7. Special Relativity

Although Newtonian mechanics gives an excellent description of Nature, it is not universally valid. When we reach extreme conditions — the very small, the very heavy or the very fast — the Newtonian Universe that we're used to needs replacing. You could say that Newtonian mechanics encapsulates our common sense view of the world. One of the major themes of twentieth century physics is that when you look away from our everyday world, common sense is not much use.

One such extreme is when particles travel very fast. The theory that replaces Newtonian mechanics is due to Einstein. It is called *special relativity.* The effects of special relativity become apparent only when the speeds of particles become comparable to the speed of light in the vacuum. The speed of light is

$$c = 299792458 \; ms^{-1}$$

This value of c is exact. It may seem strange that the speed of light is an integer when measured in meters per second. The reason is simply that this is taken to be the definition of what we mean by a meter: it is the distance travelled by light in 1/299792458 seconds. For the purposes of this course, we'll be quite happy with the approximation $c \approx 3 \times 10^8 \; ms^{-1}$.

The first thing to say is that the speed of light is fast. Really fast. The speed of sound is around 300 ms^{-1}; escape velocity from the Earth is around 10^4 ms^{-1}; the orbital speed of our solar system in the Milky Way galaxy is around 10^5 ms^{-1}. As we shall soon see, nothing travels faster than c.

The theory of special relativity rests on two experimental facts. (We will look at the evidence for these shortly). In fact, we have already met the first of these: it is simply the Galilean principle of relativity described in Section 1. The second postulate is more surprising:

- **Postulate 1:** The principle of relativity: the laws of physics are the same in all inertial frames
- **Postulate 2:** The speed of light in vacuum is the same in all inertial frames

On the face of it, the second postulate looks nonsensical. How can the speed of light look the same in all inertial frames? If light travels towards me at speed c and I run away from the light at speed v, surely I measure the speed of light as $c - v$. Right? Well, no.

This common sense view is encapsulated in the Galilean transformations that we met in Section 1.2.1. Mathematically, we derive this "obvious" result as follows: two inertial frames, S and S', which move relative to each with velocity $\mathbf{v} = (v, 0, 0)$, have Cartesian coordinates related by

$$x' = x - vt \quad , \quad y' = y \quad , \quad z' = z \quad , \quad t' = t \tag{7.1}$$

If a ray of light travels in the x direction in frame S with speed c, then it traces out the trajectory $x/t = c$. The transformations above then tell us that in frame S' the trajectory if the light ray is $x'/t' = c - v$. This is the result we claimed above: the speed of light should clearly be $c - v$. If this is wrong (and it is) something must be wrong with the Galilean transformations (7.1). But what?

Our immediate goal is to find a transformation law that obeys both postulates above. As we will see, the only way to achieve this goal is to allow for a radical departure in our understanding of time. In particular, we will be forced to abandon the assumption of absolute time, enshrined in the equation $t' = t$ above. We will see that time ticks at different rates for observers sitting in different inertial frames.

7.1 Lorentz Transformations

We stick with the idea of two inertial frames, S and S', moving with relative speed v. For simplicity, we'll start by ignoring the directions y and z which are perpendicular to the direction of motion. Both inertial frames come with Cartesian coordinates: (x, t) for S and (x', t') for S'. We want to know how these are related. The most general possible relationship takes the form

$$x' = f(x, t) \quad , \quad t' = g(x, t)$$

for some function f and g. However, there are a couple of facts that we can use to immediately restrict the form of these functions. The first is that the law of inertia holds; left alone in an inertial frame, a particle will travel at constant velocity. Drawn in the (x, t) plane, the trajectory of such a particle is a straight line. Since both S and S' are inertial frames, the map $(x, t) \mapsto (x', t')$ must map straight lines to straight lines; such maps are, by definition, linear. The functions f and g must therefore be of the form

$$x' = \alpha_1 x + \alpha_2 t \quad , \quad t' = \alpha_3 x + \alpha_4 t$$

where α_i, $i = 1, 2, 3, 4$ can each be a function of v.

Secondly, we use the fact that S' is travelling at speed v relative to S. This means that an observer sitting at the origin, $x' = 0$, of S' moves along the trajectory $x = vt$ in S shown in the figure. Or, in other words, the points $x = vt$ must map to $x' = 0$. (There is actually one further assumption implicit in this statement: that the origin $x' = 0$ coincides with $x = 0$ when $t = 0$). Together with the requirement that the transformation is linear, this restricts the coefficients α_1 and α_2 above to be of the form,

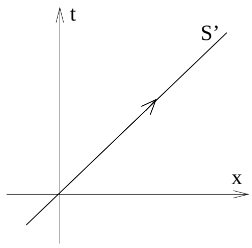

Figure 43:

$$x' = \gamma(x - vt) \tag{7.2}$$

for some coefficient γ. Once again, the overall coefficient γ can be a function of the velocity: $\gamma = \gamma_v$. (We've used subscript notation γ_v rather than the more standard $\gamma(v)$ to denote that γ depends on v. This avoids confusion with the factors of $(x - vt)$ which aren't arguments of γ but will frequently appear after γ like in the equation (7.2)).

There is actually a small, but important, restriction on the form of γ_v: it must be an even function, so that $\gamma_v = \gamma_{-v}$. There are a couple of ways to see this. The first is by using rotational invariance, which states that γ can depend only on the direction of the relative velocity $\mathbf{v}$, but only on the magnitude $v^2 = \mathbf{v} \cdot \mathbf{v}$. Alternatively, if this is a little slick, we can reach the same conclusion by considering inertial frames $\tilde{S}$ and $\tilde{S}'$ which are identical to S and S' except that we measure the x-coordinate in the opposite direction, meaning $\tilde{x} = -x$ and $\tilde{x}' = -x'$. While S is moving with velocity $+v$ relative to S', $\tilde{S}$ is moving with velocity $-v$ with respect to $\tilde{S}'$ simply because we measure things in the opposite direction. That means that

$$\tilde{x}' = \gamma_{-v} \left(\tilde{x} + v\tilde{t}\right)$$

Comparing this to (7.2), we see that we must have $\gamma_v = \gamma_{-v}$ as claimed.

We can also look at things from the perspective of S', relative to which the frame S moves backwards with velocity $-v$. The same argument that led us to (7.2) now tells us that

$$x = \gamma(x' + vt') \tag{7.3}$$

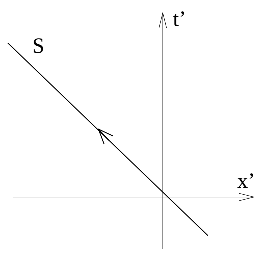

Figure 44:

Now the function $\gamma = \gamma_{-v}$. But by the argument above, we know that $\gamma_v = \gamma_{-v}$. In other words, the coefficient γ appearing in (7.3) is the same as that appearing in (7.2).

At this point, things don't look too different from what we've seen before. Indeed, if we now insisted on absolute time, so $t = t'$, we're forced to have $\gamma = 1$ and we get back to the Galilean transformations (7.1). However, as we've seen, this is not compatible with the second postulate of special relativity. So let's push forward and insist instead that the speed of light is equal to c in both S and S'. In S, a light ray has trajectory

$$x = ct$$

While, in S', we demand that the same light ray has trajectory

$$x' = ct'$$

Substituting these trajectories into (7.2) and (7.3), we have two equations relating t and t',

$$ct' = \gamma(c - v)t \quad \text{and} \quad ct = \gamma(c + v)t'$$

A little algebra shows that these two equations are compatible only if γ is given by

$$\gamma = \sqrt{\frac{1}{1 - v^2/c^2}} \tag{7.4}$$

We'll be seeing a lot of this coefficient γ in what follows. Notice that for $v \ll c$, we have $\gamma \approx 1$ and the transformation law (7.2) is approximately the same as the Galilean transformation (7.1). However, as $v \to c$ we have $\gamma \to \infty$. Furthermore, γ becomes imaginary for $v > c$ which means that we're unable to make sense of inertial frames with relative speed $v > c$.

Equations (7.2) and (7.4) give us the transformation law for the spatial coordinate. But what about for time? In fact, the temporal transformation law is already lurking in our analysis above. Substituting the expression for x' in (7.2) into (7.3) and rearranging, we get

$$t' = \gamma\left(t - \frac{v}{c^2}x\right) \tag{7.5}$$

We shall soon see that this equation has dramatic consequences. For now, however, we merely note that when $v \ll c$, we recover the trivial Galilean transformation law $t' \approx t$. Equations (7.2) and (7.5) are the *Lorentz transformations*.

7.1.1 Lorentz Transformations in Three Spatial Dimensions

In the above derivation, we ignored the transformation of the coordinates y and z perpendicular to the relative motion. In fact, these transformations are trivial. Using the above arguments for linearity and the fact that the origins coincide at $t = 0$, the most general form of the transformation is

$$y' = \kappa y$$

But, by symmetry, we must also have $y = \kappa y'$. Clearly, we require $\kappa = 1$. (The other possibility $\kappa = -1$ does not give the identity transformation when $v = 0$. Instead, it is a reflection).

With this we can write down the final form of the Lorentz transformations. Note that they look more symmetric between x and t if we write them using the combination ct,

$$\begin{aligned} x' &= \gamma\left(x - \frac{v}{c}\,ct\right) \\ y' &= y \\ z' &= z \\ ct' &= \gamma\left(ct - \frac{v}{c}x\right) \end{aligned} \tag{7.6}$$

where γ is given by (7.4). These are also known as *Lorentz boosts*. Notice that for $v/c \ll 1$, the Lorentz boosts reduce to the more intuitive Galilean boosts that we saw in Section 1. (We sometimes say, rather sloppily, that the Lorentz transformations reduce to the Galilean transformations in the limit $c \to \infty$).

It's also worth stressing again the special properties of these transformations. To be compatible with the first postulate, the transformations must take the same form if we invert them to express x and t in terms of x' and t', except with v replaced by $-v$. And, after a little bit of algebraic magic, they do.

Secondly, we want the speed of light to be the same in all inertial frames. For light travelling in the x direction, we already imposed this in our derivation of the Lorentz transformations. But it's simple to check again: in frame S, the trajectory of an object travelling at the speed of light obeys $x = ct$. In S', the same object will follow the trajectory $x' = \gamma(x - vt) = \gamma(ct - vx/c) = ct'$.

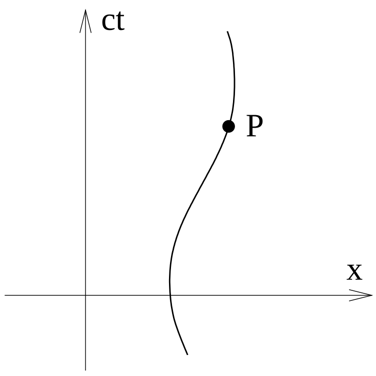

Figure 45: The worldline of a particle

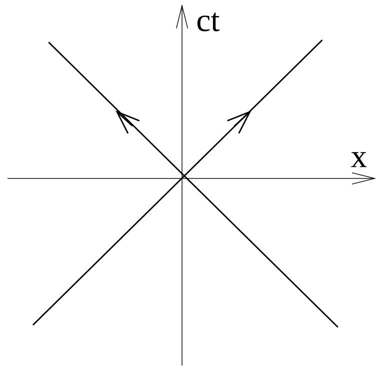

Figure 46: Light rays travel at 45°

What about an object travelling in the y direction at the speed of light? Its trajectory in S is $y = ct$. From (7.6), its trajectory in S' is $y' = ct'/\gamma$ and $x' = -vt'$. Its speed in S' is therefore $v'^2 = v_x^2 + v_y^2$, or

$$v'^2 = \left(\frac{x'}{t'}\right)^2 + \left(\frac{y'}{t'}\right)^2 = v^2 + \frac{c^2}{\gamma^2} = c^2$$

7.1.2 Spacetime Diagrams

We'll find it very useful to introduce a simple spacetime diagram to illustrate the physics of relativity. In a fixed inertial frame, S, we draw one direction of space — say x — along the horizontal axis and time on the vertical axis. But things look much nicer if we rescale time and plot ct on the vertical instead. In the context of special relativity, space and time is called *Minkowski space*. (Although the true definition of Minkowski space requires some extra structure on space and time which we will meet in Section (7.3)).

This is a spacetime diagram. Each point, P, represents an *event*. In the following, we'll label points on the spacetime diagram as coordinates (ct, x) i.e. giving the coordinate along the vertical axis first. This is backwards from the usual way coordinates but is chosen so that it is consistent with a later, standard, convention that we will meet in Section 7.3.

A particle moving in spacetime traces out a curve called a *worldline* as shown in the figure. Because we've rescaled the time axis, a light ray moving in the x direction moves at 45°. We'll later see that no object can move faster than the speed of light which means that the worldlines of particles must always move upwards at an angle steeper than 45°.

The horizontal and vertical axis in the spacetime diagram are the coordinates of the inertial frame S. But we could also draw the axes corresponding to an inertial frame S' moving with relative velocity $\mathbf{v} = (v, 0, 0)$. The t' axis sits at $x' = 0$ and is given by

$$x = vt$$

Meanwhile, the x' axis is determined by $t' = 0$ which, from the Lorentz transformation (7.6), is given by the equation

$$ct = \frac{v}{c}x$$

These two axes are drawn on the figure to the right. They can be the thought of as the x and ct axes, rotated by an equal amount towards the diagonal light ray. The fact the axes are symmetric about the light ray reflects the fact that the speed of light is equal to c in both frames.

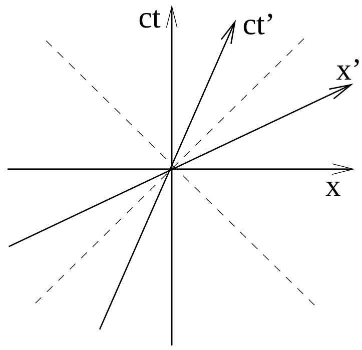

Figure 47:

7.1.3 A History of Light Speed

The first evidence that light does not travel instantaneously was presented by the Danish Astronomer Ole Rømer in 1676. He noticed that the periods of the orbits of Io, the innermost moon of Jupiter, are not constant. When the Earth is moving towards Jupiter, the orbits are a few minutes shorter; when the Earth moves away, the orbits are longer by the same amount. Rømer correctly deduced that this was due to the finite speed of light and gave a rough estimate for the value of c.

