Part II Relativity

William Royce

June 14, 2024

Part II Physics, The University of Cambridge

Abstract

> 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.

General relativity is the theory of space and time and gravity. The essence of the theory is simple: gravity is geometry. The effects that we attribute to the force of gravity are due to the bending and warping of spacetime, from falling cats, to orbiting spinning planets, to the motion of the cosmos on the grandest scale. The purpose of these lectures is to explain this.

Before we jump into a description of curved spacetime, we should first explain why Newton's theory of gravity, a theory which served us well for 250 years, needs replacing. The problems arise when we think about disturbances in the gravitational field. Suppose, for example, that the Sun was to explode. What would we see? Well, for 8 glorious minutes – the time that it takes light to reach us from the Sun – we would continue to bathe in the Sun's light, completely oblivious to the fate that awaits us. But what about the motion of the Earth? If the Sun's mass distribution changed dramatically, one might think that the Earth would start to deviate from its elliptic orbit. But when does this happen? Does it occur immediately, or does the Earth continue in its orbit for 8 minutes before it notices the change?

Of course, the theory of special relativity tells us the answer. Since no signal can propagate faster than the speed of light, the Earth must continue on its orbit for 8 minutes. But how is the information that the Sun has exploded then transmitted? Does the information also travel at the speed of light? What is the medium that carries this information? As we will see throughout these lectures, the answers to these questions forces us to revisit some of our most basic notions about the meaning of space and time and opens the to door to some of the greatest ideas in modern physics such as cosmology and black holes.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Newtonian Gravity

There is a well trodden path in physics when trying to understand how objects can influence other objects far away. We introduce the concept of a field. This is a physical quantity which exists everywhere in space and time; the most familiar examples are the electric and magnetic fields. When a charge moves, it creates a disturbance in the electromagnetic field, ripples of which propagate through space until they reach other charges. To develop a causal theory of gravity, we must introduce a gravitational field that responds to mass in some way.

It's a simple matter to cast Newtonian gravity in terms of a field theory. A particle of mass m_G experiences a force that can be written as

$$\mathbf{F} = -m_G \boldsymbol{\nabla}\Phi. \tag{1.1}$$

The quantity m_G is the *passive gravitational mass*, and it determines the gravitational force on the particle.The gravitational field $\Phi(\mathbf{r}, t)$ is determined by the surrounding matter distribution which is described by the mass density $\rho(\mathbf{r}, t)$. If the matter density is static, so that $\rho(\mathbf{r})$ is independent of time, then the gravitational field obeys

$$\nabla^2\Phi = 4\pi G\rho, \tag{1.2}$$

with Newton's constant G given by

$$G \approx 6.67 \times 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}. \tag{1.3}$$

This equation is simply a rewriting of the usual inverse square law of Newton. For example, if a mass M is concentrated at a single point we have

$$\rho(\mathbf{r}) = M\delta^{(3)}(\mathbf{r}) \implies \Phi = -\frac{GM}{r}, \tag{1.4}$$

which is the familiar gravitational field for a point mass.

The question that we would like to answer is: how should we modify (1.2) when the mass distribution $\rho(\mathbf{r})$ changes with time? Of course, we could simply postulate that (1.2) continues to hold even in this case. A change in ρ would then immediately result in a change of Φ throughout all of space. Such a theory clearly is not consistent with the requirement that no signal can travel faster than light. Our goal is to figure out how to generalise (1.2) in a manner that is compatible with the postulates of special relativity. The end result of this goal will be a theory of gravity that is compatible with special relativity: this is the general theory of relativity.

