1
Special relativity

1.1 Introduction

In this chapter, we briefly review the basics of special relativity and provide a short summary of tensor calculus. We assume that the reader is familiar with the fundamental ideas and concepts of the special theory of relativity (SR). More complete introductions to special relativity and tensor calculus can be found in, for example, *A First Course in General Relativity* by Schutz or *Gravitation* by Misner, Thorne, and Wheeler. A good summary is provided in *Radiative Processes in Astrophysics* by Rybicki and Lightman.

Practically every mechanical process that we encounter in our daily lives can be described in terms of Newtonian theory. In astrophysics, however, many systems are relativistic so that applying Newtonian physics can lead to completely wrong answers. The Lorentz factor, given by

$$\gamma = \sqrt{\frac{1}{1-(v/c)^2}}, \tag{1.1}$$

where v is the velocity and c the speed of light, quantifies the importance of special relativistic effects. In a sense, γ measures how close the velocity is to the speed of light: $\gamma = 1$ for $v = 0$ and $\gamma \to \infty$ for $v \to c$. As an example, jets that are emitted from supermassive black holes in the centers of galaxies (see Chapter 8) have Lorentz factors of up to $\sim$30, corresponding to 99.94% of the speed of light. The most violent explosions in the Universe since the Big Bang, gamma-ray bursts (see Chapter 7) accelerate material to Lorentz factors of several hundreds. Electrons spiraling around the magnetic field lines of pulsars possess Lorentz factors of $\sim 10^7$. The cosmic rays that continuously bombard the Earth's atmosphere contain protons with energies of up to $\sim 10^{20}$ eV, which correspond to Lorentz factors of $\gamma = E/m_{\rm p}c^2 = 10^{20}\,{\rm eV}/938\,{\rm MeV} \approx 10^{11}$. Therefore, neglecting special

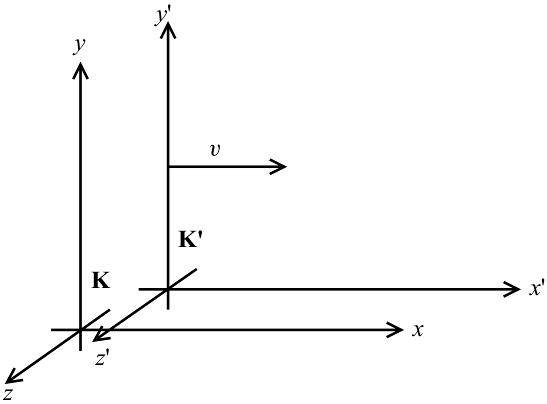

Figure 1.1 Coordinate systems used for the Lorentz transformations: the x-axes are aligned, and the relative velocity between the two frames is v.

relativity in astrophysics can lead to completely wrong interpretations of observations.

1.2 Lorentz transformations

The special theory of relativity is closely related to the notion of *inertial frames*. An inertial frame is a reference frame in which every body is either at rest or moves with a constant velocity along a straight line. In particular, a body viewed from an inertial frame accelerates only when a physical force is applied. In the absence of a net force, a body at rest remains at rest and a body in motion continues to move uniformly. In SR, the set of time and space coordinates, (ct, x, y, z),[1] labels a *space-time event* or simply an event. The coordinates of an event measured in two reference frames, $\mathbf{K}$ and $\mathbf{K}'$, that have a constant relative velocity, v, are related via a *Lorentz transformation*. Unless otherwise stated, we assume that the relative velocity is along the x-axis and that the x-axes of both frames point in the same direction (see Fig. 1.1). In this case, the Lorentz transformation between the coordinates of an event in $\mathbf{K}$ labeled with (ct, x, y, z) and coordinates of the same event in $\mathbf{K}'$ labeled with (ct', x', y', z') reads as follows:

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) \tag{1.2}$$

$$x' = \gamma(x - vt) \tag{1.3}$$

[1] We use arrows for three-vectors only; lengths of vectors are denoted just by a letter. Components of four-vectors are labeled by either super- or subscripts. Four-vectors as geometrical objects are framed by brackets.

