Special relativity

Stephen Huggett[*]

September 9, 2023

1 Preface

The general theory of relativity, discovered by Einstein in 1915, is a geometrical theory of gravitation. It is a beautiful description of our Universe as a four-dimensional geometry, whose curvature is related to the distribution of matter (such as stars and galaxies) by Einstein's equations. Astrophysical observations now verify the theory to one part in 10^{14}, and one of its many implications is the intriguing concept of a black hole.

However, we are going to study something simpler! Suppose we were in a large comfortable spaceship a very long way away from any stars, so that we could not feel any gravity. Then would it be correct to use Euclidean ideas of space and time? The answer is no; even before his general theory Einstein had already shown that a new way of thinking about space and time was needed. This is called the special theory of relativity.

These notes are of lectures given at the Summer School in Ali Nesin's "Mathematical Village" in Turkey. They are loosely based on the excellent book by Rindler [5]. I also recommend the biography of Einstein by Pais [2], and the wonderful book by Penrose [4].

2 Introduction

The most important fact about the special theory of relativity is that it is a geometrical theory. It describes spacetime as a *geometry,* thought of as a mathematical space together with a group of allowable transformations. This approach to geometry comes from Felix Klein's 1872 paper *Comparative overview of recent geometric investigations.*

Implicit in our understanding of space is that when we observe a given object from different viewpoints it will appear different while in fact remaining the same. This different appearance manifests itself in mathematics when we use coordinates to describe the object. These coordinates will be those of the

[*]http://stephenhuggett.com/index.html

position vectors of various parts of the object, such as the vertices of a polyhedron. Changing the viewpoint means changing the coordinate system, by rotation, translation, and reflection. So we are used to the idea that the spatial coordinates of an object can change, according to some transformation, while the object remains unchanged.

This also applies to time coordinates. Here the two viewpoints might be two observers who have not synchronized their clocks, in which case they will say that a given event happened at two different times. Again, we are used to this: just think of all the time zones on the Earth.

Now we think about space and time together: they make *spacetime*, which has four dimensions. A point of spacetime is an *event* (which is a point in space at an instant in time). An observer S in spacetime needs four coordinates to describe the events she sees:

$$(t, x, y, z).$$

Suppose another observer S' uses the coordinates (t', x', y', z'). In general her coordinates might be rotated, translated, reflected in space and translated in time with respect to those of S.

However, we want to concentrate on the possibility of *relative motion* between S and S'. We make the two frames of reference as nearly identical as possible, except that S' is moving relative to S. No rotations, translations, or reflections are needed: when $t = 0$ the x, y, and z axes coincide with the x', y' and z' axes and then $t' = 0$ too. Subsequently the origin O' of S' moves at constant speed v along the x axis of S. This very important arrangement, which allows us to study the effect of the relative motion, is called the *standard configuration.*

What is the relationship between the coordinates (t, x, y, z) and (t', x', y', z')?

$$t' = t \qquad x' = x - vt \qquad y' = y \qquad z' = z \tag{1}$$

We can think of this as a transformation by writing it

$$(t, x, y, z) \to (t, x - vt, y, z), \tag{2}$$

when it is called the standard *Galilean* transformation. Notice that time is unchanged. Non-accelerating frames such as S and S' are called *inertial frames.*

Exercise 1 *Denote by g_v the transformation in (2). Write out the standard Galilean transformation g_w. Compose these two transformations. Show that the standard Galilean transformations form a group.*

In 1864 James Clerk Maxwell explained light as an electromagnetic wave phenomenon[1].

[1] You may have seen Maxwell's equations for electromagnetism before: they are

$$\mathrm{div}\mathbf{E} = 0, \quad \mathrm{curl}\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \mathrm{div}\mathbf{B} = 0, \quad \mathrm{curl}\mathbf{B} = \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}.$$

Quite soon afterwards, in 1887, a peculiar property of light was discovered (inadvertently) by Michelson and Morley. They were trying to measure the motion of the Earth relative to the medium in which light waves were thought to propagate (called the aether). The orbital speed of the Earth is about 29 kilometres per second, and there must be some point on the Earth's orbit at which it moves at 29 kilometres per second through the aether. Here the apparent velocity of light would increase or decrease by 29 from its normal value of about 300000 kilometres per second. Michelson and Morley compared the velocity of light in two perpendicular directions, one tangential to the Earth's orbit, using an interferometer which was easily sensitive enough to pick up this 29 kilometres per second difference.

The following exercise is an analogy.

Exercise 2 *One day a bee flies from P to Q and back in a straight line, at an airspeed V. During this flight, the wind is blowing in the direction from P to Q at speed $v < V$. Calculate the time taken for this flight in terms of V, v, and the distance d between P and Q.*

The next day the bee repeats this flight, but with the wind blowing at right angles to the line PQ, at the same speed. Calculate the time taken for this flight.

When Michelson and Morley did their experiment they discovered that it takes the same time for light to travel across the aether and back as it does to travel the same distance down the aether and back. The conclusion of this experiment (and a great many others confirming it) is that the velocity of light is the same in all reference frames. This was a puzzle for some time.

In 1905 Einstein, who had been wondering why Maxwell's equations were not invariant under Galilean transformations, published a paper *On the Electrodynamics of Moving Bodies.* He proposed what was to become the first principle of special relativity:

> SR1: **All inertial frames are equivalent for all physical experiments**

This is simply an extension to all of physics of the principle of Newtonian relativity. This principle together with the experimental fact:

> SR2: **Light has the same velocity in all inertial frames**

leads to the *Special Theory of Relativity.* Note that this experimental fact refers to the velocity of light in a *vacuum.* We will see later that when light passes through a medium there is another effect. In this theory there is no aether!

From these two principles we shall derive the *Lorentz transformations,* which will replace (2) as the transformations between inertial frames.