Fixing this incompatibility will ultimately require a radical modification of how we think about gravity and, indeed, spacetime itself. Sticking with Newtonian gravity for the moment, it is not immediately obvious that the mass density appearing in Poisson's equation should refer to the density of the passive gravitational mass. Rather, let us also introduce the active gravitational mass m_A, so that the relevant mass density for a point particle at position $\mathbf{r}'(t)$ at time t is

$$\rho(\mathbf{r}, t) = m_A \delta^{(3)}(\mathbf{r} - \mathbf{r}'(t)). \tag{1.5}$$

For the point particle, the relevant solution of Poisson's equation is

$$\Phi(\mathbf{r}, t) = -\frac{G m_A}{|\mathbf{r} - \mathbf{r}'(t)|}. \tag{1.6}$$

It follows that the force on a test particle of passive gravitational mass $m_{G,1}$ at position $\mathbf{r}_1$ at time t due to a particle of active gravitational mass $m_{A,2}$ at position $\mathbf{r}_2$ *at the same time* t is

$$\mathbf{F}_{2 \text{ on } 1} = -G m_{G,1} m_{A,2} \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2{}^3|}. \tag{1.7}$$

Similarly, the force on the second particle due to the first is

$$\mathbf{F}_{1 \text{ on } 2} = -G m_{G,2} m_{A,1} \frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_1 - \mathbf{r}_2{}^3|}. \tag{1.8}$$

If momentum is to be conserved, i.e., $\mathbf{F}_{2 \text{ on } 1} = \mathbf{F}_{1 \text{ on } 2}$, we must have

$$m_{G,1} m_{A,2} = m_{G,2} m_{A,1}. \tag{1.9}$$

Since this must hold for arbitrary masses, we must have that the ratio of passive to active gravitational mass is the same for all particles. Thus, we can take these masses to be equal, $m_G = m_A$, for all matter (absorbing their universal ratio in the gravitational constant).

This sort of universality is not unusual in physics – a similar thing happens in electromagnetism, for example, where the passive and active electric charges are equal. However, there is a further equality of masses in Newtonian gravity that is rather more surprising: the equality of gravitational and inertial masses. A particle acted on by a force $\mathbf{F}$ experiences an acceleration such that

$$\mathbf{F} = m_I \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}, \tag{1.10}$$

where m_I is the *inertial mass*. For the gravitational force, the acceleration is

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = -\frac{m_G}{m_I} \boldsymbol{\nabla}\Phi. \tag{1.11}$$

It is an experimental fact[1] (known since Galileo's time) that the ratio m_G/m_I is the same for all particles, so we can always take $m_G = m_I$ (further absorbing their universal ratio in the gravitational constant). This means that if two particles of different composition fall freely in a gravitational field, they have the same acceleration. This is often rephrased as the *weak equivalence principle*:

> Freely-falling particles with negligible gravitational self-interaction follow the same path through space and time if they have the same initial position and velocity, independent of their composition.

This property of gravity is in striking contrast to other forces; for example, in electromagnetism the acceleration of a point particle in a given electric field depends on the ratio of the electric charge to inertial mass, which is definitely not universal.

[1] The equality of gravitational and inertial masses is now verified to the level of one part in 10^{13}.

1.2 Implications of the Equivalence Principle

Consider an observer in a free-falling, non-rotating elevator in a uniform gravitational field. Relative to this observer, free-falling particles move on straight lines at constant velocity – the effects of the uniform gravitational field have been removed and the observer perceives that the usual laws of special relativistic kinematics hold. This idea motivates an extension of the weak equivalence principle to what is known as the *strong equivalence principle*:

> In an arbitrary gravitational field, *all* the laws of physics in a free-falling, non-rotating laboratory occupying a sufficiently small region of spacetime look locally like special relativity (with no gravity).

Note how the strong equivalence principle is supposed to apply to all laws of physics, not just the dynamics of free-falling particles. Why the qualification of observations over a sufficiently small region of spacetime?

Consider the same elevator falling freely in the non-uniform gravitational field of the earth. Free particles initially at rest in the elevator will move together over time as they follow radial trajectories towards the centre of the earth. It is these tidal effects that are the physical manifestation of the gravitational field, and that cannot be removed by passing to the free-falling frame. However, for sufficiently local measurements in space and time, these tidal effects are undetectable, and physics relative to the free-falling elevator looks just like special relativistic physics in an inertial frame of reference in the absence of gravity.