By the mid 1800s, the speed of light had been determined fairly accurately using experiments involving rotating mirrors. Then came a theoretical bombshell. Maxwell showed that light could be understood as oscillations of the electric and magnetic fields. He related the speed of light to two constants, ϵ_0 and μ_0, the permittivity and permeability of free space, that arise in the theory of electromagnetism,

$$c = \sqrt{\frac{1}{\epsilon_0 \mu_0}} \tag{7.7}$$

But, as we have seen, Newtonian physics tells us that speeds are relative. If Maxwell's equations predict a value for the speed of light, it was thought that these equations must be valid only in a preferred reference frame. Moreover, this does not seem unreasonable; if light is a wave then surely there is something waving. Just as water waves need water, and sound waves need air, so it was thought that light waves need a material to propagate in. This material was dubbed the *luminiferous ether* and it was thought that Maxwell's equations must only be valid in the frame at rest with respect to this ether.

In 1881, Michelson and Morley performed an experiment to detect the relative motion of the Earth through the ether. Since the Earth is orbiting the Sun at a speed of $3 \times 10^4\ ms^{-1}$, even if it happens to be stationary with respect to the ether at some point, six months later this can no longer be the case.

Suppose that at some moment the Earth is moving in the x-direction relative to the ether with some speed v. The Newtonian addition of velocities tells us that light propagating in the x-direction should have speed $c + v$ going one way and $c - v$ going the other. The total time to travel backwards and forwards along a length L should therefore be

$$T_x = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2cL}{c^2 - v^2}$$

Meanwhile, light making the same journey in the y-direction will have to travel (by Pythagoras) a total distance of $\sqrt{L^2 + v^2(T_y/2)^2}$ on each leg of the journey. It makes this journey at speed c, meaning that we can equate

$$\frac{cT_y}{2} = \sqrt{L^2 + v^2(T_y/2)^2} \qquad \Rightarrow \qquad T_y = \frac{2L}{\sqrt{c^2 - v^2}}$$

The goal of the Michelson-Morley experiment was to measure the time difference between T_y and T_x using interference patterns of light ray making the two journeys. Needless to say, the experiment didn't work: there seemed to be no difference in the time taken to travel in the x direction and y direction.

Towards the end of the 1800s, the null result of the Michelson-Morley experiment had become one of the major problems in theoretical physics. Several explanations were proposed, including the idea that the ether was somehow dragged along with the Earth. The Dutch physicist, Hendrik Lorentz, went some way to finding the correct solution. He had noticed that Maxwell's equations had the peculiar symmetry that we now call the Lorentz transformations. He argued that if a reason could be found that would allow distances between matter to change as

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

then lengths would be squeezed in the direction parallel to the ether, explaining why no difference is seen between T_x and T_y. (We will shortly derive this contraction of lengths using special relativity). Lorentz set to work trying to provide a mechanical explanation for this transformation law.

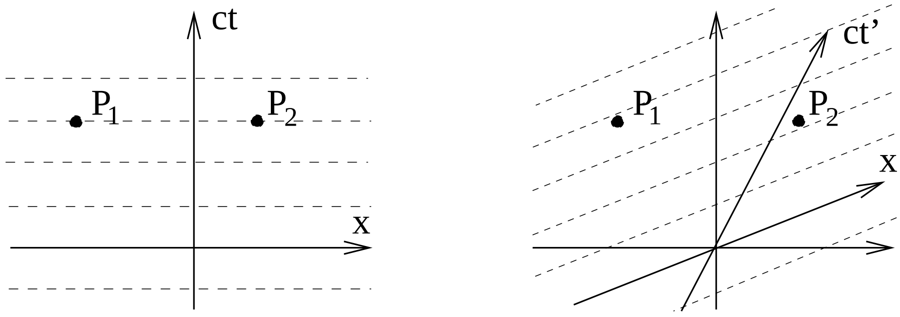

Figure 48: Simultaneity is relative

Although Lorentz had put in place much of the mathematics, the real insight came from Einstein in 1905. He understood that there is no mechanical mechanism underlying the Lorentz transformations. Nor is there an ether. Instead, the Lorentz transformations are a property of space and time themselves.

With Einstein's new take on the principle of relativity, all problems with Maxwell's equation evaporate. There is no preferred inertial frame. Instead, Maxwell's equations work equally well in all inertial frames. However, they are not invariant under the older transformations of Galilean relativity; instead they are the first law of physics to be invariant under the correct transformations (7.6) of Einstein/Lorentz relativity. It's worth pointing out that, from this perspective, we could dispense with the second postulate of relativity all together. We need only insist that the laws of physics – which include Maxwell's equations – hold in all inertial frames. Since Maxwell's equations predict (7.7), this implies the statement that the speed of light is the same in all inertial frames. But since we haven't yet seen the relationship between Maxwell's equations, light and relativity, it's perhaps best to retain the second postulate for now.

7.2 Relativistic Physics

In this section we will explore some of the more interesting and surprising consequences of the Lorentz transformations.

7.2.1 Simultaneity

We start with a simple question: how can we be sure that things happen at the same time? In Newtonian physics, this is a simple question to answer. In that case, we have an absolute time t and two events, P_1 and P_2, happen at the same time if $t_1 = t_2$. However, in the relativistic world, things are not so easy.

We start with an observer in inertial frame S, with time coordinate t. This observer sensibly decides that two events, P_1 and P_2, occur simultaneously if $t_1 = t_2$. In the spacetime diagram on the left of Figure 48 we have drawn lines of simultaneity for this observer.

But for an observer in the inertial frame S', simultaneity of events occurs for equal t'. Using the Lorentz transformation, lines of constant t' become lines described by the equation $t - vx/c^2 =$ constant. These lines are drawn on the spacetime diagram on the right of Figure 48.

The upshot of this is that two events simultaneous in one inertial frame are not simultaneous in another. An observer in S thinks that events P_1 and P_2 happen at the same time. All other observers disagree.

A Train Story

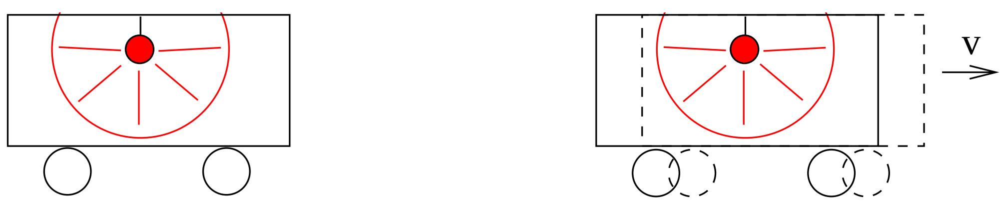

Figure 49: Lights on Trains: Simultaneity is Relative

The fact that all observers cannot agree on what events are simultaneous is a direct consequence of the fact that all observers do agree on the speed of light. We can illustrate this connection with a simple *gedankenexperiment*. (An ugly German word for "thought experiment", a favourite trick of theoretical physicists who can't be bothered to do real experiments). Consider a train moving at constant speed, with a lightbulb hanging from the middle of one of the carriages. A passenger on the train turns on the bulb and, because the bulb is equidistant from both the front and back wall of the carriage, observes that the light hits both walls at the same time.

However, a person standing on the platform as the train passes through disagrees. The light from the bulb travels at equal speed $\pm c$ to the left and right, but the back of the train is rushing towards the point in space where the light first emerged from. The person on the platform will see the light hit the back of the train first.

It is worth mentioning that although the two people disagree on whether the light hits the walls at the same time, this does not mean that they can't be friends.

A Potential Confusion: What the Observer Observes

We'll pause briefly to press home a point that may lead to confusion. You might think that the question of simultaneity has something to do with the finite speed of propagation. You don't see something until the light has travelled to you, just as you don't hear something until the sound has travelled to you. This is *not* what's going on here! A look at the spacetime diagram in Figure 48 shows that we've already taken this into account when deciding whether two events occur simultaneously. The lack of simultaneity between moving observers is a much deeper issue, not due to the finiteness of the speed of light but rather due to the constancy of the speed of light.

The confusion about the time of flight of the signal is sometimes compounded by the common use of the word *observer* to mean "inertial frame". This brings to mind some guy sitting at the origin, surveying all around him. Instead, you should think of the observer more as a Big Brother figure: a sea of clocks and rulers throughout the inertial frame which can faithfully record and store the position and time of any event, to be studied at some time in the future.

Of course, this means that there is a second question we can ask which is: what does the guy sitting at the origin actually see? Now we have to take into account both the relative nature of simultaneity *and* the issues related with the finite speed of propagation. This adds an extra layer of complexity which we will discuss in Section 7.6.

7.2.2 Causality

We've seen that different observers disagree on the temporal ordering of two events. But where does that leave the idea of causality? Surely it's important that we can say that one event definitely occurred before another. Thankfully, all is not lost: there are only some events which observers can disagree about.

To see this, note that because Lorentz boosts are only possible for $v < c$, the lines of simultaneity cannot be steeper than $45°$. Take a point P and draw the $45°$ light rays that emerge from P. This is called the *light cone.* (For once, in the figure, I've drawn this with an extra spatial dimension present to illustrate how this works in spatial dimensions bigger than one). The light cone is really two cones, touching at the point P. They are known as the future light cone and past light cone.

For events inside the light cone of P, there is no difficulty deciding on the temporal ordering of events. All observers will agree that Q occurred after P. However, for events outside the light cone, the matter is up for grabs: some observers will see R as happening after P; some before.

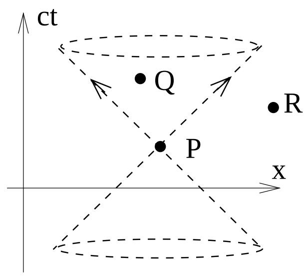

Figure 50:

This tells us that the events which all observers agree can be causally influenced by P are those inside the future light cone. Similarly, the events which can plausibly influence P are those inside the past light cone. This means that we can sleep comfortably at night, happy in the knowledge that causality is preserved, only if nothing can propagate outside the light cone. But that's the same thing as travelling faster than the speed of light.

The converse to this is that if we do ever see particles that travel faster than the speed of light, we're in trouble. We could use them to transmit information faster than light. But another observer would view this as transmitting information backwards in time. All our ideas of cause and effect will be turned on their head. You will therefore be relieved to learn that we will show in Section 7.3 why it is impossible to accelerate particles past the light speed barrier.

There is a corollary to the statement that events outside the lightcone cannot influence each other: there are no perfectly rigid objects. Suppose that you push on one end of a rod. The other end cannot move immediately since that would allow us to communicate faster than the speed of light. Of course, for real rods, the other end does not move instantaneously. Instead, pushing on one end of the rod initiates a sound wave which propagates through the rod, telling the other parts to move. The statement that there is no rigid object is simply the statement that this sound wave must travel slower than the speed of light.

Finally, let me mention that when we're talking about waves, as opposed to point particles, there is a slight subtlety in exactly what must travel slower than light. There are at least two velocities associated to a wave: the group velocity is (usually) the speed at which information can be communicated. This is less than c. In contrast, the phase velocity is the speed at which the peaks of the wave travel. This can be greater than c, but transmits no information.

7.2.3 Time Dilation

We'll now turn to one of the more dramatic results of special relativity. Consider a clock sitting stationary in the frame S' which ticks at intervals of T'. This means that the tick events in frame S' occur at $(ct'_1, 0)$ then $(ct'_1 + cT', 0)$ and so on. What are the intervals between ticks in frame S?

We can answer immediately from the Lorentz transformations (7.6). Inverting this gives

$$t = \gamma\left(t' + \frac{vx'}{c^2}\right)$$

The clock sits at $x' = 0$, so we immediately learn that in frame S, the interval between ticks is

$$T = \gamma T'$$

This means that the gap between ticks is longer in the stationary frame. A moving clock runs more slowly. But the same argument holds for any process, be it clocks, elementary particles or human hearts. The correct interpretation is that time itself runs more slowly in moving frames.

Another Train Story

Figure 51: More Lights on Trains: Time Dilation

Let's go back to our lightbulb and gedankenbahn. If the train has height h, a passenger on the train will measure time $t' = h/c$ for the light to travel from the light bulb to the middle of the floor (i.e. the point directly below the light bulb). What about for the guy on the platform? After the light turns on, the train has moved forward at speed v. To hit the same point on the floor, the light has to travel a distance $\sqrt{h^2 + (vt)^2}$. The time taken is therefore

$$t = \frac{\sqrt{h^2 + (vt)^2}}{c} \quad\Rightarrow\quad t = \frac{h}{c}\sqrt{\frac{1}{1 - v^2/c^2}} = \gamma t'$$

This gives another, more pictorial, derivation of the time dilation formula.

On Muons and Planes

Away from the world of gedankenexperiments, there are a couple of real experimental consequences of time dilation. Certainly the place that this phenomenon is tested most accurately is in particle accelerators where elementary particles routinely reach speeds close to c. The protons spinning around the LHC have $\gamma \approx 3500$. The previous collider in CERN, called LEP, accelerated electrons and positrons to $\gamma \approx 2 \times 10^5$. (Although the electrons in LEP were travelling faster than the protons in LHC, the greater mass of the protons means that there is substantially more energy in the LHC collisions).

The effect of time dilation is particularly vivid on unstable particles which live much longer in the lab frame than in their own rest frame. An early demonstration was seen in *muons* in 1941. These are heavier, unstable, versions of the electron. They decay into an electron, together with a couple of neutrinos, with a half-life of $\tau \approx 2 \times 10^{-6}\ s$. Muons are created when cosmic rays hit the atmosphere, and subsequently rain down on Earth. Yet to make it down to sea level, it takes about $t = 7 \times 10^{-6}\ s$, somewhat longer than their lifetime. Given this, why are there any muons detected on Earth at all? Surely they should have decayed. The reason that they do not is because the muons are travelling at a speed $v \approx 0.99c$, giving $\gamma \approx 10$. From the muon's perspective, the journey only takes $t' = t/\gamma \approx 7 \times 10^{-7}\ s$, somewhat less than their lifetime.

Note that elementary particles are, by definition, structureless. They're certainly not some clock with an internal machinery. The reason that they live longer can't be explained because of some mechanical device which slows down: it is time itself which is running slower.

A more direct test of time dilation was performed in 1971 by Hafele and Keating. They flew two atomic clocks around the world on commercial airliners; two more were left at home. When they were subsequently brought together, their times differed by about $10^{-7}\ s$. There are actually two contributions to this effect: the time dilation of special relativity that we've seen above, together with a related effect in general relativity due to the gravity of the Earth.

