2 The *k*-calculus

2.1 Model building

Before we start, we should be clear what we are about. The essential activity of mathematical physics, or theoretical physics, is that of **modelling** or **model building**. The activity consists of constructing a mathematical model which we hope in some way captures the essentials of the phenomena we are investigating. I think we should never fail to be surprised that this turns out to be such a productive activity. After all, the first thing you notice about the world we inhabit is that it is an extremely complex place. The fact that so much of this rich structure can be captured by what are, in essence, a set of simple formulae is to me quite astonishing. Just think how simple Newton's universal law of gravitation is; and yet it encompasses a whole spectrum of phenomena from a falling apple to the shape of a globular cluster of stars. As Einstein said, 'The most incomprehensible thing about the world is that it is comprehensible.'

The very success of the activity of modelling has, throughout the history of science, turned out to be counterproductive. Time and again, the successful model has been confused with the ultimate reality, and this in turn has stultified progress. Newtonian theory provides an outstanding example of this. So successful had it been in explaining a wide range of phenomena, that, after more than two centuries of success, the laws had taken on an absolute character. Thus it was that, when at the end of the nineteenth century it was becoming increasingly clear that something was fundamentally wrong with the current theories, there was considerable reluctance to make any fundamental changes to them. Instead, a number of artificial assumptions were made in an attempt to explain the unexpected phenomena. It eventually required the genius of Einstein to overthrow the prejudices of centuries and demonstrate in a number of simple thought experiments that some of the most cherished assumptions of Newtonian theory were untenable. This he did in a number of brilliant papers written in 1905 proposing a theory which has become known today as the **special theory of relativity**.

We should perhaps be discouraged from using words like right or wrong when discussing a physical theory. Remembering that the essential activity is model building, a model should then rather be described as good or bad, depending on how well it describes the phenomena it encompasses. Thus, Newtonian theory is an excellent theory for describing a whole range of phenomena. For example, if one is concerned with describing the motion of a car, then the Newtonian framework is likely to be the appropriate one.

However, it fails to be appropriate if we are interested in very high speeds (comparable with the speed of light) or very intense gravitational fields (such as in the nucleus of a galaxy). To put it another way: together with every theory, there should go its **range of validity**. Thus, to be more precise, we should say that Newtonian theory is an excellent theory within its range of validity. From this point of view, developing our models of the physical world does not involve us in constantly throwing theories out, since they are perceived to be wrong, or unlearning them, but rather it consists more of a process of refinement in order to increase their range of validity. So the moral of this section is that, for all their remarkable success, one must not confuse theoretical models with the ultimate reality they seek to describe.

2.2 Historical background

In 1865, James Clerk Maxwell put forward the theory of electromagnetism. One of the triumphs of the theory was the discovery that light waves are electromagnetic in character. Since all other known wave phenomena required a material medium in which the oscillations were carried, it was postulated that there existed an all-pervading medium, called the 'luminiferous ether', which carried the oscillations of electromagnetism. It was then anticipated that experiments with light would allow the absolute motion of a body through the ether to be detected. Such hopes were upset by the null result of the famous Michelson–Morley experiment (1881) which attempted to measure the velocity of the earth relative to the ether and found it to be undetectably small. In order to explain this null result, two **ad hoc** hypotheses were put forward by Lorentz, Fitzgerald, and Poincaré (1895), namely, the contraction of rigid bodies and the slowing down of clocks when moving through the ether. These effects were contained is some simple formulae called the 'Lorentz transformations'. This would affect every apparatus designed to measure the motion relative to the ether so as to neutralize exactly all expected results. Although this theory was consistent with the observations, it had the philosophical defect that its fundamental assumptions were unverifiable.

In fact, the essence of the special theory of relativity is contained in the Lorentz transformations. However, Einstein was able to derive them from two postulates, the first being called the 'principle of special relativity'—a principle which Poincaré had also suggested independently in 1904—and the second concerning the constancy of the velocity of light. In so doing, he was forced to re-evaluate our ideas of space and time and he demonstrated through a number of simple thought experiments that the source of the limitations of the classical theory lay in the concept of **simultaneity**. Thus, although in a sense Einstein found nothing new in that he rederived the Lorentz transformations, his derivation was physically meaningful and in the process revealed the inadequacy of some of the fundamental assumptions of classical thought. Herein lies his chief contribution.

2.3 Newtonian framework

We start by outlining the Newtonian framework. An **event** intuitively means something happening in a fairly limited region of space and for a short duration in time. Mathematically, we idealize this concept to become a point

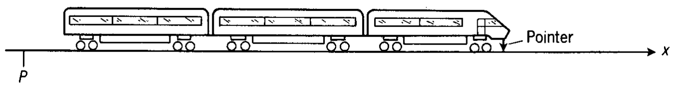

Fig. 2.1 Train travels in straight line.

in space and an instant in time. Everything that happens in the universe is an event or collection of events. Consider a train travelling from one station P to another R, leaving at 10 a.m. and arriving at 11 a.m. We can illustrate this in the following way: for simplicity, let us assume that the motion takes place in a straight line (say along the x-axis (Fig. 2.1); then we can represent the motion by a **space-time diagram** (Fig. 2.2) in which we plot the position of some fixed point on the train, which we represent by a pointer, against time. The curve in the diagram is called the **history** or **world-line** of the pointer. Notice that at Q the train was stationary for a period.

We shall call individuals equipped with a clock and a measuring rod or ruler **observers**. Had we looked out of the train window on our journey at a clock in a passing station then we would have expected it to agree with our watch. One of the central assumptions of the Newtonian framework is that two observers will, once they have synchronized their clocks, always agree about the time of an event, irrespective of their relative motion. This implies that for all observers time is an **absolute** concept. In particular, all observers can agree on an origin of time. In order to fix an event in space, an observer may choose a convenient origin in space together with a set of three Cartesian coordinate axes. We shall refer to an observer's clock, ruler, and coordinate axes as a **frame of reference** (Fig. 2.3). Then an observer is able to **coordinatize** events, that is, determine the time t an event occurs and its relative position (x, y, z).

We have set the stage with space and time; they provide the backcloth, but what is the story about? The stuff which provides the events of the universe is **matter**. For the moment, we shall idealize lumps of matter into objects called **bodies**. If the body has no physical extent, we refer to it as a **particle** or **point mass**. Thus, the role of observers in Newtonian theory is to chart the history of bodies.

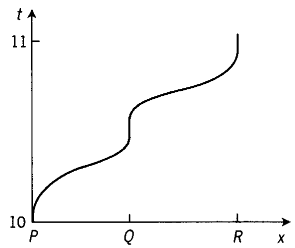

Fig. 2.2 Space-time diagram of pointer.

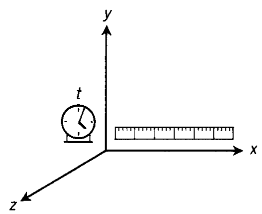

Fig. 2.3 Observer's frame of reference.

2.4 Galilean transformations

Now, relativity theory is concerned with the way different observers see the same phenomena. One can ask: are the laws of physics the same for all observers or are there preferred states of motion, preferred reference systems, and so on? Newtonian theory postulates the existence of preferred frames of reference. This is contained essentially in the first law, which we shall call N1 and state in the following form:

N1: Every body continues in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by forces acting on it.

Thus, there exists a privileged set of bodies, namely those not acted on by forces. The frame of reference of a co-moving observer is called an **inertial** frame (Fig. 2.4). It follows that, once we have found one inertial frame, then all

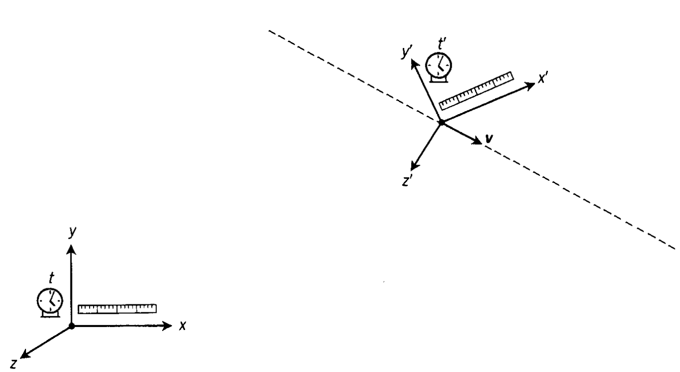

Fig. 2.4 Two observed bodies and their inertial frames.

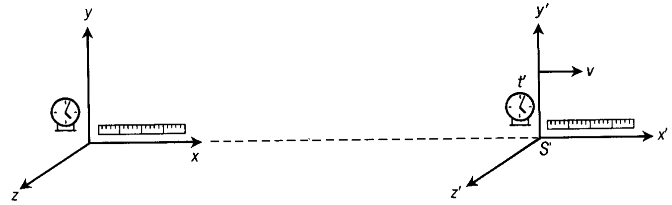

Fig. 2.5 Two frames in standard configuration at time t.

others are at rest or travel with constant velocity relative to it (for otherwise Newton's first law would no longer be true). The transformation which connects one inertial frame with another is called a **Galilean transformation**. To fix ideas, let us consider two inertial frames called S and S' in **standard configuration**, that is, with axes parallel and S' moving along S's positive x-axis with constant velocity (Fig. 2.5). We also assume that the observers synchronize their clocks so that the origins of time are set when the origins of the frames coincide. It follows from Fig. 2.5 that the Galilean transformation connecting the two frames is given by

$$x = x' + vt, \quad y = y', \quad z = z', \quad t = t'. \tag{2.1}$$

The last equation provides a manifestation of the assumption of absolute time in Newtonian theory. Now, Newton's laws hold only in inertial frames. From a mathematical viewpoint, this means that Newton's laws must be **invariant** under a Galilean transformation.

2.5 The principle of special relativity

We begin by stating the relativity principle which underpins Newtonian theory

Restricted principle of special relativity:
All inertial observers are equivalent as far as dynamical experiments are concerned.

This means that, if one inertial observer carries out some dynamical experiments and discovers a physical law, then any other inertial observer performing the same experiments must discover the same law. Put another way, these laws must be invariant under a Galilean transformation. That is to say, if the law involves the coordinates x, y, z, t of an inertial observer S, then the law relative to another observer S' will be the same with x, y, z, t replaced by x', y', z', t', respectively. Many fundamental principles of physics are statements of **impossibility**, and the above statement of the relativity principle is equivalent to the statement of the impossibility of deciding, by performing dynamical experiments, whether a body is absolutely at rest or in uniform motion. In Newtonian theory, we cannot determine the **absolute** position in space of an event, but only its position **relative** to some other event. In exactly the same way, uniform velocity has only a relative significance; we can only talk about the velocity of a body relative to some other. Thus, both position and velocity are **relative** concepts.

Einstein realized that the principle as stated above is empty because there is no such thing as a purely **dynamical** experiment. Even on a very elementary level, any dynamical experiment we think of performing involves observation, i.e. **looking**, and looking is a part of optics, not dynamics. In fact, the more one analyses any one experiment, the more it becomes apparent that practically all the branches of physics are involved in the experiment. Thus, Einstein took the logical step of removing the restriction of dynamics in the principle and took the following as his first postulate.

Postulate I. Principle of special relativity:
All inertial observers are equivalent.

Hence we see that this principle is in no way a contradiction of Newtonian thought, but rather constitutes its logical completion.

2.6 The constancy of the velocity of light

We previously defined an observer in Newtonian theory as someone equipped with a clock and ruler with which to map the events of the universe. However, the approach of the k-calculus is to dispense with the rigid ruler and use radar methods for measuring distances. (What is rigidity anyway? If a moving frame appears non-rigid in another frame, which, if either, is the rigid one?) Thus, an observer measures the distance of an object by sending out a light signal which is reflected off the object and received back by the observer. The **distance** is then simply defined as **half the time difference between emission and reception**. Note that by this method **distances** are measured in intervals of time, like the light year or the light second ($\sim 10^{10}$ cm).

Why use light? The reason is that we know that the velocity of light is independent of many things. Observations from double stars tell us that the velocity of light **in vacuo** is independent of the motion of the sources as well as independent of colour, intensity, etc. For, if we suppose that the velocity of light were dependent on the motion of the source relative to an observer (so that if the source was coming towards us the light would be travelling faster and vice versa) then we would no longer see double stars moving in Keplerian

orbits (circles, ellipses) about each other: their orbits would appear distorted; yet no such distortion is observed. There are many experiments which confirm this assumption. However, these were not known to Einstein in 1905, who adopted the second postulate purely on heuristic grounds. We state the second postulate in the following form.

Postulate II. Constancy of velocity of light:
The velocity of light is the same in all inertial systems.

