: in fact, no information has passed between the satellites. They are merely responding to information that has been transmitted from the central point at speed c. If the first satellite's antenna had failed and it had not begun broadcasting, the second satellite would have known nothing about it, and would have started its own transmission regardless. The two events are *not* physically related.

4.5 Light Cones

If we include another space dimension in our spacetime diagram, the "world line" of light then defines a cone (actually a pair of cones), with its vertex at the origin and an opening angle of 45°.

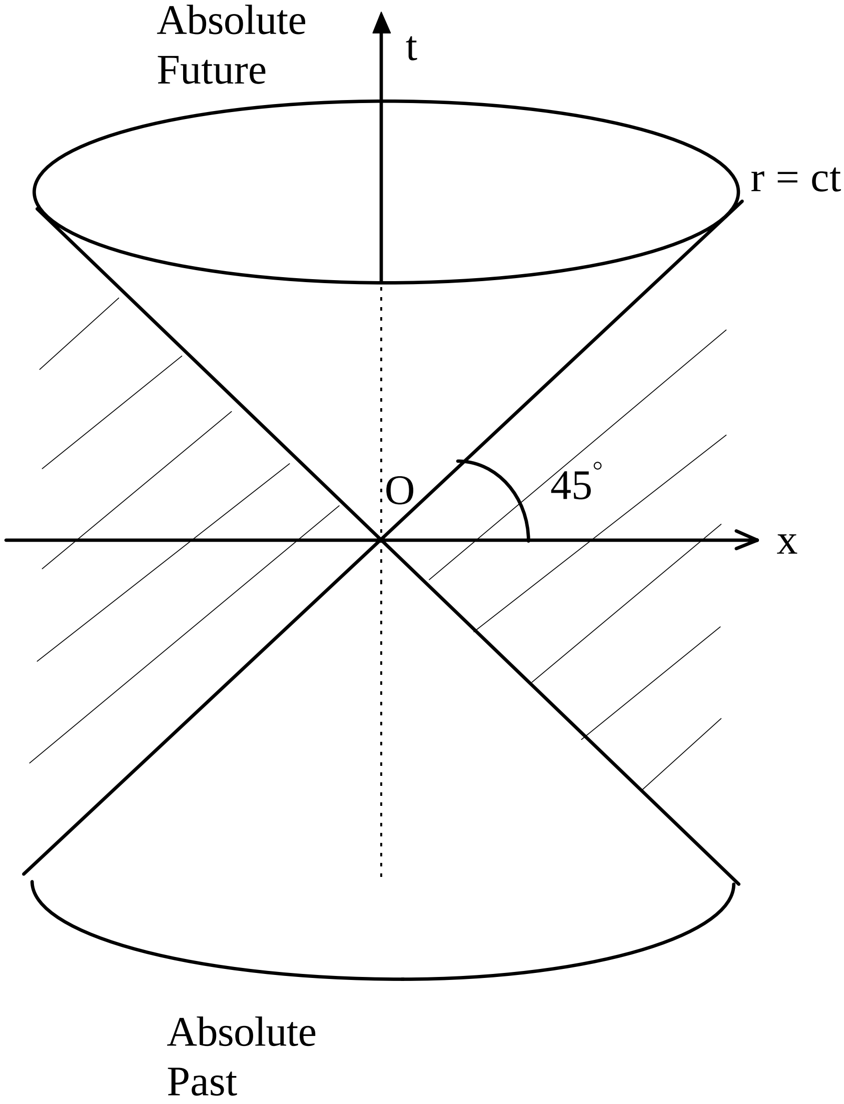

Figure 4.4: A light cone.

Events at the origin can be related to events inside the cone, since signals can travel between them at up to the speed of light; but an event at the origin *cannot* be related to events outside the light cone.

The region inside the cone and "above" the origin is therefore known as the *absolute future*, and that inside the cone below the origin is the *absolute past*.

4.6 Visual Appearance of Moving Objects

Consider a cube moving past, close to the speed of light. Naturally, it is Lorentz contracted, and this can be measured, for example, by the set of clocks (at rest and synchronised in the laboratory frame) which all read the same time at the moment that the corners of the cube pass them. In this way, the time lags for light to travel from the different corners of the cube are eliminated.

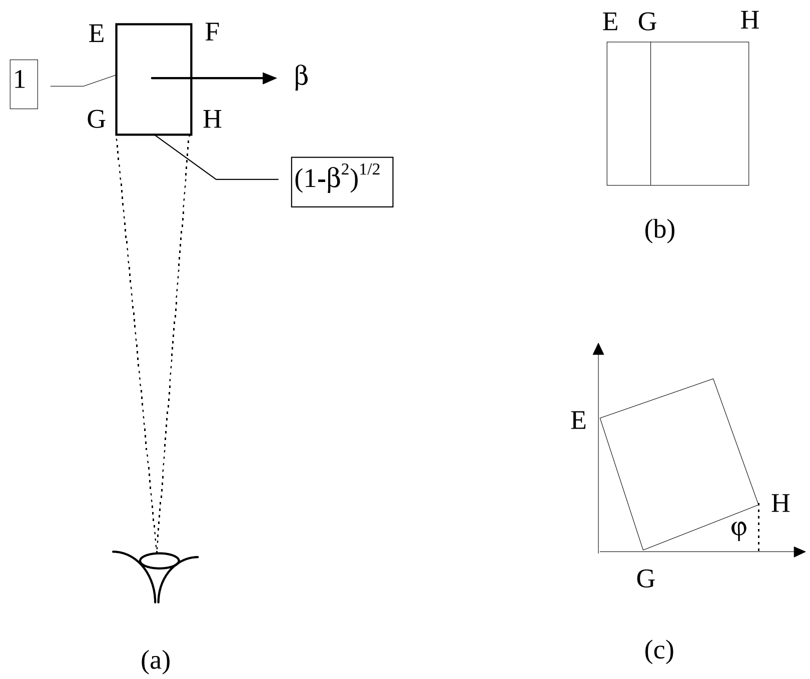

Figure 4.5: (a) A cube passing by an observer, as seen in the laboratory frame. (b) What the observer sees as he looks up. (c) How the observer interprets what he sees.

But our eye does not work like this! It can only be in one place at a time, and it registers only light that enters at the same time. Hence, what one sees may be different from the measurements made by a lattice of clocks.

When it is at 90° (see figure), photons from the nearest corners of the cube arrive simultaneously, and so one will see the normal Lorentz contraction of the bottom edge. (Here we assume that the cube is quite far away). However, light from one of the further rear corners (E in the diagram) that left that point earlier can arrive at the eye at the same time! The observer sees behind the cube at the same time — and it therefore appears to him to be rotated.

Figure (b) shows the appearance of the cube as the observer looks up at it; Figure (c) shows how the observer might interpret it as a rotation.

Different configurations can be far more complicated. A rod that is approaching an observer almost head on will appear to be *longer*, despite the Lorentz contraction, because light emitted from the rear takes some time to "catch up with" light emitted from the front.

A diagram is included to show the approximate appearance of a plane grid, moving at relativistic speeds past an observer. The observer is at unit distance in front of ("above") the origin, i.e. the line from observer to origin is perpendicular to the direction of motion. (This diagram is a rough copy of a plot from Scott, G.D., and Viner, M.R., Am. J. Phys. 33, 534, 1964).

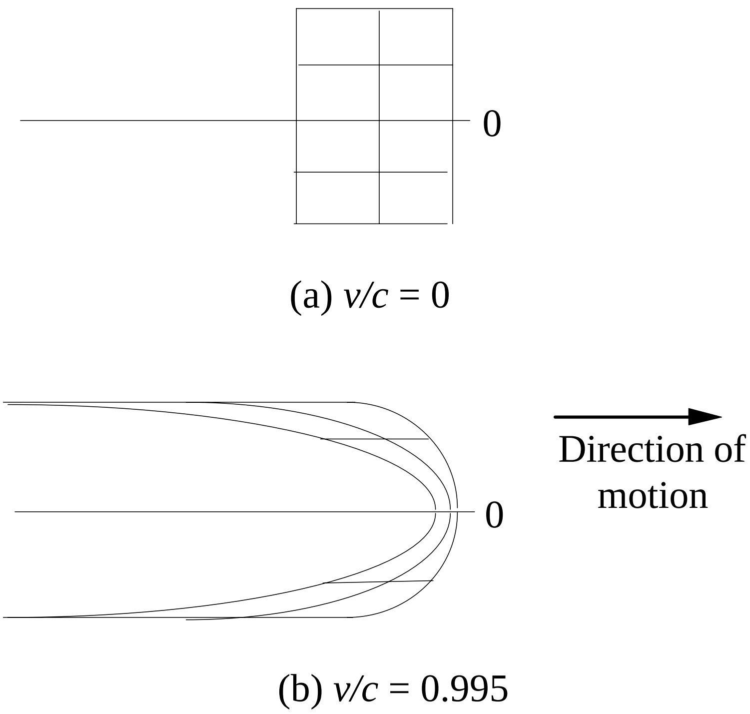

Figure 4.6: Distortion of a grid moving at relativistic velocities.

4.7 Oblique Axes

Consider a clock A moving in reference frame S. Its world line is at angle $\theta_0 = \cot^{-1} v$ to the horizontal (see figure 4.7).

In frame S' moving with the clock, the clock is at a fixed x', so its world line must be parallel to the time axis t'.

The x' axis is determined by

$$\begin{aligned} t' &= 0 \\ &\Rightarrow t = \frac{\beta x}{c}, \end{aligned}$$

and therefore it is at an angle $\theta_{x'} = \tan^{-1}(\beta/c)$ to the x axis.

Axes x', t' are *oblique* in this representation. The coordinates of an event are determined by drawing lines parallel to the axes (see diagram). This procedure reduces to dropping perpendiculars if the axes are orthogonal.

Oblique axes are essential to represent the coordinates of one event in different frames... but one *cannot* think of "distances" between points as that obtained by direct measurement with a ruler!

As an example, let's look again at the issue of simultaneity. We considered a man at the centre of a spaceship (B) emitting a flash of light to the rear (A) and the front (C) of his spaceship. In his frame of reference, the world lines are as shown in the first diagram, with the light reaching the two ends of the spacecraft simultaneously.

