1

The spacetime of special relativity

We begin our discussion of the relativistic theory of gravity by reviewing some basic notions underlying the Newtonian and special-relativistic viewpoints of space and time. In order to specify an *event* uniquely, we must assign it three spatial coordinates and one time coordinate, defined with respect to some frame of reference. For the moment, let us define such a system S by using a set of three mutually orthogonal Cartesian axes, which gives us spatial coordinates x, y and z, and an associated system of synchronised clocks at rest in the system, which gives us a time coordinate t. The four coordinates (t, x, y, z) thus label events in space and time.

1.1 Inertial frames and the principle of relativity

Clearly, one is free to label events not only with respect to a frame S but also with respect to any other frame S', which may be oriented and/or moving with respect to S in an arbitrary manner. Nevertheless, there exists a class of preferred reference systems called *inertial frames*, defined as those in which Newton's first law holds, so that a free particle is at rest or moves with constant velocity, i.e. in a straight line with fixed speed. In Cartesian coordinates this means that

$$\boxed{\frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0.}$$

It follows that, *in the absence of gravity*, if S and S' are two inertial frames then S' can differ from S only by (i) a translation, and/or (ii) a rotation and/or (iii) a motion of one frame with respect to the other at a constant velocity (for otherwise Newton's first law would no longer be true). The concept of inertial frames is fundamental to the *principle of relativity*, which states that *the laws of physics take the same form in every inertial frame*. No exception has ever been found to this general principle, and it applies equally well in both Newtonian theory and special relativity.

The Newtonian and special-relativistic descriptions differ in how the coordinates of an event P in two inertial frames are related. Let us consider two Cartesian inertial frames S and S' in *standard configuration*, where S' is moving along the x-axis of S at a constant speed v and the axes of S and S' coincide at $t = t' = 0$ (see Figure 1.1). It is clear that the (primed) coordinates of an event P with respect to S' are related to the (unprimed) coordinates in S via a *linear* transformation[1] of the form

$$\begin{aligned} t' &= At + Bx, \\ x' &= Dt + Ex, \\ y' &= y, \\ z' &= z. \end{aligned}$$

Moreover, since we require that $x' = 0$ corresponds to $x = vt$ and that $x = 0$ corresponds to $x' = -vt'$, we find immediately that $D = -Ev$ and $D = -Av$, so that $A = E$. Thus we must have

$$\begin{aligned} t' &= At + Bx, \\ x' &= A(x - vt), \\ y' &= y, \\ z' &= z. \end{aligned} \tag{1.1}$$

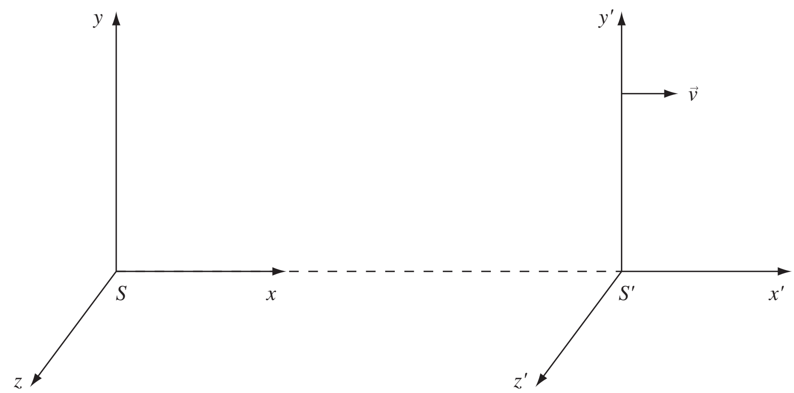

Figure 1.1 Two inertial frames S and S' in standard configuration (the origins of S and S' coincide at $t = t' = 0$).

[1] We will prove this in Chapter 5.

1.2 Newtonian geometry of space and time

Newtonian theory rests on the assumption that there exists an absolute time, which is the same for every observer, so that $t' = t$. Under this assumption $A = 1$ and $B = 0$, and we obtain the *Galilean transformation* relating the coordinates of an event P in the two Cartesian inertial frames S and S':

$$\boxed{\begin{aligned} t' &= t, \\ x' &= x - vt, \\ y' &= y, \\ z' &= z. \end{aligned}} \tag{1.2}$$

By symmetry, the expressions for the unprimed coordinates in terms of the primed ones have the same form but with v replaced by $-v$.