$$y' = y \tag{1.4}$$
$$z' = z. \tag{1.5}$$

Often, the inverse transformation from $\mathbf{K}'$ to $\mathbf{K}$ is needed. It can be easily be obtained by exchanging primed and unprimed quantities and v with $-v$:

$$t = \gamma\left(t' + \frac{vx'}{c^2}\right) \tag{1.6}$$
$$x = \gamma(x' + vt') \tag{1.7}$$
$$y = y' \tag{1.8}$$
$$z = z'. \tag{1.9}$$

1.3 Special relativistic effects

In this section, we discuss important effects that are direct consequences of the Lorentz transformations.

1.3.1 Length contraction

Consider a person sitting in frame $\mathbf{K}'$ carrying a rod oriented along the x-axis with length $L_0 \equiv x_2' - x_1'$. What length would an observer sitting in $\mathbf{K}$ measure for the same rod? Equation (1.3) yields

$$L_0 = x_2' - x_1' = \gamma(x_2 - x_1) = \gamma L, \tag{1.10}$$

where $L = x_2 - x_1$ and the ends of the rod have been measured simultaneously in each frame ($t_2' = t_1'$ and $t_2 = t_1$). This means that in its rest frame, the rod appears to be longer by a factor of γ than in a moving frame, or, seen from the frame that moves relative to the rod, its *length is contracted*.

1.3.2 Time dilation

Assume that you have a clock located at the origin of the system $\mathbf{K}'$ and that the time interval between two ticks of the clock is $T_0 = t_2' - t_1'$. Then, an observer in system $\mathbf{K}$ measures (see Eq. [1.6])

$$T = t_2 - t_1 = \gamma\left(t_2' + \frac{vx_2'}{c^2}\right) - \gamma\left(t_1' + \frac{vx_1'}{c^2}\right) = \gamma\left(t_2' - t_1'\right) = \gamma T_0, \tag{1.11}$$

as $x_2' = x_1' = 0$. This means that *the time interval appears to be stretched* by a factor of γ with respect to the object's rest frame.

This effect is observable for unstable elementary particles, for example, in accelerators or cosmic rays. To illustrate this muons produced by cosmic rays in the Earth's atmosphere can only be detected on the ground because their lifetimes are increased by this special relativistic effect. (See Exercise 1 at the end of this chapter.)

1.3.3 Proper time as a Lorentz invariant

Quantities that do not change under a Lorentz transformation are called *Lorentz invariants*. An important such quantity is the *proper time* τ defined via

$$d\tau^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2), \tag{1.12}$$

where (dt, dx, dy, dz) is measured in some arbitrary coordinate system. A clock carried by an observer at a fixed location $(dx = dy = dz = 0)$ shows $d\tau^2 = dt^2$, therefore the name *proper time*. By transforming $d\tau$ from Eq. (1.12) to a frame $\mathbf{K}'$ using Eqs. (1.2)–(1.5), we find that $d\tau' = d\tau$, that is, $d\tau$ is indeed a Lorentz invariant. So how is $d\tau$ related to the general time coordinate dt? The relation can be found by just realising that $d\tau$ measures the time increment of a resting clock. Therefore, time dilation according to Eq. (1.11) yields $\gamma d\tau = dt$. The same result can be obtained more formally by

$$d\tau = \sqrt{dt^2 - \frac{(dx^2 + dy^2 + dz^2)}{c^2}} = dt\sqrt{1 - \left(\frac{v}{c}\right)^2} = \gamma^{-1} dt. \tag{1.13}$$

1.3.4 Transformation of velocities

With the Lorentz transformation given by Eqs. (1.2)–(1.5), we can calculate how velocities transform. If the velocity between our two Lorentz frames is again v and an object has a velocity of $\vec{u} = \frac{d\vec{x}}{dt}$, where $\vec{x} = (x, y, z)$, or $\vec{u}' = \frac{d\vec{x}'}{dt'}$ in the respective frames, the relation, between $\vec{u}$ and $\vec{u}'$ can be easily found by means of Eqs. (1.6)–(1.9):

$$u_x = \frac{dx}{dt} = \frac{\gamma(dx' + v dt')}{\gamma(dt' + v dx'/c^2)} = \frac{dx'/dt' + v}{1 + (v/c^2)(dx'/dt')} = \frac{u'_x + v}{1 + v u'_x/c^2}. \tag{1.14}$$