Or stated another way: there is no overtaking of light by light in empty space. The speed of light is conventionally denoted by c and has the exact numerical value $2.997\,924\,580 \times 10^8$ ms^{-1}, but in this chapter we shall adopt **relativistic units** in which c is taken to be unity (i.e. $c = 1$). Note, in passing, that another reason for using radar methods is that other methods are totally impracticable for large distances. In fact, these days, distances from the Earth to the Moon and Venus can be measured very accurately by bouncing radar signals off them.

2.7 The *k*-factor

For simplicity, we shall begin by working in two dimensions, one spatial dimension and one time dimension. Thus, we consider a system of observers distributed along a straight line, each equipped with a clock and a flashlight. We plot the events they map in a two-dimensional space-time diagram. Let us assume we have two observers, A at rest and B moving away from A with uniform (constant) speed. Then, in a space-time diagram, the world-line of A will be represented by a vertical straight line and the world-line of B by a straight line at an angle to A's, as shown in Fig. 2.6.

A light signal in the diagram will be denoted by a straight line making an angle $\frac{1}{4}\pi$ with the axes, because we are taking the speed of light to be 1. Now, suppose A sends out a series of flashes of light to B, where the interval between the flashes is denoted by T according to A's clock. Then it is plausible to assume that the intervals of reception by B's clock are proportional to T, say kT. Moreover, the quantity k, which we call the ***k*-factor**, is

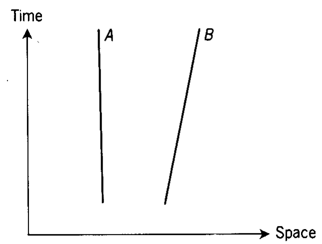

Fig. 2.6 The world-lines of observers A and B.

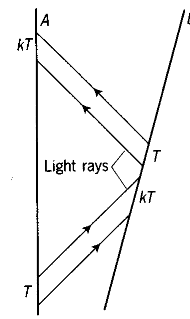

Fig. 2.7 The reciprocal nature of the k-factor.

clearly a characteristic of the motion of B relative to A. We now assume that if A and B are **inertial** observers, then k is a **constant** in time. (In fact, there is a hidden assumption here, since how do we know that B's world-line will be a **straight** line as indicated in the diagram? Strictly speaking, we are assuming that there is a **linear** relationship between the space and time coordinates of A and B.) Then the principle of special relativity requires that the relationship between A and B must be reciprocal, so that, if B emits two signals with a time lapse of T according to B's clock, then A receives them after a time lapse of kT according to A's clock (Fig. 2.7). Note that, from B's point of view, A is moving away from B with the same relative speed.

Observer A assigns coordinates to an event P by bouncing a light signal off it. So that if a light signal is sent out at a time $t = t_1$, and received back at a time $t = t_2$ (Fig. 2.8), then, according to our radar definition of distances, the coordinates of P are given by

$$(t, x) = \left(\tfrac{1}{2}(t_1 + t_2), \tfrac{1}{2}(t_2 - t_1)\right), \tag{2.2}$$

remembering that the velocity of light is 1.

We now use the k-factor to develop the k-calculus.

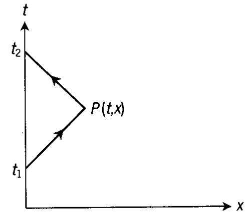

Fig. 2.8 Coordinatizing events.

2.8 Relative speed of two inertial observers

Consider the configuration shown in Fig. 2.9 and assume that A and B synchronize their clocks to zero when they cross at event O. After a time T, A sends a signal to B, which is reflected back at event P. From B's point of view, a light signal is sent to A after a time lapse of kT by B's clock. It follows from the definition of the k-factor that A receives this signal after a time lapse of $k(kT)$. Then, using (2.2) with $t_1 = T$ and $t_2 = k^2T$, we find the coordinates of P according to A's clock are given by

$$(t, x) = \left(\tfrac{1}{2}(k^2 + 1)T, \tfrac{1}{2}(k^2 - 1)T\right). \tag{2.3}$$

Thus, as T varies, this gives the coordinates of the events which constitute B's world-line. Hence, if v is the velocity of B relative to A, we find

$$v = \frac{x}{t} = \frac{k^2 - 1}{k^2 + 1}.$$

Solving for k in terms of v, and noting from the diagram that k must be greater than 1 if the observers are separating, we find

$$\mathbf{k = \left(\frac{1+v}{1-v}\right)^{\frac{1}{2}}}. \tag{2.4}$$

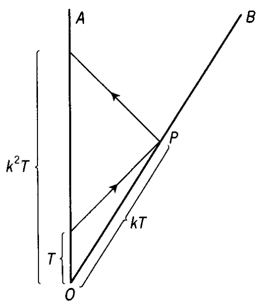

Fig. 2.9 Relating the k-factor to the relative speed of separation.

We shall see in the next chapter that this is the usual relativistic formula for the radial Doppler shift. If B is moving away from A then $k > 1$ which represents a 'red' shift, whereas if B is approaching A then $k < 1$ which represents a 'blue' shift. Note that the transformation $v \to -v$ corresponds to $k \to 1/k$. Moreover,

$$v = 0 \quad \Leftrightarrow \quad k = 1,$$

as we should expect for observers relatively at rest: once they have synchronized their clocks, the synchronization remains (Fig. 2.10).

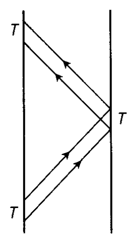

Fig. 2.10 Observers relatively at rest ($k = 1$).

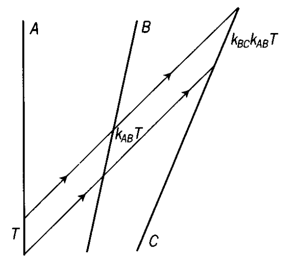

Fig. 2.11 Composition of k-factors.

2.9 Composition law for velocities

Consider the situation in Fig. 2.11, where k_{AB} denotes the k-factor between A and B, with k_{BC} and k_{AC} defined similarly. It follows immediately that

$$k_{AC} = k_{AB}k_{BC}. \tag{2.5}$$

Using (2.4), we find the corresponding **composition law** for velocities:

$$\mathbf{v_{AC} = \frac{v_{AB} + v_{BC}}{1 + v_{AB}v_{BC}}}. \tag{2.6}$$

This formula has been confirmed by Fizeau's experiment in which the speed of light in a moving fluid is measured and turns out not to be simply the sum of the speed of light and the moving fluid but rather obeys the more complicated law (2.6) to higher order. Note that, if v_{AB} and v_{BC} are small compared with the speed of light, i.e.

$$v_{AB} \ll 1, \qquad v_{BC} \ll 1,$$

then we obtain the classical Newtonian formula

$$v_{AC} = v_{AB} + v_{BC}$$

to lowest order. Although the composition law for velocities is not simple, the one for k-factors is, and in special relativity it is the k-factors which are the directly measurable quantities. Note also that, formally, if we substitute $v_{BC} = 1$, representing the speed of a light signal **relative to B**, in (2.6), then the resulting speed of the light signal relative to A is

$$v_{AC} = \frac{v_{AB} + 1}{1 + v_{AB}} = 1,$$

in agreement with the constancy of the velocity of light postulate.

From the composition law, we can show that, if we add two speeds less than the speed of light, then we again obtain a speed less than the speed of light. This does not mean, as is sometimes stated, that nothing can move faster than the speed of light in special relativity, but rather that the speed of light is a border which can not be crossed or even reached. More precisely, special relativity allows for the existence of three classes of particles.

1. Particles that move slower than the speed of light are called **subluminal** particles. They include material particles and elementary particles such as electrons and neutrons.
2. Particles that move with the speed of light are called **luminal** particles. They include the carrier of the electromagnetic field interaction, the photon, and theoretically the carrier of the gravitational field interaction, called the graviton. These are both particles with zero rest mass (see §4.5). It was thought that neutrinos also had zero rest mass, but more recent evidence suggests they may have a tiny mass.
3. Particles that move faster than the speed of light are called **superluminal** particles or **tachyons**. There was some excitement in the 1970s surrounding the possible existence of tachyons, but all attempts to detect them to date have failed. This suggests two likely possibilities: either tachyons do

not exist or, if they do, they do not interact with ordinary matter. This would seem to be just as well, for otherwise they could be used to signal back into the past and so would appear to violate causality. For example, it would be possible theoretically to construct a device which sent out a tachyon at a given time and which would trigger a mechanism in the device to blow it up **before** the tachyon was sent out!

2.10 The relativity of simultaneity

Consider two events P and Q which take place at the same time, according to A, and also at points equal but opposite distances away. A could establish this by sending out and receiving the light rays as shown in Fig. 2.12 (continuous lines). Suppose now that another inertial observer B meets A at the time these events occur **according to** A. B also sends out light rays RQU and SPV to illuminate the events, as shown (dashed lines). By symmetry $RU = SV$ and so these events are equidistant according to B. However, the signal RQ was sent before the signal SP and so B concludes that the event Q took place **well before** P. Hence, events that A judges to be simultaneous, B judges not to be simultaneous. Similarly, A maintains that P, O, and Q occurred simultaneously, whereas B maintains that they occurred in the order Q, then O, and then P.

This relativity of simultaneity lies at the very heart of special relativity and resolves many of the paradoxes that the classical theory gives rise to, such as the Michelson–Morley experiment. Einstein realized the crucial role that simultaneity plays in the theory and gave the following simple thought experiment to illustrate its dependence on the observer. Imagine a train travelling along a straight track with velocity v relative to an observer A on the bank of the track. In the train, B is an observer situated at the centre of one of the carriages. We assume that there are two electrical devices on the track which are the length of the carriage apart and equidistant from A. When the carriage containing B goes over these devices, they fire and activate two light sources situated at each end of the carriage (Fig. 2.13). From the configuration, it is clear that A will judge that the two events, when the light sources first switch on, occur **simultaneously**. However, B is travelling towards the light emanating from light source 2 and away from the light emanating from light source 1. Since the speed of light is a constant, B will see the light from source 2 before seeing the light from source 1, and so will conclude that one light source comes on **before** the other.

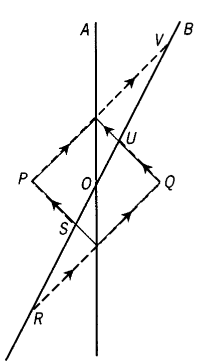

Fig. 2.12 Relativity of simultaneity.

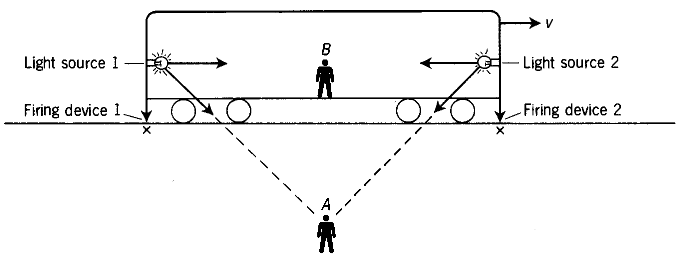

Fig. 2.13 Light signals emanating from the two sources.

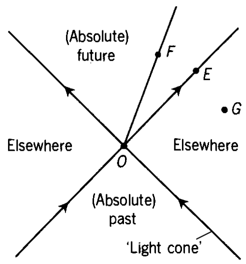

Fig. 2.14 Event relationships in special relativity.

We can now classify event relationships in space and time in the following manner. Consider any event O on A's world-line and the four regions, as shown in Fig. 2.14, given by the light rays ending and commencing at O. Then the event E is on the light ray leaving O and so occurs **after** O. Any other inertial observer agrees on this; that is, no observer sees E illuminated before A sends out the signal from O. The fact that E is illuminated (because A originally sends out a signal at O) **subsequent** to O is a manifestation of **causality**—the event O ultimately causes the event E. Similarly, the event F can be reached by an inertial observer travelling from O with finite speed. Again, all inertial observers agree that F occurs after O. Hence all the events in this region are called the **absolute future of** O. In the same way, any event occurring in the region vertically below takes place in O's **absolute past**. However, the temporal relationship to O of events in the other two regions, called **elsewhere** (or sometimes the **relative past** and **relative future**) will not be something all observers will agree upon. For example, one class of observers will say that G took place after O, another class before, and a third class will say they took place simultaneously. The light rays entering and leaving O constitute what is called the **light cone** or **null cone** at O (the fact that it is a cone will become clearer later when we take all the spatial dimensions into account). Note that the world-line of any inertial observer or material particle passing through O must lie within the light cone at O.

