Relativistic physics for teachers

Lecture given by Frank Krauss, summer term 2006,
at Dresden university of technology

Disclaimer:

This script is, to a large extent, influenced by a number of excellent books on the subject. Specifically, the following books have been used:

- R. P. Feynman: "The Feynman Lectures"

Contents

Chapter 1

What is it all about?

This chapter aims at explaining the heart and soul of special relativity in a fashion suited for high-school students. Neither higher mathematics nor any complicated physics is being used. The focus is on the basic findings, which lead to special relativity and on some of its immediate consequences.

1.1 Basic findings

1.1.1 Setting the stage: Towards relativity

In the mid 1860's, Maxwell finalised his theory of classical electrodynamics and his system of equations also included electromagnetic waves. According to Maxwell's equations, these waves have a constant velocity when travelling through the vacuum, which is independent of their frequency or the direction of propagation. The very existence of electromagnetic waves was shown by F. Hertz in the 1880's. In fact, the constance of c, the speed of light, is in the center of relativistic physics, which , however, was not fully developed until the beginning of the 20th century, when Einstein published his theory of special relativity. There were a number of reasons why it took so long that the theory of relativity was developed, mainly severe conflicts with fundamental beliefs of that time, among them:

- Mechanics a la Newton and Galilei's principle of relativity
  The basic idea of this principle is the following: Consider two **inertial systems**, $\mathcal{S}$ and $\mathcal{S}'$, where S' moves with constant velocity $\vec{v}_0$ relatively to system S. Given that at some time $t = 0$ the origin and orientation

of the two systems coincided, coordinates are related through

$$\begin{aligned} \vec{x}'(t) &= \vec{x}(t) + \vec{v}_0 \cdot t \\ \vec{v}'(t) &= \vec{v}(t) + \vec{v}_0 \,. \end{aligned} \tag{1.1}$$

In addition, these transformations (Galilei transformations) between the two inertial systems are supplemented with the assertion of constant time, i.e.

$$t' = t \,. \tag{1.2}$$

Seemingly, the laws of Newton's mechanic are untouched when going from one system to the other, after all, they deal with changes in the momentum, i.e.

$$\vec{F} = \dot{\vec{p}} = m\dot{\vec{v}} \tag{1.3}$$

and in this equation, the addition of a constant velocity has no effect. Hence, according to this principle, velocities should just add up. This has the consequence that light travelling in system S with the speed c in some direction $\vec{e}_r$ should propagate with a different speed in the primed system, namely $c\vec{e}_r + \vec{v}$. This, of course, is in striking contradiction to the predictions of Maxwell's theory.

- Existence of a medium, the (the "**ether**")
  Here, the problem is the following: Up to this point, the physicists' experience with waves was constrained to acoustic waves or waves in water, which, of course, are waves propagating through a medium - air or water. In some analogy, it was believed that there existed some kind of medium for the propagation of light and electromagnetic waves, which was called "*ether*". The speed of light was then only defined w.r.t. this medium, if existing, and thus, it should be possible to determine any observers state of motion w.r.t. this frame of reference.

The first problem has two apparent solutions: Either Maxwell's equations are valid only in a special frame of reference, or Galilei's principle of relativity and, with it, Newton's mechanics needs some refinement. The former would be quite unsatisfactory given the success of Maxwell's equations, but the latter then is seemingly enforced by the constance of the speed of light in different inertial frames. However, taking Maxwell's equations as a prime

example of physics laws leads to some improved version of the principle of relativity, namely:
All laws of nature remain valid in arbitrary inertial systems.
This automatically implies the substitution of Galilei's principle and a new theorem for the addition of velocities. This then leads to a new view on the structure of space-time.
Problem number two, the question of the existence of an ether has been answered experimentally: Seemingly no such medium exists.

The Michelson-Morley experiment

As said before, there was a number of experiments, designed to determine the absolute velocity of earth through the hypothetical ether. The most famous of these experiments was the one undertaken by Michelson and Morley in the late 1880's. It consists of a device similar to the one depicted in Fig. **??**: A light source A emits light, preferably monochromatic. This light is then split into two beams by a semi-permeable plate B; the beams are reflected by two mirrors C and D before they are recombined and hit a detector. If the distances BC and BD are identical (equal to L), then the two reflected beams are in phase and amplify. If there's some difference in distance and, hence, in the time needed for the respective routes, then interference may amplify or erase the two beams.
Now, suppose, light propagates through a medium, i.e. through "ether". To keep things easy, consider a case, where the apparatus has some relative velocity with respect to this medium, $\vec{u}$ parallel to BD. Then the time needed for the beam to reach D is given by

$$ct_{B\to D} = L + ut_{B\to D} \quad \longrightarrow \quad t_{B\to D} = \frac{L}{c-u}\,. \tag{1.4}$$

This is easy to understand: The speed of light on this route relative to the device is $c - u$. In a similar fashion, the time needed for the way back is given by

$$ct_{D\to B} = L - ut_{D\to B} \quad \longrightarrow \quad t_{D\to B} = \frac{L}{c+u}\,. \tag{1.5}$$

Thus, the total time needed for this beam is

$$t_D = t_{B\to D} + t_{D\to B} = \frac{L}{c-u} + \frac{L}{c+u} = \frac{2L}{c}\frac{1}{1-u^2/c^2} = \frac{t_0}{1-u^2/c^2}\,. \tag{1.6}$$

In a next step the time needed for the other route, through mirror C will be calculated. Here,

$$(ct_{B\to C})^2 = L^2 + (ut_{B\to C})^2 \quad \longrightarrow \quad t_{B\to C} = \frac{L}{\sqrt{c^2 - u^2}}\,. \tag{1.7}$$

Similarly, for the way back, one finds

$$(ct_{C\to B})^2 = L^2 + (ut_{C\to B})^2 \quad \longrightarrow \quad t_{C\to B} = t_{B\to C} = \frac{L}{\sqrt{c^2 - u^2}}\,, \tag{1.8}$$

Hence

$$t_C = \frac{2L}{c}\frac{1}{\sqrt{1 - u^2/c^2}} = \frac{t_0}{\sqrt{1 - u^2/c^2}} \neq t_D\,. \tag{1.9}$$

This enables the comparison of the respective times needed by the two beams on their different ways. In both cases, the numerators are identical and represent the time t_0, the beams would need if there was no additional velocity induced by the movement of the device through the medium carrying the light waves. These times are, of course, identical. However, in the very moment, there is some additional movement, the denominators kick in and change this time. As long as $u \ll c$, one may safely neglect the correction term. If, on the other hand, times would be measured precisely enough or if u was not small compared to c, this difference in runtime would show up.
Of course, this poses an experimental problem, because this precision in measurement clearly is related to the question of how to guarantee that the distances of the two mirrors to the semi-permeable plate are identical. To resolve this, Michelson and Morley rotated their apparatus by $\pi/2$ and checked for a difference in the interference pattern. In fact, Michelson and Morley oriented their device in such a way that one of the lines BC and BD was always pointing along the movement of earth around sun, which amounts to a speed of roughly 30 km/s. They did not find any sign of changes in the interference pattern at any time of the year.

First interpretations

The null result of the Michelson-Morley experiment was quite a shock for many physicists at that time. Lorentz and Fitzgerald then had the first fruitful idea to resolve the puzzle. They independently proposed that material bodies contract when moving and claimed that this contraction was in parallel to the direction of motion only. Comparing with the reasoning above, they argued that

$$L_{\parallel} = L_0 \sqrt{1 - \frac{u^2}{c^2}}\,, \tag{1.10}$$

where L_0 is the length of the body at rest, u is its speed parallel to L_0 and $L_{\parallel}$ is the new, emerging length. Taking into account this change of length in the Michelson-Morley experiment implies a change of t_D, namely

$$t_D = \frac{2L}{c} \cdot \frac{\sqrt{1 - u^2/c^2}}{1 - u^2/c^2} = \frac{2L}{c} \cdot \frac{1}{\sqrt{1 - u^2/c^2}} = t_C\,. \tag{1.11}$$

Hence, such a contraction would explain the null effect of the experiment. Now, one could argue that this was just invented to cure the apparent problems of theorists trying to explain the experiment; however, in any experiment similar problems appeared in such a way that it appeared as if nature conspired in order to hide the ether. It took some time to realise that such a perfect conspiracy by itself may be a law of nature, and it was Poincare who assumed that such a conspiracy was taking place. So he postulated that there would be no way to prove the existence of an ether and thus there would be no way of measuring “absolute” velocities.
Following Lorentz' and Fitzgerald's postulate of length contraction parallel to the direction of motion immediately implies the breakdown of validity of Galilei transformations. According to Galilei, $t' = t$, hence time is a universal and absolute entity. Having witnessed the breakdown of Galilei's principle, it is natural to ask, whether it holds true for the transformation of time. To answer this let's try to perform a Gedankenexperiment and check whether it is possible to measure absolute time. To this end, suppose space is filled with a grid which, at each point, is equipped with a clock. Suppose further, these clocks are all synchronised. Now, consider two parallel mirrors with distance L and a pulse of light being thrown back and forth between the mirrors. This may clearly serve as a clock with a period given by the time t_0 the light needs to travel between mirror A and mirror B, namely

$$t_0 = L_{\perp}/c\,. \tag{1.12}$$

Of course, this can be composed with the time measured by the clocks attached to the surrounding grid. What happens if this light-pulse clock moves in parallel to the mirror plates? If the light is emitted at A, how much time t will it need to reach B which has then moved by

$$L_{\|} = ut\,. \tag{1.13}$$

Well, the answer is that the light travels now for a time

$$t = \sqrt{L_{\perp}^2 + L_{\|}^2}c\,. \tag{1.14}$$

Hence

$$t = \frac{L_{\perp}}{\sqrt{c^2 - u^2}} = t_0 \cdot \frac{1}{\sqrt{1 - u^2/c^2}}\,, \tag{1.15}$$

or, in other words, the moving clock seems to run slower than the clock at rest. This effect is known as time dilation.
Taking together, it seems that in moving systems lengths in the direction of movement get contracted ("**length contraction**") and times get dilated ("**time dilation**").