Twin Paradox

Two twins, Luke and Leia, decide to spend some time apart. Leia stays at home while Luke jumps in a spaceship and heads at some speed v to the planet Tatooine. With sadness, Leia watches Luke leave but is relieved to see — only a time T later from her perspective — him safely reach the planet.

However, upon arrival, Luke finds that he doesn't like Tatooine so much. It is a dusty, violent place with little to do. So he turns around and heads back to Leia at the same speed v as before. When he returns, he finds that Leia has aged by $T_{\rm Leia} = 2T$. And yet, fresh faced Luke has only aged by $T_{\rm Luke} = 2T/\gamma$. We see, that after the journey, Luke is younger than Leia. In fact, for large enough values of γ, Luke could return to find Leia long dead.

This is nothing more than the usual time dilation story. So why is it a paradox? Well, things seem puzzling from Luke's perspective. He's sitting happily in his inertial spaceship, watching Leia and the whole planet flying off into space at speed v. From his perspective, it should be Leia who is younger. Surely things should be symmetric between the two?

The resolution to this "paradox" is that there is no symmetry between Luke's journey and Leia's. Leia remained in an inertial frame for all time. Luke, however, does not. When he reaches Tatooine, he has to turn around and this event means that he has to accelerate. This is what breaks the symmetry.

We can look at this in some detail. We draw the spacetime diagram in Leia's frame. Luke sits at $x = vt$, or $x' = 0$. Leia sits at $x = 0$. Luke reaches Tatooine at point P. We've also drawn two lines of simultaneity. The point Y is when Leia thinks that Luke has arrived on Tatooine. The point X is where Luke thinks Leia was when he arrived at Tatooine. As we've already seen, it's quite ok for Luke and Leia to disagree on the simultaneity of these points. Let's figure out the coordinates for X and Y.

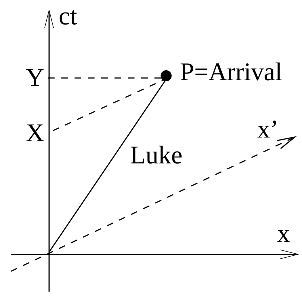

Figure 52:

Event Y sits at coordinate $(cT, 0)$ in Leia's frame, while P is at (cT, vT). The time elapsed in Luke's frame is just the usual time dilation calculation,

$$T' = \gamma\left(T - \frac{v^2 T}{c^2}\right) = \frac{T}{\gamma}$$

We can also work out the coordinates of the event X. Clearly this takes place at $x = 0$ in Leia's frame. In Luke's frame, this is simultaneous with his arrival at Tatooine, so occurs at $t' = T' = T/\gamma$. We can again use the Lorentz transformation

$$t' = \gamma\left(t - \frac{v^2 x}{c^2}\right)$$

now viewed as an equation for t given x and t'. This gives us

$$t = \frac{T'}{\gamma} = \frac{T}{\gamma^2}$$

So at this point, we see that everything is indeed symmetric. When Luke reaches Tatooine, he thinks that Leia is younger than him by a factor of γ. Meanwhile, Leia thinks that Luke is younger than her by the same factor .

Things change when Luke turns around. To illustrate this, let's first consider a different scenario where he doesn't return from Tatooine. Instead, as soon as he arrives, he synchronises his clock with a friend – let's call him Han – who is on his way to meet Leia. Now things are still symmetric. Luke thinks that Leia has aged by T/γ^2 on the outward journey; Han also thinks that Leia has aged by T/γ^2 on the inward journey. So where did the missing time go?

We can see this by looking at the spacetime diagram of Han's journey. We've again drawn lines of simultaneity. From Han's perspective, he thinks that Leia is sitting at point Z when he leaves Tatooine, while Luke is still convinced that she's sitting at point X. It's not hard to check that at point Z, Leia's clock reads $t = 2T - T/\gamma^2$.

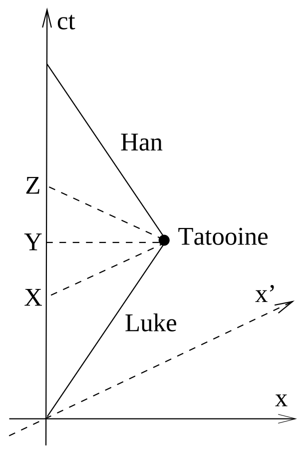

Figure 53:

From this perspective, we can also see what happens if Luke does return home. When he arrives at Tatooine, he thinks Leia is at point X. Yet, in the time he takes to turn around and head home, the acceleration makes her appear to rapidly age, from point X to point Z.

7.2.4 Length Contraction

We've seen that moving clocks run slow. We will now show that moving rods are shortened. Consider a rod of length L' sitting stationary in the frame S'. What is its length in frame S?

To begin, we should state more carefully something which seems obvious: when we say that a rod has length L', it means that the distance between the two end points *at equal times* is L'. So, drawing the axes for the frame S', the situation looks like the picture on the left. The two, simultaneous, end points in S' are P_1 and P_2. Their coordinates in S' are $(ct', x') = (0, 0)$ and $(0, L')$ respectively.

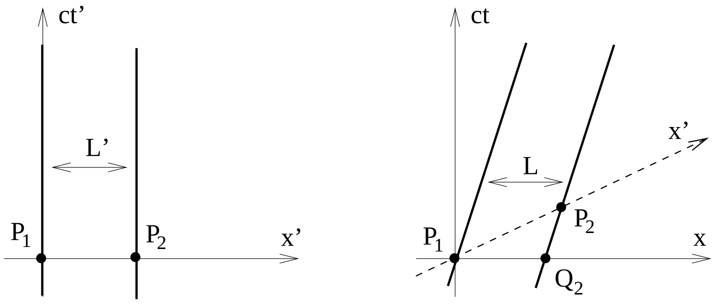

Figure 54: Length Contraction

Now let's look at this in frame S. This is drawn in right-hand picture. Clearly P_1 sits at $(ct, x) = (0, 0)$. Meanwhile, the Lorentz transformation gives us the coordinate for P_2

$$x = \gamma L' \quad \text{and} \quad t = \frac{\gamma v L'}{c^2}$$

But to measure the rod in frame S, we want both ends to be at the same time. And the points P_1 and P_2 are not simultaneous in S. We can follow the point P_2 backwards along the trajectory of the end point to Q_2, which sits at

$$x = \gamma L' - vt$$

We want Q_2 to be simultaneous with P_1 in frame S. This means we must move back a time $t = \gamma v L/c^2$, giving

$$x = \gamma L' - \frac{\gamma v^2 L'}{c^2} = \frac{L'}{\gamma}$$

This is telling us that the length L measured in frame S is

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

It is shorter than the length of the rod in its rest frame by a factor of γ. This phenomenon is known as *Lorentz contraction.*

Putting Ladders in Barns

Take a ladder of length $2L$ and try to put it in a barn of length L. If you run fast enough, can you squeeze it? Here are two arguments, each giving the opposite conclusion

- From the perspective of the barn, the ladder contracts to a length $2L/\gamma$. This shows that it can happily fit inside as long as you run fast enough, with $\gamma \geq 2$
- From the perspective of the ladder, the barn has contracted to length L/γ. This means there's no way you're going to get the ladder inside the barn. Running faster will only make things worse

What's going on? As usual, to reconcile these two points of view we need to think more carefully about the question we're asking. What does it mean to "fit a ladder inside a barn"? Any observer will agree that we've achieved this if the back end gets in the door before the front end hits the far wall. But we know that simultaneity of events is not fixed, so the word "before" in this definition suggests that it may be something different observers will disagree on. Let's see how this works.

The spacetime diagram in the frame of the barn is drawn in the figure with $\gamma > 2$. We see that, from the barn's perspective, both back and front ends of the ladder are happily inside the barn at the same time. We've also drawn the line of simultaneity for the ladder's frame. This shows that when the front of the ladder hits the far wall, the back end of the ladder has not yet got in the door. Is the ladder in the barn? Well, it all depends who you ask.

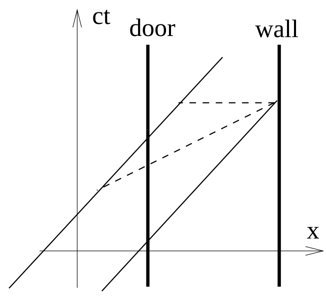

Figure 55:

7.2.5 Addition of Velocities

A particle moves with constant velocity u' in frame S' which, in turn, moves with constant velocity v with respect to frame S. What is the velocity u of the particle as seen in S?

The Newtonian answer is just $u = u' + v$. But we know that this can't be correct because it doesn't give the right answer when $u' = c$. So what is the right answer?

The worldline of the particle in S' is

$$x' = u't' \tag{7.8}$$

So the velocity of the particle in frame S is given by

$$u = \frac{x}{t} = \frac{\gamma(x' + vt')}{\gamma(t' + vx'/c^2)}$$

which follows from the Lorentz transformations (7.6). (Actually, we've used the inverse Lorentz transformations since we want S coordinates in terms of S' coordinates, but these differ only changing $-v$ to v). Substituting (7.8) into the expression above, and performing a little algebra, gives us the result we want:

$$u = \frac{u' + v}{1 + u'v/c^2} \tag{7.9}$$

Note that when $u' = c$, this gives us $u = c$ as expected.

We can also show that if $|u'| < c$ and $|v| < c$ then we necessarily have $-c < u < c$. The proof is simple algebra, if a little fiddly

$$c - u = c - \frac{u' + v}{1 + u'v/c^2} = \frac{c(c - u')(c - v)}{c^2 + u'v} > 0$$

where the last equality follows because, by our initial assumptions, each factor in the final expression is positive. An identical calculation will show you that $-c < u$ as well. We learn that if a particle is travelling slower than the speed of light in one inertial frame, it will also be travelling slower than light in all others.

7.3 The Geometry of Spacetime

> *The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.*
>
> Hermann Minkowski, 1908

We have seen that time is relative, length is relative, simultaneity is relative. Is nothing sacred anymore? Well, the answer is yes: there is one measurement that all observers will agree on.

7.3.1 The Invariant Interval

Let's start by considering a spacetime with just a single spatial coordinate, x. In frame S, two events P_1 and P_2 have coordinates (ct_1, x_1) and (ct_2, x_2). The events are separated by $\Delta t = t_1 - t_2$ in time and $\Delta x = x_1 - x_2$ in space.

We define the *invariant interval* Δs^2 as a measure of the distance between these two points:

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2$$

The advantage of the invariant interval is that it is something all observers agree upon. In frame S', we have

$$\begin{aligned}\Delta s^2 &= \gamma^2\left(c\Delta t' + \frac{v\Delta x'}{c}\right)^2 - \gamma^2\left(\Delta x' + v\Delta t'\right)^2 \\ &= \gamma^2(c^2 - v^2)\Delta t'^2 - \gamma^2\left(1 - \frac{v^2}{c^2}\right)\Delta x'^2 \\ &= c^2\Delta t'^2 - \Delta x'^2\end{aligned} \tag{7.10}$$

where, in going from the first line to the second, we see that the cross-terms $\Delta t'\Delta x'$ cancel out.

Including all three spatial dimensions, the definition of the invariant interval is

$$\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \tag{7.11}$$

which, again, is the same in all frames. (The only non-trivial part of the calculation is (7.10) above since y and z are invariant under a boost in the x direction).

The spacetime of special relativity is topologically $\mathbf{R}^4$. When endowed with the measure of distance (7.11), this spacetime is referred to as *Minkowski space*. Although topologically equivalent to Euclidean space, distances are measured differently. To stress the difference between the time and spatial directions, Minkowski space is sometimes said to have dimension $d = 1 + 3$. (For once, it's important that you don't do this sum!).

In later courses — in particular General Relativity — you will see the invariant interval written as the distance between two infinitesimally close points. In practice that just means we replace all the Δ(something)s with d(something)s.

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$

In this infinitesimal form, ds^2 is called the *line element*.

The invariant interval provides an observer-independent characterisation of the distance between any two events. However, it has a strange property: it is not positive definite. Two events whose separation is $\Delta s^2 > 0$ are said to be *timelike* separated. They are closer together in space than they are in time. Pictorially, such events sit within each others light cone.

In contrast, events with $\Delta s^2 < 0$ are said to be *spacelike* separated. They sit outside each others light cone. From our discussion in Section 7.2.2, we know that two observers can disagree about the temporal ordering of spacelike separated events. However, they agree on the ordering of timelike separated events. Note that since $\Delta s^2 < 0$ for spacelike separated events, if you insist on talking about Δs itself then it must be purely imaginary. However, usually it will be perfectly fine if we just talk about Δs^2.

Finally, two events with $\Delta s^2 = 0$ are said to be *lightlike* separated. Notice that this is an important difference between the invariant interval and most measures of distance that you're used to. Usually, if two points are separated by zero distance, then they are the same point. This is not true in Minkowski spacetime: if two points are separated by zero distance, it means that they can be connected by a light ray.

A Rotational Analogy

There's a simple analogy to understand the meaning of the invariant interval. Let's go back to consider three dimensional Euclidean space with coordinates $\mathbf{x} = (x, y, z)$. An observer measures the position of a stationary object — let's say a helicopter — and proudly announces the x and y and z coordinates of the helicopter.

Meanwhile, a second observer shares the same origin as the first, but has rotated his axes to use coordinates $\mathbf{x}' = (x', y', z')$ where $\mathbf{x}' = R\mathbf{x}$ for some rotation matrix R. He too sees the helicopter, and declares that it sits at coordinates x', y' and z'.

Of course, there's no reason why the coordinates of the two observers should agree with each other. However, there is one quantity that should be invariant: the distance from the origin (which is shared by both observers) to the helicopter. In other words, we should find that

$$s^2_{\text{Euclidean}} = x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2 \tag{7.12}$$

And, of course, this is indeed true if the rotation matrix obeys $R^T R = 1$.

The essence of special relativity is nothing more than an extrapolation of the discussion above. The Lorentz boosts can be should be thought of as a rotation between space and time. The individual spatial and temporal coordinates are different for the two observers, but there remains an invariant distance. The only thing that's different is that the time and space directions in this invariant distance (7.11) come with different minus signs. We'll now explore this relationship between boosts and rotations in some detail.

7.3.2 The Lorentz Group

We have defined the interval (7.11) as the measure of distance which is invariant under Lorentz transformations. However, it is actually better to look at things the other way: the invariant interval is the primary object. This is a property of spacetime which defines the Lorentz transformations. Let's see how the argument runs this way around.