The strong equivalence principle implies the local equivalence of a gravitational field and acceleration. In particular, it implies that a constant gravitational field is unobservable – observations in a reference frame at rest in such a field would be indistinguishable from those in a uniformly-accelerating reference frame in the absence of gravity. In special relativity, physics looks simple when referred to an inertial frame, one defined by comoving, unaccelerated observers with synchronised clocks. However, with gravity, the equivalence principle tells us that physics looks equally simple *locally* in a free-falling reference frame, suggesting that we should *define* inertial reference frames locally by free-falling observers. Acceleration should be defined relative to such local inertial frames, so that a particle acted on by no other force (and so free-falling) should be regarded as unaccelerated.

1.2.1 Gravity as Spacetime Curvature

The universality of free fall suggested to Einstein that the trajectories of free-falling particles should be determined by the local structure of spacetime, rather than by the action of a gravitational force with a mysterious universal coupling to matter.

Local inertial reference frames correspond to local systems of coordinates over spacetime so that the geometry over a small region looks like that of the spacetime of special relativity. Gravity manifests itself through our inability to extend such coordinates globally, reflecting the *curvature of spacetime.*

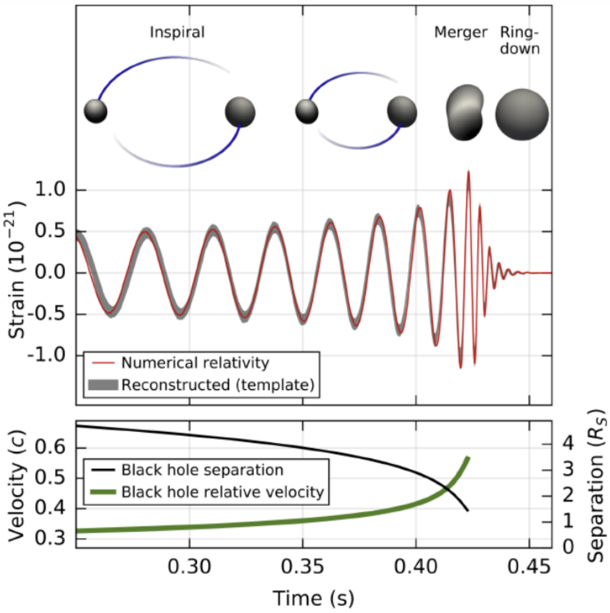

Fig. 1.1: Top: Estimated gravitational wave strain amplitude inferred from the LIGO data for their discovery event. The signal is generated from the inspiral, merger and ring-down of two massive black holes. The properties of the source can be estimated by comparing the measured waveform with detailed calculations in general relativity. Bottom: the relative speed and separation (in units of the Schwarzschild radius, $R_s = 2GM/c^2$) of the blackholes during the event. For reference, the Newtonian potential at R_S away from a mass M is $|\Phi|/c^2 = 1/2$. Figure taken from Abbot et al., Phys. Rev. Lett. 116, 061102 (2016).

General relativity abandons the idea of gravity as a force defined on the fixed spacetime of special relativity, replacing it with a geometric theory in which the geometry of spacetime determines the trajectories of free-falling particles, the geometry itself being curved by the presence of matter.

1.3 Further Motivation: Extreme Gravity

Newtonian gravity is recovered from general relativity in the limit of low relative speeds of particles, $v \ll c$, and weak gravitational fields, typically $|\Phi| \ll c^2$. Note that in situations where speeds are determined by gravity, these two regimes are generally equivalent.

To see this, consider a particle in a circular orbit of radius R around a mass M in Newtonian gravity: the speed is determined by

$$\frac{v^2}{R} = \frac{GM}{R^2}, \tag{1.12}$$

and so

$$\frac{v^2}{c^2} = \frac{GM}{Rc^2} = \frac{|\Phi|}{c^2}. \tag{1.13}$$

However, increasingly we are observing phenomena where Newtonian gravity is a very poor approximation. A striking example is the recent first detection of gravitational waves by the LIGO interferometer; see Fig. 1.1.