1.1.2 The Lorentz transformations

Basic conditions and derivation

In the previous considerations the effects of length contraction and time dilation have been found. Both are intimately connected, and it should be possible to find a transformation law for the transition from one inertial system to another one. To do so, it is necessary to list the mandatory conditions. First of all, the principle of relativity ("All laws of nature remain valid in arbitrary inertial systems.") should hold true. Then, the speed of light should be finite and constant, independent of the frame of reference. Last but not least, homogeneity of space should be enforced. One of its consequences is that in two systems S and S' bodies move linearly and homogeneous if they do not experience any force.
To keep things easy, assume that S and S' have their coordinate systems oriented in the same way and in parallel and that they have a relative velocity $\vec{u} = u\vec{e}_x$. If a body B moves with velocity constant $\vec{v}$ in S, then its velocity $\vec{v}'$ in S' should also be constant. This condition can only be fulfilled if the

transformations laws are linear, like

$$\begin{aligned} t' &= a_{11}x + a_{12}t + c_1 \\ x' &= a_{21}x + a_{22}t + c_1 \\ y' &= y \\ z' &= z\,. \end{aligned} \tag{1.16}$$

To further simplify things, impose the conditions $t'_0 = t_0 = 0$ and $x'_0 = x_0 = 0$, i.e. a common origin of the coordinate systems at some time. This enforces $c_1 = c_2 = 0$.
Consider first the movement of the origin in S', i.e. fix $x'_0 = 0$. Then

$$0 = a_{21}x + a_{22}t \quad \text{or} \quad \frac{x}{t} = u = -\frac{a_{22}}{a_{21}}\,. \tag{1.17}$$

Hence, in general

$$x' = a_{21}(u) \cdot (x - ut)\,. \tag{1.18}$$

This situation can also be viewed from the other perspective, i.e. from S'. Then S moves with velocity $-u$ w.r.t. S'. The analogous equation for x is thus

$$x = a_{21}(-u)(x' + ut')\,. \tag{1.19}$$

But, since both systems should be equivalent the coefficient a_{21} should not depend on the sign of u, leading to

$$a_{21}(u) = a_{21}(-u) =: a_{21}(\vec{u}^2)\,. \tag{1.20}$$

This notation also makes the isotropy of space manifest. Now it's time to determine a_{21}. This can be done by employing the fact that the speed of light should be constant in all frames of reference. At $t_0 = t'_0 = 0$ the two systems coincide. Suppose that at this moment, a pulse of light is emitted at the origin. In S this pulse reaches the position

$$x = ct \tag{1.21}$$

at time t and

$$x' = ct' \tag{1.22}$$

in S'. Plugging this into Eqs. (1.18) and (1.19) yields

$$\begin{aligned} x' &= a_{21}(\vec{u}^2)x \cdot (1 - u/c) \\ x &= a_{21}(\vec{u}^2)x' \cdot (1 + u/c) \end{aligned} \tag{1.23}$$

and therefore

$$1 = \left[a_{21}(\vec{u}^2)\right]^2 \left[1 - \frac{u^2}{c^2}\right] \tag{1.24}$$

or

$$a_{21}(\vec{u}^2) = \frac{1}{\sqrt{1 - u^2/c^2}} \,. \tag{1.25}$$

Then, according to Eq. (1.17),

$$a_{22} = -a_{21}\frac{x}{t} = \frac{-u}{\sqrt{1 - u^2/c^2}} \,. \tag{1.26}$$

Hence,

$$x' = \frac{x - ut}{\sqrt{1 - u^2/c^2}} \,. \tag{1.27}$$

In order to gain the non-relativistic limit, this needs to be expanded for $u/c \ll 1$ with the result

$$x' = x - ut + \frac{u^2}{c^2}(x - ut) \ldots \,. \tag{1.28}$$

now the coefficients a_{11} and a_{21} will be determined in a similar fashion. Let's start with the equation for x, cf. Eq. (1.19), and use the results for a_{21} and x'. This results in

$$t' = \frac{1}{u}\left(\frac{x}{a_{21}(\vec{u}^2)} - x'\right) = \frac{t - \frac{u}{c^2}x}{\sqrt{1 - \frac{u^2}{c^2}}} \,. \tag{1.29}$$

Again, this can be expanded in u/c, leading to

$$t' = t - \frac{u^2}{c^2}\left(\frac{x}{u} - t\right) + \ldots \,. \tag{1.30}$$

In both expansions, neglecting terms quadratic in u/c yields the Galilei transformations.

Result, summarised

The Lorentz transformations, connecting to systems S and S' with coordinate systems oriented in parallel, which move with relative constant velocity $\vec{u} = u\vec{e}_x$ and where $t_0 = t'_0 = 0$ and $x(t_0) = x_0 = x'(t'_0) = x'_0 = 0$ read

$$\begin{aligned} t' &= \frac{t - \frac{u}{c^2}x}{\sqrt{1-\frac{u^2}{c^2}}} \\ x' &= \frac{x - ut}{\sqrt{1-u^2/c^2}} \\ y' &= y \\ z' &= z \end{aligned} \tag{1.31}$$

or

$$\begin{aligned} t &= \frac{t' + \frac{u}{c^2}x'}{\sqrt{1-\frac{u^2}{c^2}}} \\ x &= \frac{x' + ut'}{\sqrt{1-u^2/c^2}} \\ y &= y' \\ z &= z'\,. \end{aligned} \tag{1.32}$$

For their derivation, the following conditions have been used:

- Homogeneity and isotropy of space;
- Constance of speed of light;
- Principle of relativity.

Adding velocities

Consider now a body moving with velocity $\vec{v} = v\vec{e}_x$ in S. When this body started at $t = 0$ at $x = 0$, then

$$x(t) = vt\,. \tag{1.33}$$

What is the velocity v' in system S' which moves with $\vec{u} = u\vec{e}_x$ relative to S. Obviously, in S'

$$v' = \frac{x'}{t'} = \frac{x - ut}{t - \frac{ux}{c^2}}\,. \tag{1.34}$$

Plugging in $x = vt$ results in the addition theorem for relativistic velocities, namely

$$v' = \frac{v - u}{1 - \frac{vu}{c^2}}\,. \tag{1.35}$$

Consider the extreme case, where $v = c$. Then

$$v' = \frac{c - u}{1 - \frac{u}{c}} = c\,. \tag{1.36}$$

In other words: c**, the speed of light, is the maximal speed**

1.2 Consequences

1.2.1 Measurements

Failure of simultaneity

Let us for the moment dwell on the equations above. Probably the most surprising term in the equation for t' is the term ux/c^2. What does this term imply. To answer this, consider a situation where two events occur at two different positions x_1 and x_2 at some identical t in S. How does this situation manifest itself in S'. The surprising answer is that equal time, simultaneity, in one frame does not imply simultaneity in another frame. To see this, calculate

$$t_1' - t_2' = \frac{\left(t - \frac{u}{c^2}x_1\right) - \left(t - \frac{u}{c^2}x_2\right)}{\sqrt{1 - \frac{u^2}{c^2}}} = \frac{\frac{u}{c^2}(x_1 - x_2)}{\sqrt{1 - \frac{u^2}{c^2}}}\,. \tag{1.37}$$

What does this imply. assume a man in a space ship S' has attached clocks at both ends of his ship and wants them to run synchronously. How can he achieve this? There are some options, for instance the following: He determines the center between both clocks and from there he sends a pulse of light, running with equal velocity towards both clocks. Of course, the pulses will arrive simultaneously and this can be used to synchronise the clocks. So far so good. But what does an observer in another system S think about the clocks running synchronously? The man in S' will be very convinced that this is the case, if his space ship is flying unaccelerated, i.e. with constant velocity. But the man in S will argue that the space ship flies, therefore the clock in the front end will run away from the pulse of light whereas the rear is catching up. In other words: The pulse has to travel a longer distance to the front than to the rear, hence the pulses won't arrive at the same time. Or, in yet other words: In order for the man in the space ship to think that his clocks run synchronously, the must run asynchronously when seen from outside the space ship.

