

Chapter 2

Special Relativity



2.1 Introduction

This chapter is included mostly for completeness, because many textbooks already explain special relativity in a deeper and more elegant way Galileo.¹

The principle of relativity is present in classical physics since its very beginning. Galileo, in his *Dialogue Concerning the Two Chief World Systems* (1632), states that the laws of physics are the same in all reference frames that are in relative uniform motion. More precisely, Galileo realised that passengers in a ship have no way to tell if the ship is moving or standing.

Shut yourself up with some friend in the main cabin below decks on some large ship, and have with you there some flies, butterflies, and other small flying animals. Have a large bowl of water with some fish in it; hang up a bottle that empties drop by drop into a wide vessel beneath it. With the ship standing still, observe carefully how the little animals fly with equal speed to all sides of the cabin. . . . When you have observed all these things carefully (though doubtless when the ship is standing still everything must happen in this way), have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still.

¹Galileo Galilei, 1564–1642, is considered the father of modern scientific method. He taught in Pisa and Padua, Italy.

Rinserratevi con qualche amico nella maggiore stanza che sia sotto coverta di alcun gran navilio, e quivi fate d'aver mosche, farfalle e simili animaletti volanti: siavi anco un gran vaso d'acqua, e dentrovi de' pescetti; sospendasi anco in alto qualche secchiello, che a goccia a goccia vada versando dell'acqua in un altro vaso di angusta bocca che sia posto a basso; e stando ferma la nave, osservate diligentemente come quelli animaletti volanti con pari velocità vanno verso tutte le parti della stanza. [...] [Osservate che avrete diligentemente tutte queste cose, benché niun dubbio ci sia mentre il vascello sta fermo non debbano succedere così: fate muovere la nave con quanta si voglia velocità; ché (pur di moto uniforme e non fluttuante in qua e in là) voi non riconoscerete una minima mutazione in tutti li nominati effetti; né da alcuno di quelli potrete comprendere se la nave cammina, o pure sta ferma.

What is new in Einstein's relativity is that, in order to make electrodynamics invariant in passing from one inertial frame to another, an additional ingredient is required: that there is one special speed, which is the speed of light in vacuum. This speed is required to be exactly the same in all inertial frames and cannot be surpassed by any particle carrying energy. This hypothesis has been proved experimentally by a famous experiment by Michelson and Morley. The Earth orbits around the sun at a variable speed of about 30 km/s with respect to the sun. This is still 0.01% of the speed of light in vacuum. In addition, the solar system orbits in one arm of our spiral galaxy, the Milky Way, at about 220 km/s with respect to the black hole at the centre of our galaxy. We also know that the galaxies next to ours are moving towards the "Great Attractor", a structure in the intergalactic space, at about 1000 km/s. This is 0.3% of the speed of light. If the speed of light changed in our reference frame on the Earth, we would notice an effect which changes daily in our laboratory, due to the Earth's rotation around its axis. Now, an effect of 0.01% may seem extremely difficult to measure directly. The light covers distance of 3 km in just 10 μ s, in vacuum, so an effect of 0.01% corresponds to a time of 1 ns, which is well within the range of modern measurements of time. The experiment could be repeated nowadays with a laser as a direct measurement, but this is not what Michelson and Morley did, back in 1887. Rather they measured a distance, on an optical bench. In terms of distance, on an optical path of only 1 m 0.01% is 10^{-4} m, or 0.1 mm. Compared to the wavelengths of Sodium light, 589.0 and 589.6 nm, this distance is about a 1000 times larger. Thanks to the short wavelength of visible light, it was possible to measure precisely that there was no variation of the speed of light in the 24-h period, or even when rotating the optical bench, which was floating on a mercury bath (this would not be allowed today). Hendrik

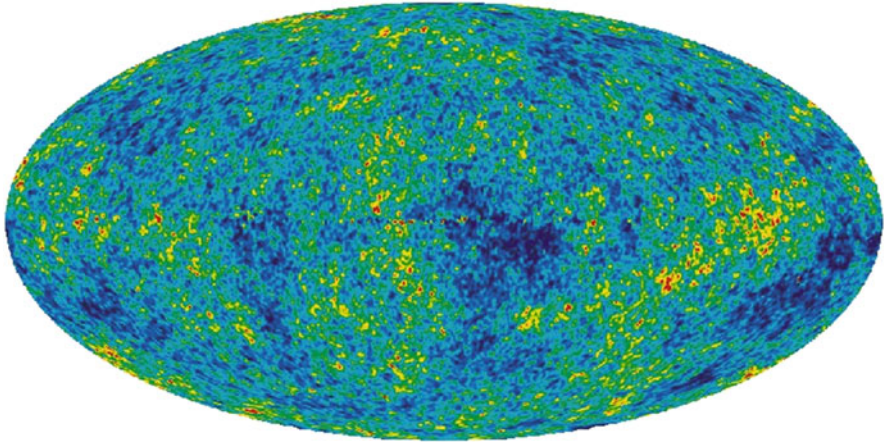


Fig. 2.1 The map of the Cosmic Microwave Background Radiation (NASA)

Lorentz² had already found and published³ the mathematical formulas to allow the speed of light to be constant; Henri Poincaré refined the formulae, which he called “Lorentz transformations”, but it was Albert Einstein, in 1905, who clarified the subject and gave the exact explanation in his article “On the electrodynamics of moving bodies”.⁴ The two principles of special relativity are:

1. The laws of physics are the same in all inertial reference frames. There is no inertial reference frame which is better than others to describe the laws of physics.
2. The speed of light is the same in all reference frames.

We’ll see that there are other effects which allow us to measure relative speeds: relative motion does not change the speed of light but changes its frequency, which corresponds to its colour, as will be shown later. This is known as red or blue shift effect: specific feature of light from distant stars appears to us at a different frequency owing to the relative speed. Nowadays, we can find one reference frame which is somehow special: the frame where the Cosmic Microwave Background Radiation (CMBR (Fig. 2.1)) is uniform and isotropic, to a first approximation; however, it is not better than others to describe the laws of physics.

²Hendrik Lorentz, 1853–1928, was professor in Leiden, the Netherlands. He was awarded the Nobel prize in 1902 for the explanation of the Zeeman effect.

Henry Poincaré, France (1854–1912) was a mathematician and theoretical physicist.

³H.A. Lorentz, *Simplified Theory of Electrical and Optical Phenomena in Moving Systems*, in: KNAW, Proceedings, 1, 1898–1899, Amsterdam, 1899, pp. 427–442;

“Electromagnetic phenomena in a system moving with any velocity smaller than that of light”, Proceedings of the Royal Netherlands Academy of Arts and Sciences, 6: 809–831 (1904).