Gravitational waves are wavelike disturbances in the geometry of spacetime, which can be detected by looking for their characteristic quadrupole distortion (i.e., a shortening in one direction and stretching in an orthogonal direction) of the two arms of a laser interferometer. Gravitational waves propagate at the speed of light and are a natural prediction of general relativity; they do not arise in Newtonian gravity where the potential responds instantly to distant rearrangements of mass.

The first LIGO signal was generated by a truly extreme astrophysical source: two merging black holes each with a mass around 30 times that of the Sun at a distance from us of around 2 Gly.As the blackholes orbited their common centre of mass, the system radiated gravitational waves causing the blackholes to spiral inwards and increase their speed until they merged to form a single black hole. Such sources probe the strong-field regime of general relativity during the merger phase and involve highly relativistic speeds (see Fig. 1.1). At its peak, the source was losing energy to gravitational waves at a rate of 3.6×10^{49}W, which is equivalent to 200 times the rest mass energy of the Sun per second!

Chapter **2**

Recap of 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

$$c = 299792458\mathrm{m\,s}^{-1} \tag{2.1}$$

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 $c \approx 3 \times 10^8\mathrm{m\,s}^{-1}$.

The first thing to say is that the speed of light is fast. Really fast. The speed of sound is around $300\mathrm{m\,s}^{-1}$; escape velocity from the Earth is around $104\mathrm{m\,s}^{-1}$; the orbital speed of our solar system in the Milky Way galaxy is around $105\mathrm{m\,s}^{-1}$. 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, the first of these is simply the Galilean principle of relativity as in classical Newtonian mechanics. The second postulate is more surprising:

- The principle of relativity: the laws of physics are the same in all inertial frames.
- 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.

2.1 Newtonian Geometry of Space and Time

Newtonian theory assumes an absolute time – the same for every observer. This common sense view is encapsulated in the Galilean transformations.. Mathematically, we derive this "obvious" result as follows: two inertial frames, S and S', in *standard configuration*: axes aligned, the same spacetime origin, which move relative to each with velocity $\mathbf{v} = (v, 0, 0)$, have Cartesian coordinates related by

$$x' = x - vt, \qquad y' = y, \qquad z' = z, \qquad t' = t \tag{2.2}$$

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 S0 the trajectory if the light ray is $x'/t' = c - nv$. 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 (2.2). 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.

For two events A and B, the Gallilieean transformation implies that

- the time difference $\Delta t = t_B - t_A$ is invariant; and
- $\Delta r^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$ is invariant for simultaneous events (since Δx, Δy, and Δz are).

Space and time are separate entities in Newtonian theory.

2.2 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

$$x' = f(x, t), \qquad t' = g(x, t), \tag{2.3}$$

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

$$x' = \alpha_1 x + \alpha_2 t, \qquad t' = \alpha_3 x + \alpha_f t, \tag{2.4}$$

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

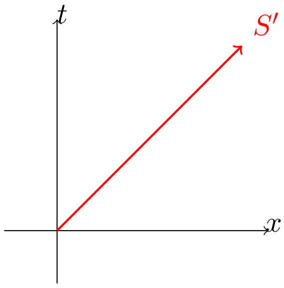

Fig. 2.1: Space-time diagram, representing the motion of a particle at the origin $x' = 0$ in S', which moves along the trajectory $x = vt$ in S.

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 Fig. (2.1). 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,

$$x' = \gamma(x - vt), \tag{2.5}$$

for some coefficient γ. Once again, the overall coefficient can be a function of the velocity: $\gamma = \gamma_v$. (We've used subscript notation v rather than the more standard (v) to denote that depends on v. This avoids confusion with the factors of $(x - vt)$ which aren't arguments of but will frequently appear after like in the equation (2.5)).