2.11 The clock paradox

Consider three inertial observers as shown in Fig. 2.15, with the relative velocity $v_{AC} = -v_{AB}$. Assume that A and B synchronize their clocks at O and that C's clock is synchronized with B's at P. Let B and C meet after a time T according to B, whereupon they emit a light signal to A. According to the k-calculus, A receives the signal at R after a time kT since meeting B. Remembering that C is moving with the opposite velocity to B (so that $k \rightarrow k^{-1}$), then A will meet C at Q after a subsequent time lapse of $k^{-1}T$. The total time that A records between events O and Q is therefore $(k + k^{-1})T$. For $k \neq 1$, this is **greater** than the combined time intervals $2T$ recorded between events OP and PQ by B and C. But should not the time lapse between the two events agree? This is one form of the so-called **clock paradox**.

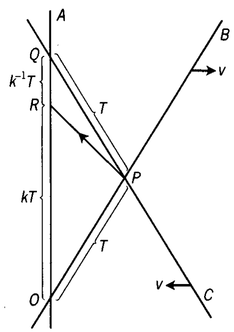

Fig. 2.15 The clock paradox.

However, it is not really a paradox, but rather what it shows is that in relativity time, like distance, is a route-dependent quantity. The point is that the $2T$ measurement is made by **two** inertial observers, not one. Some people have tried to reverse the argument by setting B and C to rest, but this is not possible since they are in relative motion to each other. Another argument says that, when B and C meet, C should take B's clock and use it. But, in this case, the clock would have to be **accelerated** when being transferred to C and so it is no longer inertial. Again, some opponents of special relativity (e.g. H. Dingle) have argued that the short period of acceleration should not make such a difference, but this is analogous to saying that a journey between two points which is straight nearly all the time is about the same length as one which is wholly straight (as shown), which is absurd (Fig. 2.16). The moral is that in special relativity time is a more difficult concept to work with than the absolute time of Newton.

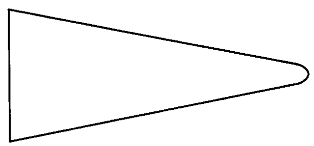

Fig. 2.16 Spatial analogue of clock paradox.

A more subtle point revolves around the implicit assumption that the clocks of A and B are 'good' clocks, i.e. that the seconds of A's clock are the same as those of B's clock. One suggestion is that A has two clocks and adjusts the tick rate until they are the same and then sends one of them to B at a very slow rate of acceleration. The assumption here is that the very slow rate of acceleration will not affect the tick rate of the clock. However, what is there to say that a clock may not be able to somehow add up the small bits of acceleration and so affect its performance. A more satisfactory approach would be for A and B to use identically constructed atomic clocks (which is after all what physicists use today to measure time). The objection then arises that their construction is based on ideas in quantum physics which is, a priori, outside the scope of special relativity. However, this is a manifestation of a point raised earlier, that virtually any real experiment which one can imagine carrying out involves more than one branch of physics. The whole structure is intertwined in a way which cannot easily be separated.

2.12 The Lorentz transformations

We have derived a number of important results in special relativity, which only involve one spatial dimension, by use of the k-calculus. Other results follow essentially from the transformations connecting inertial observers, the famous Lorentz transformations. We shall finally use the k-calculus to derive these transformations.

Let event P have coordinates (t, x) relative to A and (t', x') relative to B (Fig. 2.17). Observer A must send out a light ray at time $t - x$ to illuminate P at time t and also receive the reflected ray back at $t + x$ (check this from (2.2)). The world-line of A is given by $x = 0$, and the origin of A's time coordinate t is arbitrary. Similar remarks apply to B, where we use primed quantities for B's coordinates (t', x'). Assuming A and B synchronize their clocks when they meet, then the k-calculus immediately gives

$$t' - x' = k(t - x), \qquad t + x = k(t' + x'). \tag{2.7}$$

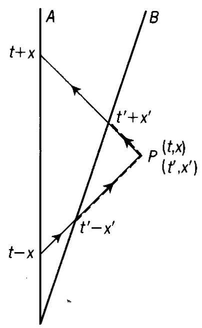

Fig. 2.17 Coordinatization of events by inertial observers.

After some rearrangement, and using equation (2.4), we obtain the so-called **special Lorentz transformation**

$$t' = \frac{t - vx}{(1 - v^2)^{\frac{1}{2}}}, \qquad x' = \frac{x - vt}{(1 - v^2)^{\frac{1}{2}}}. \tag{2.8}$$

This is also referred to as a **boost in the x-direction with speed v**, since it takes one from A's coordinates to B's coordinates and B is moving away from A with speed v. Some simple algebra reveals the result (exercise)

$$t'^2 - x'^2 = t^2 - x^2,$$

showing that the quantity $t^2 - x^2$ is an invariant under a special Lorentz transformation or boost.

To obtain the corresponding formulae in the case of three spatial dimensions we consider Fig. 2.5 with two inertial frames in standard configuration. Now, since by assumption the xz-plane ($y = 0$) of A must coincide with the $x'z'$-plane ($y' = 0$) of B, then the y and y' coordinates must be connected by a transformation of the form

$$y = ny', \tag{2.9}$$

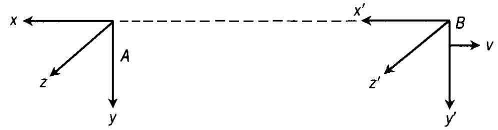

Fig. 2.18 The x- and y-axes reversed in Fig. 2.5.

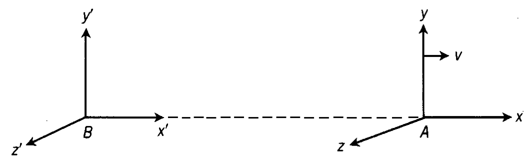

Fig. 2.19 Figure 2.18 from B's point of view.

because

$$y = 0 \quad \Leftrightarrow \quad y' = 0.$$

We now make the assumption that space is **isotropic**, that is, it is the same in any direction. We then reverse the direction of the x- and y-axes of A and B and consider the motion from B's point of view (see Figs. 2.18 and 2.19). Clearly, from B's point of view, the roles of A and B have interchanged. Hence, by symmetry, we must have

$$y' = ny. \tag{2.10}$$

Combining (2.9) and (2.10), we find

$$n^2 = 1 \quad \Rightarrow \quad n = \pm 1.$$

The negative sign can be dismissed since, as $v \to 0$, we must have $y' \to y$, in which case $n = 1$. Hence, we find $y' = y$, and a similar argument for z produces $z' = z$.

2.13 The four-dimensional world view

We now compare the special Lorentz transformation of the last section in relativistic units with the Galilean transformation connecting inertial observers in standard configuration (see Table 2.1). In a Galilean transformation, the absolute time coordinate remains invariant. However, in a

Table 2.1

Galilean transformation	Lorentz transformation
$t' = t$	$t' = \dfrac{t - vx}{(1 - v^2)^{\frac{1}{2}}}$
$x' = x - vt$	$x' = \dfrac{x - vt}{(1 - v^2)^{\frac{1}{2}}}$
$y' = y$	$y' = y$
$z' = z$	$z' = z$

Lorentz transformation, the time and space coordinates get mixed up (note the symmetry in x and t). In the words of Minkowski, '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.'

In the old Newtonian picture, time is split off from three-dimensional Euclidean space. Moreover, since we have an absolute concept of simultaneity, we can consider two simultaneous events with coordinates (t, x_1, y_1, z_1) and (t, x_2, y_2, z_2), and then the square of the Euclidean **distance** between them,

$$\sigma^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2, \tag{2.11}$$

is invariant under a Galilean transformation. In the new special relativity picture, time and space merge together into a four-dimensional continuum called **space-time**. In this picture, the square of the **interval** between any two events (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) is defined by

$$s^2 = (t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2, \tag{2.12}$$

and it is this quantity which is invariant under a Lorentz transformation. Note that we always denote the square of the interval by s^2, but the quantity s is only defined if the right-hand side of (2.12) is non-negative. If we consider two events separated infinitesimally, (t, x, y, z) and $(t + dt, x + dx, y + dy, z + dz)$, then this equation becomes

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \tag{2.13}$$

where all the infinitesimals are squared in (2.13). A four-dimensional space-time continuum in which the above form is invariant is called **Minkowski space-time** and it provides the background geometry for special relativity.

So far, we have only met a special Lorentz transformation which connects two inertial frames in standard configuration. A **full Lorentz transformation** connects two frames in general position (Fig. 2.20). It can be shown that a full Lorentz transformation can be decomposed into an ordinary spatial rotation, followed by a boost, followed by a further ordinary rotation. Physically, the first rotation lines up the x-axis of S with the velocity $\mathbf{v}$ of S'. Then a boost in

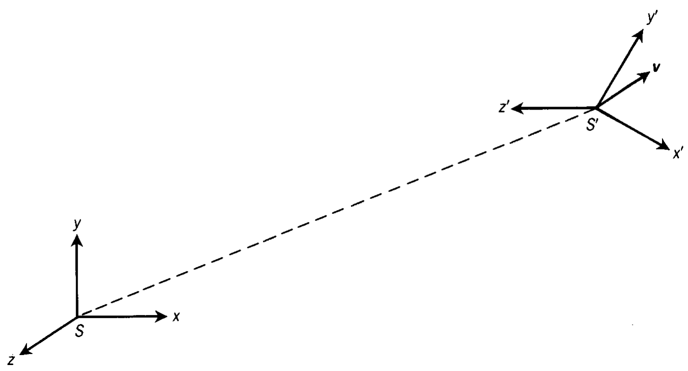

Fig. 2.20 Two frames in general position.

this direction with speed v transforms S to a frame which is at rest relative to S'. A final rotation lines up the coordinate frame with that of S'. The spatial rotations introduce no new physics. The only new physical information arises from the boost and that is why we can, without loss of generality, restrict our attention to a special Lorentz transformation.

Exercises

2.1 (§**2.4**) Write down the Galilean transformation from observer S to observer S', where S' has velocity v_1 relative to S. Find the transformation from S' to S and state in simple terms how the transformations are related. Write down the Galilean transformation from S' to S'', where S'' has velocity v_2 relative to S'. Find the transformation from S to S''. Prove that the Galilean transformations form an Abelian (commutative) group.

2.2 (§**2.7**) Draw the four fundamental k-factor diagrams (see Fig. 2.7) for the cases of two inertial observers A and B approaching and receding with uniform velocity v:

(i) as seen by A;
(ii) as seen by B.

2.3 (§**2.8**) Show that $v \to -v$ corresponds to $k \to k^{-1}$. If $k > 1$ corresponds physically to a red shift of recession, what does $k < 1$ correspond to?

2.4 (§**2.9**) Show that (2.6) follows from (2.5). Use the composition law for velocities to prove that if $0 < v_{AB} < 1$ and $0 < v_{BC} < 1$, then $0 < v_{AC} < 1$.

2.5 (§**2.9**) Establish the fact that if v_{AB} and v_{BC} are small compared with the velocity of light, then the composition law for velocities reduces to the standard additive law of Newtonian theory.

2.6 (§**2.10**) In the event diagram of Fig. 2.14, find a geometrical construction for the world-line of an inertial observer passing through O who considers event G as occurring simultaneously with O. Hence describe the world-lines of inertial observers passing through O who consider G as occurring before or after O.

2.7 (§**2.11**) Draw Fig. 2.15 from B's point of view. Coordinatize the events O, R, and Q with respect to B and find the times between O and R, and R and Q, and compare them with A's timings.

2.8 (§**2.12**) Deduce (2.8) from (2.7). Use (2.7) to deduce directly that

$$t'^2 - x'^2 = t^2 - x^2.$$

Confirm the equality under the transformation formula (2.8).

2.9 (§**2.12**) In S, two events occur at the origin and a distance X along the x-axis simultaneously at $t = 0$. The time interval between the events in S' is T. Show that the spatial distance between the events in S' is $(X^2 + T^2)^{\frac{1}{2}}$ and determine the relative velocity v of the frames in terms of X and T.

2.10 (§**2.13**) Show that the interval between two events (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) defined by

$$s^2 = (t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2$$

is invariant under a special Lorentz transformation. Deduce the Minkowski line element (2.13) for infinitesimally separated events. What does s^2 become if $t_1 = t_2$, and how is it related to the Euclidean distance σ between the two events?

The key attributes of special relativity 3

3.1 Standard derivation of the Lorentz transformations

We start this chapter by deriving again the Lorentz transformations, but this time by using a more standard approach. We shall work in non-relativistic units in which the speed of light is denoted by c. We restrict attention to two inertial observers S and S' in standard configuration. As before, we shall show that the Lorentz transformations follow from the two postulates, namely, the principle of special relativity and the constancy of the velocity of light.