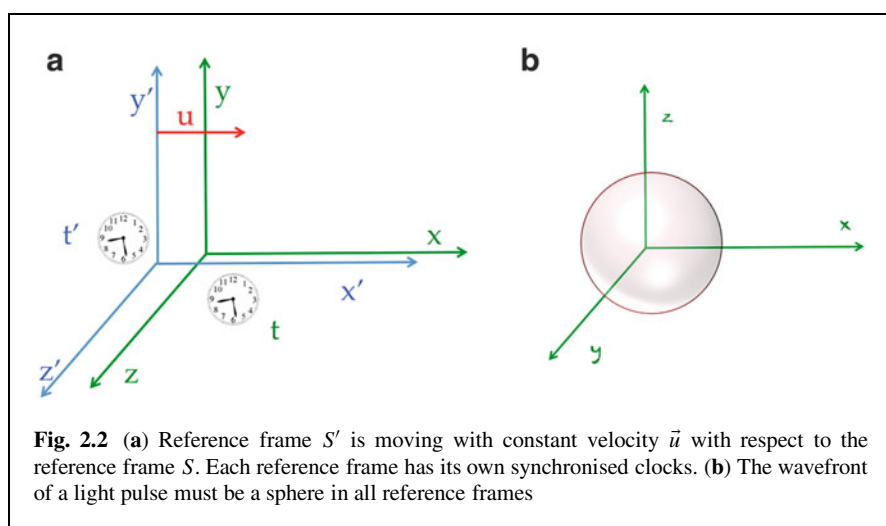
⁴“Zur Elektrodynamik bewegter Körper”, *Annalen der Physik* 17: 891 (1905).

By measuring the frequency shift in opposite directions, we can measure the velocity of the Earth with respect to this background radiation. The result is that the Earth is moving at about 390 km/s towards the Leo constellation, which is located not far from Ursa Major. If you are in the northern hemisphere, you can start from the Ursa's two stars which form the outer bowl of the Big Dipper, or Plough, and draw a straight line; to the north, you'll find Polaris, while Leo is about the same distance from Ursa Major, but at the opposite side.

2.2 The Lorentz Transformations

We'll now derive the correct way to change variables from one inertial reference frame S to another inertial reference frame S' , which is moving at constant velocity \vec{u} with respect to S , based on the principles of special relativity. We assume that each reference frame uses Cartesian coordinates, which for the system S are (x, y, z) ; the time is measured in each frame, with clocks which are at rest in that frame. In the frame S , we'll measure time t . The reference frame S' has coordinates (x', y', z', t') , where with t' (Fig. 2.2a) we mean that the time is measured with clocks which are at rest in the frame S' (Fig. 2.2a). The clocks are synchronised between the two frames: assuming that normal synchronisation occurs within each reference frame, synchronisation occurs between clocks at the same spatial position: say, a synchronisation signal is exchanged between the two clocks at the origin at the time when the origins of the two reference frames are in the same position.

The rules to correctly transform the mathematical description of a physics "event" in the frame S to a description in the frame S' are called the Lorentz transformations, after the Dutch physicist Hendrik Antoon Lorentz, who wrote these transformations well before Einstein correctly explained them.



Let's state some desirable features of these transformations. They must only depend on the relative velocity of the two frames \vec{u} . When this velocity is small, compared to the speed of light, c , the Lorentz transformations must be approximated by the Galileo–Newton transformations. Without lack of generality, we can rotate the reference frames S and S' in such a way that \vec{u} , \hat{x} and \hat{x}' are parallel to each other. The Galileo–Newton transformations are

$$x' = x - ut \quad (2.1)$$

$$t' = t \quad (2.2)$$

The Lorentz transformations must be linear, otherwise they would distort the space–time. We cannot use higher powers of x , t , or other functions, like exponentials or logarithms. A straight line in one frame must be a straight line in the other frame. A uniform motion in one reference frame must be a uniform motion in the other reference frame. So, the transformations must depend on x , y , z , t , but not on any product like xy or x^2 . Also, for $\vec{u} \rightarrow 0$, the transformation must be the identity transformation. The transformations must be symmetrical by exchange $x \leftrightarrow x'$, $y \leftrightarrow y'$, $z \leftrightarrow z'$, $\vec{u} \leftrightarrow -\vec{u}$: the two reference systems are interchangeable.

It is reasonable to expect that the directions perpendicular to the relative motion are unaffected, just like in Galileo transformations: $y' = y$ and $z' = z$. Combining all requirements from above, we can write a generic transformation for space coordinate:

$$x' = \gamma(x - ut) \quad (2.3)$$

and

$$x = \gamma(x' + ut) \quad (2.4)$$

where we need to determine the factor γ . If we require that the speed of light is the same in both reference frames (we'll call it c), the time coordinate cannot be the same in the two frames: it must transform. The same arguments as above are applied to the time coordinate: we can write a generic linear transformation which reads as $t' = b_1x + b_2t$. We have three parameters to determine: γ , b_1 , b_2 . All together, we have

$$t' = b_1x + b_2t \quad (2.5)$$

$$x' = \gamma(x - ut) \quad (2.6)$$

$$y' = y \quad (2.7)$$

$$z' = z \quad (2.8)$$

From the second postulate, we must now require that a light pulse that radiates in all directions should be described exactly in the same way in both frames. Let's assume

that at a given time $t = t' = 0$ the two reference frames overlap, and a light bulb is turned on at the origin of both frames. The wavefront of the light must be described by a sphere in both reference frames (Fig. 2.2b): observing from the frame S , we have

$$x^2 + y^2 + z^2 = c^2 t^2 \quad (2.9)$$

while in the frame S' :

$$x'^2 + y'^2 + z'^2 = c^2 t'^2 \quad (2.10)$$

We substitute (2.5)–(2.8) into (2.10), and we get

$$\gamma^2(x - ut)^2 + y^2 + z^2 = c^2(b_1x + b_2t)^2 \quad (2.11)$$

which becomes

$$(\gamma^2 - c^2 b_1^2)x^2 + y^2 + z^2 = (c^2 b_2^2 - \gamma^2 u^2)t^2 + (2\gamma^2 u + 2c^2 b_1 b_2)tx \quad (2.12)$$