Completely analogously one finds for the other components:

$$u_y = \frac{u'_y}{\gamma(1 + v u'_x/c^2)} \quad \text{and} \quad u_z = \frac{u'_z}{\gamma(1 + v u'_x/c^2)}. \tag{1.15}$$

In comparison with the x-component, in the last two equations, there is no factor γ in the nominator to cancel the one in the denominator, which is a result of the x- and y-components being unaffected by the Lorentz transformation. Of course,

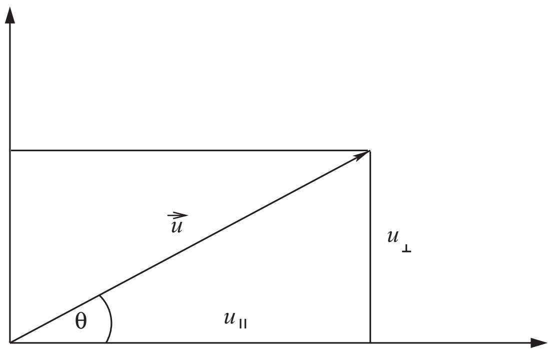

Figure 1.2 For the special relativistic transformation of velocities: splitting the velocity $\vec{u}$ into a component parallel ($u_{||}$) and perpendicular ($u_\perp$) to the velocity between the two frames, v.

for low velocities where $vu'_x \ll c^2$ and $\gamma \approx 1$, the equations reduce to the usual Galilean transformation of velocities.

If one splits up the velocity into components parallel and antiparallel to v, $u_{||}$, and $u_\perp$, the velocity transformation can be written compactly as

$$u_{||} = \frac{u'_{||} + v}{1 + \frac{vu'_{||}}{c^2}} \qquad \text{and} \qquad u_\perp = \frac{u'_\perp}{\gamma\left(1 + \frac{vu'_{||}}{c^2}\right)}. \tag{1.16}$$

Because these components transform differently, the angles appear to be different in different frames. Consider a frame **K** in which an object is moving with velocity $\vec{u}$ (see Fig. 1.2). The angle between $\vec{u}$ and the velocity between our two frames, $\vec{v}$, is given by

$$\tan\theta = \frac{u_\perp}{u_{||}} = \frac{u'_\perp/\gamma(1 + vu'_{||}/c^2)}{(u'_{||} + v)/(1 + vu'_{||}/c^2)} = \frac{u'_\perp}{\gamma(u'_{||} + v)}, \tag{1.17}$$

where we have inserted Eq. (1.16). If the primed velocity components are now expressed via the angle θ' with respect to the $\vec{v}$, $u'_{||} = u'\cos\theta'$, and $u'_\perp = u'\sin\theta'$, we have

$$\tan\theta = \frac{u'\sin\theta'}{\gamma(u'\cos\theta' + v)}. \tag{1.18}$$

This equation describes the *relativistic aberration of light*.

Aberration is an effect that also occurs at nonrelativistic speeds. Imagine standing under an umbrella, and rain is falling straight down from the sky. When you start to walk, you tilt your umbrella slightly forward to protect yourself from the rain,