If we sit at the origin in a fixed frame S, the coordinates of an event can be written as a *four vector* X. We won't denote that this is a vector by bold font or squiggly underlines (which we're really saving for three-dimensional spatial vectors). We're just getting sophisticated now and just the capital letter will have to suffice. However, we will sometimes use index notation, in which the components of the 4-vector are

$$X^\mu = (ct, x, y, z) \qquad \mu = 0, 1, 2, 3$$

Note that we write the indices running from $\mu = 0$ to $\mu = 3$ rather than starting at 1. The zeroth component of the vector is time (multiplied by c). The invariant distance between the origin and the point P can be written as an inner product, defined as

$$X \cdot X \equiv X^T \eta X = X^\mu \eta_{\mu\nu} X^\nu \tag{7.13}$$

In the first expression above we are using matrix-vector notation and in the second we have resorted to index notation, with the summation convention for both indices μ and ν. The matrix η is given by

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

This matrix is called the *Minkowski metric*. With this expression for the Minkowski metric, the inner product becomes

$$X \cdot X = c^2 t^2 - x^2 - y^2 - z^2$$

which is indeed the invariant distance (7.11) between the origin and the point X as promised.

Following our characterisation of distances using the invariant interval, a four vector obeying $X \cdot X > 0$ is said to be *timelike*; one with $X \cdot X < 0$ is said to be *spacelike*; and one with $X \cdot X = 0$ is said to be *lightlike* or, alternatively, *null*.

The Lorentz transformation can be thought of as a 4×4 matrix Λ, rotating the coordinates in frame S to coordinates in frame S', such that the four vector becomes

$$X' = \Lambda X$$

This can also be written index notation as $X'^{\mu} = \Lambda^{\mu}{}_{\nu} X^{\nu}$. The Lorentz transformations are defined to be those matrices which leave the inner product invariant. This means that

$$X' \cdot X' = X \cdot X$$

From our definition (7.13), we see that this is true only if Λ obeys the matrix equation

$$\Lambda^T \eta \Lambda = \eta \tag{7.14}$$

Let's try to understand the solutions to this. We can start by counting how many we expect. The matrix Λ has $4 \times 4 = 16$ components. Both sides of equation (7.14) are symmetric matrices, which means that the equation only provides 10 constraints on the coefficients of Λ. We therefore expect to find $16 - 10 = 6$ independent solutions.

The solutions to (7.14) fall into two classes. The first class is very familiar. Let's look at solutions of the form

$$\Lambda = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & \mathrm{R} & \\ 0 & & & \end{array} \right) \tag{7.15}$$

where R is a 3×3 matrix. These transformations change space, but leave time intact. The condition (7.14) reduces to a condition for the matrix R,

$$R^T R = 1$$

where the right-hand side is understood to be the 3×3 unit matrix. But this is something that we've seen before: it is the requirement for R to be a rotation matrix. There are three such independent matrices, corresponding to rotations about the three different spatial axes.

The remaining three solutions to (7.14) are the Lorentz boosts that have preoccupied us for much of this Section. The boost along the x axis is given by

$$\Lambda = \left(\begin{array}{cc|cc} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \tag{7.16}$$

These are precisely the Lorentz transformations (7.6). Two further solutions to (7.14) come from boosting along the y and z directions.

The set of all matrices Λ obeying (7.14) form the *Lorentz group*, denoted $O(1,3)$. You can easily check that they indeed obey all axioms of a group. Taking the determinant of both sides of (7.14), we see that $\det \Lambda^2 = 1$, so the Lorentz group splits into two pieces with $\det \Lambda = \pm 1$. The subgroup with $\det \Lambda = 1$ is called the proper Lorentz group and is denoted $SO(1,3)$.

There is one further decomposition of the Lorentz group. Any element can either flip the direction of time or leave it invariant. Those transformations which preserve the direction of time are called orthochronous. The group of proper orthochronous Lorentz transformations is denoted $SO^+(1,3)$ although people like me are usually lazy and just refer to it as $SO(1,3)$.

Rapidity

We previously derived the velocity addition law (7.9). Let's see how we get this from the matrix approach above. We can focus on the 2×2 upper-left hand part of the matrix in (7.16). We'll write this as

$$\Lambda[v] = \begin{pmatrix} \gamma & -\gamma v/c \\ -\gamma v/c & \gamma \end{pmatrix}$$

If we combine two boosts, both in the x direction, the resulting Lorentz transformation is

$$\Lambda[v_1]\Lambda[v_2] = \begin{pmatrix} \gamma_1 & -\gamma_1 v_1/c \\ -\gamma_1 v_1/c & \gamma_1 \end{pmatrix} \begin{pmatrix} \gamma_2 & -\gamma_2 v_2/c \\ -\gamma_2 v_2/c & \gamma_2 \end{pmatrix}$$

It takes a little bit of algebra, but multiplying out these matrices you can show that

$$\Lambda[v_1]\,\Lambda[v_2] = \Lambda\left[\frac{v_1 + v_2}{1 + v_1 v_2/c^2}\right]$$

which is again the velocity addition rule (7.9), now for the composition of boosts.

The algebra involved in the above calculation is somewhat tedious; the result somewhat ugly. Is there a better way to see how this works? We can get a clue from the rotation matrices R. Recall that the 2×2 matrix which rotates a plane by angle θ is

$$R[\theta] = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

If we perform two rotations in succession, we have

$$R[\theta_1]R[\theta_2] = R[\theta_1 + \theta_2]$$

But the nice addition rule only worked because we were clever in parameterising our rotation by an angle. In the case of Lorentz boosts, there is a similarly clever parameterisation. Instead of using the velocity v, we define the *rapidity* φ by

$$\gamma = \cosh\varphi$$

We can see one of the nice things about this definition if we look at

$$\sinh\varphi = \sqrt{\cosh^2\varphi - 1} = \sqrt{\gamma^2 - 1} = \frac{v\gamma}{c}$$

This is the other component of the Lorentz boost matrix. We can therefore write

$$\Lambda[\varphi] = \begin{pmatrix} \cosh\varphi & -\sinh\varphi \\ -\sinh\varphi & \cosh\varphi \end{pmatrix} \tag{7.17}$$

Looking again at the composition of two Lorentz boosts, we see that the rapidities add, just like the angles of rotation

$$\Lambda[\varphi_1]\,\Lambda[\varphi_2] = \Lambda[\varphi_1 + \varphi_2]$$

The matrix description of the Lorentz boost (7.17) shows most clearly the close relationship between rotations and boosts.

7.3.3 A Rant: Why $c = 1$

We started this section by mentioning that the speed of light, $c = 299792458\ ms^{-1}$ is exact. The only reason that this fundamental constant is exactly an integer is because the meter is defined to be the distance travelled by light in 1/299792458 seconds.

In our everyday world, the meter is a very useful unit. It is roughly the size of most things in my house. But viewed from the perspective of fundamental physics, it is rather parochial. If we're going to pick the speed of light to be an integer, we should probably pick one that is easier to remember. Like $c = 1$. We can do this by picking a different unit of length, namely

$$c = 1 \;\; (\text{light second})\, s^{-1}$$

where a light second is the distance travelled by light in one second.

There's a better way of thinking about this: the existence of a universal speed of light is Nature's way of telling us that space and time are more similar than our ancestors realised. We only labelled space and time with different units because we were unaware of the relationship between them.

We can illustrate this by going back to the rotational analogy. Suppose that you decided that all distances in the x-direction should be measured in centimeters, while all distances in the y-direction should be measured in inches. You then declared that there was a new, fundamental constant of Nature – let's call it λ – given by

$$\lambda \approx 2.54 \; cm \; (inch)^{-1}$$

Why is this a dumb thing to do? The reason it's dumb is because of the rotational symmetry of the laws of physics: different observers have different x and y coordinates and can quite happily pick the same unit of measurement for both. But we've learned in this section that there is also a symmetry between space and time. Insisting that we retain the conversion factor c in the fundamental laws of physics is no more sensible than retaining λ.

Despite my rant, in these lectures, we will retain c in all equations. (Although we will use units which allow us to set $\lambda = 1$). But in future courses, it is common practice to work with the convention $c = 1$. The equations look simpler and the only price you pay is that the units of time and space are equivalent. If, at the end of the day, you want to get your answer in terms of meters or seconds or whatever then you can always put the factors of c back in by dimensional analysis.

7.4 Relativistic Kinematics

So far, our discussion has been focussed on what the world looks like to different observers. Let's now return to the main theme of these lectures: the motion of particles. Remember that our ultimate goal is to construct laws of physics which look the same to all inertial observers. For this reason, we will start by defining some of the basic elements that go into the laws of physics: velocity, momentum and acceleration. We want to define these in such a way that they have nice transformation properties when viewed from different inertial frames.

7.4.1 Proper Time

We started these lectures in Section 1 by describing the trajectory of particle in an inertial frame in terms of a curve $\mathbf{x}(t)$ and velocity $\mathbf{u} = d\mathbf{x}/dt$. There's nothing incorrect with this description in special relativity but, as we will see, there's a much better way to parameterise the trajectory of a particle.

Let's start by considering a particle at rest at the origin of frame S' with $\mathbf{x}' = 0$. The invariant interval between two different points on the worldline of the particle is

$$\Delta s^2 = c^2 \Delta t'^2$$

We see that the invariant interval between two points on the worldline is proportional to the time experienced by the particle. But this must be true in all frames. The time experienced by the particle is called the *proper time*, τ. In all frames, it is given by

$$\Delta\tau = \frac{\Delta s}{c}$$

where Δs is real as long as the particle doesn't travel faster than the speed of light, so it sits on a timelike trajectory. (We keep promising to prove that a particle is unable to travel faster than light...we are almost there!)

Proper time provides a way to parameterise the trajectory of a particle in a manner that all inertial observers will agree on. Consider the trajectory of a general particle, not necessarily travelling in a straight line. Viewed from an inertial frame S, the worldline can be parameterised by $\mathbf{x}(\tau)$ and $t(\tau)$. This has several advantages.

For example, we can use this formulation to determine the time experienced by a particle moving along a general trajectory. Along a small segment of its trajectory, a particle experiences proper time

$$d\tau = \sqrt{dt^2 - \frac{d\mathbf{x}^2}{c^2}} = dt\sqrt{1 - \frac{1}{c^2}\left(\frac{d\mathbf{x}}{dt}\right)^2} = dt\sqrt{1 - \frac{u^2}{c^2}}$$

from which we have

$$\frac{dt}{d\tau} = \gamma \tag{7.18}$$

Note that γ here is a function of the speed, u, of the particle seen by the observer in S. From this, the total time T experienced by a particle as it travels along its worldline is simply the sum of the proper times associated to each small segment,

$$T = \int d\tau = \int \frac{dt}{\gamma} \tag{7.19}$$

7.4.2 4-Velocity

We'll now explain why it's useful to parameterise the trajectory of a particle in terms of proper time τ. We can write a general trajectory in spacetime using the 4-vector:

$$X(\tau) = \begin{pmatrix} ct(\tau) \\ \mathbf{x}(\tau) \end{pmatrix}$$

From this, we can define the *4-velocity*,

$$U = \frac{dX}{d\tau} = \begin{pmatrix} c\, dt/d\tau \\ d\mathbf{x}/d\tau \end{pmatrix}$$

Using the relationship (7.18) between the proper time of the particle τ and the observer's time t we can write this as

$$U = \frac{dt}{d\tau} \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} = \gamma \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \tag{7.20}$$

where $\mathbf{u} = d\mathbf{x}/dt$. This definition of the 4-velocity has a nice property: if an observer in frame S measures a particle's 4-velocity as U, then an observer in frame S' with $X' = \Lambda X$ will measure the 4-velocity

$$U' = \Lambda U \tag{7.21}$$

This transformation holds only because $d\tau$ is invariant, meaning that it is the same for all observers. In contrast, if we had tried to define a 4-velocity by, say, $V = dX/dt$ then both X and t would change under a Lorentz transformation and we would be left with a messy, complicated expression for V in frame S'. Our definition of U differs from V by the extra factor of γ in (7.20). This is all important!

We now have two objects which transform nicely under Lorentz transformations: the coordinates $X \to \Lambda X$ and the 4-velocity $U \to \Lambda U$. Quantities like this are called *4-vectors*. It's a name that we've already used to label points in spacetime. More generally, the *definition* of a 4-vector is any 4-component object A which transforms as $A \to \Lambda A$ under a Lorentz transformation.

Because of the simple transformation law (7.21), we can immediately import some of the things that we learned from our previous discussion of Lorentz groups. In particular, from the definition of Λ given in (7.14), we know that the inner product

$$U \cdot U = U^T \eta U$$

is invariant. It is the same for all observers: $U \cdot U = U' \cdot U'$.

Let's look at a simple example. A particle which is stationary in frame S has 4-velocity

$$U^\mu = (c, 0, 0, 0)$$

and so $U \cdot U = c^2$. But this must be true in all frames. We can check this explicitly from (7.20) (we'll take the middle equation to illustrate the point) which gives us

$$U \cdot U = \left(\frac{dt}{d\tau}\right)^2 \left(c^2 - u^2\right) = \left(\frac{dt}{d\tau}\right)^2 \frac{c^2}{\gamma^2} = c^2$$

This result also helps answer a puzzle. In Newtonian mechanics, if we want to specify the velocity, we only have to give three numbers $\mathbf{u}$. But in special relativity, the velocity is a 4-vector U. Nonetheless, we still only need specify three variables because U is not any 4-vector: it is constrained to obey $U \cdot U = c^2$.