We are used to the idea that some features of an object, such as its length, are agreed on by different observers: they arrive at the same value for length even though they started with different coordinates for the object. Mathematically, this is because length is an *invariant* of the transformations. On the other hand, consider the example of the *apparent* length of an object. This clearly varies from observer to observer; it is not an invariant, and has no physical significance. So we will use the group of Lorentz transformations to find out which features of events have physical significance in special relativity; these will be invariants of the transformations. Unsettlingly, until we get used to it, we will discover that some things which we take for granted as having physical significance are in fact observer-dependent.

The first of these is the concept of *simultaneity,* which we can analyse directly from SR1 and SR2, even before we have the Lorentz group. In Newtonian physics, given any event (t_0, x_0, y_0, z_0), all observers will agree that there is a whole $\mathbb{R}^3$ of simultaneous events (t_0, x, y, z). In special relativity this is not true; the concept of simultaneity is *not* absolute—it depends on the observer.

Consider a fast aeroplane flying overhead and suppose a camera flash goes off exactly in the middle of the cabin. The light will take the same time to reach the passengers at either end of the cabin. So in the frame of reference of the plane the two lots of passengers see the flash *simultaneously.* We are watching all this from the ground and the light travels at the same speed forward and back for us too. However, so far as we are concerned the passengers at the back of the plane are travelling towards the light whereas those at the front are travelling away. So the passengers at the back certainly see the flash *first.* Therefore simultaneity is relative.

3 Lorentz transformations

We are looking for a transformation between the frames S and S' described above which will replace (2).

The definition of inertial frames requires that free particles always move uniformly along straight lines. Therefore from SR1 our new transformation must preserve straight lines, and so it must be linear:

$$x' = ax + by + cz + dt + e. \tag{3}$$

When $x = t = 0$ we must have $x' = 0$, and so $b = c = e = 0$. If $x = vt$ then $x' = 0$. Therefore $d = -va$, and we rewrite a as γ (which is the traditional notation) and obtain

$$x' = \gamma(x - vt). \tag{4}$$

Now we argue from SR1 that we can switch $t \longleftrightarrow t'$, $x \longleftrightarrow -x'$, $y \longleftrightarrow y'$, and $z \longleftrightarrow -z'$. Equation (4) becomes

$$x = \gamma(x' + vt'). \tag{5}$$

Note that if we imposed an absolute time $t = t'$ on these equations we could deduce (by adding them) that $\gamma = 1$ and get back to (2). Instead, at this point we use SR2, which implies that $x = ct$ if and only if $x' = ct'$. Here (and from now on) c is the velocity of light in vacuum, which is 2.9979245×10^8 metres per second. Therefore from (4)

$$ct' = \gamma(c - v)t, \tag{6}$$

and from (5)

$$ct = \gamma(c+v)t'. \tag{7}$$

If we multiply (6) and (7) together and cancel tt' we get

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}. \tag{8}$$

Finally, we eliminate x' between (4) and (5) to get

$$x = \gamma(\gamma(x - vt) + vt'), \tag{9}$$

and so

$$\gamma v t' = x - \gamma^2 x + \gamma^2 vt. \tag{10}$$

Therefore

$$\begin{aligned} t' &= \gamma t + x\frac{1-\gamma^2}{\gamma v} \\ &= \gamma(t - vx/c^2). \end{aligned} \tag{11}$$

The simplification which takes place in the last equation is common with the γ function.

Exercise 3 *Demonstrate (11).*

It can be shown that $y' = y$ and $z' = z$, so the *standard Lorentz transformation* is

$$t' = \gamma(t - vx/c^2) \qquad x' = \gamma(x - vt) \qquad y' = y \qquad z' = z \tag{12}$$

where γ is as in (8).

Exercise 4 *The equations (12) give the standard Lorentz transformation for (t', x', y', z') in terms of (t, x, y, z). Calculate the inverse transformation. In other words find (t, x, y, z) in terms of (t', x', y', z').*

Note that if $(\Delta t, \Delta x, \Delta y, \Delta z)$ and $(\Delta t', \Delta x', \Delta y', \Delta z')$ are the components in S and S' respectively of a vector joining two events, then

$$\Delta t' = \gamma(\Delta t - v\Delta x/c^2) \quad \Delta x' = \gamma(\Delta x - v\Delta t) \quad \Delta y' = \Delta y \quad \Delta z' = \Delta z. \tag{13}$$

We may regard the coordinates of our events as the components of vectors. Using the index notation for vectors, we write

$$x^a = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$$

then the standard Lorentz transformation (12) can be written

$$x'^a = L^a_b x^b, \tag{14}$$

where

$$L^a_b = \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{15}$$

Exercise 5 *Show (14).*

This is only the transformation between two frames of reference in the standard configuration. To get the transformation between *any* two inertial frames we may have to perform:

- a spatial rotation and translation,
- a time translation,
- the standard Lorentz transformation (12),
- a time translation, and
- a spatial rotation and translation.

These transformations are all of the form

$$x'^a = M^a_b x^b + w^a \tag{16}$$

for various matrices M^a_b and vectors w^a. In the case of the standard Lorentz transformation (12) we have

$$M^a_b = L^a_b, \qquad w^a = (0, 0, 0, 0). \tag{17}$$

Exercise 6 *Write down examples of (a) a spatial rotation, (b) a spatial translation, and (c) a time translation and show that they can each be written in the form (16).*

But M^a_b cannot be *any* 4×4 matrix. The next exercise indicates the property these matrices will have to have.

Exercise 7 *Show that the transformation (14) preserves the value of*

$$c^2t^2 - x^2 - y^2 - z^2.$$

It follows that if $(\Delta t, \Delta x, \Delta y, \Delta z)$ and $(\Delta t', \Delta x', \Delta y', \Delta z')$ are the components in S and S' respectively of a vector joining two events,

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$$

is also invariant under Lorentz transformations.