Now, by the first postulate, if the observer S sees a **free** particle, that is, a particle with no forces acting on it, travelling in a straight line with constant velocity, then so will S'. Thus, using vector notation, it follows that under a transformation connecting the two frames

$$\boldsymbol{r} = \boldsymbol{r}_0 + \boldsymbol{u}t \quad \Leftrightarrow \quad \boldsymbol{r}' = \boldsymbol{r}'_0 + \boldsymbol{u}'t'.$$

Since straight lines get mapped into straight lines, it suggests that the transformation between the frames is **linear** and so we shall assume that the transformation from S to S' can be written in matrix form

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = L \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}, \tag{3.1}$$

where L is a 4×4 matrix of quantities which can only depend on the speed of separation v. Using exactly the same argument as we used at the end of §2.12, the assumption that space is isotropic leads to the transformations of y and z being

$$y' = y \quad \text{and} \quad z' = z. \tag{3.2}$$

We next use the second postulate. Let us assume that, when the origins of S and S' are coincident, they zero their clocks, i.e. $t = t' = 0$, and emit a flash of light. Then, according to S, the light flash moves out radially from the origin with speed c. The wave front of light will constitute a sphere. If we define the quantity I by

$$I(t, x, y, z) = x^2 + y^2 + z^2 - c^2t^2,$$

then the events comprising this sphere must satisfy $I = 0$. By the second

postulate, S' must also see the light move out in a spherical wave front with speed c and satisfy

$$I' = x'^2 + y'^2 + z'^2 - c^2t'^2 = 0.$$

Thus it follows that, under a transformation connecting S and S',

$$I = 0 \quad \Leftrightarrow \quad I' = 0, \tag{3.3}$$

and since the transformation is linear by (3.1), we may conclude

$$I = nI', \tag{3.4}$$

where n is a quantity which can only depend on v. Using the same argument as we did in §2.12, we can reverse the role of S and S' and so by the relativity principle we must also have

$$I' = nI. \tag{3.5}$$

Combining the last two equations we find

$$n^2 = 1 \quad \Rightarrow \quad n = \pm 1.$$

In the limit as $v \to 0$, the two frames coincide and $I' \to I$, from which we conclude that we must take $n = 1$.

Substituting $n = 1$ in (3.4), this becomes

$$x^2 + y^2 + z^2 - c^2t^2 = x'^2 + y'^2 + z'^2 - c^2t'^2,$$

and, using (3.2), this reduces to

$$x^2 - c^2t^2 = x'^2 - c^2t'^2. \tag{3.6}$$

We next introduce imaginary time coordinates T and T' defined by

$$T = \text{i}ct, \tag{3.7}$$

$$T' = \text{i}ct', \tag{3.8}$$

in which case equation (3.6) becomes

$$x^2 + T^2 = x'^2 + T'^2.$$

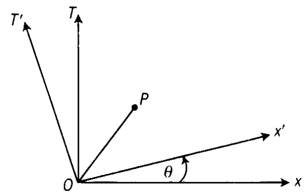

Fig. 3.1 A rotation in (x, T)-space.

In a two-dimensional (x, T)-space, the quantity $x^2 + T^2$ represents the distance of a point P from the origin. This will only remain invariant under a **rotation** in (x, T)-space (Fig. 3.1). If we denote the angle of rotation by θ, then a rotation is given by

$$x' = x\cos\theta + T\sin\theta, \tag{3.9}$$

$$T' = -x\sin\theta + T\cos\theta. \tag{3.10}$$

Now, the origin of S' ($x' = 0$), as seen by S, moves along the positive x-axis of S with speed v and so must satisfy $x = vt$. Thus, we require

$$x' = 0 \quad \Leftrightarrow \quad x = vt \quad \Leftrightarrow \quad x = vT/\text{i}c,$$

using (3.7). Substituting this into (3.9) gives

$$\tan\theta = \text{i}v/c, \tag{3.11}$$

from which we see that the angle θ is imaginary as well. We can obtain an expression for $\cos\theta$, using

$$\cos\theta = \frac{1}{\sec\theta} = \frac{1}{(1 + \tan^2\theta)^{\frac{1}{2}}} = \frac{1}{(1 - v^2/c^2)^{\frac{1}{2}}}.$$

If we use the conventional symbol β for this last expression, i.e.

$$\beta \equiv \frac{1}{(1 - v^2/c^2)^{\frac{1}{2}}},$$

where the symbol $\equiv$ here means 'is defined to be', then (3.9) gives

$$x' = \cos\theta(x + T\tan\theta) = \beta[x + ict(iv/c)] = \beta(x - vt).$$

Similarly, (3.10) gives

$$T' = ict' = \cos\theta(-x\tan\theta + T) = \beta[-x(iv/c) + ict],$$

from which we find

$$t' = \beta(t - vx/c^2).$$

Thus, collecting the results together, we have rederived the special Lorentz transformation or boost (in non-relativistic units):

$$t' = \beta(t - vx/c^2), \quad x' = \beta(x - vt), \quad y' = y, \quad z' = z. \tag{3.12}$$

If we put $c = 1$, this takes the same form as we found in §2.13.

3.2 Mathematical properties of Lorentz transformations

From the results of the last section, we find the following properties of a special Lorentz transformation or boost.

1. Using the imaginary time coordinate T, a boost along the x-axis of speed v is equivalent to an imaginary rotation in (x, T)-space through an angle θ given by $\tan\theta = iv/c$.

2. If we consider v to be very small compared with c, for which we use the notation $v \ll c$, and neglect terms of order v^2/c^2, then we regain a Galilean transformation

$$t' = t, \qquad x' = x - vt, \qquad y' = y, \qquad z' = z.$$

We can obtain this result formally by taking the limit $c \to \infty$ in (3.12).

3. If we solve (3.12) for the unprimed coordinates, we get

$$t = \beta(t' + vx'/c^2), \qquad x = \beta(x' + vt'), \qquad y = y', \qquad z = z'.$$

This can be obtained formally from (3.12) by interchanging primed and unprimed coordinates and replacing v by $-v$. This we should expect from physical reasons, since, if S' moves along the positive x-axis of S with speed v, then S moves along the negative x'-axis of S' with speed v, or, equivalently, S moves along the positive x'-axis of S' with speed $-v$.

4. Special Lorentz transformations form a **group**:
(a) The identity element is given by $v = 0$.
(b) The inverse element is given by $-v$ (as in 3 above).

(c) The product of two boosts with velocities v and v' is another boost with velocity v''. Since v and v' correspond to rotations in (x, T)-space of θ and θ', where

$$\tan\theta = iv/c \quad \text{and} \quad \tan\theta' = iv'/c,$$

then their resultant is a rotation of $\theta'' = \theta + \theta'$, where

$$iv''/c = \tan\theta'' = \tan(\theta + \theta') = \frac{\tan\theta + \tan\theta'}{1 - \tan\theta\tan\theta'},$$

from which we find

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

Compare this with equation (2.6) in relativistic units.
(d) Associativity is left as an exercise.

5. The square of the infinitesimal interval between infinitesimally separated events (see (2.13)),

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2, \tag{3.13}$$

is invariant under a Lorentz transformation.

We now turn to the key physical attributes of Lorentz transformations. Throughout the remaining sections, we shall assume that S and S' are in standard configuration with non-zero relative velocity v.

3.3 Length contraction

Consider a rod fixed in S' with endpoints x'_A and x'_B, as shown in Fig. 3.2. In S, the ends have coordinates x_A and x_B (which, of course, vary in time) given by the Lorentz transformations

$$x'_A = \beta(x_A - vt_A), \qquad x'_B = \beta(x_B - vt_B). \tag{3.14}$$

In order to measure the lengths of the rod according to S, we have to find the x-coordinates of the end points at the same time according to S. If we denote the **rest length**, namely, the length in S', by

$$l_0 = x'_A - x'_B$$

and the length in S at time $t = t_A = t_B$ by

$$l = x_B - x_A,$$

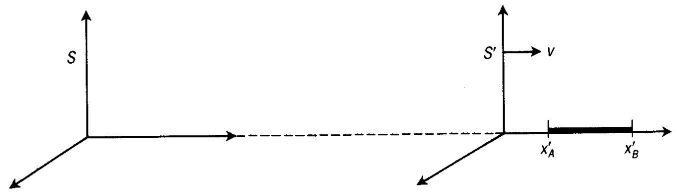

Fig. 3.2 A rod moving with velocity v relative to S.

then, subtracting the formulae in (3.14), we find the result

$$l = \beta^{-1} l_0. \tag{3.15}$$

Since

$$|v| < c \quad \Leftrightarrow \quad \beta > 1 \quad \Leftrightarrow \quad l < l_0,$$

the result shows that the length of a body in the direction of its motion with uniform velocity v is **reduced by a factor $(1 - v^2/c^2)^{\frac{1}{2}}$**. This phenomenon is called **length contraction**. Clearly, the body will have greatest length in its rest frame, in which case it is called the rest length or **proper length**. Note also that the length approaches zero as the velocity approaches the velocity of light.

In an attempt to explain the null result of the Michelson–Morley experiment, Fitzgerald had suggested the apparent shortening of a body in motion relative to the ether. This is rather different from the length contraction of special relativity, which is not to be regarded as illusory but is a very real effect. It is closely connected with the relativity of simultaneity and indeed can be deduced as a direct consequence of it. Unlike the Fitzgerald contraction, the effect is **relative**, i.e. a rod fixed in S appears contracted in S'. Note also that there are no contraction effects in directions transverse to the direction of motion.

3.4 Time dilation

Let a clock fixed at $x' = x'_A$ in S' record two successive events separated by an interval of time T_0 (Fig. 3.3). The successive events in S' are (x'_A, t'_1) and $(x'_A, t'_1 + T_0)$, say. Using the Lorentz transformation, we have in S

$$t_1 = \beta(t'_1 + vx'_A/c^2), \qquad t_2 = \beta(t'_1 + T_0 + vx'_A/c^2).$$

On subtracting, we find the time interval in S defined by

$$T = t_2 - t_1$$

is given by

$$T = \beta T_0. \tag{3.16}$$

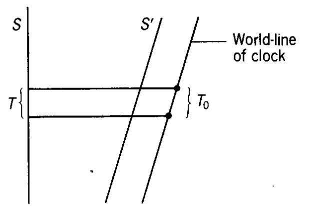

Fig. 3.3 Successive events recorded by a clock fixed in S'.

Thus, **moving clocks go slow by a factor $(1 - v^2/c^2)^{-\frac{1}{2}}$**. This phenomenon is called **time dilation.** The fastest rate of a clock is in its rest frame and is called its **proper rate**. Again, the effect has a reciprocal nature.

Let us now consider an accelerated clock. We define an **ideal clock** to be one unaffected by its acceleration; in other words, its instantaneous rate depends only on its instantaneous speed v, in accordance with the above phenomenon of time dilation. This is often referred to as the **clock hypothesis**. The time recorded by an ideal clock is called the **proper time** τ (Fig. 3.4). Thus, the proper time of an ideal clock between t_0 and t_1 is given by

$$\tau = \int_{t_0}^{t_1} \left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}} dt. \tag{3.17}$$

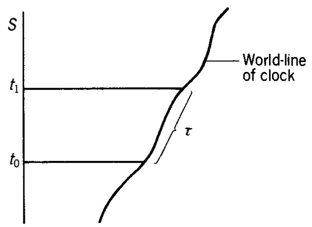

Fig. 3.4 Proper time recorded by an accelerated clock.

The general question of what constitutes a clock or an ideal clock is a non-trivial one. However, an experiment has been performed where an atomic clock was flown round the world and then compared with an identical clock left back on the ground. The travelling clock was found on return to be running slow by precisely the amount predicted by time dilation. Another instance occurs in the study of cosmic rays. Certain mesons reaching us from the top of the Earth's atmosphere are so short-lived that, even had they been travelling at the speed of light, their travel time in the absence of time dilation would exceed their known proper lifetimes by factors of the order of 10. However, these particles are in fact detected at the Earth's surface because their very high velocities keep them young, as it were. Of course, whether or not time dilation affects the human clock, that is, biological ageing, is still an open question. But the fact that we are ultimately made up of atoms, which do appear to suffer time dilation, would suggest that there is no reason by which we should be an exception.