The formula above is the equation of a sphere with radius ct only if:

$$2\gamma^2 u + 2c^2 b_1 b_2 = 0; \quad (2.13)$$

$$\gamma^2 - c^2 b_1^2 = 1; \quad (2.14)$$

$$c^2 b_2^2 - \gamma^2 u^2 = c^2 \quad (2.15)$$

Rearranging the last two equations:

$$b_1^2 = \frac{\gamma^2 - 1}{c^2} \quad (2.16)$$

$$b_2^2 = 1 + \gamma^2 \frac{u^2}{c^2} \quad (2.17)$$

and rearranging and squaring Eq. (2.13), we have

$$\gamma^4 u^2 = c^4 b_1^2 b_2^2 \quad (2.18)$$

Substituting the values of b_1^2 and b_2^2 , we obtain

$$\gamma^4 u^2 - c^4 \frac{\gamma^2 - 1}{c^2} \left(1 + \gamma^2 \frac{u^2}{c^2} \right) = 0; \quad (2.19)$$

From (2.19), we derive

$$\gamma^4 u^2 - c^2 \gamma^2 - \gamma^4 u^2 + c^2 + \gamma^2 u^2 = 0$$

$$\gamma^2(u^2 - c^2) + c^2 = 0$$

$$\gamma^2 = \frac{c^2}{c^2 - u^2} = \frac{1}{1 - \frac{u^2}{c^2}}$$

$$\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (2.20)$$

Substituting into (2.17), we have $b_2^2 = \gamma^2$, and we must choose $b_2 = +\gamma$ for continuity. This means that for very small values of u the Lorentz transformations must be well approximated by the Galileo–Newton transformations. From (2.16), we have

$$b_1^2 = \frac{\gamma^2 - 1}{c^2} = \frac{u^2}{c^4} \left(\frac{1}{1 - \frac{u^2}{c^2}} \right) \quad (2.21)$$

This time (for the same reason as above), we have to choose $b_1 = -\frac{u}{c^2}\gamma$ so that the Lorentz transformations are

$$x' = \gamma(x - ut); \quad (2.22)$$

$$y' = y; \quad (2.23)$$

$$z' = z; \quad (2.24)$$

$$t' = \gamma \left(t - \frac{u}{c^2}x \right) \quad (2.25)$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (2.26)$$

It is evident from the transformations that in the case where the velocity $u \ll c$ we have $\gamma \approx 1$, $u/c^2 \approx 0$ and we find the Galileo–Newton transformations. This allows us to keep the speed of light in vacuum constant in all reference frames. The speed of light is nowadays defined to be exactly 299792458 m/s. The price to pay to have a constant value for c in all reference frames is that time is not the same in all reference frames. Events which are seen as contemporary in one reference frame are no longer happening at the same time in another frame. However, what is preserved is causality: if one event is causing another, there is no reference frame where it may occur after its effect.

Suppose we turn on a light bulb in the middle of a train carriage at local time $t_0 = 0$. In the reference frame of the carriage where the bulb is at rest, the light will reach both ends of the carriage at the same time $t_1 = l/(2c)$, where l is the length of the carriage. If the carriage is moving at speed \vec{u} with respect to the station, the light pulse will be seen to reach the rear end at time $t'_1 = (l - u\Delta t')/c$ and the forward end of the carriage at time $t'_2 = (l + u\Delta t')/c$. Actually, with this example,

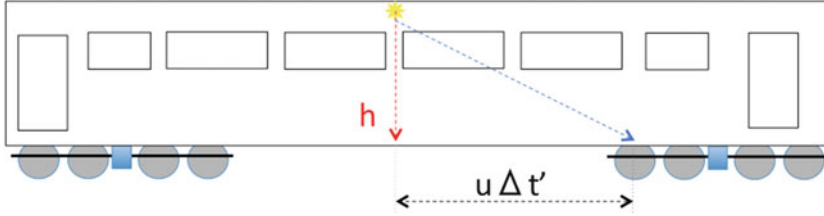


Fig. 2.3 Time intervals are different when observed in reference frames which are in motion one with respect to the other. This figure illustrates the example of the light bulb in a train carriage

we can derive the time contraction formula: suppose the bulb is on the ceiling of the carriage, at a height h from the floor. When observed within the carriage, the time needed by the light to reach the floor is simply $\Delta t = h/c$.

When observed from the station, while the light is in flight, the floor has moved by a distance $d = u\Delta t'$ (Fig. 2.3), so that the light had to go on a straight line, diagonal this time, with a length $h' = \sqrt{h^2 + u^2\Delta t'^2}$. The time to reach the floor, as seen from the station, is $\Delta t' = h'/c$. By squaring, we have

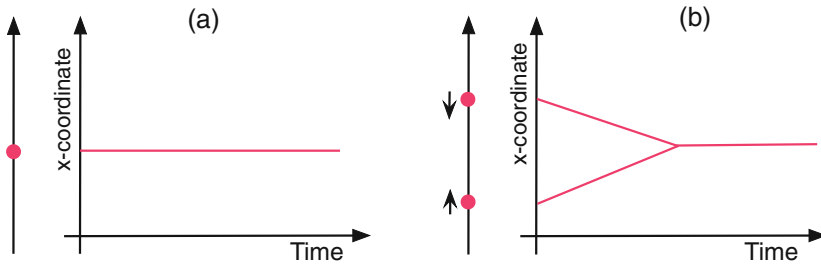
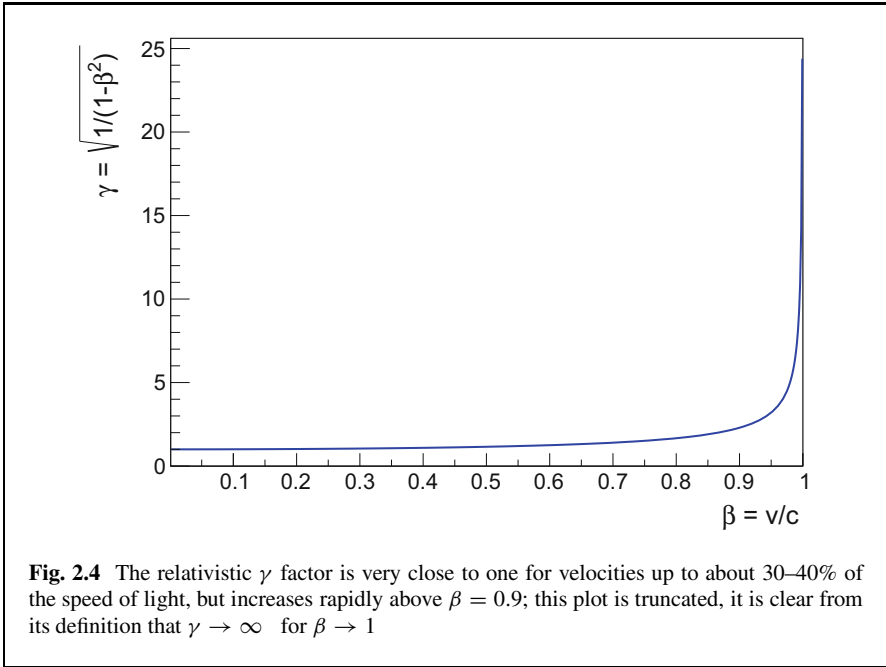
$$\Delta t'^2 = \frac{1}{c^2}(h^2 + u^2\Delta t'^2);$$

using $h^2 = c^2\Delta t^2$, we have

$$\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{u^2}{c^2}}} = \gamma\Delta t; \quad (2.27)$$

The formula above is very important for practical uses: it describes the time dilation when observing a phenomenon in a reference frame where the source of the phenomenon is not at rest.