From the Earth's point of view, with the spaceship moving past, the world lines are tilted. Light, however, still travels along its 45° world lines. We can see that the light will intercept the rear of the spaceship before it reaches the front. Clearly, in the spaceship's frame of reference, the rear of the craft is not moving, so the time axis t' is parallel to A's world line; and, since in the rocket the light reaches both ends simultaneously, the axis of constant time (x') must be parallel to the line joining the two events where the light crosses the world lines A and C.

This representation provides useful imagery, but it is not very practical for solving problems! You won't need it for your exams.

4.8 More Paradoxes

4.8.1 The Pole and Barn Paradox

A pole vaulter carries a 20-metre long pole. He runs so incredibly fast that it is Lorentz contracted to just 10 m in Earth's reference frame. He runs into a 10 m long barn; just at the moment when the pole is entirely contained inside the barn, the doors are slammed closed, trapping both runner and pole inside.

However, from the runner's point of view, it is the barn that is contracted to half its length. How can a 20-metre pole fit inside a barn that appears to be just 5 m long?

The answer, of course, lies in the failure of simultaneity. In the barn's reference frame the doors are closed at the same instant; to the poor runner, however, it appears that the exit door is closed, and his pole runs into it, *before* the rear of his pole has completely entered the barn.

4.8.2 The Thin Man and the Grid

A man (perhaps our speedy pole vaulter) is running so fast that Lorentz contraction makes him very thin. Ahead of him in the street there is a grid. A man standing beside the grid expects the thin runner to fall through one of the spaces in the grid. But to the runner, it is the grid that is contracted, and since the holes are much narrower, he does not expect to fall through them. What actually happens?

This is actually a rather subtle problem that hinges on the question of rigidity. There is, in fact, no such thing as a perfectly rigid rod. Consider a bridge from which the support at one end is suddenly removed. That end starts to fall at once. The rest of the bridge, however, stays as solid as ever, until the information that the support is missing reaches it — in this case, with a speed determined by the time required for an elastic wave to move through the steel.

Consider next a rod lying on a ledge in a rocket. The ledge suddenly collapses and the rod falls down with the acceleration of gravity. But in the Earth's frame, the end at the back of the rocket starts falling first, before the ledge at the front end has collapsed at all. The rod thus appears bent — and *is* bent — in Earth's frame.

As for the man and the grid, let us replace the problem with a "rigid" metre-long rod sliding along a table in which there is a metre-wide hole. In the frame of reference of the hole (grid), the rod (man's foot) simply falls into the hole. From the point of view of the rod (runner), though, the front end of the rod (foot) droops over the edge of the hole when the support underneath vanishes, and the rest of it comes following after.

4.8.3 Twins Paradox Revisited

Here we look again at the twins paradox, using a slightly different approach. We will not have time to go through this in the lectures; it may take a little while to understand what is happening to the clocks in each reference frame, but it is worth the effort.

Twin B leaves twin A with relative velocity v, reverses his velocity and returns to find less time has elapsed on his clock than on A's. To A, this is in agreement with time dilation, but B saw A moving with respect to him and so thinks that A's clock should show less time elapsed.

To avoid effects of accelaration, we introduce a third "twin" C who has velocity $-v$ with respect to A and who coordinates his clock with B as they cross. We now list the (x, t) coordinates of the important events in the various inertial frames:

	A's frame	B's frame	C's frame
B leaves A	$0, 0$	$0, 0$	$-2vt/\gamma, 0$
B and C cross	vt, t	$0, t/\gamma$	$0, t/\gamma$
C returns to A	$0, 2t$	$-2vt/\gamma, 2t/\gamma$	$0, 2t/g$

We see that the total time elapsed in frame A is $2t$, compared to the time elapsed in B and C which is $2t/\gamma$. Therefore, a greater time has elapsed in A than in B or C, as expected.

The paradox can be resolved by considering the time on A's clock at the instant when B and C meet. There are two events to consider: (1) B meeting C, (2) recording the time on A's clock. These events are spatially separated, and therefore subject to the usual failure of simultaneity at a distance. The crucial point is that since A is spatially separated from the B, C meeting point, the reading on A's clock will depend upon the frame of reference in which the simultaneity between the two events is assumed. Thus:

A's frame: A's time is clearly t, coincident with the point (vt, t) in A's frame at which B and C meet.

B's frame: In this frame the appropriate coordinates of A at the meeting time are $(-vt/\gamma, t/\gamma)$, which transform to $(0, t/\gamma 2)$ in A's frame; in other words, from B's point of view, the B-C crossing occurs simultaneously with A's clock reading $t/\gamma 2$. Since $t/\gamma 2 < t/\gamma$, from B's point of view A's clock is going more slowly, as expected.

C's frame: In this frame the appropriate coordinates of A are also $(-vt/\gamma, t/\gamma)$, which in this case are transformed to $(0, 2t - t/\gamma 2)$ in A's frame. Again the change in A's clock during the return journey is $t/\gamma 2 < t/\gamma$, and therefore from C's point of view also, A's clock is going more slowly.

In other words, both B and C think that A's clock is going more slowly; and they do set their clocks to agree with each other at the point at which they cross; but because A's clock is spatially separated from their meeting point, the relative velocity between B's and C's frames gives rise to a disagreement about the reading on A's clock that is simultaneous with their meeting.

A's various clock readings are:

0	Start
$t/\gamma 2$	B and C meet, according to B
t	B and C meet, according to A
$2t - t/\gamma 2$	B and C meet, according to C
$2t$	Finish

The discrepancy $2t - 2t/\gamma 2$ between B and C is of just the right value to resolve the conflict between the view held by B and C that A's clock is going more slowly and the actual result that C's clock reads less than A's at the end.

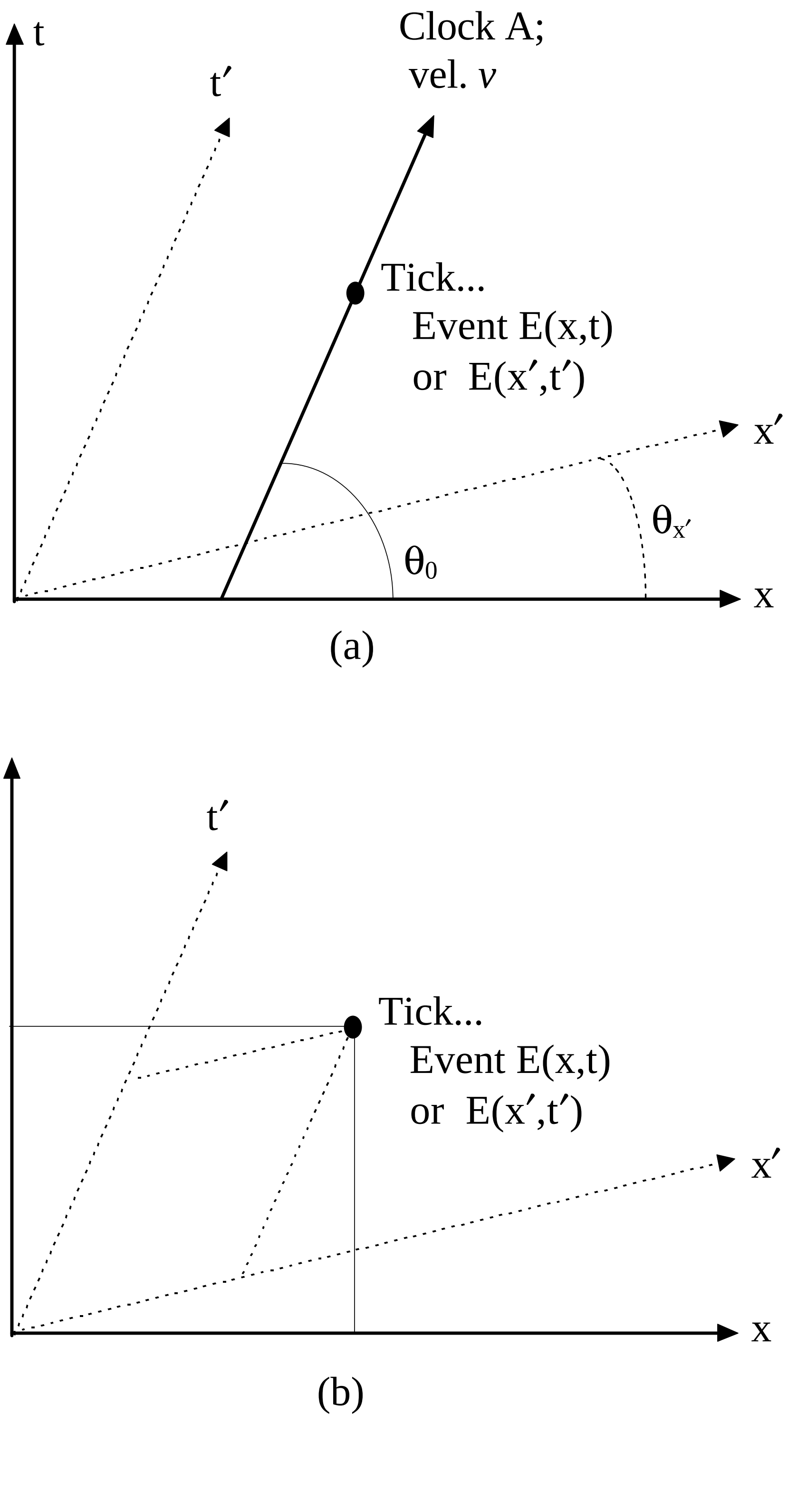

Figure 4.7: The oblique axes required to describe moving objects.

Chapter 5

Dynamics and Kinematics

5.1 Relative Velocities

A train moves at 0.8 c. A passenger fires a bullet which moves at 0.6 c relative to the train. How fast does it move relative to the ground?

A Galilean transformation would give

$$v = 0.8c + 0.6c = 1.4c.$$

This is clearly wrong...