3.5 Transformation of velocities

Consider a particle in motion (Fig. 3.5) with its Cartesian components of velocity being

$$(u_1, u_2, u_3) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \quad \text{in } S$$

and

$$(u'_1, u'_2, u'_3) = \left(\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'}\right) \quad \text{in } S'.$$

Taking differentials of a Lorentz transformation

$$t' = \beta(t - vx/c^2), \qquad x' = \beta(x - vt), \qquad y' = y, \qquad z' = z,$$

we get

$$dt' = \beta(dt - v\,dx/c^2), \qquad dx' = \beta(dx - v\,dt), \qquad dy' = dy, \qquad dz' = dz,$$

and hence

$$u'_1 = \frac{dx'}{dt'} = \frac{\beta(dx - v\,dt)}{\beta(dt - v\,dx/c^2)} = \frac{\dfrac{dx}{dt} - v}{1 - \dfrac{1}{c^2}\left(v\dfrac{dx}{dt}\right)} = \frac{u_1 - v}{1 - u_1 v/c^2}, \tag{3.18}$$

$$u'_2 = \frac{dy'}{dt'} = \frac{dy}{\beta(dt - v\,dx/c^2)} = \frac{\dfrac{dy}{dt}}{\beta\left[1 - \dfrac{1}{c^2}\left(v\dfrac{dx}{dt}\right)\right]} = \frac{u_2}{\beta(1 - u_1 v/c^2)}, \tag{3.19}$$

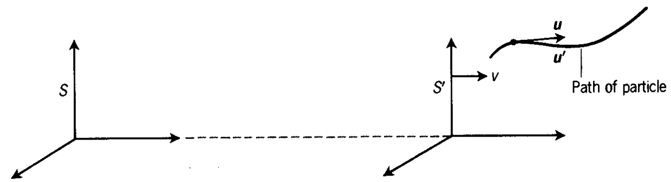

Fig. 3.5 Particle in motion relative to S and S'

$$u_3' = \frac{dz'}{dt'} = \frac{dz}{\beta(dt - v\,dx/c^2)} = \frac{\dfrac{dz}{dt}}{\beta\left[1 - \dfrac{1}{c^2}\left(v\dfrac{dx}{dt}\right)\right]} = \frac{u_3}{\beta(1 - u_1 v/c^2)}. \quad (3.20)$$

Notice that the velocity components u_2 and u_3 transverse to the direction of motion of the frame S' are affected by the transformation. This is due to the time difference in the two frames. To obtain the inverse transformations, simply interchange primes and unprimes and replace v by $-v$.

3.6 Relationship between space-time diagrams of inertial observers

We now show how to relate the space-time diagrams of S and S' (see Fig. 3.6). We start by taking ct and x as the coordinate axes of S, so that a light ray has slope $\frac{1}{4}\pi$ (as in relativistic units). Then, to draw the ct'- and x'-axes of S', we note from the Lorentz transformation equations (3.12)

$$ct' = 0 \quad \Leftrightarrow \quad ct = (v/c)x,$$

that is, the x'-axis, $ct' = 0$, is the straight line $ct = (v/c)x$ with slope $v/c < 1$. Similarly,

$$x' = 0 \quad \Leftrightarrow \quad ct = (c/v)x,$$

that is, the ct'-axis, $x' = 0$, is the straight line $ct = (c/v)x$ with slope $c/v > 1$. The lines parallel to $O(ct')$ are the world-lines of fixed points in S'. The lines parallel to Ox' are the lines connecting points at a fixed time according to S' and are called **lines of simultaneity in** S'. The coordinates of a general event P are $(ct, x) = (OR, OQ)$ relative to S and $(ct', x') = (OV, OU)$ relative to S'. However, the diagram is somewhat misleading because the length scales along the axes are not the same. To relate them, we draw in the hyperbolae

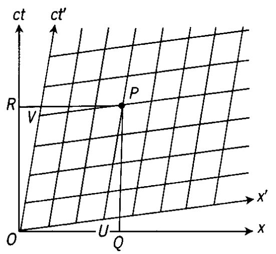

Fig. 3.6 The world-lines in S of the fixed points and simultaneity lines of S'.

$$x^2 - c^2t^2 = x'^2 - c^2t'^2 = \pm 1,$$

as shown in Fig. 3.7. Then, if we first consider the positive sign, setting $ct' = 0$, we get $x' = \pm 1$. It follows that OA is a unit distance on Ox'. Similarly, taking the negative sign and setting $x' = 0$ we get $ct' = \pm 1$ and so OB is the unit measure on Oct'. Then the coordinates of P in the frame S' are given by

$$(ct', x') = \left(\frac{OU}{OA}, \frac{OV}{OB}\right).$$

Note the following properties from Fig. 3.7.

1. A boost can be thought of as a rotation through an imaginary angle in the (x, T)-plane, where T is imaginary time. We have seen that this is equivalent, in the real (x, ct)-plane, to a skewing of the coordinate axes inwards through the same angle. (This was not appreciated by some past opponents of special relativity, who gave some erroneous counter-arguments based on the mistaken idea that a boost could be represented by a real rotation in the (x, ct)-plane.)
2. The hyperbolae are the same for all frames and so we can draw in any number of frames in the same diagram and use the hyperbolae to calibrate them.

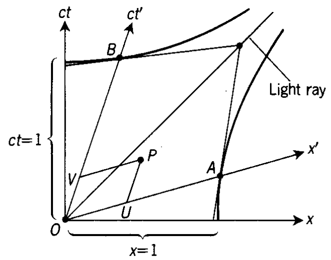

Fig. 3.7 Length scales in S and S'.

3. The length contraction and time dilation effects can be read off directly from the diagram. For example, the world-lines of the endpoints of a unit rod OA in S', namely $x' = 0$ and $x' = 1$, cut Ox in less than unit distance. Similarly world-lines $x = 0$ and $x = 1$ in S cut Ox' inside OE, from which the reciprocal nature of length contraction is evident.
4. Even A has coordinates $(ct', x') = (0, 1)$ relative to S', and hence by a Lorentz transformation coordinates $(ct, x) = (\beta v/c, \beta)$ relative to S. The quantity OA defined by

$$OA = (c^2t^2 + x^2)^{\frac{1}{2}} = \beta(1 + v^2/c^2)^{\frac{1}{2}}$$

is a measure of the calibration factor

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

3.7 Acceleration in special relativity

We start with the inverse transformation of (3.18), namely,

$$u_1 = \frac{u_1' + v}{1 + u_1' v/c^2},$$

from which we find the differential

$$du_1 = \frac{du_1'}{1 + u_1'v/c^2} - \left(\frac{u_1' + v}{(1 + u_1'v/c^2)^2}\right)\frac{v}{c^2}du_1'$$
$$= \frac{1}{\beta^2}\frac{du_1'}{(1 + u_1'v/c^2)^2}.$$

Similarly, from the inverse Lorentz transformation

$$t = \beta(t' + x'v/c^2),$$

we find the differential

$$dt = \beta(dt' + dx'v/c^2) = \beta(1 + u_1'v/c^2)\,dt'.$$

Combining these results, we find that the x-component of the acceleration transforms according to

$$\frac{du_1}{dt} = \frac{1}{\beta^3(1 + u_1'v/c^2)^3}\frac{du_1'}{dt'}. \quad (3.21)$$

Similarly, we find

$$\frac{du_2}{dt} = \frac{1}{\beta^2(1 + u_1'v/c^2)^2}\frac{du_2'}{dt'} - \frac{vu_2'}{c^2\beta^2(1 + u_1'v/c^2)^3}\frac{du_1'}{dt'}, \quad (3.22)$$

$$\frac{du_3}{dt} = \frac{1}{\beta^2(1 + u_1'v/c^2)^3}\frac{du_3'}{dt'} - \frac{vu_3'}{c^2\beta^2(1 + u_1'v/c^2)^3}\frac{du_1'}{dt'}. \quad (3.23)$$

The inverse transformations can be found in the usual way.

It follows from the transformation formulae that acceleration is not an invariant in special relativity. However, it is clear from the formulae that acceleration is an **absolute** quantity, that is, all observers agree whether a body is accelerating or not. Put another way, if the acceleration is zero in one frame, then it is necessarily zero in any other frame. We shall see that this is

Table 3.1

Theory	Position	Velocity	Time	Acceleration
Newtonian	Relative	Relative	Absolute	Absolute
Special relativity	Relative	Relative	Relative	Absolute
General relativity	Relative	Relative	Relative	Relative

no longer the case in general relativity. We summarize the situation in Table 3.1, which indicates why the subject matter of the book is 'relativity' theory.

3.8 Uniform acceleration

The Newtonian definition of a particle moving under uniform acceleration is

$$\frac{du}{dt} = \text{constant.}$$

This turns out to be inappropriate in special relativity since it would imply that $u \to \infty$ as $t \to \infty$, which we know is impossible. We therefore adopt a different definition. Acceleration is said to be **uniform** in special relativity if it has the same value in any **co-moving frame**, that is, at each instant, the acceleration in an inertial frame travelling with the same velocity as the particle has the same value. This is analogous to the idea in Newtonian theory of motion under a constant force. For example, a spaceship whose motor is set at a constant emission rate would be uniformly accelerated in this sense. Taking the velocity of the particle to be $u = u(t)$ relative to an inertial frame S, then at any instant in a co-moving frame S', it follows that $v = u$, the velocity relative to S' is zero, i.e. $u' = 0$, and the acceleration is a constant, a say, i.e. $du'/dt' = a$. Using (3.21), we find

$$\frac{du}{dt} = \frac{1}{\beta^3} a = \left(1 - \frac{u^2}{c^2}\right)^{\frac{3}{2}} a.$$

We can solve this differential equation by separating the variables

$$\frac{du}{(1 - u^2/c^2)^{\frac{3}{2}}} = a\,dt$$

and integrating both sides. Assuming that the particle starts from rest at $t = t_0$, we find

$$\frac{u}{(1 - u^2/c^2)^{\frac{1}{2}}} = a(t - t_0).$$

Solving for u, we get

$$u = \frac{dx}{dt} = \frac{a(t - t_0)}{[1 + a^2(t - t_0)^2/c^2]^{\frac{1}{2}}}.$$

Next, integrating with respect to t, and setting $x = x_0$ at $t = t_0$, produces

$$(x - x_0) = \frac{c}{a}[c^2 + a^2(t - t_0)^2]^{\frac{1}{2}} - \frac{c^2}{a}.$$

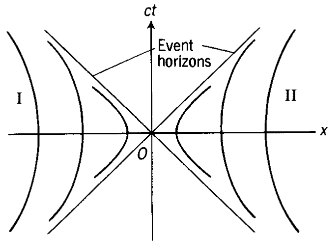

Fig. 3.8 Hyperbolic motions.

This can be rewritten in the form

$$\frac{(x - x_0 + c^2/a)^2}{(c^2/a)^2} - \frac{(ct - ct_0)^2}{(c^2/a)^2} = 1, \tag{3.24}$$

which is a hyperbola in (x, ct)-space. If, in particular, we take $x_0 - c^2/a = t_0 = 0$, then we obtain a family of hyperbolae for different values of a (Fig. 3.8). These world-lines are known as **hyperbolic motions** and, as we shall see in Chapter 23, they have significance in cosmology. It can be shown that the radar distance between the world-lines is a constant. Moreover, consider the regions I and II bounded by the light rays passing through O, and a system of particles undergoing hyperbolic motions as shown in Fig. 3.8 (in some cosmological models, the particles would be galaxies). Then, remembering that light rays emanating from any point in the diagram do so at 45°, no particle in region I can communicate with another particle in region II, and vice versa. The light rays are called **event horizons** and act as barriers beyond which no knowledge can ever be gained. We shall see that event horizons will play an important role later in this book.

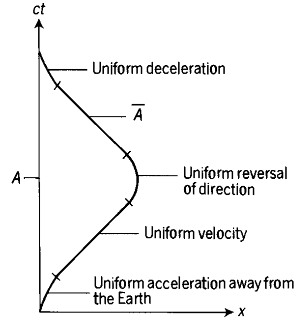

Fig. 3.9 The twin paradox.

3.9 The twin paradox

This is a form of the clock paradox which has caused the most controversy — a controversy which raged on and off for over 50 years. The paradox concerns two twins whom we shall call A and $\bar{A}$. The twin $\bar{A}$ takes off in a spaceship for a return trip to some distant star. The assumption is that $\bar{A}$ is uniformly accelerated to some given velocity which is retained until the star is reached, whereupon the motion is uniformly reversed, as shown in Fig. 3.9. According to A, $\bar{A}$'s clock records slowly on the outward and return journeys and so, on return, $\bar{A}$ will be younger than A. If the periods of acceleration are negligible compared with the periods of uniform velocity, then could not $\bar{A}$ reverse the argument and conclude that it is A who should appear to be the younger? This is the basis of the paradox.