As $\gamma \geq 1$, times are dilated: moving clocks are running slower than clocks at rest. In this case, the moving clock is the one in the station: we consider at rest the clock which is in the frame where the light bulb is at rest. The relativistic γ factor is very close to one for velocities up to about 30–40% of the speed of light, (Fig. 2.4) and this explains why we do not observe relativistic effects in everyday life. However, nuclear particles have speed comparable to the speed of light and special relativity must be used. The relativity formulas are verified in thousands of applications, including those involving particle accelerators. It is sometime useful to picture a motion in a space-time graphics, which is also called a Minkowski diagram, as shown in Fig. 2.5.



2.3 Velocity, Momentum and Energy, 4-Vectors

We have seen that the time duration of a phenomenon is larger when observed from a moving reference frame, with respect to the reference frame where the phenomenon (or its origin) is at rest. Let's consider two events recorded in the reference frame S : one with coordinate (t_1, x_1, y_1, z_1) and another with coordinates (t_2, x_2, y_2, z_2) . We can define the space–time interval as:

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \tag{2.28}$$

Table 2.1 Naming of the intervals in space–time

ds^2 value	Interval
$ds^2 = 0$	Light-like
$ds^2 > 0$	Time-like
$ds^2 < 0$	Space-like

where $\Delta x = x_2 - x_1$ etc. For two events extremely close to each other, we can define the infinitesimal space–time interval:

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad (2.29)$$

This interval is invariant by Lorentz transformations: given two events, the interval between them is the same no matter what reference frame is used to measure it.

If $c^2 dt^2 > dx^2 + dy^2 + dz^2$, the event separation is said to be *light-like* (Table 2.1). There is an ideal light ray which joins the two events.

If $c^2 dt^2 > dx^2 + dy^2 + dz^2$ (in this definition $ds^2 > 0$), the space–time interval is *time-like*.

If $c^2 dt^2 < dx^2 + dy^2 + dz^2$ (in this definition $ds^2 < 0$), the space–time interval is *space-like*. No cause–effect relationship can exist between the two events, they can happen at the same time in some reference frame.

The space–time interval is left invariant by Lorentz transformations, it is “Lorentz-invariant”. We can introduce the 4-vectors as sets of 4 quantities, or components, of which 3 are space-related and one is time-related. These mathematical objects transform according to the Lorentz transformations when changing the reference frame. As an example space–time coordinates can be expressed in the form of a 4-vector, which is indicated in boldface \mathbf{x} . We have to make sure to have homogeneous quantities: all four components must have the same physical dimensions. This is obtained by multiplying the time component by the speed of light, c .

$$\mathbf{x} = (ct, \vec{x}) = (ct, x, y, z) \quad (2.30)$$

Let’s look at velocities now, to see if they form a 4-vector. For a particle of mass m :

$$v_x = \frac{dx}{dt};$$

$$v'_x = \frac{dx'}{dt'} = \frac{\gamma(dx - udt)}{\gamma(dt - \frac{u}{c^2}dx)} = \frac{dx/dt - u}{1 - \frac{u}{c^2} \frac{dx}{dt}}$$

$$v'_x = \frac{v_x - u}{1 - \frac{u}{c^2} v_x} \quad (2.31)$$

The other components of the velocity are also affected by the Lorentz transformations, via the dt term:

$$v'_y = \frac{v_y}{\gamma(1 - \frac{u}{c^2}v_x)} \quad v'_z = \frac{v_z}{\gamma(1 - \frac{u}{c^2}v_x)} \quad (2.32)$$

The velocity (as defined above) does not transform according to the Lorentz transformations and is not part of a 4-vector. However, we do have a 4-vector if we define the velocity as a derivative with respect to the *proper time* τ $d\tau = \frac{1}{\gamma}dt$. The proper time is the time as measured in the reference frame where the particle is at rest. In this case, the 4-velocity has simply c as the “time” component. The momentum is more interesting. We define the relativistic 3-momentum (\vec{p}) of a particle as:

$$\vec{p} = m \frac{d\vec{x}}{d\tau} = \gamma_v m \vec{v} \quad (2.33)$$

The γ factor in the above formula is one of the reference frames where the particle is at rest. We can add a time-like component, which we call relativistic energy $p_0 = E = \gamma mc^2$ and show that $(p_0, \vec{p}) = (E/c, p_x, p_y, p_z)$ transforms in the same way as the space and time coordinates, i.e. is a 4-vector. We call it 4-momentum. 4-vectors can be added, can be multiplied by a numeric factor and we can define a scalar product:

$$(x_0, x_1, x_2, x_3)(y_0, y_1, y_2, y_3) = x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3. \quad (2.34)$$

The scalar product between 4-vectors can be considered as a standard scalar product when we multiply one of the two vectors by a matrix g :

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

which is called metric matrix and changes sign to the space coordinates.

The scalar products of 4-vectors are relativistic invariants, so we can define a relativistically invariant 4-vector modulus. In particular, the modulus of the 4-momentum is a relativistic invariant, and its value is left invariant by any process, like a radioactive decay. The modulus of the 4-momentum of the parent nucleus/particle is the same as the modulus of the sum of the 4-momenta of the decay products. This is the *invariant mass* of a system of particles. Let's consider a scattering process $A + B \rightarrow C + D + E$ or a α decay $A \rightarrow B + \alpha$, where A, B, C, D, E are particles or nuclei. We are used to 3-momentum conservation,

where the sum of the momentum of the right-hand side is equal to the sum of the momentum of the left-hand side.

$$\vec{p}_A + \vec{p}_B = \vec{p}_C + \vec{p}_D + \vec{p}_E \quad (2.35)$$

and

$$\vec{p}_A = \vec{p}_B + \vec{p}_\alpha \quad (2.36)$$

We write the conservation of the 4-vector energy–momentum in exactly the same way:

$$\mathbf{p}_A + \mathbf{p}_B = \mathbf{p}_C + \mathbf{p}_D + \mathbf{p}_E \quad (2.37)$$

and for a decay

$$\mathbf{p}_A = \mathbf{p}_B + \mathbf{p}_\alpha \quad (2.38)$$

This is a compact way to state the energy and momentum conservation. The equation above is valid in all inertial reference frames, provided that we transform the 4-momenta according to the Lorentz transformations. The modulus of the sum of 4-vectors has to be the same in both sides of the reaction, and it is the same in all inertial reference frames, so we say it is *invariant*.