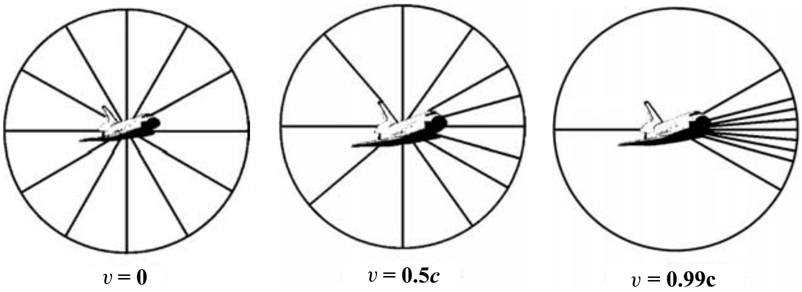

Figure 1.3 Shown are light rays from 12 fixed, distant stars as they appear to a spaceship. In the leftmost panel, the spaceship is at rest; in the middle it travels with 0.5c and in the rightmost panel with 0.99c. In the rightmost panel, almost all photons arrive from the front, even if they stem from stars that are actually behind the spacecraft (that is why some people refer to this effect as the *paranoid effect*).

the more so the faster you walk. It seems like the rain is falling from a position in the sky in front of you, rather than from directly above.

A very similar effect occurs for motion at relativistic speeds. Assume that a spaceship is at rest with respect to a set of very distant stars and that the stars are distributed isotropically, so you see the same number of stars per solid angle in each direction. As the spaceship starts to increase its speed, more and more stars seem to lie ahead of it; the stars seem to pile up in forward direction. This effect is illustrated in Fig. 1.3.

1.3.5 Relativistic beaming

This velocity transformation law leads to an interesting effect called *relativistic beaming*. Beaming plays an important role in the interpretation of observations of, say, active galactic nuclei (Chapter 8) or gamma-ray bursts (Chapter 7) (see Exercise 2) and also in radiation processes such as synchrotron radiation.

It is instructive to consider a photon that moves upward in a frame **K**′, that is, it has $\theta' = \pi/2$. If we insert this into Eq. (1.18), we see that in frame **K** the angle is given by ($u' = c$):

$$\tan\theta = \frac{u'\sin\theta'}{\gamma(u'\cos\theta' + v)} = \frac{c}{\gamma v}. \tag{1.19}$$

Therefore, for very large velocities the angle becomes very small, and for ultrarelativistic motion, $v \approx c$, we have

$$\theta \approx \tan\theta \approx \frac{1}{\gamma}, \tag{1.20}$$

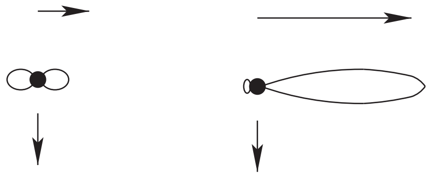

Figure 1.4 Change of a dipole pattern (acceleration perpendicular to the indicated velocity) due to relativistic beaming (left: nonrelativistic, right: relativistic velocity).

which implies that the photon is beamed in forward direction. For a source that radiates isotropically in its rest frame, half of the photons (those with angles $|\theta'| \leq \pi/2$) will be beamed into a cone with an half-opening angle given by the Lorentz factor, $\theta \approx 1/\gamma$. The effect of relativistic beaming on an electron emitting a typical dipole pattern is shown in Fig. 1.4.

1.3.6 Doppler effect

The Doppler effect describes the change of frequency as measured by an observer who moves relative to the source. As an example, think of a car that is passing you: as the car is approaching, you hear a higher frequency. Once it has passed you, the frequency is lower. In the nonrelativistic case, the frequency at the source and the observer are related by $\omega_{\text{obs}} = \omega_{\text{source}}\,(1 - v/c)^{-1}$, where v is the relative velocity.

For a rapidly moving object that emits a periodic signal, such as an electromagnetic wave, we have to apply the special relativistic version of the Doppler effect. We must account for both the previously discussed time dilation, a purely special relativistic effect, and the geometric effect that the source has moved between two pulses.