We define the *Lorentz group* to be the group of real 4×4 matrices preserving

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2, \tag{18}$$

and we define the *Poincaré group* to be the group of transformations (16), where M^a_b is in the Lorentz group. The Lorentz group sits inside the Poincaré group just as, in the plane, the orthogonal group $O(2)$ (of rotations and reflections) sits inside the Euclidean group $E(2)$ (of rotations, reflections, and translations).

We can now see that the principles SR1 and SR2 imply that any physical property must be invariant under the Poincaré group, so strictly speaking this is what we should be studying. However, we can discover most of the important properties of our new physics by restricting our attention to the *standard Lorentz transformation* (12).

For example, when $v/c \to 0$ this transformation becomes the Galilean transformation (2). This is called taking the *Newtonian limit.*

Exercise 8 *(a) Expand γ as a power series in v. (b) Draw a reasonably accurate graph of γ against v/c.*

If $v = c$, γ becomes infinite, and if $v > c$, γ becomes imaginary. Therefore inertial frames must have relative velocities less than c. In fact, in exercise 12 we will see that the speed of all physical signals is less than or equal to c.

Exercise 9 *S and S' are two frames of reference in the standard configuration. Two events in S are distinct but simultaneous. Show by varying v that there is no limit to the time separation which S' measures for these events.*

Exercise 10 *Consider the space separation S' measures for the two events referred to in exercise 9. Show that as $v \to c$ the space separation tends to infinity, and show that the space separation is a minimum when $v = 0$.*

Exercise 11 *S and S' are two frames in standard configuration. Consider the two events $P(ct_1, 0, 0, 0)$ and $Q(ct_2, 0, 0, 0)$ in S. If $t_2 > t_1$ S says that P happens before Q. Show that S' also says P happens before Q whatever the value of v. Why is this important?*

Exercise 12 *Suppose information could be sent from event P to event Q at a speed $u > c$. Suppose P and Q are both on the x axis of frame S and let the differences between their x and t coordinates be Δx and $\Delta t > 0$ respectively, with $u = \frac{\Delta x}{\Delta t}$. If S' is as usual then*

$$\Delta t' = \gamma(\Delta t - \frac{v\Delta x}{c^2}) = \gamma \Delta t(1 - \frac{vu}{c^2}).$$

Deduce that there are inertial frames in which Q precedes P.

4 Rapidity and velocity

The set of standard Lorentz transformations (12) for various values of $v \in \mathbb{R}$ forms a group (just as the standard Galilean transformations did). This is not hard to show directly, but the algebra is complicated. We can simplify it dramatically if instead of using the relative velocity we choose a new parameter.

We have already noted that

$$-1 < \frac{v}{c} < 1, \tag{19}$$

and so there is a unique ϕ, called the *rapidity,* satisfying

$$\tanh\phi = \frac{v}{c}. \tag{20}$$

Exercise 13 *What properties of the function in (20) are we using, in taking rapidity as a measure of the relative velocity?*

We discover that

$$\cosh\phi = \frac{1}{\sqrt{1-\tanh^2\phi}} = \gamma, \tag{21}$$

and

$$\sinh\phi = \sqrt{\gamma^2-1} = v\gamma/c. \tag{22}$$

Also,

$$e^{\phi} = \gamma(1+v/c) = \sqrt{\frac{c+v}{c-v}}. \tag{23}$$

Substituting (21) and (22) into the standard Lorentz transformation (12) we get:

$$ct' = -x\sinh\phi + ct\cosh\phi \qquad x' = x\cosh\phi - ct\sinh\phi \qquad y' = y \qquad z' = z \tag{24}$$

Suppose we just study the x and ct transformations, and write them in matrix form:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\phi & -\sinh\phi \\ -\sinh\phi & \cosh\phi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \tag{25}$$

Now recall the relations

$$\cosh\phi = \cos i\phi \qquad i\sinh\phi = \sin i\phi. \tag{26}$$

Using these, we can if we wish rewrite (25) as

$$\begin{pmatrix} ict' \\ x' \end{pmatrix} = \begin{pmatrix} \cos i\phi & -\sin i\phi \\ \sin i\phi & \cos i\phi \end{pmatrix} \begin{pmatrix} ict \\ x \end{pmatrix}. \tag{27}$$

Let us see next how these alternative forms of the standard Lorentz transformation can be used to make some of our calculations simpler. We start by noting that (24) implies

$$ct' + x' = e^{-\phi}(ct+x) \qquad y' = y \qquad z' = z \qquad ct' - x' = e^{\phi}(ct-x). \tag{28}$$

Exercise 14 *Use (28) to demonstrate the invariance of (18).*

We can also take derivatives of (28), and obtain the invariance of the infinitesimal version of (18):

$$c^2dt'^2 - dx'^2 - dy'^2 - dz'^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \tag{29}$$

We shall be referring to this important equation later. Next, let us see what happens when we *compose* two standard Lorentz transformations.

Exercise 15 *Prove from (25) that the composition of two standard Lorentz transformations with rapidities ϕ_1 and ϕ_2 is a standard Lorentz transformation with rapidity $\phi = \phi_1 + \phi_2$.*

Exercise 16 *Deduce from exercise 15 that standard Lorentz transformations form a group.*

Note that the group in this exercise is *not* the Lorentz group, but only a subgroup of it.