Addition of Velocities Revisited

In Section 7.2.5, we derived the rule for the addition of velocities in one-dimension. But what if the velocity of a particle is not aligned with the relative velocity between S and S'? The addition of velocities in this case is simple to compute using 4-vectors. We start with a particle in frame S travelling with 4-velocity

$$U = \begin{pmatrix} \gamma_u c \\ u\gamma_u \cos\alpha \\ u\gamma_u \sin\alpha \\ 0 \end{pmatrix}$$

Here we've added the subscript to $\gamma_u = (1 - u^2/c^2)^{-1/2}$ to distinguish it from the γ-factor arising between the two frames. Frame S' moves in the x-direction with speed v relative to S. The Lorentz boost is given in (7.16). In frame S', the 4-velocity is then

$$U' = \Lambda U = \gamma_u \begin{pmatrix} \left(1 - (uv/c^2)\cos\alpha\right)\gamma_v c \\ (u\cos\alpha - v)\gamma_v \\ u \sin\alpha \\ 0 \end{pmatrix} \equiv \begin{pmatrix} \gamma_{u'} c \\ u'\gamma_{u'} \cos\alpha' \\ u'\gamma_{u'} \sin\alpha' \\ 0 \end{pmatrix} \qquad (7.22)$$

Dividing the t and x components of this 4-vector, we recover the velocity transformation law (7.9) for the speed in the x-direction, namely

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

Meanwhile, dividing the y component by the x component gives us a formula for the angle α' that the particles trajectory makes with the x'-axis,

$$\tan\alpha' = \frac{u\sin\alpha}{\gamma_v(u\cos\alpha - v)} \tag{7.23}$$

There is an interesting application of this formula to *stellar aberration*. The apparent position of the stars changes over a 6 month period as the Earth moves around the Sun. This is *not* due to difference in position of the Earth relative to the star: the size of the Earth's orbit is puny compared to the distance to the star. Instead, it is a special relativistic effect, due to the motion of the Earth around the Sun. We can derive this effect by applying the formulae above to light. Choose coordinate system so that the position of the star — and hence the velocity of starlight — lies in the (x, y) plane, while the Earth's orbit lies in the (x, z) plane. Then we simply need to substitute $u = c$ in the above equation, while v is the component of the Earth's velocity in the x-direction.

The algebra can get somewhat messy, so it's worth walking you through the steps. It's best to return to (7.22), now with $u = u' = c$ as befits a light ray. (It's advisable to first substitute the factor of $\gamma_{u'}$ in the x and y components for the expression for $\gamma_{u'}$ given in the temporal component). This then gives us

$$\sin\alpha' = \frac{c\sin\alpha}{\gamma_v(c - v\cos\alpha)} \quad \text{and} \quad \cos\alpha' = \frac{c\cos\alpha - v}{c - v\cos\alpha}$$

We now use the trig identity $\tan(\alpha'/2) = \sin\alpha'/(1+\cos\alpha')$. Substituting the expressions above, a little algebra then results in the stellar aberration formula,

$$\tan\frac{\alpha'}{2} = \sqrt{\frac{c+v}{c-v}}\,\tan\frac{\alpha}{2}$$

Here α is the angle that the star would appear at if the Earth was at rest. Or, said another way, α is the angle at which the star appears at when the motion of the Earth is perpendicular to the direction of the star. Meanwhile, α' is the angle at which the star appears when the Earth is travelling with speed v in the x-direction. As the Earth moves around the Sun, the component of its velocity v in the x-direction changes, meaning that the angle α' at which the star appears also changes.

7.4.3 4-Momentum

The *4-momentum* is defined by

$$P = mU = \begin{pmatrix} mc\gamma \\ m\gamma\mathbf{u} \end{pmatrix} \tag{7.24}$$

where m is the mass of the particle, usually referred to as the *rest mass*. Importantly, it will turn out that P is the quantity that is conserved in the relativistic context. The spatial components give us the relativistic generalisation of the 3-momentum,

$$\mathbf{p} = m\gamma\mathbf{u} \tag{7.25}$$

Notice that as the particle approaches the speed of light, $u \to c$, the momentum diverges $p \to \infty$. Since momentum is conserved in all processes, this is really telling us that massive particles cannot break the speed of light barrier. (Here the word "massive" doesn't mean "really really big": it just means "not massless", or $m \neq 0$). This is sometimes interpreted by viewing the quantity $m\gamma$ as a velocity-dependent *relativistic mass*. In these terms, the relativistic mass of the particle diverges $m\gamma \to \infty$ as the particle approaches the speed of light. The words may be different, but the maths (and underlying physics) is the same: particles are bound by Nature's speed limit. Nothing can travel faster than the speed of light.

What is the interpretation of the time-component of the momentum 4-vector, P^0. We can get a hint of this by Taylor expanding the γ factor,

$$P^0 = \frac{mc}{\sqrt{1-u^2/c^2}} = \frac{1}{c}\left(mc^2 + \frac{1}{2}mu^2 + \ldots\right) \tag{7.26}$$

The first term is just a constant. But the second term is something familiar: it is the non-relativistic kinetic energy of the particle. This, coupled with the fact that all four components of P are conserved, strongly suggests that the right interpretation of P^0 is the energy of the particle (divided by c), so

$$P = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix} \tag{7.27}$$

To show that P^0 is indeed related to the energy in this way requires a few more techniques than we will develop in this course. The cleanest way is to use Noether's theorem – which we mentioned briefly in Section 5.1.4 – for relativistic systems and see that P^0 is the quantity that follows from time translational invariance[7].

[7]You can read more about this for particle mechanics in the *Classical Dynamics* lecture notes and, for relativistic field theories, in the *Quantum Field Theory* lecture notes. Both are available at http://www.damtp.cam.ac.uk/user/tong/teaching.html

The expansion of (7.26) shows that both the mass and the kinetic energy contribute to the energy of a particle. These combine to give

$$E = m\gamma c^2 \tag{7.28}$$

Notice that as the particle approaches the speed of light, its energy diverges. Yet again, we see a barrier to breaking the speed limit: as we approach the speed of light, the energy required to make a particle go just a little faster gets bigger and bigger.

For a stationary particle, all its energy is contained in its rest mass, giving us the famous slogan

$$E = mc^2$$

There's a nice way to rearrange (7.28), to replace the $\mathbf{u}$ in the γ factor with $\mathbf{p}$ defined in (7.25). But the algebra is laborious. Instead there's a cute trick that gives the result much more quickly: we look at the inner product $P \cdot P$. In the rest frame of the particle, $P = (mc, 0, 0, 0)$ and we have

$$P \cdot P = m^2 c^2 \tag{7.29}$$

But the inner product is an invariant, holding in any frame. From (7.27), we have

$$P \cdot P = \frac{E^2}{c^2} - \mathbf{p}^2$$

Equating these two expressions give

$$E^2 = \mathbf{p}^2 c^2 + m^2 c^4 \tag{7.30}$$

This is the generalisation of $E = mc^2$ to include the kinetic energy. This equation can also be derived the hard way by playing around with (7.28) and (7.25).

The identification $P^0 = E/c$ has dramatic consequences. In Newtonian mechanics, we boasted about the conservation of energy, but implicit in everything we did was the more elementary fact that mass is conserved. Even in the variable mass problems of Section 5.3, the mass never disappered: it just left our rocket ship. However, relativity teaches us that the conservation of mass is subsumed into the conservation of energy. There is nothing that guarantees that they are individually conserved. Just as potential energy can be converted into kinetic energy, so too can mass be converted into kinetic energy. In Japan, in 1945, this fact was vividly demonstrated.

7.4.4 Massless Particles

Until now, we built our discussion of particle trajectories on proper time. But, looking back at Section 7.4.1, proper time is only defined for time-like trajectories. This is fine for massive particles. But what about for massless particles? We can sidestep the need for proper time by looking at the invariant of the 4-momentum (7.29) which, for particles with $m = 0$, tells us that the 4-momentum must be null,

$$P \cdot P = 0$$

This means that the 4-momentum of a massless particle necessarily lies along a light ray.

This fact also allows us to clarify one of our original postulates of special relativity: that the speed of light is the same for all inertial frames. You may wonder why the propagation of light, an electromagnetic phenomenon, is singled out for special treatment. The answer is: because the photon – the particle of light – is massless. In fact, a better way of stating the postulate is to say that there is an upper speed limit in the Universe, which is the same for all inertial observers. Any massless particle must travel at this speed limit. All massive particles must go slower.

We know of only two types of massless particles in the Universe: the photon and the graviton. Both of these owe their particle-like nature to quantum mechanics (actually, this is true of all particles) and have a classical analog as light waves and gravity waves respectively. You've all seen light waves (literally!) and individual photons have been routinely measured in experiments for more than a century. In contrast there is good, but indirect, evidence for gravity waves (this is expected to change soon, with experiments up and running which are designed to directly detect gravity waves). There appears to be no hope at all of detecting an individual graviton, at least within our lifetimes.

Until the late 1990s, it was thought that neutrinos were also massless. It is now known that they have a small, but finite mass. (Actually, there's a caveat here: there are three different types of neutrino: an electron neutrino, a muon neutrino and a tau neutrino. The differences between their masses are known to be of order of 0.01 - 0.1 eV and there are constraints which limit the sum of their masses to be no greater than 0.3 eV or so. But the absolute scale of their masses has not yet been determined. In principle, one of the three neutrinos may be massless).

From (7.30), the energy and momentum of a massless particle are related by $E^2 = p^2c^2$. The four momentum takes the form

$$P = \frac{E}{c}\begin{pmatrix} 1 \\ \hat{\mathbf{p}} \end{pmatrix}$$

where $\hat{\mathbf{p}}$ is a unit vector in the direction of the particle's motion.

To get an expression for the energy, we need a result from quantum mechanics which relates the energy to the wavelength λ of the photon or, equivalently, to the angular frequency $\omega = 2\pi c/\lambda$,

$$E = \hbar\omega = \frac{2\pi\hbar c}{\lambda}$$

There's something rather nice about how this equation ties in with special relativity. Suppose that in your frame, the photon has energy E. But a different observer moves towards the light with velocity v. By the Lorentz transformation, he will measure the 4-momentum of the photon to be $P' = \Lambda P$ and, correspondingly, will see a bigger energy $E' > E$. From the above equation, this implies that he will see a smaller wavelength. But this is nothing other than Lorentz contraction.

The phenomenon of different observers observing different wavelengths of light is called the *Doppler effect*. You will derive this in the problem sheet.

Tachyons and Why They're Nonsense

It is sometimes stated that a particle which has imaginary mass, so that $m^2 < 0$, will have $P \cdot P < 0$ and so travel consistently at speeds $u > c$. Such particles are called *tachyons*. They too would be unable to cross Nature's barrier at $u = c$ and are consigned to always travel on spacelike trajectories.

Although, consistent within the framework of classical relativistic particle mechanics, the possibility of tachyons does not survive the leap to more sophisticated theories of physics. All our current best theories of physics are written in the framework of quantum field theory. Here particles emerge as ripples of fields, tied into small lumps of energy by quantum mechanics. But in quantum field theory, it is not unusual to have fields with imaginary mass $m^2 < 0$. The resulting particles do not travel faster than the speed of light. Instead, imaginary mass signals an instability of the vacuum.

7.4.5 Newton's Laws of Motion

Finally, we are in a position to write down Newton's law of motion in a manner that is consistent with special relativity: it is

$$\frac{dP^\mu}{d\tau} = F^\mu \tag{7.31}$$

where F^μ are the components of a 4-vector force. It is not difficult to anticipate that the spatial components of F should be related to the 3-vector force $\mathbf{f}$. (This is the same thing that we've been calling $\mathbf{F}$ up until now, but we'll lower its standing to a small f to save confusion with the 4-vector). In fact, we need an extra factor of γ, so

$$F = \begin{pmatrix} F^0 \\ \gamma \mathbf{f} \end{pmatrix}$$

With this factor of γ in place, the spatial components of Newton's equation (7.31) agree with the form that we're used to in the reference frame S,

$$\frac{d\mathbf{p}}{dt} = \frac{d\tau}{dt}\frac{d\mathbf{p}}{d\tau} = \frac{1}{\gamma}\frac{d\mathbf{p}}{d\tau} = \mathbf{f}$$

Similarly, a quick calculation shows that the temporal component F^0 is related to the power: the rate of change of energy with time

$$F^0 = \frac{dP^0}{d\tau} = \frac{\gamma}{c}\frac{dE}{dt}$$

With these definitions, we can derive a familiar equation, relating the change in energy to the work done. Consider a particle with constant rest mass m, so that $P \cdot P = m^2c^2$ is unchanging. Using $P^0 = m\gamma c$ and $\mathbf{p} = m\gamma\mathbf{u}$, we have

$$\frac{d}{d\tau}(P \cdot P) = 2P^0\,\frac{dP^0}{d\tau} - 2\mathbf{p} \cdot \frac{d\mathbf{p}}{d\tau} = 2\gamma^2 m\left(\frac{dE}{dt} - \mathbf{u}\cdot\mathbf{f}\right) = 0$$

All of this is just to show how the familiar laws of Newtonian physics sit within special relativity.

Electromagnetism Revisited

Ironically, equation (7.31) is rarely used in relativistic physics! The reason is that by the time we are in the relativistic realm, most of the forces that we've come across are no longer valid. The one exception is the electromagnetic force law for a particle of charge q that we met in Section 2.4. This does have a relativistic formulation, with the equation of motion given by

$$\frac{dP^\mu}{d\tau} = \frac{q}{c} G^\mu{}_\nu U^\nu$$

where U^ν is the 4-velocity of the particle and $G^\mu{}_\nu$ is the *electromagnetic tensor*, a 4×4 matrix which contains the electric and magnetic fields,

$$G^\mu{}_\nu = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ E_1 & 0 & cB_3 & -cB_2 \\ E_2 & -cB_3 & 0 & cB_1 \\ E_3 & cB_2 & -cB_1 & 0 \end{pmatrix}$$

(This tensor often goes by the name $F^\mu{}_\nu$, but we've chosen to call it G to save confusion with the force 4-vector). The spatial components of the four-vector equation gives rise to the familiar Lorentz force law (2.19). The temporal component gives the rate of work done, $dE/dt = q\mathbf{E} \cdot \mathbf{u}$.

7.4.6 Acceleration

We can construct a four-vector for acceleration simply by

$$A \equiv \frac{dU}{d\tau}$$

Note that because $U \cdot U = c^2$, we must have that A is always orthogonal to U in the Minkowski sense: $A \cdot U = 0$.

Suppose that the velocity of a particle in frame S is $\mathbf{u}$. Then, in this frame, the Newtonian notion of 3-acceleration is $\mathbf{a} = d\mathbf{u}/dt$. Recalling our expression relating time and proper time, $dt/d\tau = \gamma$, we see that the four acceleration actually depends on both $\mathbf{u}$ and $\mathbf{a}$; it is

$$A = \gamma \begin{pmatrix} \dot{\gamma} c \\ \dot{\gamma}\mathbf{u} + \gamma \mathbf{a} \end{pmatrix}$$

with $\dot{\gamma} \equiv d\gamma/dt$.