Space-time

From the Lorentz-transformations, Eq. (1.31) it is apparent that space and time coordinates mix when going from one inertial system to another one, and from the spaceship example discussed above, it is clear that the notion of simultaneity cannot be kept. Hence, it makes no sense to talk about events which are characterised separately by position in space and moment in time. Rather, one needs to talk about events happening on points in a four-dimensional space-time - mathematicians would call this object a manifold. Let us now discuss a bit the geometry of this space-time. To do so, consider a clock, moving in a system S with velocity $\vec{v} = v\vec{e}_x$ along the x-axis. Let us try to graphically represent the motion of this clock in a space-time diagram: in our case a two-dimensional plane with one time-dimension, the t-axis, and one spatial dimension, the x-axis. In such a diagram, the clock moves along a straight line, its **world line**, g which has an angle of θ with the x-axis. This angle is given by

$$\tan\theta = \frac{c}{t} \quad \text{or} \quad \theta = \text{arccot}\frac{v}{c}\,. \tag{1.38}$$

What does this look like in a system S' where the clock is at rest? Obviously, this system moves with $-\vec{v}$ with respect to S. In this system, the temporal axis t' is parallel to g, since the clock is measuring this time. t' is given by

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1-\frac{v^2}{c^2}}}\,. \tag{1.39}$$

But how is the spatial axis x' oriented. The answer is quite simple: The x'-axis can be obtained by setting $t' = 0$. The result is

$$ct = \frac{v}{c}x\,, \tag{1.40}$$

another straight line, as expected. But this implies that in the $x - t$-system, the x'-axis has an angle ϕ with respect to x, which is given by

$$\phi = \arctan\frac{v}{c}\,. \tag{1.41}$$

Hence the x'- and t'-axes of S' are within the x- and t-axes of S.

One more example: Relativistic Doppler effect

Consider a clock in S at $x_1 = x_2 = \bar{x}$, beating with a frequency of

$$\nu = \frac{1}{\Delta t} = \frac{1}{t_1 - t_2}\,, \tag{1.42}$$

where $t_{1,\,2}$ denote individual, subsequent beats. In S', moving with $\vec{u} = u\vec{e}_x$ w.r.t. S, the time interval is given by

$$\Delta\bar{t}' = t_1' - t_2' = \frac{t_1 - t_2}{\sqrt{1 - u^2/c^2}} = \frac{\Delta t}{\sqrt{1 - u^2/c^2}}\,, \tag{1.43}$$

the original interval modified by the usual time dilation. Now, let us assume that the beating clock also emits some light pulse. The extra time needed in S' for this pulse to travel from the clock is given by

$$\Delta\tau' = \frac{x_1' - x_2'}{c} = \frac{\frac{u}{c}(t_1 - t_2)}{\sqrt{1 - u^2/c^2}} = \frac{\frac{u}{c}\Delta t}{\sqrt{1 - u^2/c^2}}\,, \tag{1.44}$$

where

$$x_{1,2}' = \frac{x_{1,2} + ut_{1,2}}{\sqrt{1 - u^2/c^2}} \tag{1.45}$$

has been used. Then, taking together, the total time difference is

$$\delta t' = \Delta\bar{t}' + \Delta\tau' = \Delta t\frac{1 + u/c}{\sqrt{1 - u^2/c^2}} = \Delta t\sqrt{\frac{1 + u/c}{1 - u/c}}\,. \tag{1.46}$$

In other words, the frequency of this clock in S' will be anticipated to be

$$\nu' = \nu \cdot \sqrt{\frac{1 - u/c}{1 + u/c}}\,, \tag{1.47}$$

hence, lower. This result is known as the relativistic Doppler effect. In the non-relativistic limit, this reduces to

$$\nu' = \left(1 + \frac{u}{c}\right)\nu\,, \tag{1.48}$$

where terms of the order u^2/c^2 have been neglected.
In fact, this effect is responsible for what is known as the red-shift of stars: The spectrum of electromagnetic waves (light, etc.) emitted by the stars is shifted towards larger (blue-shift) or lower frequencies (red-shift), depending on whether the stars move towards or away from earth. Due to some significant spectral lights related to certain electron transitions in some atoms this can be measured to great accuracy. The outcome is that nearly all stars and galaxies in the universe move away from our milky way.

1.2.2 Causality: The light-cone

In the previous discussion a case was considered where two clocks were synchronised in a reference frame by emitting and receiving pulses of light. This method is nice but limited by the fact that flashes of light cannot reach any point in space-time. The limiting factor here is that the speed of light is finite. Starting from

$$(t,\, \vec{x})^T = (0,\, \vec{0})^T \tag{1.49}$$

such a pulse can only reach

$$|\vec{x}| = ct\,. \tag{1.50}$$

In other words, a pulse of light emitted at $(0,\, \vec{0})$ reaches all points $(t,\, \pm ct\vec{e}_r)$. This object is called the **light cone**. Coordinates within the light cone, i.e. such coordinates $(\tau,\, \vec{\xi})$ for which $|\vec{\xi}| \leq c\tau$, can be reached with velocities smaller than the speed of light when starting from the origin. Since there is no velocity larger than the speed of light, the region outside the light cone cannot be reached by any signal. Hence, this region is causally not connected

with the origin $(,\vec{0})$. In contrast, the region inside the light cone is causally connected.
These two regions clearly are distinct, when seen from the origin: Either they are causally connected or not. At that point it should be noted that similar reasoning holds also true, when considering negative times, i.e. the past. Again there is a light cone; events that took place inside the light cone may have a causal effect on events at the origin and events outside the light cone cannot have any effect. These regions can be separated by one equation, namely

$$c^2t^2 - \vec{x}^2 \begin{cases} < 0 & \text{outside the light cone, no causal connection} \\ > 0 & \text{inside the light cone, possible causal connection} \end{cases} . \qquad (1.51)$$

The form on the l.h.s. of the equation above will appear all over again. It can be identified with the distance s of an event from the origin,

$$s^2 = c^2t^2 - \vec{x}^2 \,. \qquad (1.52)$$

It is simple to show that this distance remains invariant under Lorentz transformations. Setting $y' = y = 0$ and $z' = z = 0$, s and s' are solely given by the t and x components, namely

$$s^2 = c^2t^2 - x^2 \qquad (1.53)$$

and

$$\begin{aligned} s'^2 &= c^2t'^2 - x'^2 \\ &= \frac{1}{1 - v^2/c^2}\left[c^2\left(t - \frac{vx}{c^2}\right)^2 - \left(x - vt\right)^2\right] \\ &= \frac{c^2t^2 - v^2t^2 + v^2/c^2x^2 - x^2}{1 - v^2/c^2} = \frac{c^2t^2(1 - v^2/c^2) - x^2(1 - v^2/c^2)}{1 - v^2/c^2} \\ &= c^2t^2 - x^2 \,. \end{aligned} \qquad (1.54)$$

This translates into the statement that the separation of causal and a-causal regions remains untouched by Lorentz-transformations, i.e. by changes of reference frame. Events that are causally connected remain connected, events that are not connected remain unconnected. The light cone itself, separating these two regimes can be described by $s^2 = 0$. According to the sign of s^2 the notion is that $s^2 > 0$ describes **time-like** distances, whereas $s^2 < 0$ describes **space-like** distances; $s^2 = 0$ describes **light-like** distances.

Chapter 2

Mathematical formulation

This chapter introduces the mathematical formulation of special relativity, most notably by means of four-vector and tensors. The general form of Lorentz transformations will be constructed, leading to rotations and boosts. Equipped with these transformations, this chapter then dwells on the symmetry of space-time as represented in the Poincare group, where the concept of generators will be introduced.

2.1 Four-vectors and Lorentz transformations

2.1.1 Introducing four-vectors and the metric tensor

Relativistic vs. non-relativistic distance

In the previous chapter, a stunning observation has been made, namely that the **distance in space-time** s of an event marked by $(t,\,\vec{x})^T$ from the origin is given by

$$s^2 = c^2t^2 - \vec{x}^2\,, \tag{2.1}$$

cf. Eq. (1.52). This automatically implies that the squared distance between two events reads

$$s^2 = c^2(t_x - t_y)^2 - (\vec{x} - \vec{y})^2\,. \tag{2.2}$$

This looks a bit surprising, doesn't it?

In order to see the difference more pronounced, remember that the distance of an event in Galilean space-time decomposed into a difference in time, and a distance in space,

$$d^2 \;=\; |\vec{x}|^2 = (x_1^2 + x_2^2 + x_3^2) = \sum_{i=1}^{3} x_i^2 \,. \tag{2.3}$$

Introducing **Einstein's convention of summing over repeated indices** this reads

$$d^2 = x_i^2 \,. \tag{2.4}$$

In this convention, for instance, the scalar product of two vectors is given by

$$\vec{x}\cdot\vec{y} \;=\; \sum_{i=1}^{3} x_i y_i = x_i y_i = x_i y_i \delta_{ij} \,. \tag{2.5}$$

The kth component of the vector product is

$$\vec{x}\times\vec{y}|_k = \sum_{i=1}^{3} x_i y_j \epsilon_{ijk} = x_i y_j \epsilon_{ijk} \,. \tag{2.6}$$

Contra- and covariant four-vectors

It would in principle be simple to introduce an object with four rather than three components, such that

$$x^\mu = (ct,\, x_1,\, x_2,\, x_3)^T \tag{2.7}$$

Such vectors conveniently are dubbed **four-vectors** in order to distinguish them from the ordinary three-vectors made up of the spatial components only. It has become a widely accepted convention that Greek indices such as μ, ν etc. denote the four components of a four-vector, i.e. $\mu,\, \nu = 0,\, 1,\, 2,\, 3$, whereas Latin indices like i or j denote the indices of a three-vector, i.e. $i,\, j = 1,\, 2,\, 3$. However, using Einsteins convention to write a four-vector's distance from the origin would yield

$$d^2 = x^\mu x^\mu = c^2 t^2 + x_1^2 + x_2^2 + x_3^2 \,, \tag{2.8}$$

i.e. a wrong sign in front of the spatial components. The cure is simple but awkward when first looking at it. In addition to x^μ another four-vector x_μ is introduced, such that by

$$x_\mu = (ct,\, -x_1,\, -x_2,\, -x_3)^T \tag{2.9}$$

the missing sign in front of the spatial components is reproduced if distances are defined through

$$d^2 = x^\mu x_\mu = x_\mu x^\mu = c^2t^2 - x_1^2 - x_2^2 - x_3^2\,. \tag{2.10}$$

These two forms of the vector are called **contravariant** (x^μ) and **covariant** (x_μ) vectors. Obviously, when **multiplying them it does not matter in which sequence they are written**, it is just important to **make sure that the relevant index to be summed over appears twice**: once as upper (contravariant) and once as lower (covariant) index. This, the position of the index, is where the difference occurs. In order to define the scalar product of four-vectors in a Lorentz-invariant fashion, it is important to use co- and contravariant vectors.