Exercise 17 *How many velocity increments of $\frac{1}{2}c$ do you have to make to get a resultant velocity of .99c?*

Exercise 18 *If $\phi = \tanh^{-1}(u/c)$ and $e^{2\phi} = z$ prove that n consecutive velocity increments of u produce a velocity*

$$\frac{c(z^n - 1)}{z^n + 1}.$$

Suppose S, S', and S'' are three inertial frames in standard configuration. If v_1 is the velocity of S' relative to S, v_2 is the velocity of S'' relative to S', and v is the velocity of S'' relative to S, then from $\tanh\phi = v/c$ and $\phi = \phi_1 + \phi_2$ we can work out the corresponding relation between v, v_1, and v_2. Using the formula

$$\tanh(\phi_1 + \phi_2) = \frac{\tanh\phi_1 + \tanh\phi_2}{1 + \tanh\phi_1 \tanh\phi_2} \tag{30}$$

we obtain

$$v = \frac{v_1 + v_2}{1 + v_1 v_2 / c^2}. \tag{31}$$

This is a *crucial* result. In particular, $v \neq v_1 + v_2$!

Exercise 19 *Use (31) with $v_2 = c$ to explain the fact that the velocity of light is not dependent on the motion of the observer.*

We can arrive at (31) in a different way. Replace the origin of the frame S'' by a particle. Then S' measures the speed of the particle to be v_2. Therefore

$$\frac{dx'}{dt'} = v_2. \tag{32}$$

So from the infinitesimal version of equation (12) with v replaced by v_1:

$$\frac{\gamma(dx - v_1 dt)}{\gamma(dt - v_1 dx/c^2)} = v_2. \tag{33}$$

Therefore

$$\frac{\frac{dx}{dt} - v_1}{1 - v_1 \frac{dx}{dt}/c^2} = v_2. \tag{34}$$

The speed of the particle according to S is $v = \frac{dx}{dt}$. Therefore

$$\frac{v - v_1}{1 - vv_1/c^2} = v_2 \tag{35}$$

and so as before

$$v = \frac{v_1 + v_2}{1 + v_1 v_2/c^2}. \tag{36}$$

If the term $v_1 v_2/c^2$ in this formula for the addition of velocities is small, then the formula just becomes $v = v_1 + v_2$ as expected. So can we think of a situation in which this term is not small? Recall that c is the velocity of light in *vacuum.* Light moves more slowly through a medium, such as water or glass. Suppose the speed of light in a liquid at rest relative to the laboratory frame S is measured to be u'. The liquid is then made to move with velocity v relative to S and the speed of light relative to S is measured to be

$$u = u' + v(1 - 1/n^2) \tag{37}$$

where n is the refractive index c/u' of the liquid. This is called the *Fresnel effect*[2].

Exercise 20 *Explain this effect using special relativity.*

We derived (36) by assuming that the particle was moving along the x axis. How do we modify our formula for the addition of velocities if the particle is moving in some other direction?

Exercise 21 *Let S and S' be in the standard configuration, and suppose that*

$$\mathbf{u} = (u_1, u_2, u_3) = (\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t})$$

and

$$\mathbf{u}' = (u_1', u_2', u_3') = (\frac{\Delta x'}{\Delta t'}, \frac{\Delta y'}{\Delta t'}, \frac{\Delta z'}{\Delta t'})$$

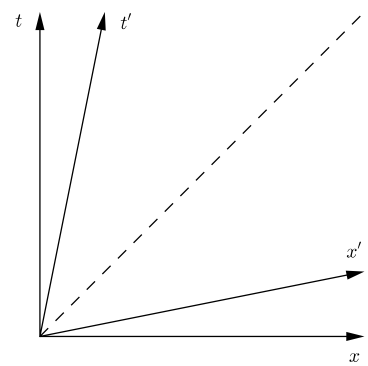
are the velocities in S and S' of a particle in uniform motion. Prove that

$$u_1' = \frac{u_1 - v}{1 - \frac{u_1 v}{c^2}}, \quad u_2' = \frac{u_2}{\gamma(1 - \frac{u_1 v}{c^2})}, \quad u_3' = \frac{u_3}{\gamma(1 - \frac{u_1 v}{c^2})}.$$

5 Minkowski diagrams

So far the y and z coordinates have been more or less irrelevant. Suppose we ignore them altogether, and choose units of time (years) and distance (lightyears) which make $c = 1$. We obtain a diagram which shows how the t' and x' axes in S' are related to the t and x axes in S.

[2] The experiment was first performed by Fizeau in 1851.



In S instants of time are lines parallel to the x-axis (joining all points with the same value for t) and a fixed point in space corresponds to a line parallel to the t-axis (x is constant). In S' instants of time are lines of constant t', which implies from (12) that $t - vx$ must be constant. In particular therefore the x'-axis is the line $t - vx = 0$, as in the figure. Also, in S' fixed points in space are lines of constant x' or in other words constant $x - vt$, so that in particular the t'-axis is $x - vt = 0$, as in the figure.

The dashed line is the line consisting of all the events in the history of a photon emitted at the origin of both frames and moving along the x and x' axes. It is called the *worldline* of the photon.

6 Length contraction

S and S' are as usual. Suppose S measures the length of a rod which is at rest in S'. To do that S has to observe the two ends of the rod (which are on the x' axis) simultaneously in S. Therefore

$$x'_1 = \gamma(x_1 - vt), \text{ and } x'_2 = \gamma(x_2 - vt), \tag{38}$$

where x'_1 and x'_2 are the ends of the rod in S' and x_1 and x_2 are the ends of the rod in S (observed at time t). Therefore

$$\ell' = \gamma\ell \tag{39}$$

where ℓ' is the length of the rod in S' and ℓ is its length in S.

$$\ell = \frac{\ell'}{\gamma} = \ell'\sqrt{1 - v^2/c^2} < \ell' \tag{40}$$

This is the *Lorentz-Fitzgerald contraction*: moving bodies contract in the direction of motion by a factor $\sqrt{1 - v^2/c^2}$. The length of a body is greatest measured in its rest frame—this length is called its *proper length*. The effect is analogous to looking at a stationary rod which has been rotated away from you and therefore looks shorter.