5.2 Lorentz Transformation of Velocities

Consider an object moving with velocity $\mathbf{u} = u_x\widehat{x} + u_y\widehat{y}$ in reference frame S. (Note that the transverse directions $\widehat{y}$ and $\widehat{z}$ are equivalent). So,

$$u_x = \frac{dx}{dt}, \quad u_y = \frac{dy}{dt}.$$

In frame S', moving with velocity v along the x axis, the same object will have velocity

$$u'_x = \frac{dx'}{dt'}, \quad u'_y = \frac{dy'}{dt'}.$$

Since

$$\begin{aligned} x &= \gamma\left(x' + \beta.ct'\right) \\ y &= y' \\ ct &= \gamma\left(ct' + \beta x'\right), \end{aligned}$$

we have

$$\begin{aligned} dx &= \gamma\left(dx' + \beta.cdt'\right) = \gamma\left(u'_x + v\right) dt' \\ dy &= dy' = u'_y.dt' \\ dt &= \gamma\left(dt' + v.dx'/c^2\right) = \gamma\left(1 + v.u'_x/c^2\right) dt'. \end{aligned}$$

Therefore,

$$\begin{aligned} u_x &= \frac{dx}{dt} = \frac{u'_x + v}{1 + v.u'_x/c^2} \qquad (5.1) \\ u_y &= \frac{dy}{dt} = \frac{u'_y}{\gamma\left(1 + v.u'_x/c^2\right)} \\ u_z &= \frac{dz}{dt} = \frac{u'_z}{\gamma\left(1 + v.u'_x/c^2\right)} \end{aligned}$$

and, correspondingly,

$$\begin{aligned} u'_x &= \frac{u_x - v}{1 - v.u_x/c^2} \qquad (5.2) \\ u'_y &= \frac{u_y}{\gamma\left(1 - v.u_x/c^2\right)} \\ u'_z &= \frac{u_z}{\gamma\left(1 - v.u_x/c^2\right)}. \end{aligned}$$

Note that velocities in the transverse direction *are* altered.

We can now answer the question about the bullet in the train; an object moving with velocity $u'_x = 0.6c$ in a frame that is itself moving at velocity $v = 0.8c$ moves (in our, unprimed, frame) at a velocity

$$u_x = \frac{u'_x + v}{1 + v.u'_x/c^2} = \frac{(0.6 + 0.8)\,c}{1 + (0.6)\,(0.8)} = 0.95c.$$

What about light itself? Suppose, instead of a bullet, the man fires a laser; the photons move at $u'_x = c$. In this case, we find that

$$u_x = \frac{(1.0 + 0.8)\,c}{1 + (1.0)\,(0.8)} = c,$$

as we expect. This is a nice cross-check.

5.3 Successive Lorentz Transformations

Consider three reference frames, S, S' and S'' — say, of someone standing on the ground; of someone moving past in a train; and of someone moving past in a train. What happens if we Lorentz transform from the first to the second frame, and then again from the second to the third? Is it consistent with transforming directly from the first to the third frame?

Let S' move along the x axis of frame S with velocity $v = \beta c$, so

$$\begin{aligned} x' &= \gamma\,(x - \beta ct) \\ ct' &= \gamma\,(ct - \beta x)\,. \end{aligned}$$

Likewise, let S'' move along the x' axis (which is the same as the x axis) with velocity $v' = \beta' c$, so

$$\begin{aligned} x'' &= \gamma'\,(x' - \beta' ct') \\ ct'' &= \gamma'\,(ct' - \beta' x')\,. \end{aligned}$$

We can substitute for x' and t' to give

$$\begin{aligned} x'' &= \gamma' \left(\gamma \left(x - \beta ct\right) - \beta' \gamma \left(ct - \beta x\right)\right) \\ &= \gamma\gamma' \left(1 + \beta\beta'\right) x - \gamma\gamma' \left(\beta + \beta'\right) ct; \\ ct'' &= \gamma' \left(\gamma \left(ct - \beta x\right) - \beta' \gamma \left(x - \beta ct\right)\right) \\ &= \gamma\gamma' \left(1 + \beta\beta'\right) ct - \gamma\gamma' \left(\beta + \beta'\right) x. \end{aligned}$$

Since, in S, the frame S'' is moving with velocity

$$v'' = \frac{v + v'}{1 + v.v'/c^2},$$

it can be shown (with a few lines of messy algebra) that the corresponding γ factor is

$$\gamma'' = \gamma\gamma' \left(1 + \beta\beta'\right).$$

Therefore, we find that

$$\begin{aligned} x'' &= \gamma'' \left(x - \beta'' ct\right) \\ ct'' &= \gamma'' \left(ct - \beta'' x\right). \end{aligned}$$

In other words, two successive Lorentz transformations give another Lorentz transformation, corresponding to the appropriate relative velocity.

5.4 Velocity Parameter

You will sometimes come across a quantity known as the *velocity parameter*. This is analagous to the idea of adding angles instead of slopes in normal geometry; two angles θ_1, θ_2 correspond to slopes $S_1 = \tan\theta_1$, $S_2 = \tan\theta_2$, but to add the angles we use

$$\theta_{\text{tot}} = \theta_1 + \theta_2$$

whereas the resulting slope has a more complicated addition law:

$$S_{\text{tot}} = \tan\theta_{\text{tot}} = \frac{\tan\theta_1 + \tan\theta_2}{1 - \tan\theta_1 \tan\theta_2}.$$

This is very similar to our formula for the addition of velocities, and indeed, we find that if we define a *velocity parameter* θ such that

$$\beta = \tanh\theta$$

then the parameter θ adds linearly like an angle. Notice that we have to use the hyperbolic tangent tanh, rather than an ordinary "tan", because of the same signs in numberator and denominator in 5.1. Tanh is defined as

$$\tanh\theta = \frac{e^{\theta} - e^{-\theta}}{e^{\theta} + e^{-\theta}}.$$

In practice, we will not make much use of this function; you won't need it for your exams.

5.5 Relativistic Dynamics

As we pointed out at the beginning, our classical laws of conservation of momentum and energy require modification in the relativistic limit. However, we can obtain laws closely related to the originals if we allow "mass" to vary with velocity in the way that we have allowed lengths and times to vary.

Let us look at a "bomb" (or particle) of mass M_0, which breaks up into two fragments, each of mass m_u that move off with equal and opposite velocities u (see diagram).

Now, let us move to the rest frame of the left-going fragment (see diagram). In that frame, the initial particle is moving

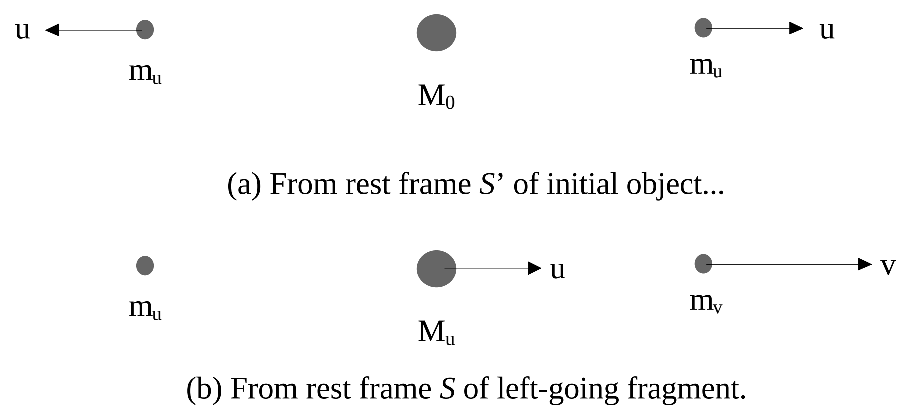

Figure 5.1: The breakup of a massive object, seen from the reference frames of (a) the parent object, (b) one of the fragments.

to the right with velocity u, whereas the other fragment is moving even faster to the right, with velocity

$$v = \frac{u+u}{1+u.u/c^2} = \frac{2u}{1+u^2/c^2},$$

according to our law of addition of velocities. Since we are allowing masses to vary with velocity, we now call the initial mass M_u; the stationary fragment has mass m_0 and the other fragment has mass m_v.

We assume that

1. something that we shall call the total "mass" is conserved;
2. "momentum" (= mass x velocity) is also conserved.

Therefore,

$$\begin{aligned} m_v + m_0 &= M_u \\ m_v.v + 0 &= M_u.u \end{aligned}$$

so

$$m_v + m_0 = M_u = m_v.v/u = m_v.\frac{2}{1+u^2/c^2}.$$

Therefore,

$$m_0 = m_v\left\{\frac{2}{1+u^2/c^2} - 1\right\} = m_v\left\{\frac{2-1-u^2/c^2}{1+u^2/c^2}\right\} = m_v.\frac{1-u^2/c^2}{1+u^2/c^2}. \tag{5.3}$$

But we need to know it in terms of v, not u. Let us consider our old friend

$$1 - v^2/c^2 = 1 - \frac{4u^2/c^2}{(1+u^2/c^2)^2} = \frac{1+2u^2/c^2+u^4/c^4-4u^2/c^2}{(1+u^2/c^2)^2} = \frac{\left(1-u^2/c^2\right)^2}{(1+u^2/c^2)^2}.$$

But this is just the square of the ratio of "stationary" to "moving" masses in 5.3. So, we finally obtain

$$m_v = \frac{m_0}{\sqrt{1-v^2/c^2}} = \gamma m_0$$

as the quantity that we called "mass" and that is conserved in this reaction. It acts just as though mass itself were increasing with velocity – and indeed, in many textbooks "relativistic mass" is defined this way, with m_0 being called the "rest mass". However, for consistency, it is best (and more common nowadays) to consider the mass to be defined as being measured at rest, and the γ is always written explicitly when required. **From now on, m or m_0 will refer to the rest mass only.** We shall drop the notation m_v and the idea of "mass" varying with velocity; so-called "relativistic mass" will be written explicitly as γm.

For the above problem, then, if we redefine the (rest) mass of the parent to be M, and the (rest) mass of each of the daughter products to be m, we find that everything is consistent if $M = 2\gamma m$.

What happens to this quantity γm at low velocities? Let us do a binomial expansion:

$$\begin{aligned}\gamma m &= m\left(1 - v^2/c^2\right)^{-1/2} \\ &= m + \frac{1}{2}m.v^2.\frac{1}{c^2} + \frac{3}{8}m\frac{v^4}{c^4}...\end{aligned}$$

The second term is the classical kinetic energy, divided by the constant c^2. This suggests that, if we multiply the entire expression by c^2, we will obtain an energy:

$$E = \gamma mc^2 = mc^2 + \frac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2}... \tag{5.4}$$

(We have dropped the subscript v from m). The first term of this expansion is called the "rest energy", and the second term is, as we noticed, the classical kinetic energy. However, we see that the true kinetic energy – literally, movement energy, so total energy minus the energy that the object has anyway when it's standing still – is

$$T = E - mc^2$$

$$= \frac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2} + ...,$$

and it actually has higher-order terms contributing; the classical expression is just a first approximation.