The metric

They are transformed into each other by a matrix, given by

$$g_{\mu\nu} = g^{\mu\nu} \tag{2.11}$$

such that

$$x_\mu = g_{\mu\nu}x^\nu \quad \text{and} \quad x^\mu = g^{\mu\nu}x_\nu\,. \tag{2.12}$$

From

$$x_\mu = g_{\mu\nu}x^\nu = g_{\mu\nu}g^{\nu\rho}x_\rho \tag{2.13}$$

it is simple to see that

$$g_{\mu\nu}g^{\nu\rho} = \delta_\mu{}^\rho\,, \tag{2.14}$$

where $\delta_\mu\rho$ denotes the components of the unity matrix in four dimensions,

$$\left(\delta_\mu{}^\rho\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =: \left(g_\mu{}^\rho\right)\,. \tag{2.15}$$

Written in matrix form g, the components $g_{mu\nu}$ can be read off

$$(g_{\mu\nu}) = (g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{2.16}$$

where the first index, μ, denotes the row and the second index ν denotes the column.
This immediately implies that the distance is given by

$$d^2 = x^\mu g_{\mu\nu} x^\nu = g_{\mu\nu} x^\mu x^\nu = g^{\mu\nu} x_\mu x_\nu\,. \tag{2.17}$$

This makes once more explicit that the matrix g can be used to lower or raise indices, i.e. to make covariant indices out of contravariant indices and vice versa. The occurrence of this matrix in the definition of the distance motivates the usage of the term **metric tensor** [1] for it.

2.1.2 Lorentz transformations represented by tensors

Construction of a simple boost

Let us now try to write Lorentz transformations as tensors - objects with Lorentz indices - acting on four vectors. Closer inspection of Eq. (1.31) shows that the transformations are linear,

$$\begin{aligned} ct' &= \frac{ct - \frac{u}{c}x}{\sqrt{1-\frac{u^2}{c^2}}} \\ x' &= \frac{x - \frac{u}{c}t}{\sqrt{1-u^2/c^2}} \\ y' &= y \\ z' &= z\,. \end{aligned} \tag{2.18}$$

This implies that they can be represented by a matrix, $\Lambda^\mu{}_\nu$ such that

$$x^\mu \longrightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu\,. \tag{2.19}$$

[1]Here the term "metric" hindsights its special role in the definition of the distance (which can be measured, hence "metric") and the term tensor reflects a property under transformations, which will be discussed later on.

Comparing coefficients, it quickly becomes apparent that for Lorentz transformations along the x-axis with velocity u_x

$$\Lambda^\mu{}_\nu(u_x) = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{u_x^2}{c^2}}} & -\frac{u_x}{c\sqrt{1-\frac{u_x^2}{c^2}}} & 0 & 0 \\ -\frac{u_x}{c\sqrt{1-\frac{u_x^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{u_x^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta_x\gamma & 0 & 0 \\ -\beta_x\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.20}$$

where

$$\beta_x = \frac{u_x}{c} \quad \text{and} \quad \gamma = \frac{1}{\sqrt{1-\beta_x^2}} \tag{2.21}$$

has been introduced. Similarly, Lorentz-transformations along the y or z-axis read

$$\begin{aligned} \Lambda^\mu{}_\nu(u_y) &= \begin{pmatrix} \frac{1}{\sqrt{1-\frac{u_y^2}{c^2}}} & 0 & -\frac{u_y}{c\sqrt{1-\frac{u_y^2}{c^2}}} & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{u_y}{c\sqrt{1-\frac{u_y^2}{c^2}}} & 0 & \frac{1}{\sqrt{1-\frac{u_y^2}{c^2}}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \gamma & 0 & -\beta_y\gamma & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_y\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ \Lambda^\mu{}_\nu(u_z) &= \begin{pmatrix} \frac{1}{\sqrt{1-\frac{u_z^2}{c^2}}} & 0 & 0 & -\frac{u_z}{c\sqrt{1-\frac{u_z^2}{c^2}}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{u_z}{c\sqrt{1-\frac{u_z^2}{c^2}}} & 0 & 0 & \frac{1}{\sqrt{1-\frac{u_z^2}{c^2}}} \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta_z\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta_z\gamma & 0 & 0 & \gamma \end{pmatrix}. \end{aligned} \tag{2.22}$$

Here, γ is to be suitably chosen as

$$\gamma = \frac{1}{\sqrt{1-\beta_{x,\,y,\,z}^2/c^2}}. \tag{2.23}$$

If the Lorentz-transformations are not along any of the axes but rather along an arbitrary vector, the transformation is thus given by

$$\Lambda^\mu{}_\nu(\vec{u}) = \begin{pmatrix} \gamma & -\gamma\vec{\beta}^T \\ -\gamma\vec{\beta} & \gamma\cdot\mathbf{1} \end{pmatrix}, \tag{2.24}$$

where the spatial components are not explicitly written out and **1** is the unity 3×3 matrix.
It should be noted that the sequence of indices matters:

$$\Lambda^{\mu}{}_{\nu}(\vec{u}) \neq \Lambda_{\nu}{}^{\mu}(\vec{u}) \,. \tag{2.25}$$

Inverse Lorentz transformation

Before proceeding, let us ask, how to construct the inverse of a given Lorentz transformation, Λ^{-1}. To this end, consider

$$x^{\mu} = \Lambda^{-1\mu}{}_{\nu}(\vec{u})\Lambda^{\nu}{}_{\rho}(\vec{u})x^{\rho} \,, \tag{2.26}$$

which should hold true for all four-vectors x and all transformations specfied by $\vec{u}$. This immediately implies that

$$\Lambda^{-1\mu}{}_{\nu}(\vec{u})\Lambda^{\nu}{}_{\rho}(\vec{u}) = \delta^{\mu\rho} \,. \tag{2.27}$$

It is simple to check that this is fulfilled by

$$\Lambda^{-1\mu}{}_{\nu}(\vec{u}) = \Lambda^{\mu}{}_{\nu}(-\vec{u}) \,. \tag{2.28}$$

Invariance of the metric

In Sec. 1.2.2 the invariance of the distance of a vectors to the origin under Lorentz-transformations has been shown, i.e.

$$s^2 = c^2t^2 - \vec{x}^2 = c^2t'^2 - \vec{x}'^2 = s'^2 \,. \tag{2.29}$$

Let us write this now in terms of four-vectors. The distance in the original, unprimed system reads

$$s^2 = x^{\mu}x_{\mu} = x^{\mu}x^{\nu}g_{\mu\nu} \,. \tag{2.30}$$

Applying a Lorentz- transformation along the x axis with velocity u_x, the four-vector in the primed system is given by

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}(u_x)x^{\nu} \tag{2.31}$$

and the distance thus is

$$s^2 = x'^{\mu}x'{}_{\mu} = x'^{\mu}x'^{\nu}g'_{\mu\nu} = x^{\rho}\Lambda^{\mu}{}_{\rho}(u_x)x^{\kappa}\Lambda^{\nu}{}_{\kappa}(u_x)g'_{\mu\nu} \,. \tag{2.32}$$

Invariance of the distance then implies that

$$g_{\rho\kappa} = \Lambda^{\mu}{}_{\rho}(u_x) g'_{\mu\nu} \Lambda^{\nu}{}_{\kappa}(u_x) \tag{2.33}$$

or, in matrix form

$$g' = \Lambda^T(-u_x) g \Lambda(-u_x)\,. \tag{2.34}$$

Here the transposition is due to the fact that the first index of $\Lambda^{\mu}{}_{\rho}(u_x)$ is contracted with the first index of $g_{\mu\nu}$. Anyways, writing this out results in

$$\begin{pmatrix} \gamma & \beta_x\gamma & 0 & 0 \\ \beta_x\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \gamma & \beta_x\gamma & 0 & 0 \\ \beta_x\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

as left hand side of the equation above. Straightforward calculation yields

$$\begin{aligned} g' &= \begin{pmatrix} \gamma^2 - \beta_x^2\gamma^2 & -\beta_x\gamma^2 + \beta_x\gamma^2 & 0 & 0 \\ -\beta_x\gamma^2 + \beta_x\gamma^2 & \beta_x^2\gamma^2 - \gamma^2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = g \end{aligned} \tag{2.35}$$

because

$$\beta_x^2 - 1 = -\gamma^{-2}\,. \tag{2.36}$$

This shows that in addition to the distance also the metric tensor g indeed is invariant under Lorentz-transformations. In fact, invariance of the one automatically implies invariance of the other.
This statement can be used in order to come to another, more formal definition of Lorentz-transformations: **Lorentz-transformations are those transformations that leave the distance s^2 and hence the metric $g_{\mu\nu}$ invariant.**