Let's now suppose that we sit in an inertial frame S' in which, at a fixed moment of time t, the particle is instantaneously at rest. Obviously, if the particle is accelerating, this will not coincide with the particle's rest frame an instant later, but momentarily this will do fine. Since $\mathbf{u}' = 0$ in this frame, the 4-acceleration is

$$A' = \begin{pmatrix} 0 \\ \mathbf{a}' \end{pmatrix}$$

with $\mathbf{a}' = d\mathbf{u}'/dt'$. (Note that you need to do a small calculation here to check that $\dot{\gamma}(u = 0) = 0$). But, since we have constructed our acceleration as a 4-vector, A and A' must be related by a Lorentz transformation. To make matters easy for ourselves, let's take both $\mathbf{u}$ and $\mathbf{a}$ to lie in the x-direction so that we can consistently ignore the y and z-directions. Then the Lorentz transformation tells us

$$A = \gamma \begin{pmatrix} \dot{\gamma} c \\ \dot{\gamma} u + \gamma a \end{pmatrix} = \begin{pmatrix} \gamma & u\gamma/c \\ u\gamma/c & \gamma \end{pmatrix} \begin{pmatrix} 0 \\ a' \end{pmatrix} = \begin{pmatrix} u\gamma a'/c \\ \gamma a' \end{pmatrix}$$

From the top component, we can determine the relationship between the accelerations a and a' seen in the two frames,

$$a \equiv \dot{u} = \left(1 - u^2/c^2\right)^{3/2} a'$$

Suppose now that the particle undergoes constant acceleration. As with everything in special relativity, we need to be more careful about what we mean by this. The natural interpretation is that the acceleration in the frame of the particle is constant. Mathematically, this means that a' is constant. In contrast, viewed from frame S, the acceleration is not constant. Indeed, for constant a', we can integrate our equation above to get u, the velocity seen in frame S as a function of time. If we assume that $u = 0$ when $t = 0$, we have

$$u = \frac{a'ct}{\sqrt{c^2 + a'^2 t^2}} \qquad \Rightarrow \qquad \gamma(t) = \sqrt{1 + \frac{a'^2 t^2}{c^2}} \tag{7.32}$$

Since $u = \dot{x}$, integrating the first of these equations once more gives us the position in the frame S as a function of time,

$$x = \frac{c}{a'} \left(\sqrt{c^2 + a'^2 t^2} - c \right) \tag{7.33}$$

where we've picked an integration constant so that $x = 0$ at time $t = 0$. We see that the particle moves on the hyperbola shown in the figure. Viewed from S, the particle approaches, but never reaches, the speed of light.

Notice that a particle at point P in the diagram can only receive information from within its own past lightcone, denoted by the red dotted lines in the figure. However, if it continues along its accelerated trajectory, it can never receive any information from the whole part of spacetime to the left of the null line $x = ct$. This part of the Universe will forever remain a mystery to an accelerated observer. The null cone, defined by, $x = ct$, which forms the boundary of the mysterious region is called the *Rindler event horizon*. It has many things in common with the event horizon of a black hole and, indeed, the Rindler horizon is often used as a toy model to understand some of the stranger aspects of black hole physics. Of course, if an accelerated observer really wants to see what's behind the horizon, it's easy: he just stops accelerating. If an observer in the background of a black hole wants to see what's behind the horizon, he must be somewhat braver.

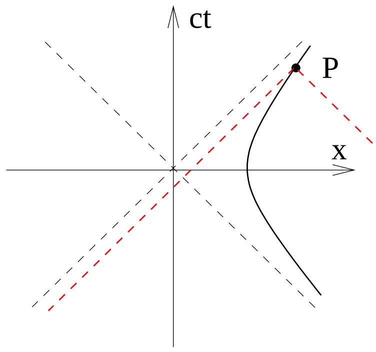

Figure 56:

We can look at what the accelerated observer feels. His time is simply the proper time of the particle. To compute this, the form of $\gamma(t)$ given in (7.32) is particularly useful. From (7.19), if time t elapses in the stationary frame S, then the particle feels

$$\tau = \int_0^t \frac{c d\tilde{t}}{\sqrt{c^2 + a'^2 \tilde{t}^2}} = \frac{c}{a'} \sinh^{-1}\left(\frac{a' t}{c}\right)$$

This analysis gives us a more quantitative way to view the twin paradox. Suppose that Luke undertakes his trip to Tatooine on a trajectory of constant acceleration. He leaves Leia at the time $t < 0$ where their worldlines intersects, arrives at Tatooine at $t = 0$ and $x = c^2/a'$, and returns back to Leia as shown. Leia experiences time t; Luke time $\tau < t$.

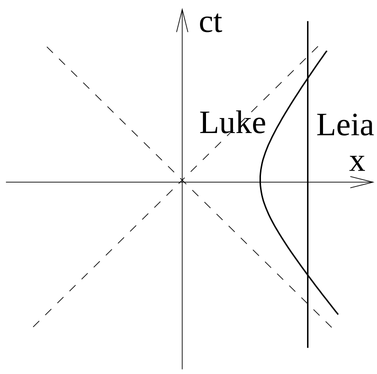

Figure 57:

Finally, we can look at how far the accelerated observer thinks he has travelled. Of course, this observer is not in an inertial frame, but at any time t we can consider the inertial frame that is momentarily at rest with respect to the accelerated particle. This allows us to simply use the Lorentz contraction formula. Using our results (7.32) and (7.33), we have

$$x' = \frac{x}{\gamma} = \frac{c^2}{a'}\left(1 - \frac{c}{\sqrt{c^2 + a'^2 t^2}}\right)$$

Curiously, $x' \to c^2/a'$ is finite as $t \to \infty$ or, equivalently, as $\tau \to \infty$. Despite all that effort, an accelerated observer doesn't think he has got very far! This again, is related to the presence of the horizon.

7.4.7 Indices Up, Indices Down

The minus signs in the Minkowski metric η means that it's useful to introduce a slight twist to the usual summation convention of repeated indices. For all the 4-vectors that we introduced above, we were careful always place the spacetime index $\mu = 0, 1, 2, 3$ as a superscript (i..e up) rather than a subscript.

$$X^\mu = (ct, \mathbf{x})$$

This is because the same object with an index down, X_μ, will mean something subtley different!

$$X_\mu = (ct, -\mathbf{x})$$

With this convention, the Minkowski inner product can be written using the usual convention of summing over repeated indices as

$$X^\mu X_\mu = c^2t^2 - \mathbf{x} \cdot \mathbf{x}$$

In contrast, writing $X^\mu X^\mu = c^2t^2 + \mathbf{x}^2$ is a dumb thing to write in the context of special relativity since it looks very different to observers in different inertial frames. In fact, we will shortly declare it illegal to write things like $X^\mu X^\mu$.

There is a natural way to think of X_μ in terms of X^μ. If we write the Minkowski metric as the diagonal matrix $\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$ then we can raise and lower indices using $\eta_{\mu\nu}$ and the summation convention, so

$$X_\mu = \eta_{\mu\nu} X^\nu$$

Moreover, we will insist that all objects with indices up and down are similarly related by contracting with η. For example, we could write the electromagnetic tensor as

$$G^{\mu\nu} = G^\mu{}_\rho \eta^{\rho\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -cB_3 & cB_2 \\ E_2 & cB_3 & 0 & -cB_1 \\ E_3 & -cB_2 & cB_1 & 0 \end{pmatrix}$$

The object $G_{\mu\nu}$ is actually somewhat more natural than $G^\rho{}_\nu$ since the former is anti-symmetric.

To raise indices back up, we need the inverse of $\eta_{\mu\nu}$ which, fortunately, is the same matrix: $\eta^{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$. We have

$$\eta^{\mu\rho}\eta_{\rho\nu} = \delta^\mu{}_\nu$$

This trick of distinguishing between indices up and indices down provides a simple formalism to ensure that all objects have nice transformation properties under the Lorentz group. We insist that, just as in the usual summation convention, repeated indices only ever appear in pairs. But now we further insist that pairs always appear with one index up and the other down. The result will be an object with is invariant under Lorentz transformations.

In future courses, you will learn that there is somewhat deeper mathematics lying behind distinguishing X^μ and X_ν: formally, these objects live in different spaces (sometimes called dual spaces). Objects such as X^μ are said to be *covariant* vectors, while X_μ is said to be a *contravariant* vector.

7.5 Particle Physics

"Oh, that stuff. We never bother with that in our work"

Ernest Rutherford, the first particle physicist, discussing relativity

Our goal in this section is to describe various relativistic phenomena that arise in particle physics. All these processes occur in the absence of external forces, so $F = 0$ and we will rely only on conservation of 4-momentum, meaning

$$\frac{dP}{d\tau} = 0$$

Of course, conservation of 4-momentum includes both conservation of 3-momentum and conservation of energy.

The calculations that follow are similar in spirit to the collision calculations of Section 5.2. Before we proceed, there are a couple of hints that may help when solving these problems. Firstly, we need to choose a frame of reference in which to calculate: the smart frame to choose is nearly always the centre of mass of the system. (Which should more correctly be called the centre of momentum frame, for it is the one with vanishing spatial 3-momentum). Secondly, you will often be presented with a situation where there is one particle with momentum P about which you know nothing. A good way to eliminate this is often to rearrange your equation so it takes the form $P = \ldots$ and then square it to get the right-hand side to be $P \cdot P = m^2c^2$. Let's now see how this works in a few examples.

7.5.1 Particle Decay

Consider a single particle with rest mass m_1 which decays into two particles with rest masses m_2 and m_3. Conservation of 4-momentum tells us

$$P_1 = P_2 + P_3$$

or, equivalently,

$$E_1 = E_2 + E_3 \quad \text{and} \quad \mathbf{p}_1 = \mathbf{p}_2 + \mathbf{p}_3$$

In the rest frame of the decaying particle, we can write (using (7.30)),

$$E_1 = m_1 c^2 = \sqrt{p_2^2 c^2 + m_2^2 c^4} + \sqrt{p_3^2 c^2 + m_3^2 c^4} \geq m_2 c^2 + m_3 c^2$$

which tells us the unsurprising result that a particle can only decay if its mass is greater than that of its decay products. In the problem sheet, you will be asked to compute the velocities v_2 and v_3 of the decay products in the centre of mass frame and show that they are given by

$$\gamma_2 = \frac{m_1^2 + m_2^2 - m_3^2}{2m_1 m_2} \quad \text{and} \quad \gamma_3 = \frac{m_1^2 + m_3^2 - m_2^2}{2m_1 m_3}$$

Here we will instead look at some slightly different problems.

An Example: Higgs Decay

The LHC has taught us that the Higgs boson has mass $m_h c^2 \approx 125\ GeV$. It mostly decays into two photons. In particle physics, photons are always denoted by γ. Do not confuse them with the Lorentz contraction factor! The "equations" in which the photon γ's appear are more like chemical reactions than true equations. The decay of the Higgs into two photons is written as

$$h \to \gamma\gamma$$

Similar decays occur for other particles, most notably the neutral pion, a meson (meaning that it is made of a quark and anti-quark) with mass $m_\pi c^2 \approx 140\ MeV$. This too decays as $\pi^0 \to \gamma\gamma$.

To be concrete (and more relevant!) we'll focus on the Higgs. Conservation of 4-momentum tells us (in, hopefully, obvious notation) that

$$P_h = P_\gamma + P'_\gamma$$

If we sit in the rest frame of the Higgs, so $P_h^\mu = (m_h c, 0)$, the photons must have equal and opposite 3-momentum, and therefore equal energy $E_\gamma = \frac{1}{2} m_h c^2$. The photons must be emitted back-to-back but, because the problem is rotationally symmetric, can be emitted at any angle.

What if we're not sitting in the rest frame of the Higgs?. Suppose that the Higgs has energy E_h and the energy of one of the photons is measured to be E_γ. What is the angle θ that this photon makes with the path of the Higgs?

We'll use the strategy that we described above. We have no information about the second photon, with 4-momentum P'_γ. So we rearrange the conservation of momentum to read $P'_\gamma = P_h - P_\gamma$. Upon squaring this, we have $P'_\gamma \cdot P'_\gamma = 0$, so

$$\begin{aligned}0 = (P_h - P_\gamma)\cdot(P_h - P_\gamma) &= P_h \cdot P_h + P_\gamma \cdot P_\gamma - 2P_h \cdot P_\gamma \\ &= m_h^2 c^2 - \frac{2E_h E_\gamma}{c^2} + 2\mathbf{p}_h \cdot \mathbf{p}_\gamma \\ &= m_h^2 c^2 - \frac{2E_h E_\gamma}{c^2} + \frac{2E_\gamma}{c}\cos\theta\sqrt{E_h^2/c^2 - m_h^2 c^2}\end{aligned}$$

where, in the last equation, we have used $E^2 = p^2c^2 + m^2c^4$ (which is just $E = pc$ for the photon). This can now be rearranged to give the answer for θ.

7.5.2 Particle Collisions

Let's now look at the physics of relativistic collisions. We'll collide two particles together, both of mass m. They will interact in some manner, preserving both energy and 3-momentum, and scatter at an angle θ.

$$P_1 + P_2 = P_3 + P_4$$

As we mentioned previously, it's easiest to see what happens in the centre of mass frame. Without loss of generality, we'll take the initial momenta to be in the x-direction. After the collision, the particles must have equal and opposite momenta, which means they must also have equal energy. This, in turn, ensures that in the centre of mass frame, the speed v after the collision is the same as before. We can choose our axes so that the initial and final momenta are given by

$$\begin{aligned}P_1^\mu &= (mc\gamma_v, mv\gamma_v, 0, 0) \quad , \quad P_2^\mu = (mc\gamma_v, -mv\gamma_v, 0, 0) \\ P_3^\mu &= (mc\gamma_v, mv\gamma_v\cos\theta, mv\gamma_v\sin\theta, 0) \quad , \quad P_4^\mu = (mc\gamma_v, -mv\gamma_v\cos\theta, -mv\gamma_v\sin\theta, 0)\end{aligned}$$

where we've put the subscript on γ_v to denote its argument. We can also look at the same collision in the *lab frame*. This refers to the situation where one of the particles is initially at rest. (Presumably in your lab). By the velocity addition formula, the other particle must start with speed

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

You can also derive this result by writing down the momenta P'_1 and P'_2 in the lab frame and equating $(P_1 + P_2)^2 = (P'_1 + P'_2)^2$

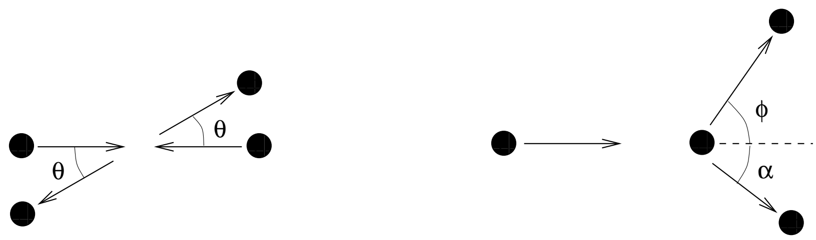

Figure 58: Collisions in the centre of mass frame on the left and the lab frame on the right

In the lab frame, the angles ϕ and α at which the particles scatter are not equal. They can be easily determined using the addition of 4-velocities that we saw in Section 7.4.2 . Set $u = -v$ in equation (7.23) and use the identity $\tan(x/2) = \sin x/(1 + \cos x)$ to get

$$\tan\phi = \frac{1}{\gamma_v}\tan\theta/2 \quad \text{and} \qquad \tan\alpha = \frac{1}{\gamma_v}\tan(\theta/2 + \pi/2)$$

One of the more interesting examples of collisions is *Compton Scattering*, in which the colour of light changes after scattering off an electron (because it changes its energy and therefore its frequency). You will derive this result in the examples sheet.