If we define a momentum

$$p = \gamma mv,$$

we see that

$$\begin{aligned} p^2 + mc^2 &= \frac{m^2v^2}{(1-\beta^2)} + m^2c^2.\frac{\left(1-\beta^2\right)}{(1-\beta^2)} \\ &= \frac{m^2v^2 + m^2c^2 - m^2v^2}{(1-\beta^2)} \\ &= \frac{m^2c^2}{(1-\beta^2)} \\ &= \gamma^2m^2c^2 = m^2c^2 = E^2/c^2, \end{aligned}$$

so

$$E^2 = p^2c^2 + m^2c^4. \tag{5.5}$$

This equation is *extremely* useful in relativity! The quantity $E^2 - p^2c^2 = m^2c^4$ just depends on the total rest energy of an object (or system), and so is independent of the reference frame in which it is measured, in just the same way as the interval S^2 that we met earlier.

Another equation that is sometimes useful:

$$p.c = \gamma mv.c = E.\beta$$

Finally, note that if a particle is in a potential of some sort, its potential energy of course must add to the expression 5.4 for the total energy.

5.6 Consequences

Equations 5.4 and 5.5 have a lot of interesting consequences...

- Equivalence of mass and energy. We postulated a "conservation of mass" law, but it has turned out that mass and energy are *the same thing*, just measured in different units (and related by the constant c^2). So the law has become equivalent to the conservation of energy.

- Notice that the rest mass is *not* conserved. When the object of mass M_0 split up, it produced two objects each of so-called "mass" (actually γ times mass) m_u; but the sum of the rest masses of the two fragments is $2m_0$, which is less than M_0; the "extra" mass in M_0 has been "converted" into kinetic energy of the fragments. The same is true in reverse; if we put two fragments of mass m_0 together gently, we end up with an object of mass $2m_0$ — which is therefore a *different* object than we get if we push the two fragments together hard, since the extra kinetic energy "turns into" extra mass.

- If an object splits up in this way, when the fragments interact with material, they deposit energy (as heat, kinetic energy and so on) until they come to rest. The total energy "liberated" then is

$$M_0c^2 - 2m_0c^2.$$

  The atomic bomb worked so well because the fragments (iodine, xenon etc) produced when the uranium atom splits are so much lighter than the initial uranium atom itself.

- The principle also works "in reverse": if you want to split up a carbon dioxide molecule into carbon and oxygen atoms, you have to add energy — so the sum of the masses of the carbon and two oxygen atoms is actually *greater* than the mass of the carbon dioxide molecule. (The sun is powered by the energy liberated when hydrogen atoms "fuse" together to make helium: it is losing mass at a rate of about 5 million tons per second).

- We see that a little mass is equivalent to a lot of energy. One gram of material is equivalent to 9×10^{13} J of energy — that's about 3 MW of power for a year, or the equivalent of 20 ktons of TNT.

- A peculiar consequence, without any particular practical use: mass can be transferred without exchange of either particles or radiation. Consider a board, on one end of which is a motor. This turns a drive belt, which itself turns a rotor at the other end of the board. Energy, and therefore mass, is transferred from the motor to the rotor, i.e., from one end of the board to the other. If it were mounted on frictionless rails, it would accelerate in the opposite direction so as to conserve momentum.

- What about light? Since

  $$E = \gamma m c^2,$$

  but $\gamma \to \infty$ as $v \to c$, the only way we can stop this from diverging is if photons have "zero rest mass": $m = 0$. They can still carry energy, though, by virtue of their momentum; from 5.5, we see that for photons,

$$E = pc.$$

What happens if you "stop" a photon? You cannot! A photon *always* travels at the speed of light!

5.7 Lorentz Transformation of Energy-Momentum

Consider a dynamic "event" like the particle breakup we saw earlier. What do the energy and momentum look like from a different coordinate system? It turns out, in fact, that they transform in a manner very similar to space and time. Here we will derive the transformation in two different ways; firstly, using familiar ideas but quite a lot of tedious algebra; and then in a much simpler way, treading on perhaps slightly less familiar territory. Let us first map out the road we shall follow.

5.7.1 The Difficult Way

This way is rather heavy in algebra, and I won't go through it in the lectures. However all of the elements are familiar, so there aren't any tricky concepts to get hold of.

- Starting point: Imagine a spaceship moving (relative to the Earth) with velocity v along the x axis, and a particle moving with velocity u.
- Notation: γ refers to the *spaceship*, not the particle; i.e. it is $\left(1 - v^2/c^2\right)^{-1/2}$ not $\left(1 - u^2/c^2\right)^{-1/2}$. The quantities u', E', γ', p' are the velocity, energy, "gamma" and momentum of the particle as seen from the spaceship. Earth's reference frame is S, the spaceship's is S'.

- Goal: to calculate E', p' in terms of E, p.
- Procedure:
  - We begin with the x component; let $u = u_x$.
  - Find u' and then γ' in terms of u and v and then substitute into

$$\begin{aligned} E' &= \gamma' m_0 c^2 \\ p'_x &= \gamma' m_0 u'. \end{aligned}$$

  - For the transverse components, we consider a particle moving in the y direction, and find a new γ' based on the new u'. From this, we can calculate p'_y (and check that the energy transformation is still valid).

Okay, here we go...

The velocity u' of the particle as seen from the spaceship is (from eqn 5.2)

$$u' = \frac{u - v}{1 - uv/c^2}.$$

Now let us calculate the energy E' of the particle as seen from the spaceship. The rest mass, of course, is the same; but we need to recalculate γ in order to find the new mass. We start with

$$u'^2 = \frac{u^2 - 2vu + v^2}{1 - 2vu/c^2 + v^2u^2/c^4}$$

so

$$\begin{aligned} 1 - u'^2/c^2 &= \frac{\big(1 - 2vu/c^2 + v^2u^2/c^4\big) - \big(u^2/c^2 - 2vu/c^2 + v^2/c}{\left(1 - vu/c^2\right)^2} \\ &= \frac{1 - u^2/c^2 - v^2/c^2 + v^2u^2/c^4}{\left(1 - vu/c^2\right)^2} \end{aligned}$$

$$= \frac{\left(1-u^2/c^2\right)\left(1-v^2/c^2\right)}{\left(1-vu/c^2\right)^2}$$

and therefore

$$\gamma' = \frac{1}{\sqrt{1-u'^2/c^2}} = \frac{1-vu/c^2}{\sqrt{(1-u^2/c^2)}\sqrt{(1-v^2/c^2)}}.$$

The energy is

$$\begin{aligned} E' &= \gamma' mc^2 \\ &= \frac{mc^2 - muv}{\sqrt{(1-u^2/c^2)}\sqrt{(1-v^2/c^2)}}. \end{aligned}$$

Since $E = mc^2/\sqrt{1-u^2/c^2}$, and the original $\gamma = 1/\sqrt{1-v^2/c^2}$, we obtain

$$\begin{aligned} E' &= \gamma E - \gamma v \frac{mu}{\sqrt{(1-u^2/c^2)}} \\ &= \gamma\left(E - vp_x\right). \end{aligned} \tag{5.6}$$

This looks familiar.... Let us now consider the momentum, firstly in the x direction.

$$\begin{aligned} p'_x &= \gamma' m' u' = \frac{E'}{c^2} u' \\ &= \frac{1}{c^2} \cdot \frac{mc^2 - muv}{\sqrt{(1-u^2/c^2)}} \sqrt{(1-v^2/c^2)} \cdot \frac{u-v}{1-vu/c^2} \\ &= \frac{m}{\sqrt{(1-u^2/c^2)}} \cdot \frac{\left(1-uv/c^2\right)}{\sqrt{(1-v^2/c^2)}} \cdot \frac{u-v}{(1-vu/c^2)} \\ &= \frac{mu - mv}{\sqrt{(1-v^2/c^2)}\sqrt{(1-u^2/c^2)}} \\ &= \gamma\left(p_x - v.E/c^2\right). \end{aligned} \tag{5.7}$$

What about transverse components? Consider a particle moving upwards with velocity u in frame S. If we transform to frame S', moving with velocity v in the x direction, we find the new velocity components are

$$\begin{aligned} u'_x &= -v \\ u'_y &= \frac{u_y}{\gamma} \end{aligned}$$

where, as usual, $\gamma = (1 - v^2/c^2)^{-1/2}$ (not to be confused with $\left(1 - u^2/c^2\right)^{-1/2}$ — we aren't transforming to the rest frame of the particle). So, the total velocity squared is

$$u'^2 = v^2 + \frac{u_y^2}{\gamma^2},$$

and

$$\begin{aligned} 1 - u'^2/c^2 &= 1 - v^2/c^2 + u_y^2/c^2(1 - v^2/c^2) \\ &= (1 - v^2/c^2)(1 - u^2/c^2). \end{aligned}$$

This gives

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

as before. Therefore, the transverse component of momentum is

$$\begin{aligned} p'_y &= \gamma' m u'_y \\ &= \frac{m}{\sqrt{(1 - u^2/c^2)}} \frac{1}{\sqrt{(1 - v^2/c^2)}} \frac{u_y}{\gamma} \\ &= \gamma m.u_y \end{aligned}$$

In other words, the transverse component of momentum is unchanged. Let's look at the energy again for this case:

$$\begin{aligned}
E' &= \gamma' mc^2 \\
&= \frac{mc^2}{\sqrt{(1-u^2/c^2)}}\frac{1}{\sqrt{(1-v^2/c^2)}} \\
&= \gamma E.
\end{aligned}$$

This time, there is *no γvp* term, so we were correct previously in writing the transformation using the x component only.

In total, then, the energy and momentum transformations are

$$\begin{aligned}
p'_x &= \gamma\left(p_x - \beta.\frac{E}{c}\right) \qquad (5.8)\\
p'_y &= p_y \\
p'_z &= p_z \\
\frac{E'}{c} &= \gamma\left(\frac{E}{c} - \beta p_x\right).
\end{aligned}$$

It is no coincidence that this set of transformations is just like those for x and t. If we replace E/c by ct and p_x by x, etc., we regain the original transformations.