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 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

$$\tilde{x}' = \gamma_{-v}(\tilde{x} + v\tilde{t}). \tag{2.6}$$

Comparing this to (2.5), 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 (2.5) now tells us that

$$x = \gamma(x' + vt'). \tag{2.7}$$

Now the function $\gamma_v = \gamma_{-v}$. But by the argument above, we know that v = v. In other words, the coefficient appearing in (2.7) is the same as that appearing in (2.5).

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 (2.2). 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

$$x = ct. \tag{2.8}$$

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

$$x' = ct'. \tag{2.9}$$

Substituting these trajectories into (2.5) and (2.7), we have two equations relating t and t',

$$ct' = \gamma(c - v)t, \qquad \text{and,} \qquad ct = \gamma(c + v)t'. \tag{2.10}$$

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

$$\boxed{\gamma = \sqrt{\frac{1}{1 - v^2/c^2}}.} \tag{2.11}$$

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 (2.5) is approximately the same as the Galilean transformation (2.2). However, as $v \to c$ we have $\gamma \to \infty$. Furthermore, becomes imaginary for $v > c$ which means that we're unable to make sense of inertial frames with relative speed $v > c$.

Equations (2.5) and (2.11) 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 (2.5) into (2.7) and rearranging, we get

$$t' = \gamma\left(t - \frac{v}{c^2}x\right). \tag{2.12}$$

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 (2.5) and (2.12) are the *Lorentz transformations.*

2.2.1 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

$$y' = \kappa y, \tag{2.13}$$

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,

$$\begin{aligned} x' &= \gamma\left(x - \frac{v}{c}ct\right), \\ y' &= y, \\ z' &= z, \\ ct' &= \gamma\left(ct - \frac{v}{c}x\right), \end{aligned} \tag{2.14}$$

where γ is given by (2.11). These are also known as Lorentz boosts. Notice that for $v/c \ll 1$, the Lorentz boosts reduce to the more intuitive Galilean boosts. (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 $x' = \gamma(x - vt) = \gamma(ct - vx/c) = ct'$

What about an object travelling in the y direction at the speed of light? Its trajectory in S is $y = ct$. From (2.14), its trajectory in S' is $y' = ct'/\gamma$ and $x' = vt'$. Its speed in S' is therefore $v'^2 = v_x^2 x + v_y^2$, or

$$v'^2 = \left(\frac{x'}{t'}\right)^2 + \left(\frac{x'}{t'}\right)^2 = v^2 + \frac{c^2}{\gamma^2} = c^2. \tag{2.15}$$

Note how time and space are mixed by the Lorentz transformation. However, for two events, the (squared) interval

$$\boxed{\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2.} \tag{2.16}$$

is *invariant* under any Lorentz transformation. In special relativity, space and time are united into a four-dimensional continuum called spacetime with invariant geometry characterised by Δs^2. The spacetime of special relativity is topologically $\mathbf{R}^4$. When endowed with the measure of distance (2.16), 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.

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\,. \tag{2.17}$$

In this infinitesimal form, $\mathrm{d}s^2$ is called the *line element*.

2.2.2 Lorentz Transformations as 4D 'Rotations'

Different Cartesian inertial frames S and S' simply relabel events in Minkowski spacetime, i.e., perform a coordinate transformation $(ct, x, y, z) \to (ct', x', y', z')$. It is often convenient to define the rapidity parameter ψ (which runs from $-\infty$ to ∞) by $v/c = \tanh\psi$, so that

$$\gamma = \cosh\psi, \qquad \text{and,} \qquad \gamma v/c = \tanh\psi \tag{2.18}$$

For S and S' in standard configuration, we have

$$\begin{aligned} ct' &= ct\cosh\psi - x\sinh\psi \\ x' &= -ct\sinh\psi + x\cosh\psi \\ y' &= y \\ z' &= z \end{aligned} \tag{2.19}$$

These are like a rotation in the $ct - x$ plane, but with hyperbolic rather than trigonometric functions. The hyperbolic functions are necessary to ensure the invariance of Δs^2 given the minus signs in its definition.