Rotations

In addition to the Lorentz-transformations, which mix spatial and temporal components of the four-vectors there are also rotations which map the spatial components onto each other. Since they also leave the metric invariant (remember: rotations leave spatial distances invariant!) they can be understood as Lorentz-transformations, given by

$$\begin{aligned}
\Lambda^\mu{}_\nu(\theta_x) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_x & \sin\theta_x \\ 0 & 0 & -\sin\theta_x & \cos\theta_x \end{pmatrix}\,, \\
\Lambda^\mu{}_\nu(\theta_y) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_y & 0 & -\sin\theta_y \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta_y & 0 & \cos\theta_y \end{pmatrix}\,, \\
\Lambda^\mu{}_\nu(\theta_z) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_z & \sin\theta_z & 0 \\ 0 & -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,.
\end{aligned} \tag{2.37}$$

Obviously, these transformations map $\vec{x}$ onto another vector, $\vec{x'}$, which is rotated by θ_x, θ_y, or θ_z around the x-, y-, or z-axis, respectively. In order to distinguish the different types of Lorentz-transformations, the ones in the previous section, mixing temporal and spatial components, are called **boosts** or **Lorentz-boosts**, the kind discussed here are just the rotations.

Lorentz transformation, once again

Consider again Eq. (2.36). Written in another way, it can be cast into

$$\gamma^2 - \gamma^2\beta^2 = 1\,. \tag{2.38}$$

This has some similarity with

$$\begin{aligned}
\cosh^2\eta - \sinh^2\eta &= \left(\frac{e^\eta + e^{-\eta}}{2}\right)^2 - \left(\frac{e^\eta - e^{-\eta}}{2}\right)^2 \\
&= \frac{(e^{-2\eta} - 2 + e^{2\eta}) - (e^{-2\eta} + 2 + e^{2\eta})}{4} = 1\,.
\end{aligned} \tag{2.39}$$

Expressing η through γ and β yields

$$\tanh\eta = \frac{\sinh\eta}{\cosh\eta} = \frac{\beta\gamma}{\gamma} = \beta \tag{2.40}$$

or

$$\cosh\eta = \gamma\,. \tag{2.41}$$

η is called the **rapidity** related to this Lorentz-boost.

Sequence of boosts

Having at hand the matrix representation of the simple boosts along the x-, y-,nd z-axes, it is simple to calculate the effect of sequential boosts, for instance of two of them. This can be done by realising that

$$x''^{\mu} = \Lambda^{\mu}{}_{\nu}(\vec{u}_1)x'^{\nu} = \Lambda^{\mu}{}_{\nu}(\vec{u}_1)\Lambda^{\nu}{}_{\rho}(\vec{u}_2)x^{\rho}\,, \tag{2.42}$$

implying that such sequences are given by the multiplication of two Lorentz-transformations. Thus, the following transformation matrix emerges

$$\Lambda^{\mu}{}_{\rho}(\vec{u}_1\otimes\vec{u}_2) = \Lambda^{\mu}{}_{\nu}(u_1\vec{e}_x)\Lambda^{\nu}{}_{\rho}(u_2\vec{e}_x)\,, \tag{2.43}$$

and it is interesting to check its properties.
Let us now consider two cases:

1. $\vec{u}_1 \parallel \vec{u}_2 \parallel \vec{e}_x$ (two boosts along the same direction, here the x-axis): The overall transformation is given by

$$\begin{aligned}
&\left(\Lambda^{\mu}{}_{\rho}(\vec{u}_1\otimes\vec{u}_2)\right)\\
&= \begin{pmatrix} \gamma_1 & -\beta_1\gamma_1 & 0 & 0\\ -\beta_1\gamma_1 & \gamma_1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix} \gamma_2 & -\beta_2\gamma_2 & 0 & 0\\ -\beta_2\gamma_2 & \gamma_2 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\\
&= \begin{pmatrix} \gamma_1\gamma_2(1+\beta_1\beta_2) & -\gamma_1\gamma_2(\beta_1+\beta_2) & 0 & 0\\ -\gamma_1\gamma_2(\beta_1+\beta_2) & \gamma_1\gamma_2(1+\beta_1\beta_2) & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}
\end{aligned} \tag{2.44}$$

Straightforward calculation shows that this indeed is a boost along the x-axis with velocity u_3. The relevant velocities can be added,

$$\beta_3 = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}\,, \tag{2.45}$$

in full accordance with the finding of Chapter 1, cf. Eq. (1.35). It is trivial to check that also γ_3 complies with this added velocity.

2. $\vec{u}_1 \parallel \vec{e}_x$ and $\vec{u}_2 \parallel \vec{e}_y$ (two boosts along orthogonal directions): Here, the overall transformation is given by

$$\begin{aligned}
&\left(\Lambda^\mu{}_\rho(\vec{u}_1 \otimes \vec{u}_2)\right) \\
&= \begin{pmatrix} \gamma_1 & -\beta_1\gamma_1 & 0 & 0 \\ -\beta_1\gamma_1 & \gamma_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \gamma_2 & 0 & -\beta_2\gamma_2 & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_2\gamma_2 & 0 & \gamma_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma_1\gamma_2 & -\gamma_1\beta_1 & -\gamma_1\gamma_2\beta_1 & 0 \\ -\gamma_1\gamma_2\beta_1 & \gamma_1 & \gamma_1\gamma_2\beta_1\beta_2 & 0 \\ -\gamma_2\beta_2 & 0 & \gamma_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned} \tag{2.46}$$

It can be shown that this expression is identical to one which emerges after a sequence of one boost and one rotation. But before addressing this issue in the next section, it is worthwhile to note that here the sequence of actions matters. If, instead first the boost along the x- and then the one along the y-axis is performed, the resulting transformation is given by

$$\begin{aligned}
&\left(\Lambda^\mu{}_\rho(\vec{u}_2 \otimes \vec{u}_1)\right) \\
&= \begin{pmatrix} \gamma_2 & 0 & -\beta_2\gamma_2 & 0 \\ 0 & 1 & 0 & 0 \\ -\beta_2\gamma_2 & 0 & \gamma_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \gamma_1 & -\beta_1\gamma_1 & 0 & 0 \\ -\beta_1\gamma_1 & \gamma_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma_1\gamma_2 & -\gamma_1\gamma_2\beta_1 & -\gamma_2\beta_2 & 0 \\ -\gamma_1\beta_1 & \gamma_1 & 0 & 0 \\ -\gamma_1\gamma_2\beta_1 & \gamma_1\gamma_2\beta_1\beta_2 & \gamma_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
\end{aligned} \tag{2.47}$$

clearly different from the result obtained before. Closer inspection reveals that in this case

$$\left(\Lambda^{\mu}{}_{\rho}(\vec{u}_2 \otimes \vec{u}_1)\right) = \left(\Lambda^{\mu}{}_{\rho}(\vec{u}_1 \otimes \vec{u}_2)\right)^T . \tag{2.48}$$

Note that this behaviour of the boosts is quite analogous to the case of rotations: Two rotations around the same axis can be replaced by a third one around this axis, and there also the sequence does not matter. If instead, the rotations are around different axis, the sequence matters.

2.2 Further reading: Infinitesimal transformations

2.2.1 Lorentz group

Definition

Let us now **define general Lorentz-transformations**: It is the set of all transformations which can be obtained through **arbitrary sequential application of rotations and boosts around or along the three axes**. Stated in more mathematical fashion: It is the set of all transformations Λ which can be applied on four-vectors x and leave their norm invariant.

$$\left\{\Lambda : x^{\mu} \to x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu},\, x^2 = x^{\mu}x_{\mu} = x'^{\mu}x'_{\mu} = x'^2\right\} . \tag{2.49}$$

In order to gain some more insight into the general structure of Lorentz-transformations let us state a few of their properties:

- Closure:
  Applying two Lorentz-transformations in sequence yields a new one. It suffices to check this for any sequence of two: two boosts, two rotations, or one boost and one rotation.

- Existence of identity:
  This is simple. The corresponding matrix is just the four-dimensional unity matrix.

- Existence of inverse element:
  It suffices to show that an inverse element exists for all individual building blocks, rotations and boosts around or along the three axes. This is trivial.

- Associativity:
  Associativity of Lorentz transformations is guaranteed through the associativity of matrix multiplication.

These properties imply that the Lorentz transformations form a group. Since, however, the sequence of the application of two Lorentz-transformations matters, commutativity is not realised and thus this is not an Abelian group. In fact, since the elements of this group are characterised by continuous parameters, he Lorentz group is a **Lie group**.