7 Time dilation

Fix a clock in S' (in other words do not change its spatial coordinates in S') and consider two events t'_1 and t'_2 at the clock. The inverse transformation to (12) says in particular that

$$t = \gamma(t' + vx'/c^2) \tag{41}$$

and therefore

$$t_1 = \gamma(t'_1 + vx'/c^2) \text{ and } t_2 = \gamma(t'_2 + vx'/c^2). \tag{42}$$

So the time interval measured in S is

$$t_1 - t_2 = \gamma(t'_1 - t'_2) > t'_1 - t'_2. \tag{43}$$

Therefore S says that the clock in S' is going slow.

This effect has been observed in the study of cosmic rays. The cosmic rays collide with the atmosphere about 20km up and produce particles called muons. The rest frame mean life of muons has been measured to be 2.2×10^{-6}s so even at the velocity of light (3×10^8m/s) they could only travel a distance of about

$$3 \times 10^8 \times 2 \times 10^{-6} = 600\text{m}.$$

However, these muons are observed at sea level!

This puzzle is resolved by the time dilation effect, which increases the lifetimes of muons by a factor γ and allows them to reach sea level without even travelling very close to the speed of light.

Exercise 22 *What is the* muon's *explanation of the fact that it reaches sea level?*

The satellites used in the global positioning system move at 14,000 km per hour, in high orbits (20,000 km above ground level). So their clocks lose 7 microseconds per day through time dilation. (They also gain 45 microseconds per day because they are higher in the Earth's gravitational field: this is an effect of *general* relativity.) The GPS system requires nanosecond accuracy in the clocks, so it has to allow for these relativistic effects.

8 Proper acceleration

Consider twins A and B. A stays at home while B rushes off in a spaceship to the Lesser Magellanic Cloud and back. B's watch will go slow according to A because B is moving relative to A. Therefore when he gets back he will have aged less than A. In other words A is older than B. Suppose you argue that A and B are symmetrical because A is certainly moving relative to B and so according to B it is A's watch which is going slow. Then B says that he will be older than A when he returns. Each cannot be older than the other so, apparently, we have a paradox.

It is resolved simply by pointing out that A and B are *not* symmetrical—A has been at rest in an inertial frame all the time while B had to accelerate.

Here is a preparatory exercise, before we look carefully at an example of acceleration.

Exercise 23 *Prove the following:*

$$\gamma v = c\sqrt{\gamma^2 - 1}, \quad c^2 d\gamma = \gamma^3 v dv, \quad d(\gamma v) = \gamma^3 dv.$$

Consider a particle accelerating through the frame S and let S' be an instantaneous rest frame of the particle. The velocity of S' relative to S is v_1, the velocity of the particle is v_2 relative to S' and v relative to S. So momentarily $v_2 = 0$ and $v = v_1$, but v_1 is constant whereas v and v_2 vary. From equation (31)

$$v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}. \tag{44}$$

Therefore, after a little calculation

$$\frac{dv}{dv_2} = \frac{\gamma^{-2}(v_1)}{(1 + \frac{v_1 v_2}{c^2})^2}. \tag{45}$$

Note also that

$$\gamma^{-2}(v) = \frac{\gamma^{-2}(v_1)\gamma^{-2}(v_2)}{(1 + \frac{v_1 v_2}{c^2})^2}. \tag{46}$$

Therefore

$$\frac{dv}{dv_2} = \frac{\gamma^{-2}(v)}{\gamma^{-2}(v_2)}. \tag{47}$$

At the particular instant when $v_2 = 0$ we obtain

$$dv = \gamma^{-2}(v) dv_2. \tag{48}$$

Also, time dilation says that at that instant

$$dt' = \frac{dt}{\gamma(v_1)} \tag{49}$$

$$= \frac{dt}{\gamma(v)}. \tag{50}$$

We define the *proper acceleration* α to be the acceleration of the particle relative to its instantaneous rest frame, so in our case

$$\alpha = \frac{dv_2}{dt'}. \tag{51}$$

However, from (48) and (50) we deduce that

$$\alpha = \gamma^3(v)\frac{dv}{dt} \tag{52}$$

$$= \frac{d}{dt}(v\gamma(v)). \tag{53}$$

Now, as a particular example, suppose α is constant. Then we can integrate (53) to get $v\gamma(v) = \alpha t$ (choosing $v = 0$ when $t = 0$). Solve this for v to obtain

$$v = \frac{dx}{dt} = \frac{c\alpha t}{\sqrt{c^2 + \alpha^2 t^2}} \tag{54}$$

which we can integrate again to get

$$x^2 - c^2 t^2 = \frac{c^4}{\alpha^2} \tag{55}$$

(ignoring the constant of integration). Therefore the worldline of the particle is a hyperbola in the Minkowski diagram. This is called *hyperbolic motion.*

Exercise 24 *Given that g, the acceleration of gravity at the Earth's surface, is $\sim 980\text{cm/sec}^2$, and that a year has $\sim 3.2 \times 10^7$ seconds, verify that, in units of years and light years, $g \approx 1$.*

A spaceship moves from rest in an inertial frame S with constant proper acceleration g. Using (50) and (52) show that its Lorentz factor relative to S is given by $\gamma = \cosh(\tau)$, where τ is time measured by its own clock.

Find its Lorentz factor relative to S when its own clock indicates times τ equals 1 day, 1 year, 10 years.

Exercise 25 *Consider the same spaceship as in the previous exercise. Using (53), find expressions for v then x in terms of t and hence calculate the distances and times travelled in S corresponding to τ equals 1 day, 1 year, 10 years.*

If the spaceship accelerates for 10 years of its own time, then decelerates for 10 years, and then repeats the whole manoeuvre in the reverse direction, what is the total time elapsed in S during the spaceship's absence?