$$|\mathbf{p}_B + \mathbf{p}_\alpha|^2 = |\mathbf{p}_A|^2 = m_A^2 c^2 \quad (2.39)$$

From Eq. (2.39) above, we see that this modulus, in case of a decay, is the square of the mass of the initial particle, and therefore it is called *invariant mass*. In particle physics, a method to detect short-lived particles, like the Z^0 or the Higgs boson, is to calculate the invariant mass of the decay products: the obtained values cluster around the mass of the particle which decays (see Fig. 6.10). In case of a decay of a particle “A” into N particles, the sum has to be extended to all the decay products:

$$A \rightarrow \text{particle}_1 + \cdots + \text{particle}_N \quad (2.40)$$

$$|(E_A/c, p_{x_A}, p_{y_A}, p_{z_A})|^2 = \left| \sum_{i=1}^N (E_i/c, p_{x_i}, p_{y_i}, p_{z_i}) \right|^2 = m_A^2 c^2. \quad (2.41)$$

2.4 Relativistic Energy

We have defined the relativistic energy of a particle of mass m moving at velocity \vec{v} as $p_0 \equiv E \equiv \gamma_v m c^2$. We need to show that it has something to do with the standard

definition of energy in classical physics. Taking the limit $v \ll c$:

$$E = mc^2\gamma = mc^2(1 - v^2/c^2)^{-1/2} \quad (2.42)$$

$$\approx mc^2\left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) \quad (2.43)$$

$$= mc^2 + \frac{1}{2}mv^2 \quad (2.44)$$

we find that it correctly corresponds to the particle kinetic energy plus a constant term, which is fine because only energy differences are important. In the case when $\vec{v} = 0$, we have that the energy at rest of a particle is its mass:

$$E = mc^2 \quad (2.45)$$

This quantity is also the modulus of the energy–momentum 4-vector $\mathbf{p} = (E/c, \vec{p})$, and therefore it is a relativistic invariant. It is the *invariant mass*. From the equation above, we can measure the mass in terms of energy. A process that normally occurs in high-energy physics experiments is the transformation of kinetic energy of particles into mass of new particles and their kinetic energy. This process also occurs naturally when cosmic rays hit the atmosphere. Another example is given by high-energy photons: they have no mass, but they carry energy. In the electric field of a nucleus, they can convert into a pair of charged particles, which have mass. We'll see later the fusion and fission processes, which convert the mass of atomic nuclei into kinetic energy, and ultimately thermal energy.

The masses of subnuclear particles are typically measured in MeV/c^2 or in GeV/c^2 . A self-consistent unit system, which is used in theoretical physics, chooses $c = 1$, a dimensionless constant (and also $\hbar = 1$, but this constant will be introduced later). In many text books, this convention is used, and both masses and momenta are measured in units of eV .

It is also evident that the relativistic 3-momentum $\vec{p} = \gamma m\vec{v} \rightarrow m\vec{v}$ becomes the non-relativistic 3-momentum when $v \ll c$ and $\gamma_v \approx 1$.

For free particles, it is fair to assume that the energy and momentum are conserved. Suppose now that we also deal with forces which conserve energy and momentum. Then, from the conservation of our 4-momentum we can show that our new definition of energy is really what we need in terms of Newton's law in some of its various forms.

$$\frac{d}{dt}\mathbf{pp} = 0 = \frac{d}{dt}\left(\frac{E^2}{c^2} - \vec{p} \cdot \vec{p}\right) = \quad (2.46)$$

$$= \frac{2E}{c^2} \frac{dE}{dt} - 2\vec{p} \cdot \frac{d\vec{p}}{dt} = \quad (2.47)$$

$$= 2m\gamma \frac{dE}{dt} - 2m\gamma \vec{v} \cdot \frac{d\vec{p}}{dt} \text{ because } E = m\gamma c^2 \text{ and } \vec{p} = \gamma m \vec{v} \quad (2.48)$$

$$= 2m\gamma \left(\frac{dE}{dt} - \vec{v} \cdot \frac{d\vec{p}}{dt} \right) \quad (2.49)$$

So, we have demonstrated that for conservative forces, with our relativistic definition of energy we indeed have

$$\frac{dE}{dt} = \vec{v} \cdot \frac{d\vec{p}}{dt} \quad (2.50)$$

The relativistic *kinetic energy* for our free particle is obtained from the relativistic total energy by subtracting the energy of the particle at rest:

$$E_k = \gamma_v mc^2 - mc^2 = (\gamma_v - 1)mc^2 \quad (2.51)$$

The modulus of the energy–momentum 4-vector is Lorentz invariant. Calculating it in the particle rest frame, it is

$$E^2 - \vec{p}\vec{p}c^2 = m^2c^4 \quad (2.52)$$

For massless particles (i.e. the *photon*, the only⁵ massless particle we know of), $E^2 = p^2c^2$. It will be shown later that the energy of a photon is given by its frequency: $E_\gamma = h\nu$. The photon carries a momentum which is $p_\gamma = \frac{h\nu}{c}$. Massless particles in vacuum can only travel at the speed of light. While the *photon* is the only particle with exactly zero mass, other particles, the *neutrinos*, are massless with a good approximation. Neutrinos are produced in beta decays, we do not know exactly their mass, but we know it is very low.

2.5 Doppler Effect

At this point, we need to know what happens to photons when observed from different reference frames. An electromagnetic plane wave can be described by:

$$A(\vec{x}, t) = A_0 \sin(\vec{k}\vec{x} - \omega t) . \quad (2.53)$$

⁵The other particle which we believe to be massless is the *gluon*, the messenger of the strong force, which is confined within *mesons* and *baryons* and cannot travel freely, because of the “colour confinement”.