2.2.3 More Complicated Lorentz Transformations

More generally, the relation between two Cartesian inertial frames S and S' can differ from that for the standard configuration since[1]:

- the $4D$ origins may not coincide, i.e., the event at $ct = x = y = z = 0$ may not be at $ct' = x' = y' = z' = 0$;
- the relative velocity of the two frames may be in an arbitrary direction in S, rather than along the x-axis; and
- the spatial axes in S and S' may not be aligned, e.g., the components of the relative velocity in S' may not be minus those in S.

We can always deal with the origins not coinciding (known as inhomogeneous Lorentz transformations or Poincaré transformations) by appropriate temporal and spatial displacements. We can find the form of the remaining Lorentz transformation in the general case by decomposing as follows.

1. Apply a purely spatial rotation in the frame S to align the new x-axis with the relative velocity of the two frames.
2. Apply a standard Lorentz transformation as in Eq. (2.12) and (2.5).
3. Apply a spatial rotation in the transformed coordinates to align the axes with those of S'.

[1] Lorentz transformations can be considered more formally as linear transformations that preserve the interval Δs^2. In this case, the definition admits transformations that are not continuously connected to the identity; i.e., parity transformations and/or time reversal. We shall not consider such transformations further.

Given a reference frame S, the *Lorentz boost* of this frame for a general relative velocity $\mathbf{v}$ is obtained by rotating the spatial axes of S so that the relative velocity is along the new x-axis, applying the standard Lorentz transformation, and applying the inverse spatial rotation in the transformed frame.If the relative velocity is along the original x-axis, this reduces to the standard Lorentz transformation.

More generally, reference frames connected by a Lorentz boost have their spatial axes as aligned as possible given the relative velocity of the frames, i.e., they are generated by hyperbolic "rotations" in the plane defined by the ct-axis and the relative velocity.

2.2.4 The Interval

As we have seen, the interval is invariant under Lorentz transformations. This is particularly transparent using the hyperbolic form of the standard transformation:

$$\begin{aligned}\Delta s^2 &= c^2(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 \\ &= [(c\Delta t)\cosh\psi - (\Delta x)\sinh\psi]^2 - [-(c\Delta t)\sinh\psi + (\Delta x)\cosh\psi]^2 - \Delta y^2 - \Delta z^2 \\ &= c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2\end{aligned} \tag{2.20}$$

The interval is also invariant under more general Lorentz transformations since a shift in origin does not alter the differences $(c\Delta t, \Delta x, \Delta y, \Delta z)$, and rotations preserve the spatial interval $\Delta x^2 + \Delta y^2 + \Delta z^2$.

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. 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.

2.2.5 Space-Time 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

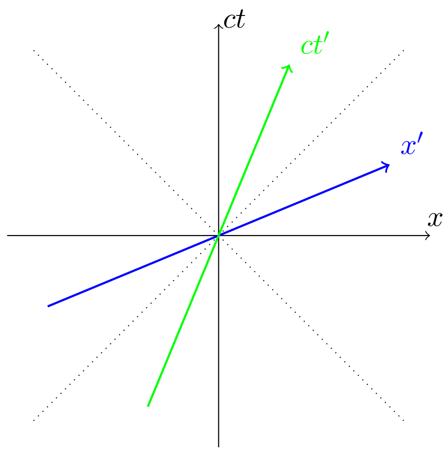

Fig. 2.2: Space-time diagram with axes corresponding to an inertial frame S' moving with a relative velocity. They can be the thought of as the x and ct axes, rotated by an equal amount towards the diagonal light ray. The fact the axes are symmetric about the light ray reflects the fact that the speed of light is equal to c in both frames.

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*.

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.

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 (2.14), is given by the equation $ct = \frac{v}{c}x$.

These two axes are drawn on the Fig. (2.2). 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.