9 The metric

Suppose x^a and y^a are the position vectors in a frame S of two events. We define the *scalar product* between x^a and y^a to be

$$x^a y^b g_{ab} = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3, \tag{56}$$

where the matrix g_{ab}, called the *metric*, is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{57}$$

Now let $x^a = (ct, x, y, z)$. We saw in exercise 7 that $x^a x^b g_{ab}$ is invariant under Lorentz transformations. This scalar invariant is called the square of the *magnitude* of x^a.

It is extremely important to note that it is perfectly possible for $x^a x^b g_{ab}$ to be negative. We will explore this in the next section.

Exercise 26 *Show that for any three x^a, y^a and z^a*

$$x^a y^b g_{ab} = y^a x^b g_{ab}$$

and

$$x^a(y^b + z^b)g_{ab} = x^a y^b g_{ab} + x^a z^b g_{ab}.$$

Exercise 27 *Using the fact that*

$$(x^a + y^a)(x^b + y^b)g_{ab} = x^a x^b g_{ab} + 2x^a y^b g_{ab} + y^a y^b g_{ab}$$

show that $x^a y^b g_{ab}$ is invariant under Lorentz transformations.

Exercise 28 *Show that for any two x^a and y^a*

$$d(x^a y^b g_{ab}) = dx^a y^b g_{ab} + x^a dy^b g_{ab}.$$

Finally in this section, we say that x^a and y^a are *orthogonal* if $x^a y^b g_{ab} = 0$.

10 Causality

When we first drew the Minkowski diagram we ignored both y and z. We will draw it again, this time only ignoring z.

The vectors are of various types. A vector x^a is said to be

$$\begin{array}{lll} \text{(a)} & \textit{timelike} & \text{if} \quad x^a x^b g_{ab} > 0 \\ \text{(b)} & \textit{spacelike} & \text{if} \quad x^a x^b g_{ab} < 0 \\ \text{(c)} & \textit{null} & \text{if} \quad x^a x^b g_{ab} = 0 \end{array} \tag{58}$$

Timelike or null vectors can point either into the future ($x^0 > 0$) or into the past ($x^0 < 0$). The null vectors lie along the *null cone*, whose equation in our picture is

$$(ct)^2 = x^2 + y^2. \tag{59}$$

This classification of the vectors into these various types is completely independent of the frame of reference in which the components of the vectors are calculated, because $x^a x^b g_{ab}$ is invariant under Lorentz transformations.

Exercise 29 *Why are we allowed to use the condition $x^0 > 0$ in the definition of a future-pointing vector?*

Exercise 30 *A flash of light at the origin can be thought of as an expanding sphere of photons; the radius of this sphere is increasing at c. Describe this using a Minkowski diagram.*

Suppose x^a is a timelike vector: we will show that there is a frame S' in which the spatial components of x^a are all zero. We can start with $x^a = (ct, x, 0, 0)$, having performed a spatial rotation if necessary. Because x^a is a timelike vector, we know that

$$(ct)^2 - x^2 > 0 \tag{60}$$

and so $x/t < c$. Hence we can choose our frame S' in standard position with respect to S, with $v = x/t$. Then $x' = \gamma(x - vt) = 0$, as required. In the case that x^a is future pointing, we can think of it as defining the worldline of a particle moving at $v < c$, with S' as the particle's rest frame.

Conversely, if there is a frame S' in which all the spatial components of x^a are zero, then certainly $x'^a x'^b g_{ab} > 0$ in S', and hence $x^a x^b g_{ab} > 0$ in S, making x^a a timelike vector.

Exercise 31 *Show that x^a is spacelike if and only if there is a frame S' in which the temporal part of x^a is zero, which means that in S' the two ends of the vector x^a are* simultaneous *events.*

Exercise 32 *Choose vectors x^a and t^a in S pointing along the x' and t' axes in the Minkowski diagram. Prove that $x^a t^b g_{ab} = 0$ (which implies that these axes are still orthogonal in the Minkowski sense).*

Exercise 33 *The set of events considered by any inertial observer to be simultaneous at his time $t = t_0$ is said to be the observer's* instantaneous 3-space *$t = t_0$. Show that the join of any two events in such a space is orthogonal to the observer's worldline, and that, conversely, any two events whose join is orthogonal to the observer's worldline are considered simultaneous by him.*

11 Proper time

The *proper time* interval $\Delta\tau$ between two events is defined as follows:

$$\Delta\tau^2 = \frac{\Delta s^2}{c^2}, \tag{61}$$

where

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2, \tag{62}$$

which is invariant.

Let $\mathcal{L}$ be the worldline of a particle moving with constant velocity. This worldline is a straight line in spacetime which lies inside the light cone of any event on the line. The proper time is a natural parameter along $\mathcal{L}$. It is defined as the time measured by a clock carried by the particle, or in other words the time t' in the *rest frame* S' of the particle. All observers *agree* on the value of τ.

Consider two events on $\mathcal{L}$. In S, the coordinates of the vector joining them are $(c\Delta t, \Delta x, \Delta y, \Delta z,)$. In the rest frame S', these coordinates are $(c\Delta\tau, 0, 0, 0)$.