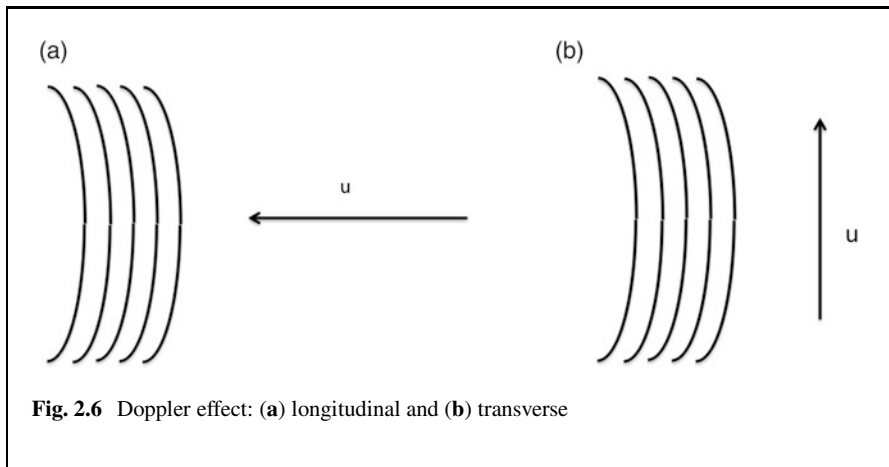
We rename $k_0 = \omega/c$ the quantity (k_0, \vec{k}) must be a 4-vector if we require that the phase must be the same in all moving frames: $\vec{k}\vec{x} - k_0ct = \vec{k}'\vec{x}' - k'_0ct'$. This is a *light-like* 4-vector. Suppose we have a plane wave along the x -axis, the same direction as the relative velocity of the two reference frames \vec{u} .

$$k'_x = \gamma_u \left(k_x \pm \frac{u}{c} k_0 \right); \text{ we can call } \beta = u/c, k = k_0 = 2\pi c\nu \tag{2.54}$$

$$\nu' = \gamma\nu(1 \pm \beta) = \nu \frac{1 \pm \beta}{\sqrt{1 - \beta^2}} = \nu \sqrt{\frac{1 \pm \beta}{1 \mp \beta}} \tag{2.55}$$

When we observe the light from a star, we can measure its speed relative to the Earth by measuring the frequency of characteristic emission lines. In general, this is shifted towards lower frequencies, indicating that stars are moving away from us. As the red colour is located at the lower end of the colour frequency spectrum, this is called *red shift*. The relativistic Doppler effect is completely symmetrical if we exchange source and observer, as it should be. In the acoustic Doppler effect, the motion is relative to the medium, so it does make a difference whether the source is moving or the observer (listener) is moving (Fig. 2.6(a)).

In addition to the longitudinal Doppler effect, there is also the relativistic transverse Doppler effect: the frequency changes also when the observer is moving parallel to the optical wavefront (Fig. 2.6(b)).



2.6 Group Theory in a Nutshell

The 4-vectors are mathematical objects “living” in a vector space, which is called Minkowski space.⁶ We can add them, multiply by a scalar number and/or make scalar numbers out of them. Then, we have several *mappings* of 4-vectors onto 4-vectors: we can rotate them in space, with normal rotations, we can translate in space and in time, we can “boost” them with velocity \vec{u} , i.e. we can observe them from a reference frame with velocity \vec{u} with respect to the previous one. In general, all these operations on 4-vectors can be parametrised by one or more parameters: rotations are parametrised by the angle θ or ϕ , Lorentz boosts by a velocity \vec{u} and so on. We indicate with $T(m)$ such generic transformations. The sets \mathcal{G} of these transformations become some mathematical objects, and they can start living their own mathematical life, independently of the vector space where we initially defined them. It is reasonable to require that all these transformation must be invertible, meaning that we can always return to the initial coordinate system; that we can combine these transformation one after the other and still obtain a transformation; and that when combining several transformations, we can associate two or more of them, while keeping the same order. The fact that exists a transformation that leaves the system unchanged is almost trivial, but required. If the above conditions are satisfied, the set of transformations is said to form a *group*. More formally:

Even more formally, let \mathcal{G} be a set of transformations and \circ a *composition* operation

$$1) \forall S, T \in \mathcal{G}, (S \circ T) \in \mathcal{G}$$

$$2) \forall R, S, T \in \mathcal{G}, (R \circ S) \circ T = R \circ (S \circ T)$$

$$3) \exists I \in \mathcal{G} \ni \forall T \in \mathcal{G}, I \circ T = T$$

$$4) \forall T \in \mathcal{G} \exists T^{-1} \ni T \circ T^{-1} = I$$

we say that $(\mathcal{G}; \circ)$ is a *group*.

- 1) We can define a *composition* of transformations (“ \circ ”) to combine any two of them: we can apply them one after the other, and the result is still the result of a transformation of the set;
- 2) The composition of transformations is associative;
- 3) There is an identity transformation, which, when composed, with any of the others, leaves them invariant;
- 4) For all transformations, there is an inverse transformation, which is still part of the set;

⁶After the mathematician Hermann Minkowski, 1864–1909. He taught at Königsberg, Bonn, Göttingen and Zürich, where Einstein was one of his students.

If, in addition, the composition of transformations commute, i.e. $S \circ T = T \circ S$, the group is said to be *commutative* or *Abelian*.⁷ If the transformations, which are the elements of the group, depend on some continuous parameters, for instance the rotation angles, the group is called a *Lie group*.⁸ Transformations can be represented by matrices. In case of Lorentz transformations, we have 4×4 matrices of \mathbb{R} real numbers. Rotations in the three-dimensional space are represented by 3×3 matrices.

Lorentz transformations for boosts along the x coordinate have the following matrix form:

$$b_x = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\beta = u/c$.

The set made of the Lorentz transformations and the rotations form a group, which is called the *Lorentz group* \mathcal{L} . While rotations form a subgroup of \mathcal{L} , generic Lorentz transformations don't. However, Lorentz transformations along each of the three axes do form, each in its own, a subgroup of \mathcal{L} . A mathematical digression: given a square matrix M its transposed matrix is obtained by swapping its rows with its columns: $M = (M_{ij})$, $M^T = M_{ji}$. A matrix O is called *orthogonal* if $O^T O = \mathbb{1}$ is the diagonal unit matrix. A matrix is said to be *special* if its determinant is equal to 1. The orthogonal matrices in three dimensions with determinant 1 represent the rotation group, which is called $SO(3)$. The Lorentz group is also indicated as $SO^+(1, 3)$. Rotations depend on three parameters, and Lorentz transformations also depend on three parameters, so an element of the Lorentz group is specified by six real parameters.

Analogously to real matrices, complex matrices $N \times N$ can be transposed and conjugated $(M^\dagger)_{ij} = \bar{M}_{ji}$. Here, $\bar{}$ indicates complex conjugation: $z = (a + ib)$; $\bar{z} = (a - ib)$. A matrix U is said to be *unitary* if $UU^\dagger = U^\dagger U = \mathbb{1}$. If, in addition, the determinant is equal to 1, the group they form is called $SU(N)$, indicating special unitary matrices $N \times N$. Groups can be defined in an abstract way, independently of the vector space and the transformations where we started from. They can act on several *vector spaces*. Given a vector space, a group \mathcal{G} has a certain *representation* in that space: a set of invertible matrices *represent* all the transformations of that group.