2.2.6 Causality and the Lightcone

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,

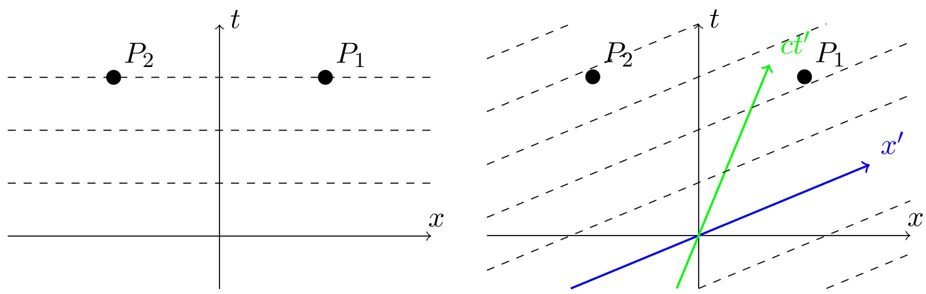

Fig. 2.3: Simultaneity is relative.

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 Fig. 2.3 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 Fig. 2.3.

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.

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 A and draw the 45° light rays that emerge from A. This is called the *light cone*. In more than a single spatial dimension, the light cone is really two cones, touching at the point A. They are known as the future light cone and past light cone.

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

This tells us that the events which all observers agree can be causally influenced by A are those inside the future light cone. Similarly, the events which can plausibly influence A 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. We will show later 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.

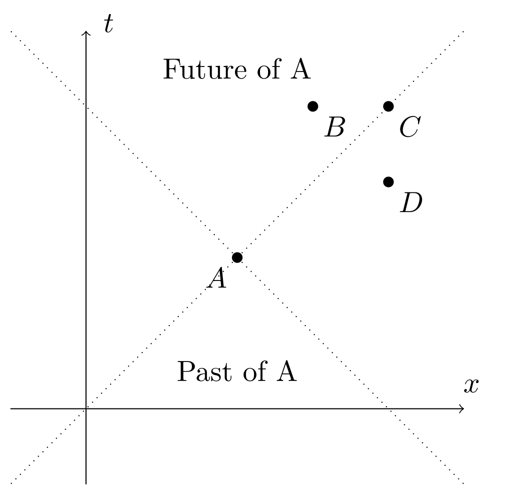

Fig. 2.4: Lightcone structure around the event A. Events B and A are separated by a timelike interval, and B lies in the forward lightcone of A. The events could be causally connected. Events C and A are separated by a null (or lightlike) interval and could be connected by a light signal. Events D and A are separated by a spacelike interval and cannot be causally connected.

2.2.6.1 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.

2.3 Length Contraction and Time Dilation

2.3.1 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 (2.14). Inverting this gives

$$t = \gamma\left(t' + \frac{vx'}{c^2}\right). \tag{2.21}$$

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

$$T = \gamma T' \tag{2.22}$$

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. Note that, throughout this course, we shall consider only *ideal clocks* – clocks that are unaffected by acceleration – for example, the half-life of a decaying particle.

2.3.2 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 left diagram in Fig. 2.5. The two, simultaneous, end points in S' are P' and P'. Their coordinates in S' are $(ct', x') = (0, 0)$ and $(0, L')$ respectively.

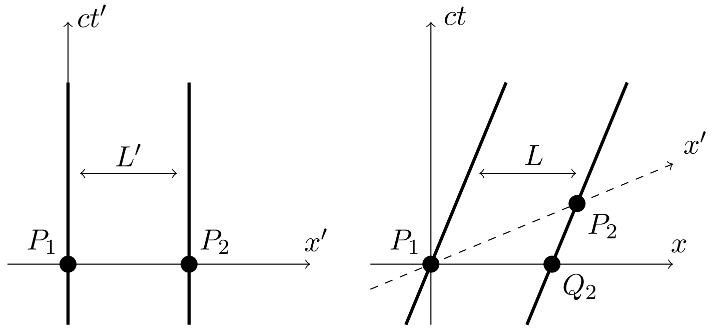

Fig. 2.5: Length contraction.