The first equation in (1.2) is clearly valid for any two inertial frames S and S' and shows that the time coordinate of an event P is the same in all inertial frames. The second equation leads to the 'common sense' notion of the addition of velocities. If a particle is moving in the x-direction at a speed u in S then its speed in S' is given by

$$u'_x = \frac{dx'}{dt'} = \frac{dx'}{dt} = \frac{dx}{dt} - v = u_x - v.$$

Differentiating again shows that the acceleration of a particle is the same in both S and S', i.e. $du'_x/dt' = du_x/dt$.

If we consider two events A and B that have coordinates (t_A, x_A, y_A, z_A) and (t_B, x_B, y_B, z_B) respectively, it is straightforward to show that both the time difference $\Delta t = t_B - t_A$ and the quantity

$$\boxed{\Delta r^2 = \Delta x^2 + \Delta y^2 + \Delta z^2}$$

are *separately invariant* under any Galilean transformation. This leads us to consider space and time as separate entities. Moreover, the invariance of Δr^2 suggests that it is a geometric property of space itself. Of course, we recognise Δr^2 as the square of the distance between the events in a three-dimensional Euclidean space. This defines the *geometry* of space and time in the Newtonian picture.

1.3 The spacetime geometry of special relativity

In special relativity, Einstein abandoned the postulate of an absolute time and replaced it by the postulate that *the speed of light c is the same in all inertial* *frames*.[2] By applying this new postulate, together with the principle of relativity, we may obtain the *Lorentz transformations* connecting the coordinates of an event P in two different Cartesian inertial frames S and S'.

Let us again consider S and S' to be in standard configuration (see Figure 1.1), and consider a photon emitted from the (coincident) origins of S and S' at $t = t' = 0$ and travelling in an arbitrary direction. Subsequently the space and time coordinates of the photon in each frame must satisfy

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

Substituting the relations (1.1) into this expression and solving for the constants A and B, we obtain

$$\boxed{\begin{aligned} ct' &= \gamma(ct - \beta x), \\ x' &= \gamma(x - \beta ct), \\ y' &= y, \\ z' &= z, \end{aligned}} \tag{1.3}$$

where $\beta = v/c$ and $\gamma = (1-\beta^2)^{-1/2}$. This Lorentz transformation, also known as a *boost* in the x-direction, reduces to the Galilean transformation (1.2) when $\beta \ll 1$. Once again, symmetry demands that the unprimed coordinates are given in terms of the primed coordinates by an analogous transformation in which v is replaced by $-v$.

From the equations (1.3), we see that the time and space coordinates are in general mixed by a Lorentz transformation (note, in particular, the symmetry between ct and x). Moreover, as we shall see shortly, if we consider two events A and B with coordinates (t_A, x_A, y_A, z_A) and (t_B, x_B, y_B, z_B) in S, it is straightforward to show that the *interval* (squared)

$$\boxed{\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2} \tag{1.4}$$

is *invariant* under any Lorentz transformation. As advocated by Minkowski, these observations lead us to consider space and time as united in a four-dimensional continuum called *spacetime*, whose *geometry* is characterised by (1.4). We note that the spacetime of special relativity is non-Euclidean, because of the minus signs in (1.4), and is often called the *pseudo-Euclidean* or *Minkowski geometry*. Nevertheless, for any fixed value of t the spatial part of the geometry remains Euclidean.

[2] The reasoning behind Einstein's proposal is discussed in Appendix 1A.

We have arrived at the familiar viewpoint (to a physicist!) where the physical world is modelled as a four-dimensional spacetime continuum that possesses the Minkowski geometry characterised by (1.4). Indeed, many ideas in special relativity are most simply explained by adopting a four-dimensional point of view.

1.4 Lorentz transformations as four-dimensional 'rotations'

Adopting a particular (Cartesian) inertial frame S corresponds to labelling events in the Minkowski spacetime with a given set of coordinates (t, x, y, z). If we choose instead to describe the world with respect to a different Cartesian inertial frame S' then this corresponds simply to relabelling events in the Minkowski spacetime with a new set of coordinates (t', x', y', z'); the primed and unprimed coordinates are related by the appropriate Lorentz transformation. Thus, describing physics in terms of different inertial frames is equivalent to performing a *coordinate transformation* on the Minkowski spacetime.

Consider, for example, the case where S' is related to S via a spatial rotation through an angle θ about the x-axis. In this case, we have

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

Clearly the inverse transform is obtained on replacing θ by $-\theta$.