Particle Creation

Just as mass can be converted into kinetic energy, so kinetic energy can be converted into mass through the creation of new particles. Roughly speaking, this is the way we discover new particles of Nature.

Suppose we collide two particles, each of mass m. After the collision, we hope to be left with these two particles, together with a third of mass M. How fast must the original two particles collide?

Conservation of momentum gives us

$$P_1 + P_2 = P_3 + P_4 + P_5$$

where $P_1^2 = P_2^2 = P_3^2 = P_4^2 = m^2c^2$, while $P_5^2 = M^2c^2$. Let's work in the centre of mass frame of the colliding particles, each of which has speed v. In this case, we have

$$(P_1 + P_2)^2 = 4m^2\gamma_v^2c^2 = (P_3 + P_4 + P_5)^2 \tag{7.34}$$

Since we're in the centre of mass frame, the final momenta must take the form $P_3 + P_4 + P_5 = ((E_1 + E_2 + E_3)/c, \mathbf{0})$ so that

$$(P_3 + P_4 + P_5)^2 = \frac{1}{c^2}(E_1 + E_2 + E_3)^2 \geq \frac{1}{c^2}(2mc^2 + Mc^2)^2$$

where, for each particle, we've used the fact that $E = \sqrt{m^2c^4 + p^2c^2} \geq m^2c^2$. Substituting this into (7.34) gives

$$4m^2\gamma_v^2c^2 \geq 4m^2c^2 + M^2c^2 + 4Mmc^2 \quad \Rightarrow \quad \gamma \geq 1 + \frac{M}{2m} \tag{7.35}$$

This makes sense. The amount of minimum amount of kinetic energy per particle is $T = \gamma mc^2 - mc^2 = \frac{1}{2}Mc^2$. With this minimum amount, the two colliding particles can combine their kinetic energies to form the new particle. After the collision, all three particles are then at rest.

It's worth mentioning another way to do the above computation. Suppose that you hadn't noticed that the three-momentum of $P_3 + P + 4 + P_5$ vanished and instead expanded out the right-hand side of (7.34) to end up with nine terms. Things are a bit harder this way, but all is not lost. We can apply a Cauchy-Schwarz-like inequality to each of these terms. For any massive particles with 4-momenta P and Q, such that $P^2 = m_1^2c^2$ and $Q^2 = m_2^2c^2$, we necessarily have $P \cdot Q \geq m_1m_2c^2$. It is simplest to prove this by working in a frame in which one particle is stationary. Then we have

$$P \cdot Q = \begin{pmatrix} m_1c \\ 0 \end{pmatrix} \cdot \begin{pmatrix} E_2/c \\ p_2 \end{pmatrix} = m_1E_2 = m_1\sqrt{m_2^2c^4 + p_2^2c^2} \geq m_1m_2c^2$$

Applied to (7.34) this once again gives (7.35).

What if we re-do this experiment in the lab frame, in which of the original particles is at rest and the other has speed u? Now we have $P_1 = (m\gamma_uc, m\gamma_uu)$ and $P_2 = (mc, 0)$, so

$$(P_1 + P_2)^2 = P_1^2 + P_2^2 + 2P_1 \cdot P_2 = 2m^2c^2 + 2m^2\gamma_uc^2$$

But we don't have to compute $(P_3 + P_4 + P_5)^2$ again since the beauty of taking the square of the 4-momenta is that the result is frame independent. We have

$$2m^2c^2 + 2m^2\gamma_uc^2 \geq 4m^2c^2 + M^2c^2 + 4Mmc^2 \quad \Rightarrow \quad \gamma_u \geq 1 + \frac{2M}{m} + \frac{M^2}{2m^2}$$

Now we see it's not so easy to create a particle. It's certainly not enough to give the incoming particle kinetic energy $T = \frac{1}{2}Mc^2$ as one might intuitively expect. Instead, if you want to create very heavy particles, $M \gg m$, you need to give your initial particle a kinetic energy of order $T \approx M^2c^2/2m$. This scales quadratically with M, rather than the linear scaling that we saw in the centre of mass frame. The reason for this simple: there's no way that the end products can be at rest. The need to

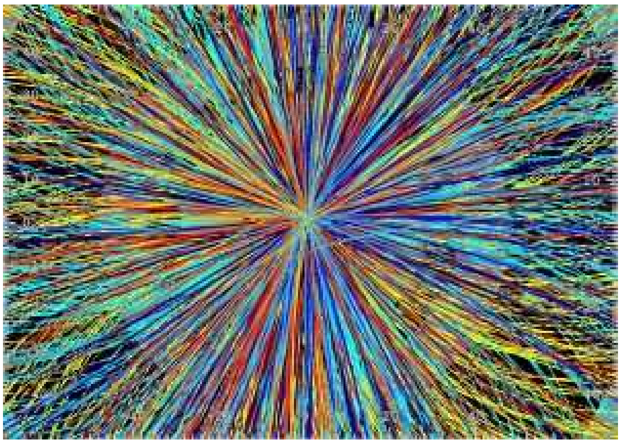

Figure 59:

conserve momentum means that much of the kinetic energy of the incoming particle goes into producing kinetic energy of the outgoing particles. This is the reason that most particle accelerators have two colliding beams rather than a single beam and a stationary target.

The LHC primarily collides protons in its search to discover new elementary particles. However, for one month a year, it switches to collisions of lead nuclei in an attempt to understand a new form of matter known as the quark-gluon plasma. Each lead nuclei contains around 200 protons and neutrons. The collision results in a dramatic demonstration of particle creation, with the the production of many thousands of particles – protons, neutrons, mesons and baryons. Here's a very pretty picture. It's one of the first collisions of lead nuclei at LHC in 2010, shown here in all its glory by the ALICE detector.

7.6 Spinors

In this final section, we return to understand more of the mathematical structure underlying spacetime and the Lorentz group. Ultimately, the new structure that we will uncover here has very important implications for the way the Universe works. But we will also see a nice application of our new tools.

Let's start by recalling our definition of the Lorentz group. We introduced elements of the group as 4×4 real matrices satisfying

$$\Lambda^T \eta \Lambda = \eta$$

where $\eta = \text{diag}(1, -1, -1, -1)$ is the diagonal Minkowski metric. Elements with $\det\Lambda = 1$ define the group $SO(1,3)$. If we further restrict to elements with the upper-left component $\Lambda^0{}_0 > 0$, which ensures that the transformation does not flip the direction of time, then we have the sub-group $SO^+(1,3)$. As we will now see, there's some rather beautiful subtleties associated with this group.

7.6.1 The Lorentz Group and $SL(2, \mathbf{C})$

The Lorentz group $SO^+(1,3)$ is (almost) the same as the rather different looking group $SL(2, \mathbf{C})$, the group of 2×2 complex matrices with determinant one. We will start by providing the map between these two groups, and explaining what the word "almost" means.

Before we talk about Lorentz transformations, let's first go back to think about the points in Minkowski space themselves. So far, we've been labelling these by the 4-vector $X^\mu = (ct, x, y, z)$. But there is alternative way of labelling these points, not by a 4-vector but instead by a 2×2 Hermitian matrix. Given a 4-vector X, we can write down such a matrix $\hat{X}$ by

$$\hat{X} = \begin{pmatrix} ct + z & x - iy \\ x + iy & ct - z \end{pmatrix}$$

which clearly satisfies $\hat{X} = \hat{X}^\dagger$. Moreover, this is the most general form of a 2×2 Hermitian matrix. This means that there is a one-to-one map between 4-vectors X and 2×2 Hermitian matrices. We can equally well take the latter to define a point in Minkowski space.

We learned earlier that Minkowski space comes equipped with an inner product structure on 4-vectors. The inner product $X \cdot X$ measures the distance in spacetime between the origin and the point X. But this is very natural in terms of the matrix language: it is simply the determinant

$$X \cdot X = \det \hat{X} = c^2t^2 - x^2 - y^2 - z^2$$

With this new way of labelling points in Minkowski space using the matrices $\hat{X}$, we can return to think about Lorentz transformations. Recall that, by definition, a Lorentz transformation is a linear map which preserves the inner-product structure on Minkowski space. Let's consider a general matrix $A \in SL(2, \mathbf{C})$. We can use this to define a linear map

$$\hat{X} \ \rightarrow \ \hat{X}' = A\hat{X}A^\dagger \tag{7.36}$$

By construction, if $\hat{X} = \hat{X}^\dagger$ then we also have $\hat{X}' = (\hat{X}')^\dagger$, so $\hat{X}'$ also defines a point in Minkowski space. Moreover,

$$\det \hat{X}' = \det(A\hat{X}A^\dagger) = \det A \; \det X \; \det A^\dagger = \det X$$

where the last equality follows because $\det A = 1$. This means that the map (7.36) preserves the inner product on Minkowski space and therefore defines a Lorentz transformation.

We may wonder if all Lorentz transformations can be implemented by suitable choices of A. The answer is yes. We'll exhibit the map explicitly below, but first let's just count the dimension of the two groups to make sure we stand a chance of it working. A general complex 2×2 matrix has 4 complex entries. The requirement that its determinant is 1 reduces this to 3 complex parameters, or 6 real parameters. This agrees with the dimension of the Lorentz group: $6 = 3$ rotations $+ 3$ boosts.

Although the dimensions of $SO^+(1,3)$ and $SL(2,\mathbf{C})$ are equal, they are not quite the same groups. In some sense, $SL(2,\mathbf{C})$ is twice as big. The reason is that the matrices A and $-A$ both implement the same Lorentz transformation in (7.36). We say that $SL(2,\mathbf{C})$ is the double cover of $SO^+(1,3)$ or, alternatively,

$$SO^+(1,3) \cong SL(2,\mathbf{C})/\mathbf{Z}_2$$

Mathematically, there is a 2:1 group homomorphism between $SL(2,\mathbf{C})$ and $SO^+(1,3)$. The word "homomorphism" means that the group structure is preserved under this map. The existence of this double cover leads to some quite extraordinary consequences. But, before we get to these, let's first just look at how the map works in more detail.

Rotations

We've seen that points in Minkowski space can be written as a 4-vector X or Hermitian matrix $\hat{X}$. Meanwhile, Lorentz transformations act as $X \rightarrow \Lambda X$ or $\hat{X} \rightarrow A\hat{X}A^\dagger$. Here we would like to be more explicit about which matrices A correspond to the different Lorentz transformations.

We start with rotations. By definition, these are the transformations which leave time untouched. From (7.36), this means that we want matrices A which map $\hat{X} = ct\,1$ (where 1 here is the unit 2×2 matrix) to itself. In other words, rotations should obey

$$AA^\dagger = 1$$

But such matrices are familiar unitary matrices. We learn that rotations sit in the subgroup $A \in SU(2) \subset SL(2,\mathbf{C})$. You may be used to thinking of the rotation group as $SO(3)$ rather than $SU(2)$. But these are almost the same thing: $SU(2)$ is the double cover of $SO(3)$,

$$SO(3) \cong SU(2)/\mathbf{Z}_2$$

Let's see how this equivalence between matrices $R \in SO(3)$ matrices and $A \in SU(2)$ works. For rotations around the x-axis, we have

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \quad \longleftrightarrow \quad A = \pm \begin{pmatrix} \cos(\theta/2) & i\sin(\theta/2) \\ i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$$

To see this, you just need to substitute the matrix A into the map (7.36) and check that it reproduces the same rotation as the matrix R. Note the $\pm$ possibility on A which reflects the fact that $SL(2, \mathbf{C})$ is the double cover of the Lorentz group. This is also related to the fact that the angle in A is $\theta/2$ rather than θ: we will return to this shortly. For rotations about the y-axis, we have

$$R = \begin{pmatrix} \cos\theta & 1 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 1 & \cos\theta \end{pmatrix} \quad \longleftrightarrow \quad A = \pm \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$$

Finally, for rotations about the z-axis, we have

$$R = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \longleftrightarrow \quad A = \pm \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}$$

There's a somewhat nicer way of writing these matrices which makes their structure clearer. To see this, we first need to introduce the *Pauli matrices*,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.37}$$

Together with the unit matrix, these form a basis of 2×2 Hermitian matrices. They have the nice property that $\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k$. In general, a rotation by angle θ around an axis with unit vector $\vec{n}$ is associated to the unitary matrix

$$A = \pm \exp\left(\frac{i\theta}{2}\, n^i\sigma^i\right) \tag{7.38}$$

Of course, the discussion above also tells us how the rotations fit within the Lorentz group. The matrix A remains unchanged, while the Lorentz transformation Λ is constructed by embedding the orthogonal matrix R in the lower-right block as shown in (7.15).