A representation of a transformation is *completely reducible* if all matrices are the form of block matrices, which are zero in the off-diagonal blocks. In this case,

⁷After Niels Henrik Abel, 1802–1829 Norwegian mathematician. Also, the Abel prize for mathematics is named after him.

⁸After M. Sophus Lie, (pron. as Lee) 1842–1899, Norwegian mathematician, who taught in Oslo and Leipzig.

the vector space, originally of dimension N_0 , is divided into invariant subspaces each of dimension N_1 and N_2 , with $N_0 = N_1 + N_2$. The representation \mathbf{R}_0 is said to be the *direct sum* of two representations \mathbf{R}_1 and \mathbf{R}_2 , and we write

$$\mathbf{R}_0 = \mathbf{R}_1 \oplus \mathbf{R}_2 \quad \text{e.g. } \mathbf{4} = \mathbf{3} \oplus \mathbf{1} \quad (2.56)$$

where the boldface number indicate the dimension of the representation. The number of parameters is still the same as in the original group which is represented. We can form the *direct product* of two or more groups, which is a group depending on a number of parameters given by the sum of the parameters of each group. A representation of this group is also a block matrix, in which each block is a representation of the corresponding sub-group. We write in this case

$$\mathbf{R} = \mathbf{R}_1 \otimes \mathbf{R}_2 \quad (2.57)$$

We have introduced an empty space, some free particles at rest or in uniform motion and mathematical objects like 4-vectors and their transformations, which form groups. In these ideal conditions, many quantities are conserved. The interesting part comes when we introduce the fundamental *interactions*.

2.7 Symmetries and Conserved Quantities

In 1915, Emmy Nöther (Fig. 2.7) proved a fundamental theorem. In general, a physical system with interacting particles can be described by equations in the Lagrangian or Hamiltonian formulation. It is beyond the scope of this book to introduce these, but some of the readers may be familiar with them. The important issue is that given a system, its Lagrangian (or Hamiltonian) equation fully describes its evolution. If these equations don't change when we perform any of the coordinate transformations which are elements of a continuous group with N parameters, then our physical system has N independent conserved quantities. These are also called "*first integrals*".

This theorem applies equally well to spinning tops and elementary particles. Its derivation is beyond the scope of this course, but beautiful proofs of it can be found in several textbooks. Some notable examples are: if the Lagrangian of our system is independent of time, then the energy is conserved. If the Lagrangian equation remains invariant under translations in space, then the momentum is conserved; if it does not depend on the particular orientation, then the angular momentum is conserved. This theorem is very important and has inspired the theories of fundamental interactions: the interactions can be introduced starting from the relativistic quantum mechanical Lagrangian equation of non-interacting particles and requiring its invariance with respect to groups of local transformations. A transformation is local when the parameters of the transformation are a smooth



Fig. 2.7 Amalie Emmy Nöther in 1930, in Koenigsberg. She was born in Germany in 1885 to a Jewish family, and she taught in Goettingen and for a short period in Moscow. In 1933, she had to move to the USA to continue her activity as a mathematician. She prematurely died 2 years later

An example, a representation of the group $SO(2)$ in four dimensions, which is a direct sum has the following matrix form:

$$R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{pmatrix}$$

Note that there is only one parameter, θ .

$$\mathbf{4} = \mathbf{2} \oplus \mathbf{2} \quad SO(2)$$

An example, a representation of $SO(2) \otimes SO(2)$, which is a direct product of representations has the following matrix form:

$$R(\theta, \phi) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos \phi & -\sin \phi \\ 0 & 0 & \sin \phi & \cos \phi \end{pmatrix}$$

Note that there are two parameters, θ and ϕ .

$$\mathbf{4} = \mathbf{2} \otimes \mathbf{2} \quad SO(2) \otimes SO(2)$$

function of the space coordinates, rather than constant quantities. In the simple example of Fig. 2.8, this means $\phi = \phi(\mathbf{x})$. To make the Lagrangian equation invariant, we need to introduce new terms to it, and these terms correspond to interactions. These theories are called *gauge* theories.

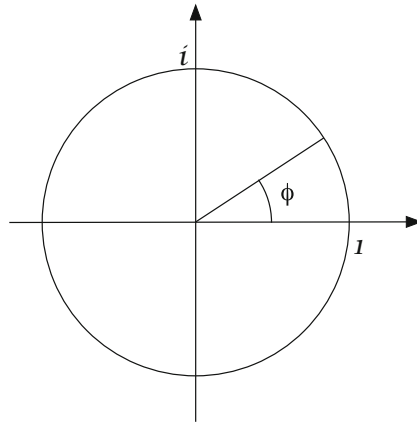


Fig. 2.8 Phase transformations depend on one \mathbb{R} eal parameter, ϕ . They act on complex functions $\psi(x) \rightarrow \psi'(x) = e^{i\phi}\psi(x)$ multiplying them by $e^{i\phi}$. They form a unitary and abelian group called $U(1)$. If the parameter ϕ is a constant, we have a *global* transformation

2.8 Problems

- 2.1 A *heavy-flavoured B-meson* is a particle with a mass of approximately $5 \text{ GeV}/c^2$. For what values of momentum is its γ factor $\gamma \geq 3$.
- 2.2 Three spaceships fly in triangular formation at 100 km from each other at a constant speed of $0.8c$ with respect to the space station Alpha, which can be considered to be an inertial system with good approximation. How is the radio communication among the three spaceships affected?
- 2.3 In the example in Chap. 2, a ^{90}Sr source produces β rays, which are electrons, with a kinetic energy distribution which extends up to 546 keV. This is called the *end-point* of the spectrum. What is the corresponding relativistic electron velocity?
- 2.4 A particle with mass of $125 \text{ GeV}/c^2$ decays into two γ 's. Calculate the energy of the two gammas and their relative direction in the reference frame where the initial particle is at rest.
- 2.5 Prove that the space–time interval $ds^2 = c^2t^2 - dx^2 - dy^2 - dz^2$ is invariant by Lorentz transformations.
- 2.6 A photon can be described as a plane wave $A(\vec{x}, t) = A_0 \sin(\vec{k}\vec{x} - \omega t)$ with phase velocity $\frac{\omega}{|\vec{k}|} = c$. The angular frequency is related to the frequency by $\omega = 2\pi\nu$. The photon can also be treated as a massless particle, with 4-momentum $(E_\gamma/c, E_\gamma, 0, 0)$. The relation between energy and frequency is given to be $E_\gamma = h\nu$, where h is Planck's constant, which will be introduced

in the next chapter. Show that the Doppler effect formulas (Eq. (2.55)) can be obtained by Lorentz-transforming the photon 4-momentum, and using $E_\gamma = h\nu$.