Properties

Let us go back to the invariance of the metric, cf. Eqs. (2.33) and (2.34). Taking the determinant of the second equation,

$$g' = \Lambda^T g \Lambda \Longrightarrow \det(g') = \det(\Lambda^T g \Lambda) = \det(g)\,[\det(\Lambda)]^2 = -1 \tag{2.50}$$

implies that

$$[\det(\Lambda)]^2 = 1 \quad \text{or} \quad \det(\Lambda) = \pm 1\,. \tag{2.51}$$

Using now the 00-component of Eq. (2.33),

$$g'_{00} = 1 = \Lambda^{\mu}{}_{0} g_{\mu\nu} \Lambda^{\nu}{}_{0} = (\Lambda^{0}{}_{0})^2 - \sum_{i=1}^{3} (\Lambda^{i}{}_{0})^2 \tag{2.52}$$

immediately implies that

$$(\Lambda^{0}{}_{0})^2 \geq 1 \tag{2.53}$$

or

$$\Lambda^{0}{}_{0}) \geq 1 \quad \text{or} \quad \Lambda^{0}{}_{0}) \leq 1\,. \tag{2.54}$$

Thus the Lorentz-transformations fall in either of the four sets:

- $\det(\Lambda) = +1$ and $\Lambda^0{}_0) \geq 1$
- $\det(\Lambda) = -1$ and $\Lambda^0{}_0) \geq 1$
- $\det(\Lambda) = -1$ and $\Lambda^0{}_0) \leq -1$
- $\det(\Lambda) = +1$ and $\Lambda^0{}_0) \leq -1$

Using the =-sign for the 00-component yields four transformations which can be assigned a clear meaning, namely

- $\det(\Lambda) = +1$ and $\Lambda^0{}_0) = 1$:

$$\Lambda_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad \text{the identity} \, . \tag{2.55}$$

- $\det(\Lambda) = -1$ and $\Lambda^0{}_0) = 1$

$$\Lambda_S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \quad \text{the space-inversion} \, . \tag{2.56}$$

- $\det(\Lambda) = -1$ and $\Lambda^0{}_0) = -1$

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad \text{the time-reversal} \, . \tag{2.57}$$

- $\det(\Lambda) = +1$ and $\Lambda^0{}_0) = -1$

$$\Lambda_{ST} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \quad \text{the total inversion} \, . \tag{2.58}$$

These four special transformations in fact form an **Abelian subgroup of the Lorentz group**.

The last property concerns the number of free parameters of the elements of the Lorentz group. In principle, the transformation matrices are real 4×4-matrices, hence, naively one would expect that there are 16 parameters needed to describe them. However, the symmetry of Eq. (2.33),

$$g_{\rho\kappa} = \Lambda^{\mu}{}_{\rho}(u_x) g'_{\mu\nu} \Lambda^{\nu}{}_{\kappa}(u_x) \tag{2.59}$$

under exchange of ρ and κ plus the constraint of the determinant imply that after all there are only 6 parameters which can independently be chosen. As will be seen later, they are connected with boost parameters along the three axes and rotation angles around them.

2.2.2 Boosts and rotations

Will be continued

Chapter 3

Relativistic mechanics

This chapter deals with relativistic mechanics. First of all, the basic objects, such as mass of a point particle and its position are introduced and their transformation properties are discussed. Then Newton's mechanics, based on the connection of forces acting on point particles and their corresponding change of momentum will be reformulated in a relativistic way. This is highlighted by a few examples, before some care will be taken in discussing the Lorentz force, related to the effect of electric and magnetic fields on charged particles.

3.1 Basic objects

3.1.1 Mass

$E = mc^2$

In 1905, Einstein published his five papers on the "Electrodynamics of Moving Bodies", the third of which is titled "Special Relativity". It is in this publication where he firstly deduced the indisputably most famous equation of physics,

$$E = mc^2 \,. \tag{3.1}$$

In the following, this formula will be motivated here, using some basic knowledge of quantum mechanics.
To this end, assume an excited photon to fall back into its ground state under emission of two photons. In the atom's rest frame, the two photons have the

frequencies $\nu_{1,2}$ and thus they have energies

$$\epsilon_{1,2} = 2\pi\hbar\nu_{1,2} \,. \tag{3.2}$$

To keep things simple let us assume that in the atom's rest frame the two energies are identical, i.e.

$$\epsilon_1 = \epsilon_2 = \epsilon \quad \text{and} \quad \nu_1 = \nu_2 = \nu \tag{3.3}$$

and that the photons are emitted back to back. This implies that the atom does not experience any recoil. Denoting the energy of the excited state by $E_{\rm exc}$ and the ground state energy by $E_{\rm gr}$ the two energies are thus related through

$$E_{\rm exc} = E_{\rm gr} + 2\epsilon \,. \tag{3.4}$$

Now let us consider the same decay in another frame, moving with velocity v along the direction of one of the two photons. The frequencies of the two photons are then Doppler-shifted, such that in this frame

$$\nu'_{1,2}(v) = \nu'_{\pm}(v) = \frac{1 \pm \beta}{\sqrt{1+\beta^2}}\nu \,, \tag{3.5}$$

where $\beta = v/c$. Here the $\pm$ sign refers to the photon moving in the direction of v (-) or against it, respectively (+) - the photons are just red- and blue-shifted. The two energies are thus given by

$$\epsilon'_{1,2}(v) = \epsilon'_{\pm}(v) = \frac{1 \pm \beta}{\sqrt{1+\beta^2}}\epsilon \tag{3.6}$$

and the energies in this frame are related by

$$E'_{\rm exc}(v) = E'_{\rm gr}(v) + \epsilon'_{+}(v) + \epsilon'_{-}(v) = E'_{\rm gr}(v) + \frac{2\epsilon}{\sqrt{1+\beta^2}} \,. \tag{3.7}$$

Thus the energy difference of excited and ground state in the two systems reads

$$\Delta E = E_{\rm exc} - E_{\rm gr} = 2\epsilon \tag{3.8}$$

and

$$\Delta E'(v) = E'_{\rm exc}(v) - E'_{\rm gr}(v) = \frac{2\epsilon}{\sqrt{1+\beta^2}} \,, \tag{3.9}$$

respectively - a difference given by a factor of $1/\sqrt{1+\beta^2}$! Subtracting both equations and reshuffling yields

$$\begin{aligned}\Delta E'(v) - \Delta E &= [E'_{\mathrm{exc}}(v) - E'_{\mathrm{gr}}(v)] - [E_{\mathrm{exc}} - E_{\mathrm{gr}}] \\ &= [E'_{\mathrm{exc}}(v] - E_{\mathrm{exc}}) - [E'_{\mathrm{gr}}(v) - E_{\mathrm{gr}}] = 2\epsilon\left(\frac{1}{\sqrt{1+\beta^2}} - 1\right) . \quad (3.10)\end{aligned}$$

An observer in the primed system may measure the kinetic energy of the atom before and after the photons have been emitted. The kinetic energy of course is just given by the effect of motion, i.e. it is given by the difference of the respective energies in the primed and unprimed system:

$$E_{\mathrm{kin,exc}} = E'_{\mathrm{exc}}(v) - E_{\mathrm{exc}} \quad (3.11)$$

and

$$E_{\mathrm{kin,0}} = E'_{\mathrm{gr}}(v) - E_{\mathrm{gr}} , \quad (3.12)$$

respectively. Thus,

$$E_{\mathrm{kin,exc}} - E_{\mathrm{kin,0}} = 2\epsilon\left(\frac{1}{\sqrt{1+\beta^2}} - 1\right) . \quad (3.13)$$

This is awkward, since it seems that the emission of the photons changed the kinetic energy of the atom although by construction they did not transfer any net momentum in the atom's rest frame.
To clarify this situation, consider the non-relativistic limit, i.e. $v \ll c$. Then, keeping only leading terms in v/c

$$E_{\mathrm{kin,exc}} = \frac{1}{2} m_{\mathrm{exc}} v^2 \quad (3.14)$$

and

$$E_{\mathrm{kin,0}} = \frac{1}{2} m_{\mathrm{gr}} v^2 . \quad (3.15)$$

Since v is fixed through the situation, the only chance to accommodate for the change in kinetic energy is by assuming a change of the atom's mass through the emission of the photons. In the same limit

$$2\epsilon\left(\frac{1}{\sqrt{1+\beta^2}} - 1\right) = 2\epsilon\left(\frac{1}{2}\beta^2\right) = \epsilon\left(\frac{v}{c}\right)^2 \quad (3.16)$$

and hence

$$\Delta m = m_{\rm exc} - m_{\rm gr} = \frac{2\epsilon}{c^2}\,. \tag{3.17}$$

This implies that energies and masses are connected through a factor c^2. To see this better, assume that the ground state of the atom is massless, clearly unphysical, but a good way to highlight the idea. Then the mass of the excited state is given by its energy times c^2. In turn this implies that any particle at rest **must** have a well-defined **energy at rest**, namely

$$E = m_0 c^2\,, \tag{3.18}$$

where m_0 denotes the **rest mass**.
This can now be used to find the energies in the primed system. Let us consider for instance the ground state energy. From Eq. (3.17)

$$m_{\rm exc}c^2 - m_{\rm gr}c^2 = 2\epsilon \tag{3.19}$$

and Eq. (3.9),

$$E'_{\rm exc}(v) - E'_{\rm gr}(v) = \frac{2\epsilon}{\sqrt{1+\beta^2}}\,,$$

it is quite straightforward to conclude that

$$m'_{\rm gr}(v) = \frac{m_{\rm gr}}{\sqrt{1+\beta^2}}\,, \tag{3.20}$$

i.e. that the mass of a physical object changes when moving.
To summarise: Physical objects have a **well-defined rest mass**, m_0. It is a property of the object in question and thus independent of the choice of inertial system. However, the mass appearing in energy balances etc. is dependent on the choice of inertial system and thus scales like Eq. (3.20).