Boosts

The Pauli matrices also provide a simple way to describe the $A \in SL(2, \mathbf{C})$ corresponding to Lorentz boosts. A boost with rapidity φ in the direction $\vec{n}$ is associated to

$$A = \pm \exp\left(-\frac{\varphi}{2}\, n^i \sigma^i\right) \tag{7.39}$$

Unlike rotations, these matrices are not unitary. This ensures that they affect the time component. Again, you can check that this reproduce the Lorentz boosts of the form (7.17) simply by substituting this expression for A into the map (7.36). For example, a boost in the z-direction is given by the matrix

$$A = \begin{pmatrix} e^{-\varphi/2} & 0 \\ 0 & e^{+\varphi/2} \end{pmatrix} \quad \Rightarrow \quad A\hat{X}A^\dagger = \hat{X}' = \begin{pmatrix} e^{-\varphi}(t+z) & x - iy \\ x + iy & e^{+\varphi}(t-z) \end{pmatrix}$$

This tells us that x and y are left unchanged, while $t' + z' = e^{-\varphi}(t+z)$ and $t' - z' = e^{+\varphi}(t-z)$. Doing the algebra gives

$$t' = \cosh\varphi\, t - \sinh\varphi\, z \quad , \quad z' = \cosh\varphi\, z - \sinh\varphi\, t$$

which indeed agrees with the usual form of the Lorentz transformation (7.17) written in terms of the rapidity.

7.6.2 What the Observer Actually Observes

There's a rather nice application of the above formalism. In Section 7.2, when we first encountered relativistic phenomena such as length contraction, we stressed that different observers ascribe different coordinates to spacetime events. But this is not the same thing as what the observer actually sees, for this also involves the time that the light took to travel from the event to the observer. So this leaves open the question: what *does* an observer observe? What do Lorentz contracted objects really look like? As we will now show, writing the Lorentz group as $SL(2, \mathbf{C})$ gives a wonderfully elegant way to answer this question. Moreover, what we will find is somewhat surprising.

What an observer actually sees are, of course, light rays. As objects move through Minkowski space, they emit light which then propagates to the position of the observer. We've drawn this in the diagrams, both of which have the observer placed at the origin of Minkowski space. We've also drawn the future and past lightcones emitted from the origin.

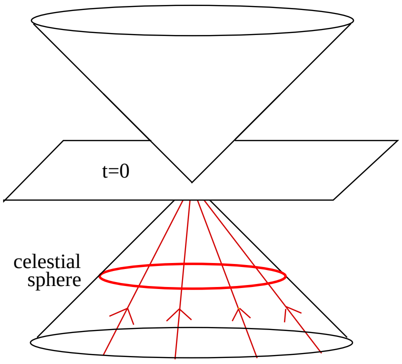

Figure 60: The celestial sphere for one observer...

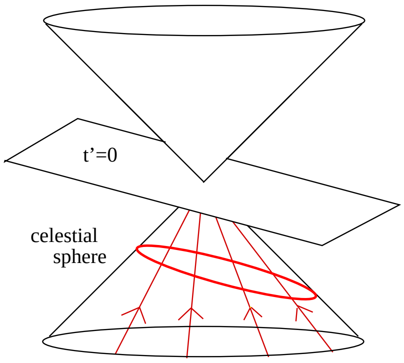

Figure 61: ...and for another

In the left-hand figure, the observer is assumed to be stationary with time coordinate t. At each fixed moment in time, t, the light rays form a sphere $\mathbf{S}^2$. This is drawn as the red circle in the past lightcone of the diagram. If we assume that no other object comes between this sphere and the observer, then the light rays intersecting the sphere are a good representation of what the observer actually sees. If he takes a snapshot of everything around him with some really super-dooper fancy camera, he would record the image on this sphere. This is sometimes given the name of the *celestial sphere*, reflecting the fact that this is how we should think of viewing the night sky (at least if the Earth wasn't obscuring half of it).

Let's now look at what an observer in a different inertial frame sees. This is shown in the right-hand figure. This second observer will also take a snapshot using his fancy camera as he passes through the origin. But this new observer's celestial sphere is given by null rays that sit at $t' =$ constant. Although it's no longer obvious from the picture, we know that the space defined by the intersection of light rays with the constant t' hyperplane must still be a sphere simply because all inertial observers are equivalent. However, this new celestial sphere is clearly tilted with respect to the previous one.

The four light rays drawn in the figure intersect both celestial spheres. These light rays therefore provide a map between what the two observers see. This is a map between the two celestial spheres, $\mathbf{S}^2 \to \mathbf{S}^2$. Our goal is to construct this map.

This is where our new mathematical formalism comes in. Any point on a light ray is, by definition, at vanishing distance from the origin when measured in the Minkowski metric. Equivalently, the 2×2 Hermitian matrix $\hat{X}$ describing this point must have vanishing determinant. But there's a nice way to write down such matrices with zero determinant. We introduce a two-component complex vector, ξ_α with $\alpha = 1, 2$. Then we write

$$\hat{X} = \xi\xi^\dagger = \begin{pmatrix} |\xi_1|^2 & \xi_1\xi_2^\dagger \\ \xi_2\xi_1^\dagger & |\xi_2|^2 \end{pmatrix}$$

which, by construction, obeys $\det \hat{X} = 0$. It's simple to check that the most general Hermitian matrix $\hat{X}$ with $\det \hat{X} = 0$ can be written in this way. Note, however, that there's a redundancy in this description, since if we rotate both components of ξ by a phase, so that $\xi \to e^{i\beta}\xi$, then $\hat{X}$ remains unchanged.

An Aside: The Hopf Map

In our new notation, the celestial sphere at constant time t is simply given by

$$\xi^\dagger\xi = |\xi_1|^2 + |\xi_2|^2 = \text{constant} \tag{7.40}$$

There's actually some interesting maths in this statement. It's obvious that given two complex variables ξ_1 and ξ_2, the equation (7.40) defines a 3-dimensional sphere $\mathbf{S}^3$. What's perhaps less obvious, but nonetheless true, is that if we identify all points on $\mathbf{S}^3$ related by $\xi \to e^{i\beta}\xi$, then we get a 2-dimensional sphere $\mathbf{S}^2$. In mathematical language, we say that $\mathbf{S}^3/U(1) \cong \mathbf{S}^2$.

It's simple to write directly the map $\mathbf{S}^3 \to \mathbf{S}^2$. Given a complex 2-vector, ξ, obeying $\xi^\dagger\xi = 1$, you can define 3 real numbers k^i by

$$k^i = \xi^\dagger\sigma^i\xi$$

where σ^i are the three Pauli matrices (7.37). Then a little algebra shows that $k^ik^i = 1$. In other words, k^i gives a point on $\mathbf{S}^2$. This is map from $\mathbf{S}^3$ to $\mathbf{S}^2$ is called the *Hopf map*.

Back to the Real World

Let's now use these new objects ξ to construct the map between the two celestial spheres. A nice fact is that Lorentz transformations act in a natural way on the two-component ξ. To see this, recall that

$$\hat{X}' = \xi'\xi'^\dagger = A\xi\xi^\dagger A^\dagger$$

But we can view this as a transformation of ξ itself. We have simply the $SL(2, \mathbf{C})$ transformation

$$\xi' = A\xi \tag{7.41}$$

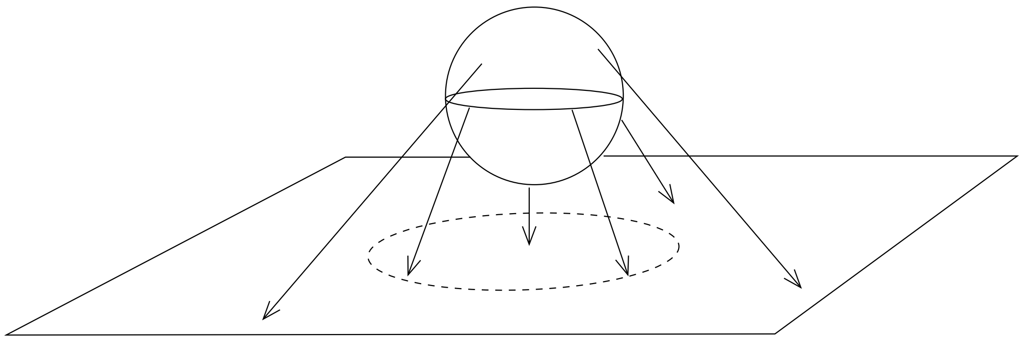

Figure 62: The stereographic projection. The southern hemisphere is mapped to inside the dotted circle; the northern hemisphere is mapped to outside this circle.

However, this is not quite our mapping. We can start with a celestial sphere defined by (7.40) and act with a Lorentz transformation. The trouble is that the resulting space we get remains the first celestial sphere, just written in the second observer's coordinates. We still need to propagate the light rays forward and backwards so that they intersect the second celestial sphere.

To avoid this complication, it's best to think about these celestial spheres in a slightly different way. Rather than saying that they are defined at constant time, let's instead define them as equivalence classes of light rays. This means that we lose the information about where we are along the light ray: we only keep the information about which light ray we're talking about. Mathematically, this is very simple: to each ξ we associate a single complex number $\omega \in \mathbf{C}$ by

$$\omega = \frac{\xi_1}{\xi_2}$$

The map from the celestial sphere $\mathbf{S}^2 \to \mathbf{C}$ is known as *stereographic projection* and is shown in the figure. Strictly speaking, ω parameterises $\mathbf{C} \cup \{\infty\}$, with the point at infinity included to accommodate the point $\xi_2 = 0$, which is the North pole of the celestial sphere. This extended complex plane is called the *Riemann sphere.*

Now the light rays seen by the first observer are labelled by $\omega \in \mathbf{C}$ and form a celestial sphere. The light rays seen by the second observer are labelled by $\omega' = \xi'_1/\xi'_2$ and form a different celestial sphere. A Lorentz transformation $A \in SL(2, \mathbf{C})$ acts on ξ as (7.41) which, in terms of ω, reads

$$\omega' = \frac{a\omega + b}{c\omega + d} \quad \text{where } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } ad - bc = 1 \tag{7.42}$$

This transformation on the complex plane is known as a *Möbius transformation.* It's simple to see that Möbius transformations form a group. In fact, from what we've seen above, you shouldn't be surprised to learn that the group of Möbius transformations is $SL(2, \mathbf{C})$, up to a discrete $\mathbf{Z}_2$ identification.

Suppose now that the first observers sees an object on his celestial sphere that traces out some shape. After stereographic projection, that will result in a shape on the complex plane (perhaps passing through the point at infinity). This appears to the second observer to be transformed by (7.42). Upon taking the inverse stereographic projection, we will learn what shape the second observer really sees.

To make progress, we should look at a simple example. And the simplest example is for an object which is itself a sphere. This means that, when stationary with respect to the first observer, the outline of the object looks like a circle. What does the second observer see? To answer this, I'll need to invoke some simple facts about stereographic projection and Möbius transformations. Although I won't prove them, they are among the most basic properties of these transformations and will be proven in next year's Geometry course. The facts are:

- The stereographic projection maps circles on the sphere to circles or lines on the plane.
- Möbius transformations map circles and lines on the plane to circles or lines on the plane.

Hiding behind these facts is the statement that both maps are conformal, meaning that they preserve angles. But, for us, the upshot is that a circle on the first celestial sphere is mapped under a Lorentz transformation to a circle on the second.

Let's pause to take this in. The first observer saw an object which had the shape of a circle. Based on the arguments of Lorentz contraction, you might expect that the second observer sees a squashed circle, maybe an ellipse. Yet this is not what happens. Instead, the second observer also sees a circle! The effects of the time of flight of light completely eliminate the Lorentz contraction. This fact was only realised more than 50 years after Einstein's formulation of special relativity when it was discovered independently by Terrell and Penrose. It is sometimes said to be the "invisibility of the Lorentz contraction". Note that it doesn't mean that the effects of Lorentz contraction that we discussed before are not real. It just means that you don't get to see them if you take a picture of a sphere. Moreover, if you look more closely you find that there are things that change. For example, if you paint a picture on the surface of the sphere, this will appear deformed to the other observer.

7.6.3 Spinors

Finally, we're in a position to explain what the title of this section means. A *spinor* is simply a two-dimensional complex vector ξ which, under a Lorentz transformation $A \in SL(2, \mathbf{C})$, changes as $\xi \;\rightarrow\; A\xi$.

(Some confusing caveats: ξ defined in this way is known as a *Weyl spinor*. In fact, strictly speaking, it is known as a left-handed Weyl spinor. For reasons that I won't go into here, we can also define something called a right-handed Weyl spinor by exchanging $\varphi \rightarrow -\varphi$ in the definition of the boosts (7.39). Then combining a left-handed Weyl spinor together with a right-handed Weyl spinor gives a four component complex object that is called a *Dirac Spinor*. See, I told you it would be confusing!)

We've already seen how spinors can be used to describe light rays. But this is not their only use; they have much more a life of their own. Before I describe this, let me firstly explain a property that makes it very surprising that spinors have any real relevance in the world. This harks back to the observation that $SL(2, \mathbf{C})$ is the double cover of the Lorentz group. Suppose that there is some object in the Universe that is actually described by a spinor. This means, in particular, that the state of the object with ξ is different from the state of the object with $-\xi$. What happens when we rotate this object? Well, we've already seen how to enact a rotation using $SL(2, \mathbf{C})$ matrices: they are given by (7.38). Except if we're acting on spinors we need to make a decision: do we pick $+A$ or do we pick $-A$? Because, unlike the action on Minkowski space, these two different matrices will result in different states ξ after a rotation. It doesn't actually matter which choice we pick, as long as we make one. So let's decide that a rotation about an axis n^i acts on a spinor by

$$\xi \;\rightarrow\; \exp\left(\frac{i\theta}{2} n^i \sigma^i\right) \xi$$

This all seems fine. The surprise comes when we look at what happens if we rotate the spinor by 2π. It doesn't come back to itself. Instead, after a rotation by 2π we find $\xi \rightarrow -\xi$. We have to rotate by 4π to get the spinor to return to itself!

Wouldn't it be astonishing if there were objects in the Universe which had this property: that you could rotate them and find that they didn't come back to themselves. This is even more astonishing when you realise that rotating an object is the same thing as walking around it. If such objects existed, you would be able to circle them once and see that the object sits in a different state just because you walked around it. How weird would that be?

Well, such objects exist. What's more, they're the same objects that you and I are made of: electrons and protons and neutrons. All of these particles carry a little angular momentum whose direction is described by a spinor rather than a vector. This means that Nature makes use of all the pretty mathematics that we've introduced in this section. The symmetry group of the Universe we live in is not the Lorentz group $SO^+(1,3)$. Instead, it is the double cover $SL(2,\mathbf{C})$. And the basic building blocks of matter have subtle and wonderful properties. Turn an electron 360^o and it isn't the same; turn it 720^o and you're back to where you started. If you want to learn more about this, you can find deeper explanations in the lecture notes on Quantum Field Theory.