Assume a source moves at velocity v and emits in its rest frame $\mathbf{K}'$ pulses at a period T' and frequency $\omega_{\text{source}} = \omega' = 2\pi/T'$. What is the frequency observed by an observer in frame $\mathbf{K}$? First, the observer sees the time interval stretched because of the relativistic time dilation, $\Delta t = \gamma T' = \gamma(2\pi/\omega')$. In addition, the source has moved the distance l from 1 to 2 between two pulses (see Fig. 1.5). The light emitted at point 2 has to travel a shorter distance than the light coming from point 1. Therefore, the observed period, the time between the two arriving light pulses as measured by an observer in $\mathbf{K}$, will be Δt minus the time it took to travel the

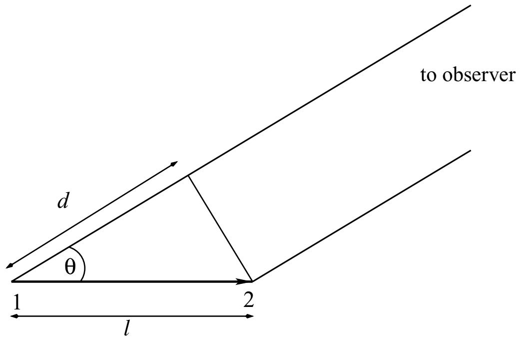

Figure 1.5 A rapidly moving, periodically emitting source travels during one period from point 1 to 2.

extra distance $d = l \cos\theta = v\Delta t \cos\theta$. The observer will measure

$$\Delta t_{\rm obs} = \Delta t - \frac{d}{c} = \Delta t \left[1 - \left(\frac{v}{c}\right)\cos\theta\right] \tag{1.21}$$

and interpret this as a frequency

$$\begin{aligned}\omega_{\rm obs} &= \frac{2\pi}{\Delta t_{\rm obs}} = \frac{2\pi}{\Delta t\ [1-(v/c)\cos\theta]}\\ &= \frac{\omega_{\rm source}}{\gamma\ [1-(v/c)\cos\theta]} = \mathcal{D}\cdot\omega_{\rm source},\end{aligned} \tag{1.22}$$

where $\mathcal{D}$ is called the *Doppler factor*. This is the *relativistic Doppler formula*. The γ in the denominator accounts for the relativistic time dilation; the second term in the bracket corrects for the light-travel effect that occurs also in the nonrelativistic case. In nonrelativistic physics, motions perpendicular to the line of sight do not cause a frequency shift. This is different in the relativistic case. Even for $\theta = \pi/2$, a frequency shift occurs: $\omega_{\rm obs} = \omega_{\rm source}/\gamma$. As we had seen before, the Lorentz factor occurs because of relativistic time dilation. This purely relativistic effect is called the *transverse Doppler effect*.

1.4 Basics of tensor calculus

A tensor is the generalization of the concept of a vector and can be thought of as a set of numbers, for example, a matrix, with a well-defined behavior under a change of the coordinate basis.

1.4.1 The metric tensor

An important example of a tensor is the metric tensor. It can be used to measure distances via a *scalar product*, which associates a number with two vectors. The metric tensor essentially determines how to assign the number to the vectors.

Let us start with the well-known scalar product of real-valued, three-dimensional vectors. Consider two vectors $\vec{x} = (x^1, x^2, x^3)$ and $\vec{y} = (y^1, y^2, y^3) \in \mathbb{R}^3$, where, as usual, the components are just the projections onto the basis vectors, for example, $\vec{x} = (x^1, x^2, x^3) = x^1\hat{e}_1 + x^2\hat{e}_2 + x^3\hat{e}_3$. We use the *Einstein summation convention* according to which we sum over a repeated upper and lower index, that is, the expression $a_j b^j$ stands for $\sum_{j=1}^3 a_j b^j$. Of course, j is just a *dummy index*: $a_j b^j$ is exactly the same as $a_i b^i$. To avoid conflicts, it is sometimes necessary to rename dummy indices. For example, $\left(\sum a_i x^i\right)\left(\sum b_i y^i\right)$ should be written as $a_i x^i b_j y^j$. With this rule, the vector $\vec{x}$ can be written as

$$\vec{x} = x^i \hat{e}_i. \tag{1.23}$$

The scalar product of the vectors $\vec{x}$ and $\vec{y}$ is then given as

$$\vec{x} \cdot \vec{y} = (x^i \hat{e}_i) \cdot (y^j \hat{e}_j) = (\hat{e}_i \cdot \hat{e}_j) x^i y^j. \tag{1.24}$$