From equation (62) we obtain

$$\frac{\Delta t}{\Delta \tau} = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} = \gamma(u) \tag{63}$$

where

$$u^2 = \left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2 + \left(\frac{\Delta z}{\Delta t}\right)^2. \tag{64}$$

Consider the vector

$$\frac{1}{\Delta \tau}\left(c\Delta t, \Delta x, \Delta y, \Delta z\right) = \frac{\Delta t}{\Delta \tau}\left(c, \frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t}\right), \tag{65}$$

which from (63) can be written

$$\gamma(u)\left(c, \mathbf{u}\right). \tag{66}$$

This is called the *4-velocity* of the particle with 3-velocity

$$\mathbf{u} = \left(\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t}\right).$$

Exercise 34 *Show that the components of any 4-velocity are* $(c, 0, 0, 0)$ *in its rest frame, and deduce that for any 4-velocity* v^a, $v^a v^b g_{ab} = c^2$. *Note that 4-velocities must therefore be timelike and future pointing.*

All of the above remains true if the 3-velocity $\mathbf{u}$ is varying. In this case the worldline is a curve, just as in equation (55). We simply take the limit $\Delta t \to 0$ and replace Δ by the differential d, so that (63) is replaced by

$$\frac{dt}{d\tau} = \gamma(u). \tag{67}$$

Now, the 4-velocity is the tangent to the wordline.

Exercise 35 *Find the unit tangent to the worldline.*

Exercise 36 *For any two 4-velocities* v^a *and* w^a *prove that* $v^a w^b g_{ab} = c^2\gamma(u)$, *where* $\mathbf{u}$ *is the 3-velocity of the second particle in the rest frame of the first. Deduce that*

$$v^a w^b g_{ab} = vw \cosh\phi$$

where v *and* w *are the magnitudes of* v^a *and* w^a, *and* ϕ *is the relative rapidity of the particles.*

12 The Doppler effect

Let a light source P travelling through an inertial frame S have instantaneous velocity $\mathbf{u}$ and radial velocity component u_r relative to the origin O of S. Let the proper time between successive pulses (or crests of the light waves) be $d\tau$.

In S this is $\gamma(u)d\tau$ (by time dilation). The next pulse is not only emitted later but has farther to travel, by $\gamma d\tau u_r$. So these pulses arrive at O a time

$$dt = \gamma d\tau + \gamma d\tau u_r/c$$

apart.

The quantities $d\tau$ and dt are inversely proportional, respectively, to the proper frequency ν_0 of P and its frequency ν as observed at O. Therefore

$$\frac{\nu_0}{\nu} = \frac{1+\frac{u_r}{c}}{\sqrt{1-\frac{u^2}{c^2}}} = 1 + \frac{u_r}{c} + \frac{1}{2}\frac{u^2}{c^2} + O(u^3/c^3). \tag{68}$$

The first two terms are the classical formula. The rest contribute an extra effect due to time dilation. If $u_r > 0$ we get an extra redshift, while if $u_r < 0$ we get a reduced blueshift, because the u^2/c^2 term partly cancels the u_r/c term.

Completely unlike the classical case, there is a transverse Doppler effect, as can see by putting $u_r = 0$ in (68). This is a redshift due *purely* to time dilation.

If the direction of motion of P relative to O is radial, then $u_r = u$ and (68) can be written

$$\frac{\nu_0}{\nu} = \sqrt{\frac{c+u}{c-u}}. \tag{69}$$

Exercise 37 *Describe a situation in which the classical and relativistic formulae predict Doppler shifts in opposite directions—meaning that one of them predicts a redshift while the other predicts a blueshift.*

Finally in this section, we consider the thermal Doppler effect. Radioactive nuclei in a crystal emit light rays (γ rays) of characteristic frequencies, called "spectral lines". In a hot crystal, the nuclei move thermally in a rapid and random way about their average positions. In the Doppler formula (68) the first order term u_r/c averages out to zero (because u_r is random) and causes only a broadening of the spectral lines. The second order term u^2/c^2 causes a *shift* in the spectral lines—a pure time dilation effect. This shift was experimentally observed by Pound and Rebka in 1960.

13 Aberration

If two observers in standard configuration measure the angle which an incoming ray of light makes with their x (or x') axis, then they will obtain different results. This is called the *aberration* effect. It also occurs in Newtonian theory—imagine running through vertically falling rain—but as with the Doppler effect there is a new formula in special relativity.

Suppose an incoming ray of light makes an angle which is measured in S to be α with the x axis and let α' be the corresponding angle in S'. Choose the y axis so that the ray is in the xy plane. The 3-velocity of the ray is

$$\mathbf{u} = (-c\cos\alpha, -c\sin\alpha, 0). \tag{70}$$

Now we use the velocity transformation formulae worked out in exercise (21):

$$u'_1 \equiv -c\cos\alpha' = \frac{u_1 - v}{1 - \frac{u_1 v}{c^2}} = \frac{-c\cos\alpha - v}{1 + \frac{v\cos\alpha}{c}}, \tag{71}$$

$$u'_2 \equiv -c\sin\alpha' = \frac{u_2}{\gamma(1 - \frac{u_1 v}{c^2})} = \frac{-c\sin\alpha}{\gamma(1 + \frac{v\cos\alpha}{c})}. \tag{72}$$

Hence

$$\cos\alpha' = \frac{\cos\alpha + v/c}{1 + \frac{v\cos\alpha}{c}}, \qquad \sin\alpha' = \frac{\sin\alpha}{\gamma(1 + \frac{v\cos\alpha}{c})}. \tag{73}$$