Now let's look at this in frame S, illustrated in Fig. 2.5. Clearly P_1 sits at $(ct, x) = (0, 0)$. Meanwhile, the Lorentz transformation gives us the coordinate for P_2

$$x = \gamma L', \qquad \text{and,} \qquad t = \frac{\gamma v L'}{c^2}. \tag{2.23}$$

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

$$x = \gamma L' - vt. \tag{2.24}$$

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

$$x = \gamma L' - \frac{\gamma v^2 L'}{c^2} = \frac{L'}{\gamma}. \tag{2.25}$$

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

$$L = \frac{L'}{\gamma} \tag{2.26}$$

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*. The rod suffers no contraction in the y- and z-directions (i.e., perpendicular to its velocity)

But hold on! If we look at the lengths of the rod, marked in the two diagrams in Fig. 2.5, it seems that we have got it the wrong way round: the length in frame S' definitely looks longer than in frame S. This is a trap: lengths in space-time diagrams are not like lengths in the more familiar x-y plane and we must rely on our calculations.[2]

[2]Lengths are shorter the closer the inclination to the 45° of the null cone. This is because instead of the Euclidean norm (Pythagoras), one must use the norm $|(ct, x)| = (c^2t^2 - x^2)^{1/2}$.

2.4 Paths in spacetime

2.4.1 Minkowski spacetime line element

2.4.2 Particle worldlines and proper time

2.4.3 Doppler effect

2.4.4 Addition of velocities

2.5 Acceleration in special relativity

CHAPTER **3**

Manifolds and Coordinates

3.1 Concept of a manifold

3.2 Coordinates

3.2.1 Curves and surfaces

3.2.2 Coordinate transformations

3.2.3 Einstein summation convention

3.3 Local geometry of Riemannian manifolds

3.3.1 The metric

3.3.2 Intrinsic and extrinsic geometry

3.4 Lengths and volumes

3.4.1 Lengths along curves

3.4.2 Volumes of regions

3.4.3 Invariance of the volume element

3.5 Local Cartesian coordinates

3.5.1 Proof of existence of local Cartesian coordinates

3.6 Pseudo-Riemannian manifolds

3.7 Topology of manifolds

Chapter 4

Vector Tensor Algebra

4.1 Scalar and vector fields on manifolds

4.1.1 Scalar fields

4.1.2 Vector fields and tangent spaces

4.1.3 Vectors as differential operators

4.1.4 Dual vector fields

4.2 Tensor fields

4.2.1 Tensor equations

4.2.2 Elementary operations with tensors

4.2.3 Quotient theorem

4.3 Metric tensor

4.3.1 Inverse metric

4.4 Scalar products of vectors revisited

Chapter 5

Vector and Tensor Calculus on Manifolds

5.1 Covariant derivatives

5.1.1 Derivatives of scalar fields

5.1.2 Covariant derivatives of tensor fields

5.1.3 The connection

5.1.3.1 Extension to other tensor fields

5.1.4 The metric connection

5.1.4.1 Other useful properties of the metric connection

5.1.5 5 Relation to local Cartesian coordinat

5.1.6 Divergence, curl and the Laplacian

5.2 Intrinsic derivative of vectors along a curv

5.3 Parallel transpor

5.3.1 Properties of parallel transport

5.4 Geodesic curves

5.4.1 Tangent vectors

5.4.2 Stationary property of non-null geodesics

5.4.3 Relation to parallel transport

5.4.4 Alternative “Lagrangian” procedure

5.4.5 Conserved quantities along geodesics

Chapter **6**

Chapter

Chapter **7**

Chapter

Chapter 8

Chapter

Chapter **9**

Chapter

Chapter **10**

Chapter

Chapter 11

Chapter

Chapter **12**

Chapter

Chapter **13**

Chapter

Appendix A

Code