The close similarity between the 'boost' (1.3) and an ordinary spatial rotation can be highlighted by introducing the *rapidity* parameter

$$\psi = \tanh^{-1}\beta.$$

As β varies from zero to unity, ψ ranges from 0 to ∞. We also note that $\gamma = \cosh\psi$ and $\gamma\beta = \sinh\psi$. If two inertial frames S and S' are in standard configuration, we therefore have

$$\boxed{\begin{aligned} ct' &= ct\cosh\psi - x\sinh\psi, \\ x' &= -ct\sinh\psi + x\cosh\psi, \\ y' &= y, \\ z' &= z. \end{aligned}} \tag{1.5}$$

This has essentially the same form as a spatial rotation, but with hyperbolic functions replacing trigonometric ones. Once again the inverse transformation is obtained on replacing ψ by $-\psi$.

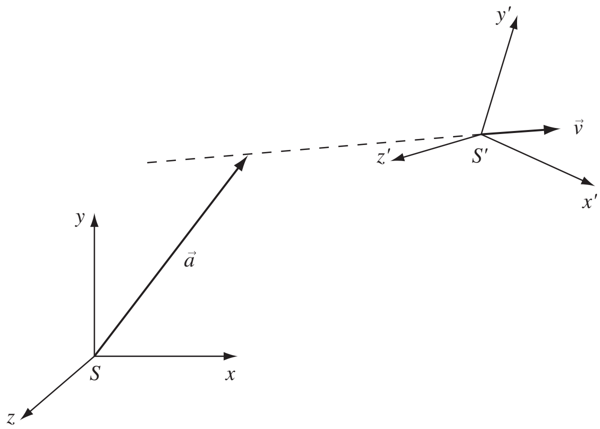

Figure 1.2 Two inertial frames S and S' in general configuration. The broken line shown the trajectory of the origin of S'.

In general, S' is moving with a constant velocity $\vec{v}$ with respect to S in an arbitrary direction[3] and the axes of S' are rotated with respect to those of S. Moreover, at $t = t' = 0$ the origins of S and S' need not be coincident and may be separated by a vector displacement $\vec{a}$, as measured in S (see Figure 1.2).[4] The corresponding transformation connecting the two inertial frames is most easily found by decomposing the transformation into a displacement, followed by a spatial rotation, followed by a boost, followed by a further spatial rotation. Physically, the displacement makes the origins of S and S' coincident at $t = t' = 0$, and the first rotation lines up the x-axis of S with the velocity $\vec{v}$ of S'. Then a boost in this direction with speed v transforms S into a frame that is at rest with respect to S'. A final rotation lines up the coordinate frame with that of S'. The displacement and spatial rotations introduce no new physics, and the only special-relativistic consideration concerns the boost. Thus, without loss of generality, we can restrict our attention to inertial frames S and S' that are in standard configuration, for which the Lorentz transformation is given by (1.3) or (1.5).

1.5 The interval and the lightcone

If we consider two events A and B having coordinates (t'_A, x'_A, y'_A, z'_A) and (t'_B, x'_B, y'_B, z'_B) in S', then, from (1.5), the interval between the events is given by

[3] Throughout this book, the notation $\vec{v}$ is used specifically to denote three-dimensional vectors, whereas $\boldsymbol{v}$ denotes a general vector, which is most often a 4-vector.

[4] If $\vec{a} = \vec{0}$ then the Lorentz transformation connecting the two inertial frames is called *homogeneous*, while if $\vec{a} \neq \vec{0}$ it is called *inhomogeneous*. Inhomogeneous transformations are often referred to as *Poincaré transformations*, in which case homogeneous transformations are referred to simply as Lorentz transformations.

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

Thus the interval is invariant under the boost (1.5) and, from the above discussion, we may infer that Δs^2 is in fact invariant under *any* Poincaré transformation. This suggests that the interval is an underlying geometrical property of the spacetime itself, i.e. an invariant 'distance' between events in spacetime. It also follows that the sign of Δs^2 is defined invariantly, as follows:

for $\Delta s^2 > 0$, the interval is timelike;
for $\Delta s^2 = 0$, the interval is null or lightlike;
for $\Delta s^2 < 0$, the interval is spacelike.