The resolution rests on the fact that the accelerations, however brief, have immediate and finite effects on $\bar{A}$ but not on A who remains inertial throughout. One striking way of seeing this effect is to draw in the simultaneity lines of $\bar{A}$ for the periods of uniform velocity, as in Fig. 3.10. Clearly, the period of uniform reversal has a marked effect on the simultaneity lines. Another way of looking at it is to see the effect that the periods of acceleration have on shortening the length of the journey as viewed by $\bar{A}$. Let us be specific: we assume that the periods of acceleration are T_1, T_2, and T_3, and that, after the period T_1, $\bar{A}$ has attained a speed $v = \sqrt{3}c/2$. Then, from A's viewpoint, during the period T_1, $\bar{A}$ finds that more than half the outward journey has been accomplished, in that $\bar{A}$ has transferred to a frame in which the distance between the Earth and the star is more than halved by length contraction. Thus, $\bar{A}$ accomplishes the outward trip in about half the time which A ascribes to it, and the same applies to the return trip. In fact, we could use the machinery of previous sections to calculate the time elapsed in both the periods of uniform acceleration and uniform velocity, and we would again reach the conclusion that on return $\bar{A}$ will be younger than A. As we have said before, this points out the fact that in special relativity time is a route-dependent quantity. The fact that in Fig. 3.9 $\bar{A}$'s world-line is longer

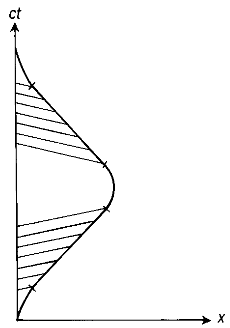

Fig. 3.10 Simultaneity lines of $\bar{A}$ on the outward and return journeys.

than A's, and yet takes **less** time to travel, is connected with the Minkowskian metric

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

and the **negative** signs which appear in it compared with the positive signs occurring in the usual three-dimensional Euclidean metric.

3.10 The Doppler effect

All kinds of waves appear lengthened when the source recedes from the observer: sounds are deepened, light is reddened. Exactly the opposite occurs when the source, instead, approaches the observer. We first of all calculate the **classical** Doppler effect.

Consider a source of light emitting radiation whose wavelength in its rest frame is λ_0. Consider an observer S relative to whose frame the source is in motion with radial velocity u_r. Then, if two successive pulses are emitted at time differing by dt' as measured by S', the distance these pulses have to travel will differ by an amount $u_r dt'$ (see Fig. 3.11). Since the pulses travel with speed c, it follows that they arrive at S with a time difference

$$\Delta t = dt' + u_r dt'/c,$$

giving

$$\Delta t/dt' = 1 + u_r/c.$$

Now, using the fundamental relationship between wavelength and velocity, set

$$\lambda = c\Delta t \quad \text{and} \quad \lambda_0 = c\,dt'.$$

We then obtain the **classical Doppler formula**

$$\lambda/\lambda_0 = 1 + u_r/c. \tag{3.25}$$

Let us now consider the special relativistic formula. Because of time dilation (see Fig. 3.3), the time interval between successive pulses according to S is $\beta dt'$ (Fig. 3.12). Hence, by the same argument, the pulses arrive at S with a time difference

$$\Delta t = \beta\,dt' + u_r \beta\,dt'/c$$

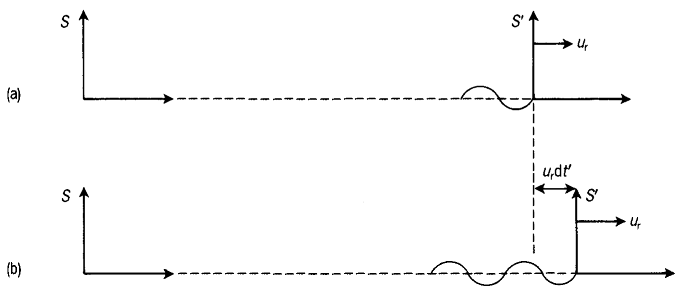

Fig. 3.11 The Doppler effect: (a) first pulse; (b) second pulse.

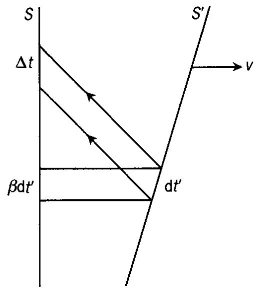

Fig. 3.12 The special relativistic Doppler shift.

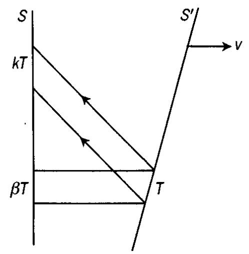

Fig. 3.13 The radial Doppler shift k.

and so this time we find that the **special relativistic Doppler formula** is

$$\frac{\lambda}{\lambda_0} = \frac{1 + u_r/c}{(1 - v^2/c^2)^{\frac{1}{2}}}. \tag{3.26}$$

If the velocity of the source is purely radial, then $u_r = v$ and (3.26) reduces to

$$\frac{\lambda}{\lambda_0} = \left(\frac{1 + v/c}{1 - v/c}\right)^{\frac{1}{2}}. \tag{3.27}$$

This is the **radial Doppler shift**, and, if we set $c = 1$, we obtain (2.4), which is the formula for the ***k*-factor.** Combining Figs. 2.7 and 3.12, the radial Doppler shift is illustrated in Fig. 3.13, where dt' is replaced by T. From equation (3.26), we see that there is also a change in wavelength, even when the radial velocity of the source is zero. For example, if the source is moving in a circle about the origin of S with speed v (as measured by an instantaneous co-moving frame), then the **transverse Doppler shift** is given by

$$\frac{\lambda}{\lambda_0} = \frac{1}{(1 - v^2/c^2)^{\frac{1}{2}}}. \tag{3.28}$$

This is a purely relativistic effect due to the time dilation of the moving source. Experiments with revolving apparatus using the so-called 'Mössbauer effect' have directly confirmed the transverse Doppler shift in full agreement with the relativistic formula, thus providing another striking verification of the phenomenon of time dilation.

Exercises

3.1 (§**3.1**) S and S' are in standard configuration with $v = \alpha c$ ($0 < \alpha < 1$). If a rod at rest in S' makes an angle of 45° with Ox in S and 30° with $O'x$ in S', then find α.

3.2 (§**3.1**) Note from the previous question that perpendicular lines in one frame need not be perpendicular in another frame. This shows that there is no obvious meaning to the phrase 'two inertial frames are parallel', unless their relative velocity is along a common axis, because the axes of either frame need not appear rectangular in the other. Verify that the Lorentz transformation between frames in standard configuration with relative velocity $\boldsymbol{v} = (v, 0, 0)$ may be written in vector form

$$\boldsymbol{r}' = \boldsymbol{r} + \left(\frac{\boldsymbol{v}\cdot\boldsymbol{r}}{v^2}(\beta - 1) - \beta t\right)\boldsymbol{v}, \qquad t' = \beta\left(t - \frac{\boldsymbol{v}\cdot\boldsymbol{r}}{c^2}\right),$$

where $\boldsymbol{r} = (x, y, z)$. The formulae are said to comprise the 'Lorentz transformation without relative rotation'. Justify this name by showing that the formulae remain valid when the frames are not in standard configuration, but are parallel in the sense that the same rotation must be applied to each frame to bring the two into standard configuration (in which case $\boldsymbol{v}$ is the velocity of S' relative to S, but $\boldsymbol{v} = (v, 0, 0)$ no longer applies).

3.3 (§**3.1**) Prove that the first two equations of the special Lorentz transformation can be written in the form

$$ct' = -x\sinh\phi + ct\cosh\phi, \qquad x' = x\cosh\phi - ct\sinh\phi,$$

where the **rapidity** ϕ is defined by $\phi = \tanh^{-1}(v/c)$. Establish also the following version of these equations:

$$ct' + x' = e^{-\phi}(ct + x),$$
$$ct' - x' = e^{\phi}(ct - x),$$
$$e^{2\phi} = (1 + v/c)/(1 - v/c).$$

What relation does ϕ have to θ in equation (3.11)?

3.4 (§3.1) **Aberration** refers to the fact that the direction of travel of a light ray depends on the motion of the observer. Hence, if a telescope observes a star at an inclination θ' to the horizontal, then show that **classically** the 'true' inclination θ of the star is related to θ' by

$$\tan\theta' = \frac{\sin\theta}{\cos\theta + v/c},$$

where v is the velocity of the telescope relative to the star. Show that the corresponding relativistic formula is

$$\tan\theta' = \frac{\sin\theta}{\beta(\cos\theta + v/c)}.$$

3.5 (§3.2) Show that special Lorentz transformations are associative, that is, if $O(v_1)$ represents the transformation from observer S to S', then show that

$$\big(O(v_1)O(v_2)\big)O(v_3) = O(v_1)\big(O(v_2)O(v_3)\big).$$

3.6 (§3.3) An athlete carrying a horizontal 20-ft-long pole runs at a speed v such that $(1 - v^2/c^2)^{-\frac{1}{2}} = 2$ into a 10-ft-long room and closes the door. Explain, in the athlete's frame, in which the room is only 5 ft long, how this is possible. [Hint: no effect travels faster than light.] Show that the minimum length of the room for the performance of this trick is $20/(\sqrt{3} + 2)$ ft. Draw a space-time diagram to indicate what is going on in the rest frame of the athlete.

3.7 (§3.5) A particle has velocity $\boldsymbol{u} = (u_1, u_2, u_3)$ in S and $\boldsymbol{u}' = (u_1', u_2', u_3')$ in S'. Prove from the velocity transformation formulae that

$$c^2 - u^2 = \frac{c^2(c^2 - u'^2)(c^2 - v^2)}{(c^2 + u_1'v)^2}.$$

Deduce that, if the speed of a particle is less than c in any one inertial frame, then it is less than c in every inertial frame.

3.8 (§3.7) Check the transformation formulae for the components of acceleration (3.21)–(3.23). Deduce that acceleration is an absolute quantity in special relativity.

3.9 (§3.8) A particle moves from rest at the origin of a frame S along the x-axis, with constant acceleration α (as measured in an instantaneous rest frame). Show that the equation of motion is

$$\alpha x^2 + 2c^2x - \alpha c^2t^2 = 0,$$

and prove that the light signals emitted after time $t = c/\alpha$ at the origin will never reach the receding particle. A standard clock carried along with the particle is set to read zero at the beginning of the motion and reads τ at time t in S. Using the clock hypothesis, prove the following relationships:

$$\frac{u}{c} = \tanh\frac{\alpha\tau}{c}, \qquad \left(1 - \frac{u^2}{c^2}\right)^{-\frac{1}{2}} = \cosh\frac{\alpha\tau}{c},$$

$$\frac{\alpha t}{c} = \sinh\frac{\alpha\tau}{c}, \qquad x = \frac{c^2}{\alpha}\left(\cosh\frac{\alpha\tau}{c} - 1\right).$$

Show that, if $T^2 \ll c^2/\alpha^2$, then, during an elapsed time T in the inertial system, the particle clock will record approximately the time $T(1 - \alpha^2T^2/6c^2)$.

If $\alpha = 3g$, find the difference in recorded times by the spaceship clock and those of the inertial system

(a) after 1 hour;
(b) after 10 days.

3.10 (§3.9) A space traveller $\bar{A}$ travels through space with uniform acceleration g (to ensure maximum comfort). Find the distance covered in 22 years of $\bar{A}$'s time. [Hint: using years and light years as time and distance units, respectively, then $g = 1.03$]. If on the other hand, $\bar{A}$ describes a straight double path $XYZYX$, with acceleration g on XY and ZY, and deceleration on YZ and YX, for 6 years each, then draw a space-time diagram as seen from the Earth and find by how much the Earth would have aged in 24 years of $\bar{A}$'s time.

3.11 (§3.10) Let the relative velocity between a source of light and an observer be u, and establish the **classical** Doppler formulae for the frequency shift:

source moving, observer at rest: $\nu = \dfrac{\nu_0}{1 + u/c}$,

observer moving, source at rest: $\nu = (1 - u/c)\nu_0$,

where ν_0 is the frequency in the rest frame of the source. What are the corresponding relativistic results?

3.12 (§3.10) How fast would you need to drive towards a red traffic light for the light to appear green? [Hint: $\lambda_{\text{red}} \approx 7 \times 10^{-5}$ cm, $\lambda_{\text{green}} \approx 5 \times 10^{-5}$ cm.]

4 The elements of relativistic mechanics

4.1 Newtonian theory

Before discussing relativistic mechanics, we shall review some basic ideas of Newtonian theory. We have met Newton's first law in § 2.4, and it states that a body not acted upon by a force moves in a straight line with uniform velocity. The second law describes what happens if an object changes its velocity. In this case, something is causing it to change its velocity and this something is called a **force**. For the moment, let us think of a force as something tangible like a push or a pull. Now, we know from experience that it is more difficult to push a more massive body and get it moving than it is to push a less massive body. This resistance of a body to motion, or rather change in motion, is called its **inertia**. To every body, we can ascribe, at least at one particular time, a number measuring its inertia, which (again for the moment) we shall call its **mass** m. If a body is moving with velocity $\boldsymbol{v}$, we define its **linear momentum** $\boldsymbol{p}$ to be the product of its mass and velocity. Then Newton's second law (N2) states that the force acting on a body is equal to the rate of change of linear momentum. The third law (N3) is less general and talks about a restricted class of forces called **internal** forces, namely, forces acting on a body due to the influence of other bodies in a system. The third law states that the force acting on a body due to the influence of the other bodies, the so-called **action**, is equal and opposite to the force acting on these other bodies due to the influence of the first body, the so-called **reaction**. We state the two laws below.