Recall that the interval

$$S^2 = c^2t^2 - x^2 - y^2 - z^2$$

was an invariant under Lorentz transformations. If we make the same substitutions here, and then divide by c^2, we obtain

$$E^2 - p^2c^2 = \text{const} = \mathrm{m_0^2c^4}.$$

The invariant corresponding to the interval is therefore the rest energy of the object, which is naturally the same whichever frame you calculate it in.

5.7.2 The Easy Way

Now we shall derive the same transformation, but rather more quickly and easily. This is the way it will be done in the lecture. Recall that the interval

$$S^2 = c^2t^2 - \left(x^2 + y^2 + z^2\right)$$

is invariant under Lorentz transformations — it's the same in all inertial reference frames. Now, for a particle travelling a short distance $u\Delta t$ in time Δt,

$$\Delta x^2 + \Delta y^2 + \Delta z^2 = u^2\Delta t^2,$$

so the interval between the start and the end of the journey is

$$\begin{aligned} \Delta S^2 &= c^2\Delta t^2 - u^2\Delta t^2 \\ &= c^2\left(1 - \beta^2\right)\Delta t^2 \\ &= c^2 \cdot \frac{\Delta t^2}{\gamma_u^2}. \end{aligned} \tag{5.9}$$

(The time

$$\tau = \frac{\Delta t}{\gamma_u}$$

is the proper time — the time as measured in the particle's own rest frame, and so it is quite natural that the quantity (5.9) is the same from whatever frame it is calculated).

Knowing that $\Delta t/\gamma$ and m_0 are both invariant, we construct

$$k = \frac{\gamma_u m_0}{\Delta t} = k',$$

which must also be invariant under the Lorentz transformation — i.e., it's the same quantity in all frames: $k = k'$.

Let us go back to our particle, and see how its short journey looks from another reference frame. Under the usual transformation, we find

$$\begin{aligned}\Delta x' &= \gamma\left(\Delta x - \beta . c\Delta t\right)\\ \Delta y' &= \Delta y\\ c\Delta t' &= \gamma\left(c\Delta t - \beta . \Delta x\right)\end{aligned}$$

where, of course, β and γ refer to the velocity of the new reference frame rather than the particle.

Multiplying throughout by the Lorentz-invariant quantity k (or k', which is equivalent), we find

$$\begin{aligned}k'\Delta x' &= \gamma\left(k\Delta x - \beta . ck\Delta t\right)\\ k'\Delta y' &= k\Delta y\\ ck'\Delta t' &= \gamma\left(ck\Delta t - \beta . k\Delta x\right).\end{aligned} \tag{5.10}$$

But

$$k\Delta x = \gamma_u m_0 \frac{\Delta x}{\Delta t} = \gamma_u m_0 u_x = p_x,$$

and, likewise, $k\Delta y = p_y$, $k'\Delta x' = p'_x$ etc. Also,

$$\begin{aligned}ck\Delta t &= \gamma_u m_0 = \frac{E}{c^2}\\ ck'\Delta t' &= \frac{E'}{c^2}.\end{aligned}$$

So, from (5.10), the Lorentz transformations for energy and momentum are

$$\begin{aligned}p'_x &= \gamma\left(p_x - \beta . \frac{E}{c}\right)\\ p'_y &= p_y\\ p'_z &= p_z\end{aligned}$$

$$\frac{E'}{c} = \gamma\left(\frac{E}{c} - \beta p_x\right) \qquad (5.11)$$

as we found above. Again, it is easy to show that

$$E^2 - p^2c^2,$$

which is simply the rest mass squared times c^2, is invariant under the Lorentz transformation.

5.7.3 Matrix Notation

Using matrices, the transformation (5.11) may be written as

$$\begin{pmatrix} p'_x \\ p'_y \\ p'_z \\ \frac{E'}{c} \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \\ \frac{E}{c} \end{pmatrix}.$$

5.8 Example: Collision Threshold Energies

Consider the reaction

$$p + p \longrightarrow p + n + \pi^+,$$

in which an incoming proton p (mass ≈ 938 MeV/c^2) of (total) energy E_p hits a target proton at rest, to create a proton, a neutron (also of mass ≈ 938 MeV/c^2) and a positive pion (mass 139 MeV/c^2). What is the minimum (threshold) kinetic energy that the incident proton requires for this reaction to take place?

- The Trap to Avoid: the simple, obvious answer — that the sum of the masses afterwards is 139 MeV greater than the sum of the masses before, so this is the energy required — is *wrong*. This is because the incident proton carries some momentum, and so the outgoing particles must carry some momentum too.

- The Way to Solve it: A standard trick for many relativity "collision" problems — see, for example, our earlier "derivation" of the energy-momentum transformations — is to transform into the *centre-of-momentum* (or *zero momentum*) frame (note, this is no longer necessarily the "centre of mass"!). In this frame the two protons rush together with equal and opposite momenta, carrying just enough energy to make the three final particles (but not to give any of them any kinetic energy). Now, you can use conservation of energy and momentum, solve the resulting equations, and transform back to the lab frame; but the easy way is to use the fact that $E^2 - p^2c^2$ is invariant: it is the same in the laboratory and in the centre-of-momentum system.

In the lab frame, looking at the system before the collision, we have

$$\begin{aligned}\text{Total energy} &= E_{\text{tot}} = E_p + m_p c^2 \\ \text{Total momentum} &= p_p = \left(\frac{E_p^2}{c^2} - m_p^2 c^2\right)^{1/2} .\end{aligned}$$

In the centre-of-momentum system, the two protons come together with equal and opposite momenta, to give a total of zero momentum; and, as we are at threshold, the particles that are created in the collision are created at rest. Therefore,

$$\begin{aligned} \text{Total energy} &= E'_{tot} = m_p c^2 + m_n c^2 + m_\pi c^2 \\ \text{Total momentum} &= 0. \end{aligned}$$

Now we simply equate $E^2 - p^2c^2$ in the two frames:

$$\left(E_p + m_p c^2\right)^2 - \left(\frac{E_p^2}{c^2} - m_p^2 c^2\right) c^2 = \left(m_p c^2 + m_n c^2 + m_\pi c^2\right)^2 - 0.$$

Expanding the brackets, we obtain

$$\begin{aligned} E_p^2 + 2E_p m_p c^2 + m_p^2 c^4 - E_p^2 + m_p^2 c^4 &= \left(m_p + m_n + m_\pi\right)^2 c^4 \\ 2E_p m_p c^2 + 2m_p^2 c^4 &= \left(m_p + m_n + m_\pi\right)^2 c^4 \\ E_p &= \frac{\left(m_p + m_n + m_\pi\right)^2 c^2}{2m_p} - m \end{aligned}$$

Thus, the *total* energy of the incoming proton is 1226 MeV; this includes 938 MeV "locked up" in the rest mass energy m_0c^2, and, therefore, a kinetic energy of 288 MeV.

5.9 Kinematics: Hints for Problem Solving

At this stage, it's worthwhile summarising some useful formulae.

1. **Conservation of Mass-Energy.** The total mass-energy is always conserved. Add up $E = \gamma mc^2$ for all of the particles, and it is the same before as after the collision.

2. **Conservation of Momentum.** Momentum is also conserved; but remember, this is $p = \gamma mv$; the momentum of each particle is the usual classical mv times the relativistic factor γ.

3. **Invariance of** $E^2 - p^2c^2$. Do you want to change reference frames at all? For instance, to find out what's going on in the c.m. frame? Remember that $E^2 - p^2c^2$ is the same in all inertial frames, and it applies to the whole system as well as to individual particles. Furthermore, in the c.m. frame, the total momentum is zero (by definition). Very, very useful!

These three together will do most of your dirty work for you. To go further, e.g. to find the easiest ways out when particles leave the collision at various angles, we need four-vectors; but they come later.

5.10 Doppler Shift Revisited

From equation 5.8 we can see that, if we receive a photon of energy $E = h\nu$ that was emitted by a source moving with velocity $u = \beta c$ in the x direction, the energy of the photon in the source's frame of reference is

$$\begin{aligned} E' &= \gamma\,(E - \beta c p_x) \\ &= \gamma\,(E - \beta.E.\cos\theta) \\ &= \gamma E\,(1 - \beta\cos\theta)\,. \end{aligned}$$

If the source is coming directly towards us, $\theta = 0°$, so we obtain

$$\begin{aligned} E' &= \gamma E\,(1 - \beta) \\ &= E\frac{(1 - \beta)}{\sqrt{1 - \beta^2}} \end{aligned}$$

$$= E\sqrt{\frac{1-\beta}{1+\beta}},$$

which is consistent with the expression we obtained for the Doppler shift earlier, where we only considered this limited case. Be careful! The result we obtained then referred to the received frequency in terms of the source frequency, not the other way around, and so the primed and unprimed frames are swapped in this discussion.

5.11 Aberration of Light Revisited

Consider a telescope on a "stationary" Earth, pointing at an angle θ to capture light from a star (see diagram). Now let the Earth move with velocity v in the x direction. As before, we will have to tilt the telescope to a new angle, θ', so that photons can travel down the telescope tube to the bottom.