There is another way of writing these relations. Use the trigonometric identity

$$\tan\frac{\alpha}{2} = \frac{\sin\alpha}{1 + \cos\alpha} \tag{74}$$

to get

$$\tan\frac{\alpha'}{2} = \sqrt{\frac{c - v}{c + v}}\tan\frac{\alpha}{2} \tag{75}$$

$$= e^{-\phi}\tan\frac{\alpha}{2}. \tag{76}$$

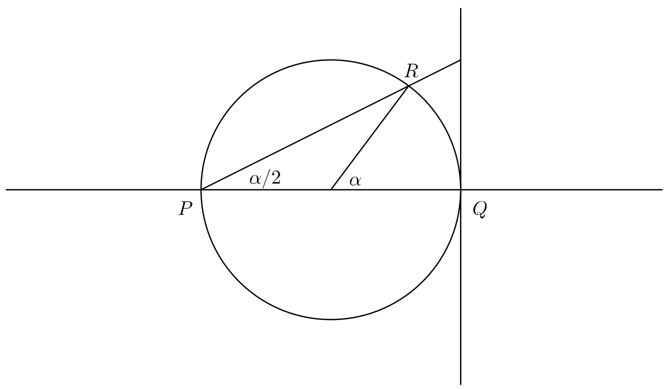
Suppose you are in a spaceship in interstellar space. Let S be the rest frame of the nearby stars (ignoring the fact that they are probably moving relative to each other) and let S' be the rest frame of your spaceship. If the x' axis points out of the front of the spaceship, then the frame S' has speed v with respect to S, and these two frames are in the standard configuration. What do you see? Consider a star for which $\alpha = 90^o$. Then from (73) $\cos\alpha' = v/c$, so if v is equal to $\frac{c\sqrt{3}}{2}$ for example, we have $\cos\alpha' = \frac{\sqrt{3}}{2}$, and so $\alpha' = 30^o$. This means that *half* the stars can be seen within 30^o of the forward direction. Of course $\frac{c\sqrt{3}}{2}$ is quite a high speed! Note that the colours of these stars will also have been affected by the relative motion; how?

Exercise 38 *Two momentarily coincident observers travel towards a small and distant object. To one observer that object looks twice as large as to the other, so that in equation* (73) $\alpha' = 2\alpha$, *and* α *is small. Find their relative velocity.*

Roger Penrose [3] discovered an ingenious way of determining the visual distortion of moving objects. Consider two observers S and S' in the standard configuration at the moment they coincide. Draw a spatial unit sphere around each observer's spatial origin O. All that the observer sees can be mapped onto this sphere: it is the observer's sky.

Consider any single light ray received from the object. We apply the formula (75) to this ray, where α is the angle between the ray and the x (or x') axis. Let R be the position on the sky at which S observes this ray, and let P and Q be the intersections of the sky with the x axis, with $\overrightarrow{PQ}$ the direction of motion.

Consider also the tangent plane to the sphere at point Q, which we can think of as the observers' screen.



Since the angle at the circumference is half of the angle at the centre of the circle we see that $2\tan\alpha/2$ gives the stereographic projection of the sky point R onto the screen.

It then follows from (75) that whatever the two momentarily coincident observers S and S' see, the images on their screens are *identical* except for scale, and the rescaling is given by $\sqrt{\frac{c-v}{c+v}}$.

Thus for an object moving at speed $-v$ (parallel to the x axis) let S be its rest frame so then S' is the observer's rest frame. To construct its appearance for the observer:

1. take its silhouette on the sphere as it would appear at *rest,*
2. project stereographically from P onto the screen,
3. shrink the screen image by $\sqrt{\frac{c-v}{c+v}}$ through point Q, and
4. invert the stereographic projection from P to get the silhouette as seen by the observer.

Consider a moving sphere. It always has a circular outline on the sky in its rest frame. But it is an important property of any stereographic projection that circles on the sky sphere are mapped to circles on the screen plane and vice versa. Thus by the above analysis the moving sphere will *always* appear circular to any observer.

14 $SL(2,\mathbb{C})$

In fact, the best way to understand aberration is to study the relationship between the group $SL(2,\mathbb{C})$ and the Lorentz group. Recall first that $SL(2,\mathbb{C})$ is the group of 2×2 matrices with complex entries and unit determinant.

Given a vector x^a, we may define a Hermitian matrix $\Psi(x^a)$ by

$$\Psi(x^a) = \frac{1}{\sqrt{2}} \begin{bmatrix} x^0 + x^3 & x^1 + ix^2 \\ x^1 - ix^2 & x^0 - x^3 \end{bmatrix}. \tag{77}$$

Clearly Ψ is a one-one correspondence between 2×2 Hermitian matrices and vectors in Minkowski space. Furthermore, the determinant of the matrix is half the square of the magnitude of the vector:

$$|\Psi(x^a)| = \frac{1}{2} x^a x^b g_{ab}.$$

If we multiply the matrix $\Psi(x^a)$ on the left by an element of $SL(2,\mathbb{C})$ and on the right by its conjugate transpose

$$\Psi(x^a) \to M\Psi(x^a)M^*$$

then the result will be another Hermitian matrix and the determinant will be unchanged. This process therefore defines a linear transformation on the vector x^a preserving its magnitude, which is a Lorentz transformation. Thus we have a map from $SL(2,\mathbb{C})$ to the Lorentz group. The following properties of this map may be readily established:

- it is a group homomorphism;
- it is into the component $L_+^\uparrow$ of the Lorentz group which preserves time and space orientation;
- its kernel consists of $\pm I$ in $SL(2,\mathbb{C})$, where I is the identity matrix.

Since $SL(2,\mathbb{C})$ is also a six-parameter group the map is necessarily onto $L_+^\uparrow$ and so is a two-to-one isomorphism.

In fact the night skies of two inertial observers S and S' will be related by an element of $SL(2,\mathbb{C})$. This is the underlying reason for the aberration effects in the previous section. Penrose's work has made brilliant use of this observation, which has ramifications throughout special and general relativity and is a key ingredient of twistor geometry [1].

References

[1] S A Huggett and K P Tod 1994 *An introduction to twistor theory*, second edition, Cambridge University Press.

[2] A Pais 1982 *'Subtle is the Lord . . . '*, Oxford University Press.

[3] R Penrose 1959 *The apparent shape of a relativistically moving sphere*, Proc. Camb. Phil. Soc. **55**, 4, 137–139.

[4] R Penrose 2004 *The Road to Reality*, Jonathan Cape.

[5] W Rindler 2006 *Relativity: Special, General, and Cosmological*, second edition, Oxford University Press.