N2: The rate of change of momentum of a body is equal to the force acting on it, and is in the direction of the force.

N3: To every action there is an equal and opposite reaction.

Then, for a body of mass m with a force $\boldsymbol{F}$ acting on it, Newton's second law states

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \frac{\mathrm{d}(m\boldsymbol{v})}{\mathrm{d}t}. \tag{4.1}$$

If, in particular, the mass is a constant, then

$$F = m\frac{\mathrm{d}v}{\mathrm{d}t} = ma \tag{4.2}$$

where a is the acceleration.

Now, strictly speaking, in Newtonian theory, all observable quantities should be defined in terms of their measurement. We have seen how an observer equipped with a frame of reference, ruler, and clock can map the events of the universe, and hence measure such quantities as position, velocity, and acceleration. However, Newton's laws introduce the new concepts of force and mass, and so we should give a prescription for their measurement. Unfortunately, any experiment designed to measure these quantities involves Newton's laws themselves in its interpretation. Thus, Newtonian mechanics has the rather unexpected property that the operational definitions of force and mass which are required to make the laws physically significant are actually contained in the laws themselves.

To make this more precise, let us discuss how we might use the laws to measure the mass of a body. We consider two bodies isolated from all other influences other than the force acting on one due to the influence of the other and vice versa (Fig. 4.1). Since the masses are assumed to be constant, we have, by Newton's second law in the form (4.2),

$$F_1 = m_1 a_1 \quad \text{and} \quad F_2 = m_2 a_2.$$

In addition, by Newton's third law, $F_1 = -F_2$. Hence, we have

$$m_1 a_1 = -m_2 a_2. \tag{4.3}$$

Fig. 4.1 Measuring mass by mutually induced accelerations.

Therefore, if we take one standard body and define it to have **unit** mass, then we can find the mass of the other body, by using (4.3). We can keep doing this with any other body and in this way we can calibrate masses. In fact, this method is commonly used for comparing the masses of elementary particles. Of course, in practice, we cannot remove all other influences, but it may be possible to keep them almost constant and so neglect them.

We have described how to use Newton's laws to measure mass. How do we measure force? One approach is simply to use Newton's second law, work out ma for a body and then read off from the law the force acting on m. This is consistent, although rather circular, especially since a force has independent properties of its own. For example, Newton has provided us with a way for working out the force in the case of gravitation in his **universal law of gravitation** (UG).

UG: Two particles attract each other with a force directly proportional to their masses and inversely proportional to the distance between them.

If we denote the constant of proportionality by G (with value 6.67×10^{-11} in m.k.s. units), the so-called Newtonian constant, then the law is (see Fig. 4.2)

$$F = -G\frac{m_1 m_2}{r^2}\hat{r}, \tag{4.4}$$

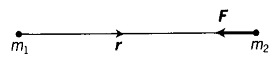

Fig. 4.2 Newton's universal law of gravitation.

where a hat denotes a unit vector. There are other force laws which can be stated separately. Again, another independent property which holds for certain forces is contained in Newton's third law. The standard approach to defining force is to consider it as being **fundamental**, in which case force laws can be stated separately or they can be worked out from other considerations. We postpone a more detailed critique of Newton's laws until Part C of the book.

Special relativity is concerned with the behaviour of material bodies and light rays **in the absence of gravitation**. So we shall also postpone a detailed consideration of gravitation until we discuss general relativity in Part C of the book. However, since we have stated Newton's universal laws of gravitation in (4.4), we should, for completeness, include a statement of Newtonian gravitation for a distribution of matter. A distribution of matter of mass density $\rho = \rho(x, y, z, t)$ gives rise to a gravitational potential ϕ which satisfies **Poisson's equation**

$$\nabla^2\phi = 4\pi G\rho \tag{4.5}$$

at points inside the distribution, where the Laplacian operator ∇^2 is given in Cartesian coordinates by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

At points external to the distribution, this reduces to **Laplace's equation**

$$\nabla^2\phi = 0. \tag{4.6}$$

We assume that the reader is familiar with this background to Newtonian theory.

4.2 Isolated systems of particles in Newtonian mechanics

In this section, we shall, for completeness, derive the conservation of linear momentum in Newtonian mechanics for a system of n particles. Let the ith particle have constant mass m_i and position vector r_i relative to some arbitrary origin. Then the ith particle possesses linear momentum p_i defined by $p_i = m_i\dot{r}_i$, where the dot denotes differentiation with respect to time t. If F_i is the total force on m_i, then, by Newton's second law, we have

$$F_i = \dot{p}_i = m_i\ddot{r}_i. \tag{4.7}$$

The total force F_i on the ith particle can be divided into an external force F_i^{ext} due to any external fields present and to the resultant of the internal forces. We write

$$F_i = F_i^{\text{ext}} + \sum_{j=1}^{n} F_{ij},$$

where F_{ij} is the force or the ith particle due to the jth particle and where, for convenience, we define $\boldsymbol{F}_{ii} = \mathbf{0}$. If we sum over i in (4.7), we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n} \boldsymbol{p}_i = \sum_{i=1}^{n} \frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} = \sum_{i=1}^{n} \boldsymbol{F}_i^{\mathrm{ext}} + \sum_{i,j=1}^{n} \boldsymbol{F}_{ij}.$$

Using Newton's third law, namely, $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$, then the last term is zero and we obtain $\dot{\boldsymbol{P}} = \boldsymbol{F}^{\mathrm{ext}}$, where $\boldsymbol{P} = \sum_{i=1}^{n} \boldsymbol{p}_i$ is termed the **total linear momentum** of the system and $\boldsymbol{F}^{\mathrm{ext}} = \sum_{i=1}^{n} \boldsymbol{F}_i^{\mathrm{ext}}$ is the **total external force** on the system. If, in particular, the system of particles is **isolated**, then

$$\boldsymbol{F}^{\mathrm{ext}} = \mathbf{0} \quad \Rightarrow \quad \boldsymbol{P} = \boldsymbol{c},$$

where $\boldsymbol{c}$ is a constant vector. This leads to the law of the **conservation of linear momentum** of the system, namely,

$$\boldsymbol{P}_{\mathrm{initial}} = \boldsymbol{P}_{\mathrm{final}}. \tag{4.8}$$

4.3 Relativistic mass

The transition from Newtonian to relativistic mechanics is not, in fact, completely straightforward, because it involves at some point or another the introduction of *ad hoc* assumptions about the behaviour of particles in relativistic situations. We shall adopt the approach of trying to keep as close to the non-relativistic definition of energy and momentum as we can. This leads to results which in the end must be confronted with experiment. The ultimate justification of the formulae we shall derive resides in the fact that they have been repeatedly confirmed in numerous laboratory experiments in particle physics. We shall only derive them in a simple case and state that the arguments can be extended to a more general situation.

It would seem plausible that, since length and time measurements are dependent on the observer, then mass should also be an observer-dependent quantity. We thus assume that a particle which is moving with a velocity $\boldsymbol{u}$ relative to an inertial observer has a mass, which we shall term its **relativistic mass**, which is some function of $\boldsymbol{u}$, that is,

$$m = m(\boldsymbol{u}), \tag{4.9}$$

where the problem is to find the explicit dependence of m on $\boldsymbol{u}$. We restrict attention to motion along a straight line and consider the special case of two equal particles colliding **inelastically** (in which case they stick together), and look at the collision from the point of view of two inertial observers S and S' (see Fig. 4.3). Let one of the particles be at rest in the frame S and the other possess a velocity u before they collide. We then assume that they coalesce and that the combined object moves with velocity U. The masses of the two particles are respectively $m(0)$ and $m(u)$ by (4.9). We denote $m(0)$ by m_0 and term it the **rest mass** of the particle. In addition, we denote the mass of the combined object by $M(U)$. If we take S' to be the **centre-of-mass frame**, then it should be clear that, relative to S', the two equal particles collide with equal and opposite speeds, leaving the combined object with mass M_0 at rest. It follows that S' must have velocity U relative to S.

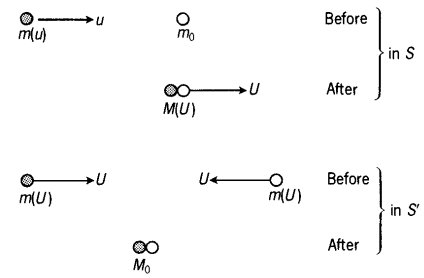

Fig. 4.3 The inelastic collision in the frames S and S'.

We shall assume both conservation of relativistic mass and conservation of linear momentum and see what this leads to. In the frame S, we obtain

$$m(u) + m_0 = M(U), \qquad m(u)u + 0 = M(U)U,$$

from which we get, eliminating $M(U)$,

$$m(u) = m_0\left(\frac{U}{u - U}\right). \tag{4.10}$$

The left-hand particle has a velocity U relative to S', which in turn has a velocity U relative to S. Hence, using the composition of velocities law, we can compose these two velocities and the resultant velocity must be identical with the velocity u of the left-hand particle in S. Thus, by (2.6) in non-relativistic units,

$$u = \frac{2U}{(1 + U^2/c^2)}.$$

Solving for U in terms of u, we obtain the quadratic

$$U^2 - \left(\frac{2c^2}{u}\right)U + c^2 = 0,$$

which has roots

$$U = \frac{c^2}{u} \pm \left[\left(\frac{c^2}{u}\right)^2 - c^2\right]^{\frac{1}{2}} = \frac{c^2}{u}\left[1 \pm \left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}}\right].$$

In the limit $u \to 0$, this must produce a finite result, so we must take the negative sign (check), and, substituting in (4.10), we find finally

$$m(u) = \gamma m_0, \tag{4.11}$$

where

$$\gamma = (1 - u^2/c^2)^{-\frac{1}{2}}. \tag{4.12}$$

This is the basic result which relates the relativistic mass of a moving particle to its rest mass. Note that this is the same in structure as the time dilation formula (3.16), i.e. $T = \beta T_0$, where $\beta = (1 - v^2/c^2)^{-\frac{1}{2}}$, except that time dilation involves the factor β which depends on the velocity v of the frame S' relative to S, whereas γ depends on the velocity u of the particle relative to S. If we plot m against u, we see that relativistic mass increases without bound as u approaches c (Fig. 4.4).

It is possible to extend the above argument to establish (4.11) in more general situations. However, we emphasize that it is not possible to derive the result a priori, but only with the help of extra assumptions. However it is produced, the only real test of the validity of the result is in the experimental arena and here it has been extensively confirmed.

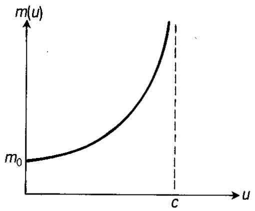

Fig. 4.4 Relativistic mass as a function of velocity.

4.4 Relativistic energy

Let us expand the expression for the relativistic mass, namely,

$$m(u) = \gamma m_0 = m_0(1 - u^2/c^2)^{-\frac{1}{2}},$$

in the case when the velocity u is small compared with the speed of light c. Then we get

$$m(u) = m_0 + \frac{1}{c^2}(\tfrac{1}{2}m_0u^2) + O\left(\frac{u^4}{c^4}\right), \tag{4.13}$$

where the final term stands for all terms of order $(u/c)^4$ and higher. If we multiply both sides by c^2, then, apart from the constant m_0c^2, the right-hand side is to first approximation the classical kinetic energy (k.e.), that is,

$$mc^2 = m_0c^2 + \tfrac{1}{2}m_0u^2 + \cdots \simeq \text{constant} + \text{k.e.} \tag{4.14}$$

We have seen that relativistic mass contains within it the expression for classical kinetic energy. In fact, it can be shown that the conservation of relativistic mass leads to the conservation of kinetic energy in the Newtonian approximation. As a simple example, consider the collision of two particles with rest mass m_0 and $\bar{m}_0$, initial velocities v_1 and $\bar{v}_1$, and final velocities v_2 and $\bar{v}_2$, respectively (Fig. 4.5). Conservation of relativistic mass gives

$$m_0(1 - v_1^2/c^2)^{-\frac{1}{2}} + \bar{m}_0(1 - \bar{v}_1^2/c^2)^{-\frac{1}{2}} = m_0(1 - v_2^2/c^2)^{-\frac{1}{2}} + \bar{m}_0(1 - \bar{v}_2^2/c^2)^{-\frac{1}{2}}. \tag{4.15}$$

If we now assume that $v_1, v_2, \bar{v}_1$, and $\bar{v}_2$ are all small compared with c, then we find (exercise) that the leading terms in the expansion of (4.15) give

$$\tfrac{1}{2}m_0v_1^2 + \tfrac{1}{2}\bar{m}_0\bar{v}_1^2 = \tfrac{1}{2}m_0v_2^2 + \tfrac{1}{2}\bar{m}_0\bar{v}_2^2, \tag{4.16}$$

which is the usual conservation of energy equation. Thus, in this sense, conservation of relativistic mass includes within it conservation of energy. Now, since energy is only defined up to the addition of a constant, the result

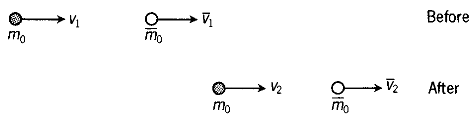

Fig. 4.5 Two colliding particles.