Measuring the mass

The natural question now arises which mass is actually measured? Is it the rest mass, commonly denoted by m_0 or is it $m(v) = m/\sqrt{1+\beta^2} = E(v)/c^2$? As long as the measurement apparatus moves in parallel with the body, this is not complicated to answer: The measurement will then yield m.

The situation naturally is by far more complicated when the apparatus is at rest and the body moves. Investigating this situation in more detail lead Einstein to his theory of general relativity - gravitation acts on the total energy of a body and not on its rest mass. Hence, the weigh will measure $E(v)/c^2$ rather them m.

3.1.2 Kinematic quantities

Four-momentum

Recall that the position of an event in four-dimensional space-time is given by a contravariant four-vector

$$x^\mu = (ct,\, \vec{x})^T = (ct,\, x_1,\, x_2,\, x_3)^T\,. \tag{3.21}$$

The norm of this four-vector, i.e. its distance to the origin reads

$$|x|^2 = x^2 = x^\mu x_\mu = c^2t^2 - \vec{x}^2 \tag{3.22}$$

and is invariant under Lorentz-transformations. Accordingly, the distance of two events

$$s^2 = (x-y)^2 = (x^\mu - y^\mu)(x_\mu - y_\mu) \tag{3.23}$$

is also a Lorentz-invariant quantity.
In the following a four-momentum corresponding to this way of defining positions will be constructed. Naively, the spatial components of momentum are given though

$$\vec{p} = m_0\vec{v}\,, \tag{3.24}$$

clearly dependent on the velocity and thus on the frame. Similarly, the energy is also frame-dependent,

$$E(v) = \frac{m_0c^2}{\sqrt{1+v^2/c^2}}\,, \tag{3.25}$$

cf. Eq. (3.20).
In order to construct a contravariant four-momentum p^μ (or its covariant counterpart p_μ) it makes perfect sense to **demand** that the invariant quantity

is related to the other invariant related to the physical object, namely its rest mass m_0:

$$p^\mu p_\mu := m_0^2 c^2 \,, \tag{3.26}$$

where the factor c^2 is mandatory for dimensional reasons. Of course the spatial components should be given by $\vec{p}$, so the only remaining question is what is the 0-component of the four-momentum? In order to answer this, let us go to the object's rest frame, where of course $\vec{p} = 0$. There

$$p^\mu p_\mu = p_0^2 - \vec{p}^2 = m_0^2 c^2 \,, \tag{3.27}$$

leading to

$$p_0 = m_0 c = m_0 c^2/c = E_0/c \,. \tag{3.28}$$

Here the positive sign has been chosen, ensuring that the 0-component of the four-momentum is strictly positive. In order to find the four-momentum of a moving object, a Lorentz-transformation has to be applied. Consider a particle moving with $\vec{v} = v\vec{e}_z$ along the positive z-axis. Then the Lorentz-transformation that has to be applied is $\Lambda(-\vec{v})$,

$$p^\mu \longrightarrow \Lambda^\mu{}_\nu(-\vec{v}) p^\nu \tag{3.29}$$

and thus

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} 1 & 0 & 0 & \frac{v}{c} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{v}{c} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} m_0 c \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{m_0 c}{\sqrt{1-\frac{v^2}{c^2}}} \\ 0 \\ 0 \\ \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix} \tag{3.30}$$

must be calculated. Hence, in the general case, this results in

$$p^\mu = (E/c, \vec{p})^T \,, \tag{3.31}$$

where

$$E = E(v) = \frac{m_0 c^2}{\sqrt{1-\frac{v^2}{c^2}}} \tag{3.32}$$

and

$$\vec{p} = \frac{m_0 \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3.33}$$

depend on the velocity. The invariant p^2 is simple to calculate:

$$p^2 = p^\mu p_\mu = E^2/c^2 - \vec{p}^2 = \frac{m_0^2(c^2 - v^2)}{1 - v^2/c^2} = m_0^2 c^2 \tag{3.34}$$

as demanded.

Four-velocity

The four-momentum can be used to define a four-velocity u^μ through

$$p^\mu = m_0 u^\mu \,. \tag{3.35}$$

Obviously,

$$u^\mu = \left(\frac{c}{\sqrt{1 - v^2/c^2}}, \frac{\vec{v}}{\sqrt{1 - v^2/c^2}} \right)^T \,. \tag{3.36}$$

Since both p^2 and m_0 are invariants, also u^2 is a Lorentz-invariant quantity, namely

$$u^2 = u^\mu u_\mu = \frac{c^2 - v^2}{1 - v^2/c^2} = c^2 \,. \tag{3.37}$$

Causality dictates that both four-momentum and four-velocity are time-like, i.e. both $p^2 \geq 0$ and $u^2 \geq 0$[1] and therefore their 0-components need both to be larger than 0.
What is the interpretation of this four-velocity? In order to see this, **define** the **eigen-time of an object** τ through

$$s^2 = x^\mu x_\mu =: c^2 \tau^2 \,, \tag{3.38}$$

[1] If this was not the case, one could maybe use an object with space-like four-momentum as messenger between two events, thus ignoring c as maximal speed of signal transmission.

where, as before

$$x^\mu = (ct,\, \vec{x})^T\,. \tag{3.39}$$

Let us now convince ourselves that the four-velocity is the derivative of x w.r.t. the eigentime. To this end, use

$$\mathrm{d}x^\mu = (c\mathrm{d}t,\, \mathrm{d}\vec{x}) \tag{3.40}$$

and

$$d\tau = \mathrm{d}s/c\,. \tag{3.41}$$

Then,

$$u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau = \frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} ct \\ \vec{x} \end{pmatrix} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\begin{pmatrix} c \\ \mathrm{d}\vec{x}/\mathrm{d}t \end{pmatrix} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\begin{pmatrix} c \\ \vec{v} \end{pmatrix}. \tag{3.42}$$

Using the fact that

$$\mathrm{d}\tau^2 = \frac{1}{c^2}\left(c^2\mathrm{d}t^2 - \mathrm{d}\vec{x}^2\right) = \left(\mathrm{d}t^2 - \frac{\mathrm{d}\vec{x}^2}{c^2}\right) \tag{3.43}$$

and thus

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{\gamma}\,. \tag{3.44}$$

Hence,

$$u^\mu = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\begin{pmatrix} c \\ \vec{v} \end{pmatrix}. \tag{3.45}$$

Four-acceleration

Having defined four-velocity through differentiation of four-position w.r.t. the eigentime, it seems natural also to define four-acceleration through differentiation w.r.t. eigentime,

$$b^\mu = \mathrm{d}u^\mu/\mathrm{d}\tau\,. \tag{3.46}$$

However, before calcuating this explicitly, consider

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(u^\mu u_\mu\right) = \frac{\mathrm{d}u^2}{\mathrm{d}\tau} = \frac{\mathrm{d}c^2}{\mathrm{d}\tau} = 0 \tag{3.47}$$

and use the fact that

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(u^\mu u_\mu\right) = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}u_\mu + u^\mu\frac{\mathrm{d}u_\mu}{\mathrm{d}\tau} = 2\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}u_\mu = 2b^\mu u_\mu\,. \tag{3.48}$$

This immediately implies that four-acceleration and -velocity are orthogonal to each other,

$$b^\mu u_\mu = 0\,. \tag{3.49}$$

Returning to the definition of four-acceleration, Eq. (3.46), and actually differentiating yields

$$b^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = \left(\frac{\frac{\vec{v}\cdot\vec{a}}{c}}{\left(1-\frac{v^2}{c^2}\right)^2},\ \frac{\vec{a}}{1-\frac{v^2}{c^2}} + \frac{\vec{v}}{c}\frac{\frac{\vec{v}\cdot\vec{a}}{c}}{\left(1-\frac{v^2}{c^2}\right)^2}\right)^T\,, \tag{3.50}$$

where $\vec{a} = \mathrm{d}\vec{v}/\mathrm{d}t$ has been used. Clearly, four-acceleration has a spatial component pointing into the direction of the original three-velocity.

3.1.3 Four-momentum conservation

Before entering the realm of mechanics, it is worthwhile to study a few kinematical situations and in particular the effect of energy and momentum conservation, which can be nicely summarised as four-momentum conservation in the relativistic case. To this end, let us consider various production, decay and scattering processes. To simplify life, from now one, **a system of units will be used, in which** $c = 1$. This implies that mass, momentum and energy are characterised by the "same unit", usually taken to be one electron volt (eV), i.e. a unit of energy. Masses therefore are often given in units of eV/c^2, and momenta are then measured in eV/c. Also, velocites are now given in ratios of c, i.e. they have no physical unit and are just numbers in the range $v \in [0,\,1]$.