Now,

$$\tan\theta = \frac{p_y}{p_x}, \ \tan\theta' = \frac{p'_y}{p'_x}.$$

Using equations 5.8, we see that

$$\tan\theta' = \frac{p'_y}{p'_x} = \frac{p_y}{\gamma\left(p_x - v.E/c^2\right)}. \tag{5.12}$$

Since these are photons (and, as we have drawn it, moving in the $-x$, $-y$ direction),

$$\begin{aligned} p_x &= -E/c.\cos\theta \\ p_y &= -E/c.\sin\theta. \end{aligned}$$

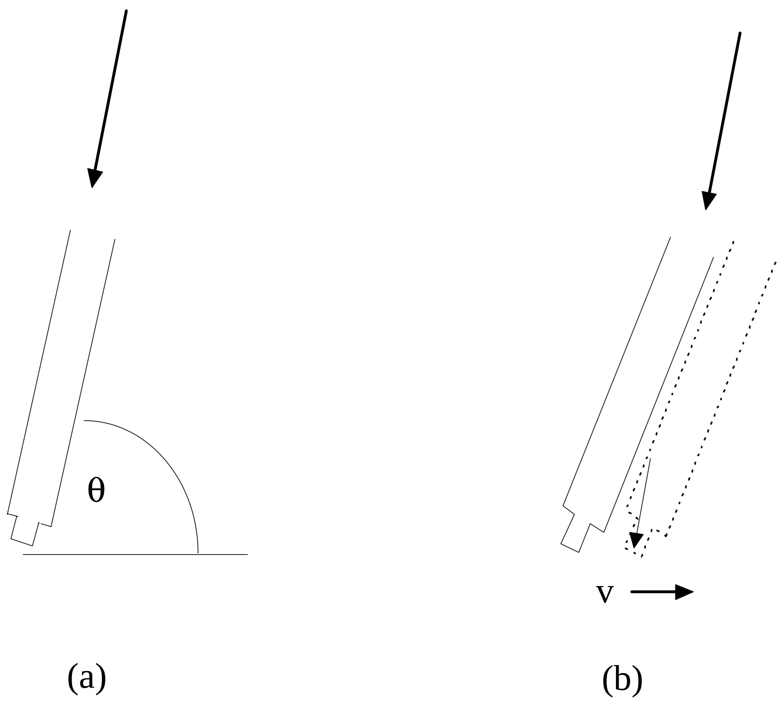

Figure 5.2: Aberration of starlight.

Therefore,

$$\tan\theta' = \frac{\sin\theta}{\gamma\left(\cos\theta + \beta\right)}$$

Equation 5.12 is the relativistic formula for the aberration of light.

5.12 Natural Units

Before continuing, let us introduce *natural units*, whereby

$$c = \hbar = 1.$$

This just means that, instead of using the artificial unit of a metre for distance (originally defined as a ten-millionth of the distance from the pole to the equator; now the distance that light travels in 1/299792458 s), we use the "light-second"; c, if you like, is then one light-second-per-second. This seems reasonable to do, as we have seen how space and time are intertwined. The velocity u now becomes equivalent to β, and the Lorentz transformations become

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

where, as usual, $\gamma = \left(1 - u^2\right)^{-1/2}$. The interval is now

$$S^2 = t^2 - x^2 - y^2 - z^2.$$

It is easier to write the equations this way, and easier to work with them — and very simple to go back, if we have to, to equations with c all over the place. Just look at the dimensions: for example, with the expression $1 - u^2$, we know that we cannot subtract a velocity squared, which has units, from the pure number 1, so it is clear that we have to divide u^2 by c^2 in order to make it unitless.

The same trick works with energy and mass, which are also obviously different aspects of the same thing. The equations for energy and momentum are therefore

$$\begin{aligned} E &= m = \gamma m_0 \\ \mathbf{p} &= m\mathbf{v} = \gamma m_0 \mathbf{v}, \end{aligned}$$

and they transform as

$$\begin{aligned} p'_x &= \gamma\,(p_x - uE) \\ p'_y &= p_y,\ p'_z = p_z \\ E' &= \gamma\,(E - up_x)\,. \end{aligned}$$

The similarity between the space-time and the mass-energy Lorentz transformations is now even more striking. We shall explore the similarity further in the next chapter — although c will creep back in!

Chapter 6

Four-Vectors

We are used to the idea of vectors in three-dimensional space; quantities like momentum, which we write **p** (in bold) to remind us that it has three components; we are free to rotate or translate the axes as we like, but, even though the individual components of the vector are all mixed up amongst each other, the magnitude of the momentum stays the same. For example, if we rotate by an angle θ, the new coordinate system becomes

$$
\begin{aligned}
x' &= x\cos\theta + y\sin\theta \\
y' &= y\cos\theta - x\sin\theta,
\end{aligned}
$$

and we see how the new x' is a mixture of the old x and y. However, the length squared $l^2 = x^2 + y^2$ is invariant. This is very similar to the way in which Lorentz transformations treat space and time.

Is there, therefore, some way in which we can put the separate components of space and time together to come up with a function that is invariant under Lorentz transformations? There is, in fact; it is called a *four-vector*, written A_μ, where μ indicates any of the four components. For spacetime,

$$x_1 = x$$

$$\begin{aligned} x_2 &= y \\ x_3 &= z \\ x_4 &= ct, \end{aligned}$$

and the transformation is

$$\begin{aligned} x'_1 &= \gamma x_1 - \gamma\beta x_4 \\ x'_2 &= x_2,\ x'_3 = x_3 \\ x'_4 &= \gamma x_4 - \gamma\beta x_1. \end{aligned}$$

The interval is then

$$S^2 = x_4^2 - \left(x_1^2 + x_2^2 + x_3^2\right).$$

There are different conventions; sometimes, the time component is called x_0 instead of x_4, and sometimes this component is written with an i:

$$x_4 = ict,$$

just so that the components will add in quadrature like normal vectors (so, e.g., the interval S is given by $-S^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$). We shall *not* be using that convention, but you should be aware that it exists.

6.1 Scalar Products

The scalar product of three-dimensional vectors is written as

$$\mathbf{u}.\mathbf{v} = u_x v_x + u_y v_y + u_z v_z,$$

where the dot indicates the cosine of the angle between the vectors. This product is invariant when we rotate axes in space. When dealing with four-vectors, we denote a scalar product by

$$\sum_{\mu=1}^{4} A_\mu B_\mu = A_4 B_4 - A_1 B_1 - A_2 B_2 - A_3 B_3.$$

This quantity is invariant under Lorentz transformations — it is the same in all inertial frames of reference. It is therefore often known as a (*Lorentz*) *scalar.*

Often, the $\sum$ is omitted; using Einstein's convention of *summation over repeated indices*, we can write

$$A_\mu B_\mu$$

which indicates that we should sum over all pairs of terms where the indices are the same (remembering to be careful with the signs). This is just like using $\mathbf{u}.\mathbf{v}$ as shorthand for the summation over the three pairs of components for three-vectors. Note that the scalar product of the spacetime four-vector with itself gives the interval. Another convention is that four-vectors may be written as

$$A = (\mathbf{A}, A_4)\,,$$

where $\mathbf{A}$ represents the three spatial components. The scalar product is then

$$\sum_{\mu=1}^{4} A_\mu B_\mu = A_4 B_{4.} - \mathbf{A}.\mathbf{B}. \tag{6.1}$$

In the case of the spacetime fourvector x, the scalar product $x \cdot x$ is of course just the interval S^2.

Note that the invariance of scalar products is *extremely useful* for solving lots of relativity problems...

6.2 Energy-Momentum Four-Vector

Consider the interval

$$S^2 = -\sum_\mu x_\mu^2 = c^2t^2 - x^2 - y^2 - z^2.$$

In a short time Δt, a particle moving with velocity u travels a distance

$$\Delta x^2 + \Delta y^2 + \Delta z^2 = u^2 \Delta t^2,$$

and so the element of interval between the beginning and end of that short journey is

$$\begin{aligned}\Delta S^2 &= c^2\Delta t^2 - u^2\Delta t^2 \\ &= c^2\left(1-\beta^2\right)\Delta t^2 \\ &= c^2.\Delta t^2/\gamma^2 \\ &= c^2.\Delta\tau^2,\end{aligned}$$

where τ is the proper (or Lorentz invariant) time. We saw this earlier, when we were looking at the transformation of energy and momentum.

Now, ΔS is invariant, and so if we introduce a function

$$U_\mu = c\frac{\Delta x_\mu}{\Delta S} = \frac{\Delta x_\mu}{\Delta \tau},$$

it will transform in the same manner as Δx_μ. As we make the time Δt shorter and shorter, in the usual way in calculus, we end up with

$$U_\mu = \frac{dx_\mu}{d\tau}.$$

The quantity U_μ is a four-vector, called the *four-velocity* because its components are

$$U_\mu = \gamma\left(u_x, u_y, u_z, c\right).$$

Notice that the scalar product of U_μ with itself is

$$\begin{aligned} U_\mu U_\mu &= c^2 \left(\frac{dx_\mu}{dS}\right)^2 \\ &= -c^2 \end{aligned}$$

(since $dS^2 = -\sum_\mu x_\mu^2$); this is, of course, invariant under Lorentz transformations.

We can now construct the *four-momentum* with components

$$\begin{aligned} P_\mu &= mU_\mu \\ P &= \left(\mathbf{p}, \frac{E}{c}\right), \end{aligned}$$

where E is an "arbitrarily" chosen letter that represents the quantity γmc^2, from the fourth component of U.

We therefore know immediately that the four-momentum will transform in the way we have seen,

$$\begin{aligned} p'_x &= \gamma\left(p_x - \beta.E/c\right) \\ p'_y &= p_y \\ p'_z &= p_z \\ E'/c &= \gamma\left(E/c - \beta p_x\right). \end{aligned} \tag{6.2}$$

Furthermore, the Lorentz invariant

$$P_\mu \cdot P_\mu = \frac{E^2}{c^2} - p^2 = m^2c^2$$

emerges trivially. It was at this point before that we realised that it makes sense to identify

$$E = \gamma mc^2$$

with the total energy of the body. The four-vector notation has thus allowed us to derive the Lorentz transformation for energy and momentum in a straightforward way, without lots of tedious algebra in studying collisions.

6.3 Examples

6.3.1 Collision Threshold Energies

We earlier considered the reaction

$$p + p \longrightarrow p + n + \pi^+,$$

in which an incoming proton p (mass ≈ 938 MeV$/c^2$) of energy E_0 hits a target proton at rest, to create a proton, a neutron (also of mass ≈ 938 MeV$/c^2$) and a positive pion (mass 139 MeV$/c^2$), and we derived the minimum kinetic energy that the incident proton required for the reaction to take place. Let us quickly do the same derivation, but using four-vectors.

The total four-momentum in the lab frame is

$$P = \left(\mathbf{p}_p, \frac{E_p + m_p c^2}{c} \right).$$

The relativistic invariant is

$$P \cdot P = \frac{\left(E_p + m_p c^2\right)^2}{c^2} - p^2,$$

which we know is the same in all frames; $P \cdot P = P' \cdot P'$.