Note that it is important here to use two different summation indices to distinguish the two sums. If we know which numbers are assigned to the products of the basis vectors, $\hat{e}_i \cdot \hat{e}_j$, we have, via Eq. (1.24), a rule for the scalar products of general vectors. This is the information contained in the metric tensor. Therefore, one defines the components of the metric tensor as

$$g_{ij} \equiv \hat{e}_i \cdot \hat{e}_j, \tag{1.25}$$

and we can now write

$$\vec{x} \cdot \vec{y} = g_{ij} x^i y^j \equiv G(\vec{x}, \vec{y}). \tag{1.26}$$

Of course, our Cartesian basis vectors in $\mathbb{R}^3$ are unit vectors, that is, of length unity, and they are mutually perpendicular to each other. Thus, the metric tensor is in this case just the unit matrix

$$g = (g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.27}$$

At this stage, it may seem somewhat cumbersome to write the simple scalar product in this way, but this allows a very smooth transition to the special relativistic case.

In relativity, four-vectors with one time and three space components play a prominent role. We have already encountered an important four-vector at the beginning of this chapter, the space-time point (ct, x, y, z), which in tensor calculus is simply written as (x^0, x^1, x^2, x^3).

We follow here the convention that Latin indices refer to the space components and run from 1 to 3, whereas Greek indices refer to space-time components and run from 0 to 3. Therefore, $a_j b^j = \sum_{j=1}^{3} a_j b^j$ is different from $a_\mu b^\mu = \sum_{\mu=0}^{3} a_\mu b^\mu$. Like in $\mathbb{R}^3$, the components of a four-vector are just its projections onto the basis vectors:

$$X^\mu = (x^0, x^1, x^2, x^3) = x^0\hat{e}_0 + x^1\hat{e}_1 + x^2\hat{e}_2 + x^3\hat{e}_3 = x^\mu\hat{e}_\mu. \tag{1.28}$$

As in the three-dimensional case, the *metric tensor* can be thought of as a machine with two input slots that produces a number out of two vectors via a scalar product. It is symmetric

$$G(u, v) = G(v, u) \tag{1.29}$$

and linear

$$G(u, \xi v + \psi w) = \xi G(u, v) + \psi G(u, w), \tag{1.30}$$

where u, v, w can be either three- or four-vector vectors and $\xi, \psi \in \mathbb{R}$. The *components of the metric tensor* are again defined as the scalar products of the basis vectors:

$$g_{\mu\nu} \equiv G(\hat{e}_\mu, \hat{e}_\nu) = g_{\nu\mu}, \tag{1.31}$$

where at the last equal sign we have used the symmetry property. With these conventions, the scalar product of the vectors u and v is given as

$$G(u, v) = G(u^\alpha\hat{e}_\alpha, v^\beta\hat{e}_\beta) = u^\alpha v^\beta G(\hat{e}_\alpha, \hat{e}_\beta) = u^\alpha v^\beta g_{\alpha\beta}, \tag{1.32}$$

where we have made use of Eqs. (1.30) and (1.31). The inverse matrix of $g_{\alpha\beta}$ is denoted $g^{\alpha\beta}$ and fulfills

$$g^{\alpha\lambda} g_{\lambda\beta} = \delta^\alpha{}_\beta, \tag{1.33}$$

where $\delta^\alpha{}_\beta$ is the *Kronecker delta*,[2] which has the value of 1 for $\alpha = \beta$ and 0 otherwise. Generally, a tensor $T^{\mu\nu}$ is said to be *symmetric* if $T^{\mu\nu} = T^{\nu\mu}$ and *antisymmetric* if $T^{\mu\nu} = -T^{\nu\mu}$. It can easily be shown that if a tensor is (anti-) symmetric in one coordinate system, this is also true in any other coordinate system.

The special relativistic scalar product is similar to the scalar product of Eq. (1.26). We have already seen an example of such a special relativistic scalar product (see

[2] After the German mathematician Leopold Kronecker (1823–1891).