This embodies the standard lightcone structure shown in Figure 1.3. Events A and B are separated by a timelike interval, A and C by a lightlike (or null) interval and

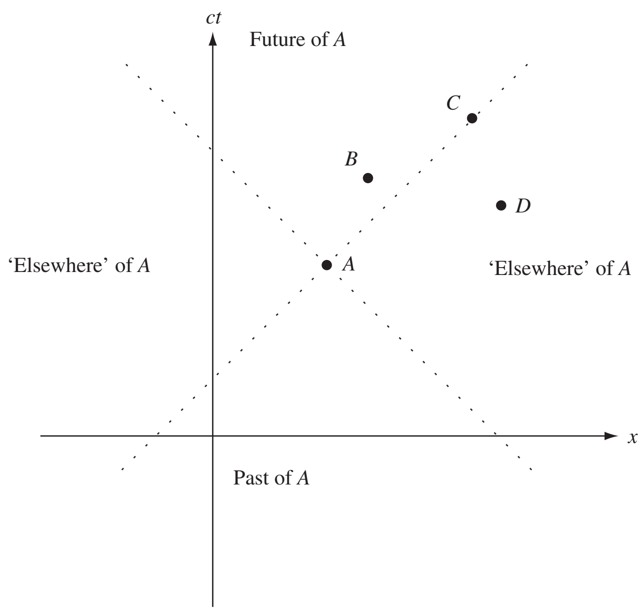

Figure 1.3 Spacetime diagram illustrating the lightcone of an event A (the y- and z- axes have been suppressed). Events A and B are separated by a timelike interval, A and C by a lightlike (or null) interval and A and D by a spacelike interval.

A and D by a spacelike interval. The geometrical distinction between timelike and spacelike intervals corresponds to a physical distinction: if the interval is timelike then we can find an inertial frame in which the events occur at the same spatial coordinates and if the interval is spacelike then we can find an inertial frame in which the events occur at the same time coordinate. This becomes obvious when we consider the spacetime diagram of a Lorentz transformation; we shall do this next.

1.6 Spacetime diagrams

Figure 1.3 is an example of a *spacetime diagram*. Such diagrams are extremely useful in illustrating directly many special-relativistic effects, in particular coordinate transformations on the Minkowski spacetime between different inertial frames. The spacetime diagram in Figure 1.4 shows the change of coordinates of an event A corresponding to the standard-configuration Lorentz transformation (1.5). The x'-axis is simply the line $t' = 0$ and the t'-axis is the line $x' = 0$. From the Lorentz-boost transformation (1.3) we see that the angle between the x- and x'- axes is the same as that between the t- and t'- axes and has the value

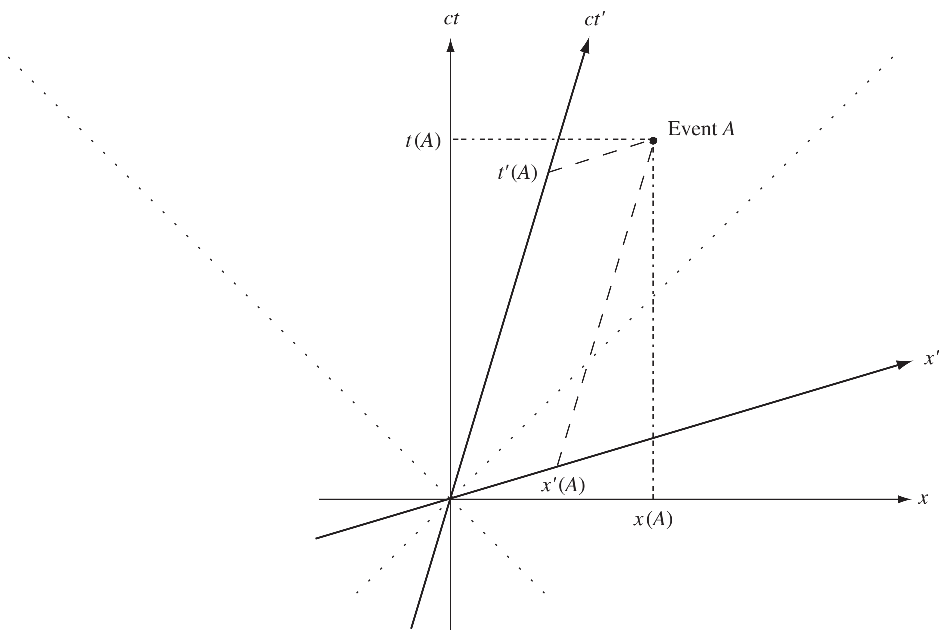

Figure 1.4 Spacetime diagram illustrating the coordinate transformation between two inertial frames S and S' in standard configuration (the y- and z-axes have been suppressed). The worldlines of the origins of S and S' are the axes ct and ct' respectively.

$\tan^{-1}(v/c)$. Moreover, we note that the t- and t'- axes are also the *worldlines* of the origins of S and S' respectively.

It is important to realise that the coordinates of the event A in the frame S' are *not* obtained by extending perpendiculars from A to the x'- and t'- axes. Since the x'-axis is simply the line $t' = 0$, it follows that lines of simultaneity in S' are parallel to the x'-axis. Similarly, lines of constant x' are parallel to the t'-axis. The same reasoning is equally valid for obtaining the coordinates of A in the frame S but, since the x- and t- axes are drawn as orthogonal in the diagram, this is equivalent simply to extending perpendiculars from A to the x- and t- axes in the more familiar manner.