Now, if we move to the centre-of-momentum frame, the four-momentum vector (evaluated after the collision) is

$$P' = (0, (m_p + m_n + m_\pi)c).$$

So, we have

$$\begin{aligned} P \cdot P &= P' \cdot P' \\ \frac{\left(E_p + m_p c^2\right)^2}{c^2} - p^2 &= (m_p + m_n + m_\pi)^2 c^2. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{\left(E_p + m_p c^2\right)^2}{c^2} - \frac{\left(E_p^2 - m_p^2 c^4\right)}{c^2} &= (m_p + m_n + m_\pi)^2 c^2 \\ 2m_p c^2 E_p + 2m_p^2 c^4 &= (m_p + m_n + m_\pi)^2 c^2 \end{aligned}$$

and, just as we found before,

$$E_p = \frac{(m_p + m_n + m_\pi)^2}{2m_p} - m_p c^2.$$

This is the total energy of the incoming proton; its *kinetic* energy is $T = E_p - m_p c^2$. For the masses given above, the total energy is 1226 MeV; the kinetic energy is 288 MeV.

This looks complicated, but it isn't too bad if you work through it step by step... it's a good example to bear in mind as you work through the problems. Use of four-vectors doesn't make solving this particular problem any easier, but it is useful to become familiar with the terminology so that it can later be applied to more complex problems.

6.3.2 Matter-Antimatter annihilation

A positron and an electron can annihilate to produce a pair of photons. Consider the case of a positron at rest in the laboratory and an electron colliding with it, as shown in Figure 6.1. What is the energy of one of the photons, as a function of its angle with respect to the incoming electron?

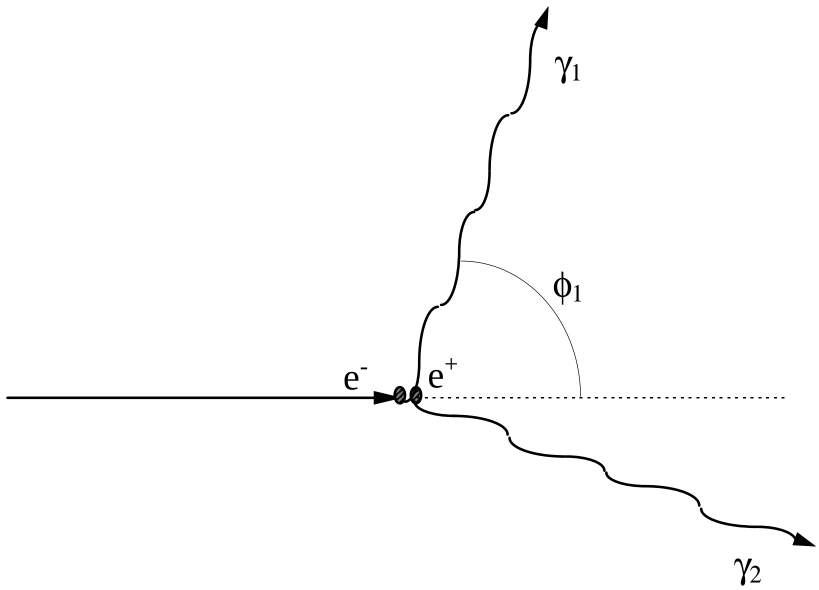

Figure 6.1: Annihilation of an electron and a positron to produce a pair of gamma rays.

This is the kind of problem where four-vectors can really be useful; because it is no longer one-dimensional, we have several coordinates to consider at once, and, just as with ordinary vectors in Newtonian mechanics, the use of four-vectors here makes the algebra rather simpler than it would otherwise be.

Conservation of the energy-momentum four-vector gives

$$P_+ + P_- = Q_{\gamma_1} + Q_{\gamma_2}.$$

Here we have used p for the particle four-momentum, and q for the photon four-momentum. We are not interested in photon 2, so (note this trick!) we take its four-vector to one side, and then "square" both sides:

$$\begin{aligned} Q_{\gamma_2}^2 &= P_+ \cdot P_+ + P_- \cdot P_- + Q_{\gamma_1} \cdot Q_{\gamma_1} \\ &\quad +2P_+ \cdot P_- - 2P_+ \cdot Q_{\gamma_1} - 2P_- \cdot Q_{\gamma_1}. \end{aligned}$$

To evaluate this, we use equation (6.1), together with the facts that

- For a photon, $Q_\gamma^2 = E^2/c^2 - p^2 = 0$ (i.e., it has zero rest mass).
- The positron is at rest, so $\mathbf{p}_+ = 0$ and $E_+ = m_0c^2$.

Therefore, we have

$$0 = m_0^2c^2 + m_0^2c^2 + 0 + \\ 2\left(\frac{m_0c^2}{c}\right)\left(\frac{E_-}{c}\right) - 2\left(\frac{m_0c^2}{c}\right)\left(\frac{E_{\gamma_1}}{c}\right) - 2\left(\frac{E_-}{c}\right)\left(\frac{E_{\gamma_1}}{c}\right) -$$

If ϕ is the angle between the electron and the photon,

$$\mathbf{p}_-.\mathbf{q}_{\gamma_1} = p_- \cdot q_{\gamma_1}\cos\phi = p_-\frac{E_{\gamma_1}}{c}\cos\phi.$$

So, multiplying through by $c^2/2$,

$$0 = m_0^2c^4 + m_0c^2E_- - m_0c^2E_{\gamma_1} - E_- \cdot E_{\gamma_1} + p_-cE_{\gamma_1}\cos\phi,$$

which may be rearranged to give

$$E_{\gamma_1}\left(E_- + m_0c^2 - p_-c\cos\phi\right) = m_0c^2\left(E_- + m_0c^2\right)$$

or

$$E_{\gamma_1} = \frac{m_0c^2\left(E_- + m_0c^2\right)}{(E_- + m_0c^2) - p_-c\cos\phi}.$$

Thus, for a given energy E_- and momentum p_- of the incident electron, the photon's energy is at a maximum if it goes forward and at a minimum if it goes backwards. Of course, the situation is symmetric, and the same equation will apply to either photon.

Notice the trick we used at the start of isolating the quantity that was not of interest. If we had had both photons on the same side of the equation, we would have had to deal with the dot product between them; as it is, the only angle that entered was between one of the photons and the incoming particle.

Chapter 7

Relativity and Electromagnetism

7.1 Magnetic Field due to a Current

The magnetic force on a charge is

$$\mathbf{F} = q\left(\mathbf{v} \times \mathbf{B}\right). \tag{7.1}$$

But as we know, the velocity is not absolute. What happens to this force if we move into the reference frame where the charged particle is at rest? Does it matter which frame we are in when we measure the magnetic field?

Let us try applying our knowledge of relativity to the simple case of a current-carrying wire to see what we can find out about the relationship between electricity and magnetism. This example is taken directly from the Feynman lectures in physics.

We will consider a wire as a lattice of stationary positive charges, of density ρ_+, with negative charges of density ρ_- moving through it at an average velocity v_- (figure 7.1). Outside the wire, at a distance r, there is a negative charge q_-, moving with velocity v; for simplicity, we will let $v = v_-$. We could treat the more complicated case of an arbitrary velocity; but we won't do so here. We shall be looking at two reference frames; in one, S, the wire is at rest; in the second, S', the charge is at rest. In the S frame, there is a magnetic force on the particle, and if it were moving freely we would see it curve in towards the wire. But what about the S' frame? Clearly, since $\mathbf{v}' = 0$, there is no magnetic force at all; does the charge, therefore, move in towards the wire or not? And if it does, what would make it do so?

In the S frame, the force on the particle is given by the Lorentz equation 7.1. The magnitude of the magnetic field is

$$B = \frac{\mu_0 I}{2\pi r}.$$

The current I is the amount of charge passing any given point per second. If this charge is enough to fill a "cylinder" of the wire of length x, then

$$\begin{aligned} I &= \frac{dq_-}{dt} \\ &= \rho_- A \frac{dx}{dt} \\ &= \rho_- A v_-. \end{aligned}$$

So, the magnetic force on the charge has magnitude

$$F_B = q v_0 \frac{\mu_0}{2\pi r} \rho_- A v_-,$$

and since we are considering the simple case where $v = v_-$, we have

$$F_B = q \frac{\mu_0}{2\pi r} \rho_- A v^2.$$

Notice that, as the wire is uncharged, the positive and negative charges cancel each other out, so we have

$$\rho_- = -\rho_+.$$

Now, let us look at the S' frame. Here, the negative charges are all at rest; but the positive charges of the lattice are moving past with speed $v'_+ = -v$.

Firstly, we have to establish what happens to charge when we change reference frames. The answer is: nothing at all. The total charge of a particle remains the same, whatever its speed in our reference frame. Suppose we take a block of metal, and heat it up; because the electrons have a different mass than the protons, their speeds will change by different amounts. If the charge on each particle depended on the speed, the charges of the electrons and protons would no longer balance, and the block would become charged. Likewise, the net charge of a piece of material would change in a chemical reaction. As we have never seen any such effects, we have to conclude that charge is independent of velocity.

But charge *density* is not. If we take a length L_0 of wire, in which there is a charge density ρ_0 of stationary charges, it contains a total charge

$$Q_0 = \rho_0 . L_0 . A.$$

If we now observe these same charges from a moving frame, they will all be found in a piece of the material with length

$$L = L_0 \sqrt{1 - v^2/c^2}.$$

Therefore, since the charge Q_0 and the transverse dimension A are unchanged, we find

$$Q_0 = \rho_0 . L_0 . A = \rho . . A L_0 \sqrt{1 - v^2/c^2},$$

and so

$$\rho = \frac{\rho_0}{\sqrt{1 - v^2/c^2}}.$$

Thus, the charge *density* of a moving *distribution* of charges varies in the same way as the length or the relativistic mass of a particle.