(4.14) suggest that we regard the **energy** E of a particle as given by

$$E = mc^2. \tag{4.17}$$

This is one of the most famous equations in physics. However, it is not just a mathematical relationship between two different quantities, namely energy and mass, but rather states that energy and mass are **equivalent** concepts. Because of the arbitrariness in the actual value of E, a better way of stating the relationship is to say that a change in energy is equal to a change in relativistic mass, namely,

$$\Delta E = \Delta mc^2$$

Using conventional units, c^2 is a large number and indicates that a small change in mass is equivalent to an enormous change in energy. As is well known, this relationship and the deep implications it carries with it for peace and war, have been amply verified. For obvious reasons, the term m_0c^2 is termed the **rest energy** of the particle. Finally, we point out that conservation of linear momentum, using relativistic mass, leads to the usual conservation law in the Newtonian approximation. For example (exercise), the collision problem considered above leads to the usual conservation of linear momentum equation for slow-moving particles:

$$m_0v_1 + \bar{m}_0\bar{v}_1 = m_0v_2 + \bar{m}_0\bar{v}_2. \tag{4.18}$$

Extending these ideas to three spatial dimensions, then a particle moving with velocity $\boldsymbol{u}$ relative to an inertial frame S has relativistic mass m, energy E, and linear momentum $\boldsymbol{p}$ given by

$$m = \gamma m_0, \quad E = mc^2, \quad \boldsymbol{p} = m\boldsymbol{u}. \tag{4.19}$$

Some straightforward algebra (exercise) reveals that

$$(E/c)^2 - p_x^2 - p_y^2 - p_z^2 = (m_0c)^2, \tag{4.20}$$

where m_0c is an invariant, since it is the same for all inertial observers. If we compare this with the invariant (3.13), i.e.

$$(ct)^2 - x^2 - y^2 - z^2 = s^2,$$

then it suggests that the quantities $(E/c, p_x, p_y, p_z)$ transform under a Lorentz transformation in the same way as the quantities (ct, x, y, z). We shall see in Part C that the language of tensors provides a better framework for discussing transformation laws. For the moment, we shall assume that energy and momentum transform in an identical manner and quote the results. Thus, in a frame S' moving in standard configuration with velocity v relative to S, the transformation equations are (see (3.12))

$$E' = \beta(E - vp_x), \quad p'_x = \beta(p_x - vE/c^2), \quad p'_y = p_y, \quad p'_z = p_z. \tag{4.21}$$

The inverse transformations are obtained in the usual way, namely, by interchanging primes and unprimes and replacing v by $-v$, which gives

$$E = \beta(E' + vp'_x), \quad p_x = \beta(p'_x + vE'/c^2), \quad p_y = p'_y, \quad p_z = p'_z. \tag{4.22}$$

If, in particular, we take S' to be the instantaneous rest frame of the particle, then $\boldsymbol{p}' = \boldsymbol{0}$ and $E' = E_0 = m_0c^2$. Substituting in (4.22), we find

$$E = \beta E' = \frac{m_0c^2}{(1 - v^2/c^2)^{\frac{1}{2}}} = mc^2,$$

where $m = m_0(1 - v^2/c^2)^{-\frac{1}{2}}$ and $\boldsymbol{p} = (\beta vE'/c^2, 0, 0) = (mv, 0, 0) = m\boldsymbol{v}$, which are precisely the values of the energy, mass, and momentum arrived at in (4.19) with $\boldsymbol{u}$ replaced by $\boldsymbol{v}$.

4.5 Photons

At the end of the last century, there was considerable conflict between theory and experiment in the investigation of radiation in enclosed volumes. In an attempt to resolve the difficulties, Max Planck proposed that light and other electromagnetic radiation consisted of individual 'packets' of energy, which he called **quanta**. He suggested that the energy E of each quantum was to depend on its frequency ν, and proposed the simple law, called **Planck's hypothesis**,

$$E = h\nu, \tag{4.23}$$

where h is a universal constant known now as **Planck's constant**. The idea of the quantum was developed further by Einstein, especially in attempting to explain the photoelectric effect. The effect is to do with the ejection of electrons from a metal surface by incident light (especially ultraviolet) and is strongly in support of Planck's quantum hypothesis. Nowadays, the quantum theory is well established and applications of it to explain properties of molecules, atoms, and fundamental particles are at the heart of modern physics. Theories of light now give it a dual wave–particle nature. Some properties, such as diffraction and interference, are wavelike in nature, while the photoelectric effect and other cases of the interaction of light and atoms are best described on a particle basis.

The particle description of light consists in treating it as a stream of quanta called **photons**. Using equation (4.19) and substituting in the speed of light, $u = c$, we find

$$m_0 = \gamma^{-1}m = (1 - u^2/c^2)^{\frac{1}{2}}m = 0, \tag{4.24}$$

that is, the rest mass of a photon must be zero! This is not so bizarre as it first seems, since no inertial observer ever sees a photon at rest — its speed is always c — and so the rest mass of a photon is merely a notional quantity. If we let $\hat{\boldsymbol{n}}$ be a unit vector denoting the direction of travel of the photon, then

$$\boldsymbol{p} = (p_x, p_y, p_z) = p\hat{\boldsymbol{n}},$$

and equation (4.20) becomes

$$(E/c)^2 - p^2 = 0.$$

Taking square roots (and remembering c and p are positive), we find that the energy E of a photon is related to the magnitude p of its momentum by

$$E = pc. \tag{4.25}$$

Finally, using the energy–mass relationship $E = mc^2$, we find that the relativistic mass of a photon is non-zero and is given by

$$m = p/c. \tag{4.26}$$

Combining these results with Planck's hypothesis, we obtain the following formulae for the energy E, relativistic mass m, and linear momentum $\boldsymbol{p}$ of the photon:

$$E = h\nu, \quad m = h\nu/c^2, \quad \boldsymbol{p} = (h\nu/c)\hat{\boldsymbol{n}}. \tag{4.27}$$

It is gratifying to discover that special relativity, which was born to reconcile conflicts in the kinematical properties of light and matter, also includes their mechanical properties in a single all-inclusive system.

We finish this section with an argument which shows that Planck's hypothesis can be derived directly within the framework of special relativity. We have already seen in the last chapter that the radial Doppler effect for a moving source is given by (3.27), namely

$$\frac{\lambda}{\lambda_0} = \left(\frac{1 + v/c}{1 - v/c}\right)^{\frac{1}{2}},$$

where λ_0 is the wavelength in the frame of the source and λ is the wavelength in the frame of the observer. We write this result, instead, in terms of frequency, using the fundamental relationships $c = \lambda\nu$ and $c = \lambda_0\nu_0$, to obtain

$$\frac{\nu_0}{\nu} = \left(\frac{1 + v/c}{1 - v/c}\right)^{\frac{1}{2}}. \tag{4.28}$$

Now, suppose that the source emits a light flash of total energy E_0. Let us use the equations (4.22) to find the energy received in the frame of the observer S. Since, recalling Fig. 3.11, the light flash is travelling along the negative x-direction of both frames, the relationship (4.25) leads to the result $p'_x = -E_0/c$, with the other primed components of momentum zero. Substituting in the first equation of (4.22), namely,

$$E = \beta(E' + vp'_x),$$

we get

$$E = \beta(E_0 - vE_0/c) = \frac{E_0(1 - v/c)}{(1 - v^2/c^2)^{\frac{1}{2}}} = E_0\left(\frac{1 - v/c}{1 + v/c}\right)^{\frac{1}{2}},$$

or

$$\frac{E_0}{E} = \left(\frac{1 + v/c}{1 - v/c}\right)^{\frac{1}{2}}. \tag{4.29}$$

Combining this with equation (4.28), we obtain

$$\frac{E_0}{\nu_0} = \frac{E}{\nu}.$$

Since this relationship holds for **any** pair of inertial observers, it follows that the ratio must be a universal constant, which we call h. Thus, we have derived Planck's hypothesis $E = h\nu$.

We leave our considerations of special relativity at this point and turn our attention to the formalism of tensors. This will enable us to reformulate special relativity in a way which will aid our transition to general relativity, that is, to a theory of gravitation consistent with special relativity.

Exercises

4.1 (§**4.1**) Discuss the possibility of using force rather than mass as the basic quantity, taking, for example, a standard weight at a given latitude as the unit of force. How should one then define and measure the mass of a body?

4.2 (§**4.3**) Show that, in the inelastic collision considered in §4.3, the rest mass of the combined object is greater than the sum of the original rest masses. Where does this increase derive from?

4.3 (§**4.3**) A particle of rest mass $\bar{m}_0$ and speed u strikes a stationary particle of rest mass m_0. If the collision is perfectly inelastic, then find the rest mass of the composite particle.

4.4 (§**4.4**) (i) Establish the transition from equation (4.15) to (4.16).
(ii) Establish the Newtonian approximation equation (4.18).

4.5 (§**4.4**) Show that (4.19) leads to (4.20). Deduce (4.21).

4.6 (§**4.4**) Newton's second law for a particle of relativistic mass m is

$$\boldsymbol{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{u}).$$

Define the work done $\mathrm{d}E$ in moving the particle from $\boldsymbol{r}$ to $\boldsymbol{r} + \mathrm{d}\boldsymbol{r}$. Show that the rate of doing work is given by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}(m\boldsymbol{u})}{\mathrm{d}t}\cdot\boldsymbol{u}.$$

Use the definition of relativistic mass to obtain the result

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{m_0}{(1-u^2/c^2)^{3/2}}u\frac{\mathrm{d}u}{\mathrm{d}t} \qquad \left[\text{Hint: } \boldsymbol{u}\cdot\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = u\frac{\mathrm{d}u}{\mathrm{d}t}\right].$$

Express this last result in terms of $\mathrm{d}m/\mathrm{d}t$ and integrate to obtain

$$E = mc^2 + \text{constant}.$$

4.7 (§**4.4**) Two particles whose rest masses are m_1 and m_2 move along a straight line with velocities u_1 and u_2, measured in the same direction. They collide inelastically to form a new particle. Show that the rest mass and velocity of the new particle are m_3 and u_3, respectively, where

$$m_3^2 = m_1^2 + m_2^2 + 2m_1m_2\gamma_1\gamma_2(1 - u_1u_2/c^2),$$

$$u_3 = \frac{m_1\gamma_1u_1 + m_2\gamma_2u_2}{m_1\gamma_1 + m_2\gamma_2},$$

with

$$\gamma_1 = (1 - u_1^2/c^2)^{-\frac{1}{2}}, \qquad \gamma_2 = (1 - u_2^2/c^2)^{-\frac{1}{2}}.$$

4.8 (§**4.4**) A particle of rest mass m_0, energy e_0, and momentum p_0 suffers a head on elastic collision (i.e. masses of particles unaltered) with a stationary mass M. In the collision, M is knocked straight forward, with energy E and momentum P, leaving the first particle with energy e and p. Prove that

$$P = \frac{2p_0M(e_0 + Mc^2)}{2Me_0 + M^2c^2 + m_0^2c^2}$$

and

$$p = \frac{p_0(m^2c^2 - M^2c^2)}{2Me_0 + M^2c^2 + m_0^2c^2}.$$

What do these formulae become in the classical limit?

4.9 (§**4.4**) Assume that the formulae (4.19) hold for a tachyon, which travels with speed $v > c$. Taking the energy to be a measurable quantity, then deduce that the rest mass of a tachyon is imaginary and define the real quantity μ_0 by $m_0 = \mathrm{i}\mu_0$.

If the tachyon moves along the x-axis and if we assume that the x-component of the momentum is a real positive quantity, then deduce

$$m = \frac{v}{|v|}\alpha\mu_0, \qquad p = \mu_0|v|\alpha, \qquad E = mc^2,$$

where $\alpha = (v^2/c^2 - 1)^{-\frac{1}{2}}$.

Plot E/m_0c^2 against v/c for both tachyons and subluminal particles.

4.10 (§**4.5**) Two light rays in the (x, y)-plane of an inertial observer, making angles θ and $-\theta$, respectively, with the positive x axis, collide at the origin. What is the velocity v of