2.9 Solutions

Solution 2.1 We can use more than one formula as a starting point: the definition of momentum $p = \gamma\beta mc$ but we see that this depends on β as well. The modulus of the energy–momentum 4-vector is a good alternative: $E^2 - p^2c^2 = m^2c^4$; but, we also know that $E = \gamma mc^2$, and the formula only depends on p and m :

$$E^2 = m^2c^4 + p^2c^2 = \gamma^2 m^2c^4$$

$$\gamma^2 = \frac{m^2c^4 + p^2c^2}{m^2c^4} = 1 + \frac{p^2}{m^2c^2}$$

$$p^2 = (\gamma^2 - 1)m^2c^2$$

Setting $\gamma = 3$, we have $p \geq 2\sqrt{2}mc = 2 \times 1.414 \times 5 \text{ GeV}/c^2 \times c = 14.1 \text{ GeV}/c$.

Solution to 2.2 The spaceships are flying in formation so we can find a reference frame where the three spaceships are at a constant separation distance from each other. This frame is moving at a constant speed with respect to an inertial reference frame, and therefore it is in turn an inertial reference frame. The radio communications among these ships occur at the speed of light in vacuum, which is the same in all inertial reference frames. As the relative distance does not change, no Doppler effect is present.

Solution to 2.3 From Eq. (2.51):

$$E_k = (\gamma - 1)mc^2 \Rightarrow \gamma = 1 + \frac{E_k}{m_e c^2} = 1 + \frac{546}{511} = 2.068$$

From Eq. (2.26):

$$\gamma = \sqrt{\frac{1}{1 - \beta^2}} \Rightarrow \beta = \sqrt{1 - 1/\gamma^2} = \sqrt{1 - \frac{1}{4.28}} = 0.48$$

With a proper relativistic calculation, the electrons from ^{90}Sr have a maximum speed which is about half of the speed of light.

Solution to 2.4 In the reference frame where the initial particle is at rest, its total energy is $E = mc^2$. The initial momentum is zero, so also in the final state the total momentum must be zero. Therefore, the two photons must be emitted along the same straight line, but in opposite directions and with the same modulus of

3-momentum. For massless particles, $E = |p|c$ so the two photons have also the same energy:

$$p_{\gamma 1} = p_{\gamma 2}; \quad E_{\gamma 1} = E_{\gamma 2} = p_{\gamma} c$$

The invariant mass of the initial state is $E = 125 \text{ GeV}$. In the final state:

$$E' = |(E_{\gamma}, p_{\gamma}, 0, 0) + (E_{\gamma}, -p_{\gamma}, 0, 0)| = 2 E_{\gamma}$$

$E' = E$ and therefore $E_{\gamma} = 1/2 \cdot 125 = 62.5 \text{ GeV}$.

Solution to 2.5 It is important to state the problem correctly. In this case, we cannot use trivially the time dilation and length contraction formulae, because they have implicitly built-in the hypothesis that the time interval is measured at the same place and the length is measured at the same time. Here, we have two independent events: $A = (ct_a, x_a, y_a, z_a)$ and $B = (ct_b, x_b, y_b, z_b)$. We can safely assume $y_a = y_b$ and $z_a = z_b$.

$$\Delta s^2 = c^2(t_b - t_a)^2 - (x_b - x_a)^2$$

$$(\Delta s^2)' = c^2(t'_b - t'_a)^2 - (x'_b - x'_a)^2$$

and we use the Lorentz transformations equation to express the primed quantities as a function of the coordinates in the non-primed reference frame. We can set $c(t_b - t_a) = a$ and $(x_b - x_a) = b$. The expression we obtain is

$$(\Delta s')^2 = \gamma^2 [a - (u/c)b]^2 - \gamma^2 [b - (u/c)a]^2$$

Developing the squares, we obtain

$$\begin{aligned} (\Delta s')^2 &= \gamma^2 \left[a^2 - (u/c)^2 a^2 - b^2 + (u/c)^2 b^2 \right] \\ &= \gamma^2 \left[a^2 (1 - (u/c)^2) - b^2 (1 - (u/c)^2) \right] \end{aligned}$$

Recalling that $1 - \frac{u^2}{c^2} = \frac{1}{\gamma^2}$, we obtain:

$$(\Delta s')^2 = a^2 - b^2 = c^2(t_b - t_a)^2 - (x_b - x_a)^2 = (\Delta s)^2$$

and therefore the *interval* is a Lorentz-invariant quantity.

Solution to 2.6 We assume that the reference frame is moving with velocity $\vec{u} = (u, 0, 0)$ in the same direction as the photon, which we assume along the \hat{x} axis.

The photon 4-momentum is

$$\mathbf{p}_\gamma = (E_\gamma/c, E_\gamma, 0, 0) = (pc, p, 0, 0);$$

$$p'_x = \gamma \left(p_x - \frac{u}{c^2} E \right); \quad E = c p$$

$$p'_x = \gamma \left(p_x - \frac{u}{c} p_x \right) = p_x \frac{1 - u/c}{\sqrt{1 - (u/c)^2}}$$

Renaming $\beta = u/c$, we have

$$p'_x = p_x \sqrt{\frac{(1 - \beta)^2}{1 - \beta^2}} = p_x \sqrt{\frac{(1 - \beta)(1 - \beta)}{(1 + \beta)(1 - \beta)}} = p_x \sqrt{\frac{1 - \beta}{1 + \beta}}$$

The photon described in the initial frame has a frequency ν such that $p = h\nu/c$; in the primed reference frame, the photon has a frequency $\nu' = cp'/h$. So, replacing momentum with frequency we have

$$\nu' = \nu \sqrt{\frac{1 - \beta}{1 + \beta}}, \text{ which is the Doppler effect formula, Eq. (2.55).}$$

Bibliography and Further Reading

- W.G. Dixon, *Special Relativity* (Cambridge University Press, Cambridge, 1978)
 V. Faraoni, *Special Relativity* (Springer, Berlin, 2013)
 H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, 1980)
 S.P. Puri, *Special Theory of Relativity* (Pearson, New Delhi/Dorling Kindersley, London, 2013)
 W. Rindler, *Introduction to Special Relativity*, 2nd edn. (Oxford University Press, Oxford, 1991)
 W.G.B. Rosser, *Introductory Special Relativity* (Taylor & Francis, New York, 1991)
 J. Rosen, *Symmetry Rules* (Springer, Berlin, 2008)