What does all this imply for our wire? If we look at the *positive* charge density in the S' frame, we find

$$\rho'_+ = \frac{\rho_+}{\sqrt{1 - v^2/c^2}}.$$

For the *negative* charges, though, it's the other way around. They are at rest in S', and moving in S; they have their "rest density" in S'. Therefore,

$$\rho'_- = \rho_- \sqrt{1 - v^2/c^2}.$$

So, we find that the charges no longer balance — when we transfer to the moving frame, the wire becomes electrically charged! This is why the exterior negative charge is attracted to it. The total charge density is

$$\begin{aligned} \rho' &= \rho'_+ + \rho'_- \\ &= \frac{\rho_+}{\sqrt{1 - v^2/c^2}} + \rho_- \sqrt{1 - v^2/c^2}. \end{aligned}$$

Since the wire is uncharged in S, we have $\rho_- = -\rho_+$, and therefore

$$\begin{aligned} \rho' &= \rho_+ \left\{ \frac{1}{\sqrt{1 - v^2/c^2}} - \frac{1 - v^2/c^2}{\sqrt{1 - v^2/c^2}} \right\} \\ &= \rho_+ \frac{v^2/c^2}{\sqrt{1 - v^2/c^2}}. \end{aligned}$$

For a charged wire, with charge density ρ, the electric field a distance r away is

$$E = \frac{\rho.A}{2\pi\varepsilon_0 r}.$$

Therefore, with our charge density ρ', the force $F' = E'q$ acting on the charge in the S' frame is

$$F' = \rho_+ \frac{v^2/c^2}{\sqrt{1 - v^2/c^2}} \frac{A}{2\pi\varepsilon_0 r} q.$$

Let us transform this force back into the S frame. We know that

$$F = \frac{dp}{dt},$$

and we have seen that *transverse* momenta are unchanged by Lorentz transformations. Therefore, we expect transverse *forces* to transform like the inverse of time; in other words, the force as seen in the S frame is

$$\begin{aligned} F &= F'\sqrt{1 - v^2/c^2} \\ &= \rho_+ v^2/c^2 \frac{A}{2\pi\varepsilon_0 r} q. \end{aligned}$$

Comparing this with our earlier expression for the magnetic force in S,

$$F_B = q\frac{\mu_0}{2\pi r}\rho_- A v^2,$$

we see that the magnitudes of these forces are identical provided that

$$c^2 = \frac{1}{\sqrt{\mu_0\varepsilon_0}},$$

which is Maxwell's well-known result for the speed of propagation of electromagnetic waves. So, the magnetic force in one reference frame is seen as a purely electrostatic force in another.

It is perhaps not surprising to learn that magnetism and electricity are not independent forces. Just as space and time are united, and mass and energy, so electricity and magnetism are just different aspects of the same underlying *electromagnetic* field, and we have to treat both together as a complete entity.

7.2 Lorentz Transformation of Electromagnetic Fields

We shall not derive the Lorentz transformation of electromagnetic fields; we simply state the result, namely

$$\begin{aligned} \mathbf{E}'_{\|} &= \mathbf{E}_{\|} \\ \mathbf{E}'_{\perp} &= \gamma\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)_{\perp} \end{aligned}$$

and

$$\begin{aligned} \mathbf{B}'_{\|} &= \mathbf{B}_{\|} \\ \mathbf{B}'_{\perp} &= \gamma\left(\mathbf{B} - \mathbf{v} \times \mathbf{E}/c^2\right)_{\perp}, \end{aligned}$$

where $\|$ and $\perp$ indicate the parallel and perpendicular components respectively of the fields. This means that a purely magnetic (or purely electric) field in one frame of reference looks like a mixture of electric and magnetic fields in another. This should not be surprising, having seen the example of the current-carrying wire.

7.3 Electromagnetic Waves

We have already seen how momentum and energy transform from one Lorentz frame to another:

$$\begin{aligned} p'_x &= \gamma \left(p_x - \beta \frac{E}{c} \right) \\ \frac{E'}{c} &= \gamma \left(\frac{E}{c} - \beta p_x \right). \end{aligned}$$

For light waves, the momentum is

$$\mathbf{p} = \hbar \mathbf{k} \tag{7.2}$$

where $\mathbf{k}$ is the wavenumber, $|\mathbf{k}| = 2\pi/\lambda$, and the energy carried by the wave is

$$E = \hbar \omega,$$

where the angular frequency ω is related to the wavenumber by $\omega = c\,|\mathbf{k}|$. From this, it follows that

$$\begin{aligned} \hbar k'_x &= \hbar \gamma \left(k_x - \beta \frac{w}{c} \right) \\ \hbar \frac{\omega'}{c} &= \hbar \gamma \left(\frac{\omega}{c} - \beta k_x \right). \end{aligned} \tag{7.3}$$

7.3.1 Phase

The phase of a (monochromatic) wave is given by

$$\phi = \omega t - \mathbf{k} \cdot \mathbf{r}.$$

If we restrict ourselves to waves travelling in the x direction, we have

$$\phi = \omega t - k_x x.$$

If we now use the transformations (7.3) and the familiar Lorentz transformation, we find that the phase ϕ' in the primed frame is

$$\begin{aligned}
\phi' &= \omega' t' - k_x' x' \\
&= \gamma\left(\frac{\omega}{c} - \beta k_x\right)\gamma\left(ct - \beta x\right) - \gamma\left(k_x - \beta\frac{w}{c}\right)\gamma\left(x - \beta ct\right) \\
&= \gamma^2\left(\omega t + \beta^2 k_x x - k_x x - \beta^2 \omega t\right) \\
&= \gamma^2\left(1 - \beta^2\right)\left(\omega t - k_x x\right) \\
&= \omega t - k_x x = \phi.
\end{aligned}$$

Therefore, the phase is an invariant. This should come as no surprise; the crest of a wave, for example, is a well-defined object that should not depend upon one's frame of reference. In fact, we have our argument slightly reversed; it is usual to begin with the invariance of the phase, and to show that the de Broglie postulate (7.2) is consistent with relativity.

It is trivial to extend this argument to a wave travelling in three dimensions.

7.3.2 Wave four-vector

By looking at the transformation (7.3), one can see immediately that it is possible to construct a four-vector with components $(c\mathbf{k}, w)$. The invariant is then $w^2 - c^2k^2 = 0$ (which arises because the photon has zero rest mass).

7.3.3 Doppler Shift

The latter of equations (7.3) is of course identical with our earlier equation for the Doppler shift:

$$\begin{aligned} \omega' &= \gamma\left(w - \beta c \cdot \frac{\omega}{c}\cos\theta\right) \\ &= \omega\gamma\left(1 - \beta\cos\theta\right). \end{aligned}$$

Thus, if a source of frequency ω moves away from us at speed $u = u_x$, we will detect waves of frequency

$$\begin{aligned} \omega' &= \omega\gamma\left(1 - \beta\right) \\ &= \omega\sqrt{\frac{1-\beta}{1+\beta}}. \end{aligned}$$

Thus, the frequency drops, as expected for a receding source. If it moves towards us, $\beta \to -\beta$, and

$$\omega' = \omega\sqrt{\frac{1+\beta}{1-\beta}},$$

again as expected.

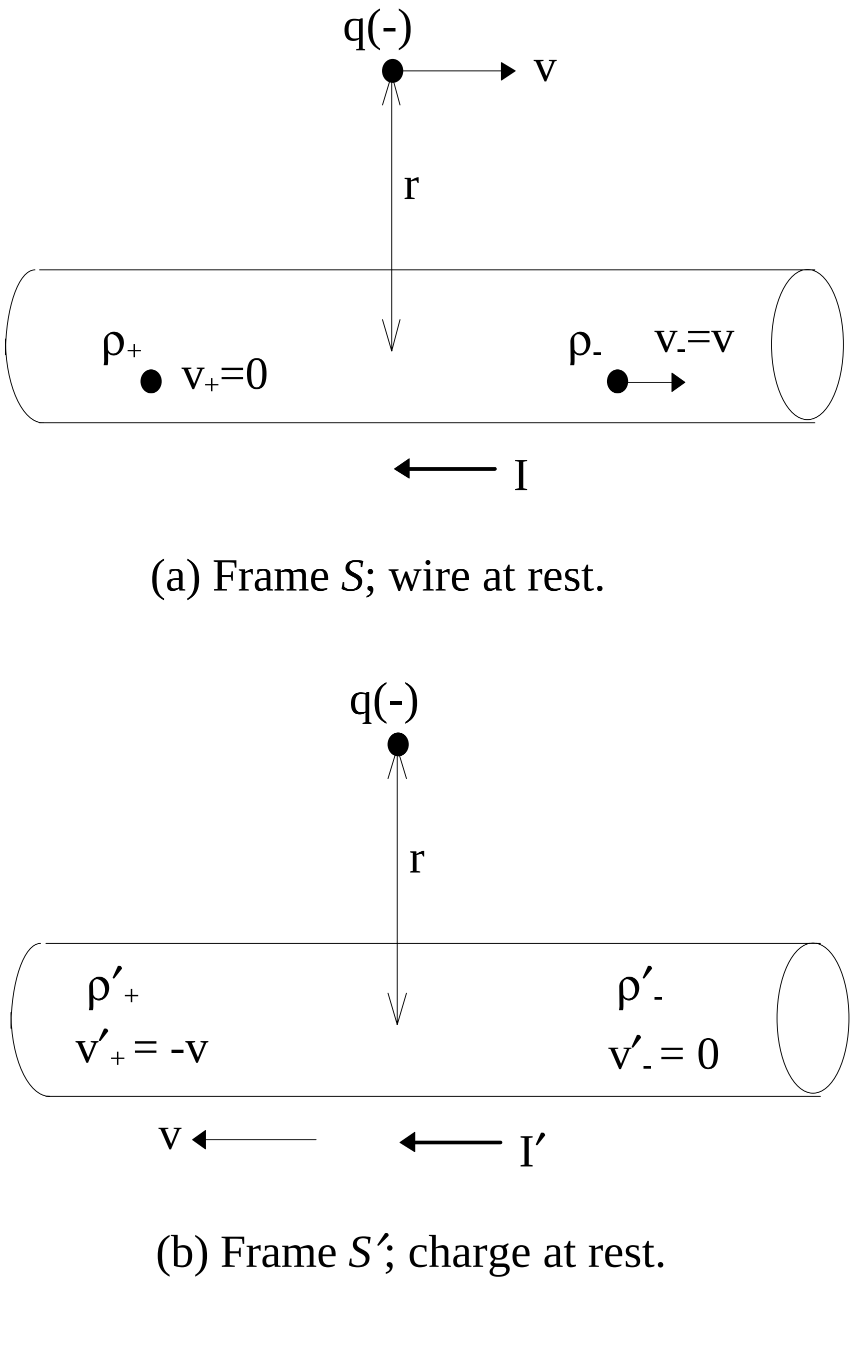

Figure 7.1: A current-carrying wire, seen in its rest frame (a) and in the rest frame (b) of the negative charges that constitute the current.