The concept of simultaneity is simply illustrated using a spacetime diagram. For example, in Figure 1.5 we replot the events in Figure 1.3, together with the x'- and t'- axes corresponding to a Lorentz boost in standard configuration at some velocity v. We see that the events A and D, which are separated by a spacelike interval, lie on a line of constant t' and so are *simultaneous* in S'. Evidently, A and D are not simultaneous in S; D occurs at a later time than A. In a similar way, it is straightforward to find a standard-configuration Lorentz boost such that the events A and B, which are separated by a timelike interval, lie on a line of constant x' and hence occur at the same spatial location in S'.

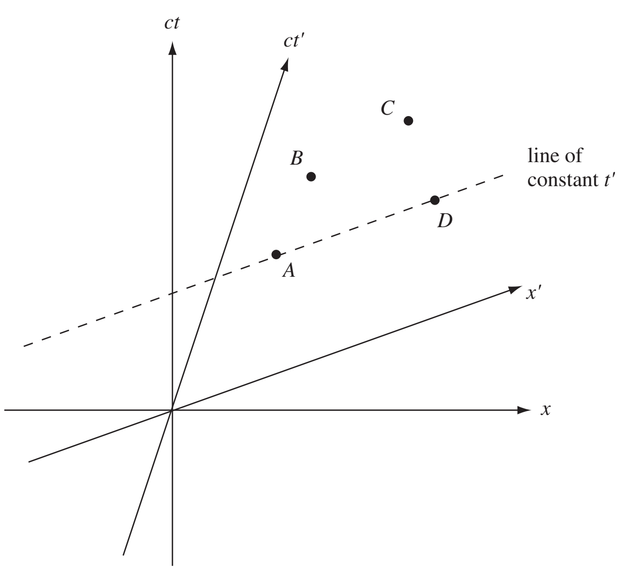

Figure 1.5 The events illustrated in figure 1.3 and a Lorentz boost such that A and D are simultaneous in S'.

1.7 Length contraction and time dilation

Two elementary (but profound) consequences of the Lorentz transformations are length contraction and time dilation. Both these effects are easily derived from (1.3).

Length contraction

Consider a rod of *proper length* ℓ_0 at rest in S' (see Figure 1.6); we have

$$\ell_0 = x'_B - x'_A.$$

We want to apply the Lorentz transformation formulae and so find what length an observer in frame S assigns to the rod. Applying the second formula in (1.3), we obtain

$$x'_A = \gamma(x_A - vt_A),$$
$$x'_B = \gamma(x_B - vt_B),$$

relating the coordinates of the ends of the rod in S' to the coordinates in S. The observer in S measures the length of the rod at a *fixed* time $t = t_A = t_B$ as

$$\ell = x_B - x_A = \frac{1}{\gamma}\left(x'_B - x'_A\right) = \frac{\ell_0}{\gamma}.$$

Hence in S the rod appears contracted to the length

$$\boxed{\ell = \ell_0\left(1 - v^2/c^2\right)^{1/2}.}$$

If a rod is moving relative to S in a direction perpendicular to its length, however, it is straightforward to show that it suffers no contraction. It thus follows that the volume V of a moving object, as measured by simultaneously noting the positions of the boundary points in S, is related to its *proper volume* V_0 by $V = V_0(1 - v^2/c^2)^{1/2}$. This fact must be taken into account when considering densities.

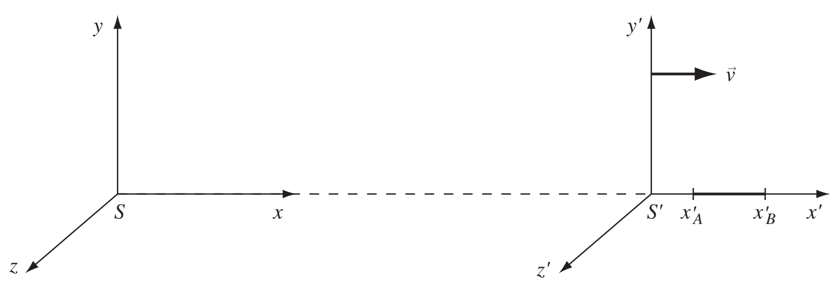

Figure 1.6 Two inertial frames S and S' in standard configuration. A rod of proper length ℓ_0 is at rest in S'.