Two-body decays

To start with, consider the case, where a particle decays into two other particles. For obvious reasons, such a process can be investigated in the rest frame of the decaying particle and only later it can be boosted into any other system, where the decaying particle and hence the decay products move. Four-momentum conservation in any case implies that

$$P^\mu = p_1^\mu + p_2^\mu \tag{3.51}$$

or

$$\left(\begin{array}{c} M \\ \vec{0} \end{array}\right) = \left(\begin{array}{c} E_1 \\ \vec{p}_1 \end{array}\right) + \left(\begin{array}{c} E_2 \\ \vec{p}_2 \end{array}\right) . \tag{3.52}$$

Simple inspection shows that in the center-of-mass (c.m.) or rest frame of the decaying particle,

$$\vec{p}_1 = -\vec{p}_2 = \vec{p} \tag{3.53}$$

and

$$M = E_1 + E_2 > m_1 + m_2 . \tag{3.54}$$

Hence, the decay products individually and together **must be lighter than the deaying particle**. Also, in the c.m. frame of the decay, the two emerging particles are oriented back-to-back, i.e. the angle between their trajectories is π. Plugging in that

$$p^2 = E^2 - \vec{p}^2 = m^2 \tag{3.55}$$

cf. Eq. (3.34), and thus

$$E^2 = \vec{p}^2 + m^2 \tag{3.56}$$

for each of the two particles. This, together with $M = E_1 + E_2$ and the identity of the two momenta up to a sign yields

$$M^2 = 2\vec{p}^2 + m_1^2 + m_2^2 \tag{3.57}$$

and thus

$$|\vec{p}|^2 = \frac{(M^2 - m_1^2 - m_2^2)^2 - 4m_1^2 m_2^2}{4M^2} =: \lambda^2(M,\, m_1,\, m_2) . \tag{3.58}$$

In the limit where both decay products have vanishing mass, therefore

$$|\vec{p}| = \lambda(M,\,0,\,0) = M/2\,. \tag{3.59}$$

However, the energies of the two outgoing particles read

$$E_{1,2} = \frac{M^2 \pm (m_1^2 - m_2^2)}{2M}\,. \tag{3.60}$$

Examples for such reactions are the decay of charged pions into a muon and its neutrino, of a neutral pion into two photons, or of a Λ-baryon into a nucleon and a pion,

$$\pi^- \longrightarrow \mu^- + \nu_\mu\,,\quad \pi^0 \longrightarrow 2\gamma\,,\quad \text{or}\quad \Lambda \to p\pi^-,\, n\pi^0 \tag{3.61}$$

Let us, for the moment, focus on the first case, $\pi^- \longrightarrow \mu^- + \nu_\mu$. Since the neutrinos do interact only weakly (after all, they have no electrical charge and they are practically massless), the experimental signature of such a process will look like the trajectory of a charged particle having a kink. This kink marks the place, where the pion actually decayed. It is clear that the maximal angle of this kink is totally determined by the velocity of the incoming pion. To see this, convince yourself that the deflection angle θ' of the muon trajectory w.r.t. the pion trajectory is given by

$$\tan\theta' = \frac{p'_\perp}{p'_\|} = \frac{p'_\perp}{|\vec{p}'|^2 - {p'_\perp}^2}\,, \tag{3.62}$$

where $p'_\perp$ is the component orthogonal to the reference trajectory and $p'_\|$ is the parallel component. Obviously, the deflection angle is maximal if $p_\perp$ is maximal; this is the case, if in the rest frame of the decay the complete momentum of the muon is orthogonal w.r.t. the pion trajectory. Using that the neutrino is massless then yields in the (unprimed) c.m. system for the muon

$$\begin{aligned} E &= \frac{M_\pi^2 + m_\mu^2}{2M_\pi} \\ \vec{p} = p_\perp \vec{e}_\perp &= \frac{M_\pi^2 - m_\mu^2}{2M_\pi}\vec{e}_\perp\,. \end{aligned} \tag{3.63}$$

Boosts along the $\|$-direction do not change the $\perp$-component of the momentum, therefore

$$\begin{aligned} E' &= \gamma E \\ p'_\perp &= p_\perp \\ p'_\| &= \beta\gamma E \end{aligned} \tag{3.64}$$

and thus

$$\tan\theta'_{\max} = \frac{p'_{\perp}}{p'_{\parallel}} = \frac{\sqrt{1-v^2}}{v}\,\frac{M_\pi^2 - m_\mu^2}{M_\pi^2 + m_\mu^2}\,. \tag{3.65}$$

Plugging in numbers, $M_\pi \approx 139$ MeV$/c^2$ and $m_\mu \approx 105$ MeV$/c^2$ yields

$$\tan\theta'_{\max} = \frac{p'_{\perp}}{p'_{\parallel}} \approx 0.27 \cdot \sqrt{\frac{1}{v^2} - 1}\,. \tag{3.66}$$

It is clear however, that, whatever the actual masses, the deflection angle goes to 0 if v approaches 1, i.e. the speed of light.

Further experimental signatures in decays

Let us consider two more important practical examples:

1. Mass distribution:
   Assume that there is a so far unknown particle, which then should be found by investigating its decay products. A prime example for this is the so far unfound Higgs particle, which, among other decay channels, also has a small probability to decay into two photons, $H \to \gamma\gamma$. In such decays the **invariant mass of the two-photon system** is given by

   $$m_{\gamma\gamma}^2 = (p_1 + p_2)^2 = P^2 = M_H^2\,, \tag{3.67}$$

   identical to the mass of the Higgs boson.

   At the Large Hadron Collider (LHC), where the search for this particle will start in 2007, there are of course other sources for the production of two photons, resulting in a wide background spectrum of values for $m_{\gamma\gamma}$. The signal will then manifest itself as a "spike" on top of this background. This is illustrated in Fig. 3.1.

2. Energy distribution:
   To see how this works, assume that the decay is isotropic in the rest frame of the decaying particle. Then, in this system the number of particles per unit spatial angle reads

   $$\mathrm{d}N = \frac{1}{4\pi}\mathrm{d}\Omega = \frac{\sin\theta}{4\pi}\mathrm{d}\phi\mathrm{d}\theta = \frac{1}{2}\mathrm{d}\cos\theta\,, \tag{3.68}$$

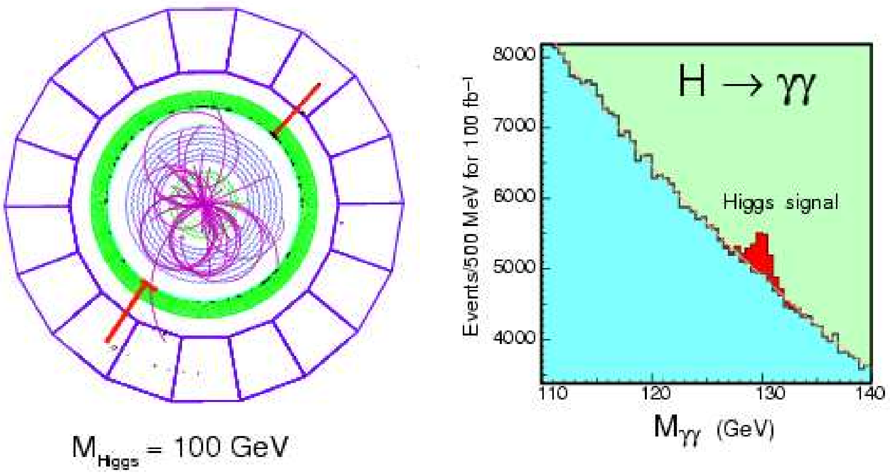

Figure 3.1: Possible signature for Higgs boson production and its subsequent decay into two photons, as seen by a detector (left: in simulation, right: the resulting invariant mass spectrum of two phorons as expected at the LHC, superimposed by a still fictitous Higgs-peak).

where the two angles can be understood as the polar and azimuthal angle w.r.t. the axis of motin of the decaying particle, and where the integration over the azimuthal angle has been performed. In the c.m. frame

$$\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{p_\perp}{p_\parallel} = \frac{|\vec{p}|\sin\theta}{|\vec{p}|\cos\theta}\,, \tag{3.69}$$

and hence

$$p_\parallel = |\vec{p}|\cos\theta\,. \tag{3.70}$$

Using the same transformations as before, cf. Eq. (3.64) (but this time with non-vanishing parallel component) yields

$$\begin{aligned}
E' &= \gamma E + \beta\gamma p_\parallel = \gamma E\left(1 + \frac{\beta|\vec{p}|}{E}\cos\theta\right)\\
p'_\perp &= p_\perp\\
p'_\parallel &= \beta\gamma E + \gamma p_\parallel = \gamma E\left(\beta + \frac{|\vec{p}|}{E}\cos\theta\right)\,.
\end{aligned} \tag{3.71}$$

This implies that the energy in the (primed) lab system ranges between

$$E' \in \left[\gamma E\left(1-\frac{\beta|\vec{p}|}{E}\right),\ \gamma E\left(1+\frac{\beta|\vec{p}|}{E}\right)\right], \tag{3.72}$$

where energy and momentum in the c.m. sysmte have been calculated before. Therefore, having fixed the energy and momentum in the c.m. system

$$\mathrm{d}E' = \gamma\beta|\vec{p}|\cdot \mathrm{d}\cos\theta \tag{3.73}$$

and thus

$$\mathrm{d}N' = \frac{1}{2\gamma\beta|\vec{p}|}\,\mathrm{d}E' . \tag{3.74}$$

Elastic scattering

Let us specify this for the elastic scattering of two particles, say 1 and 2. Elastic scattering implies that both particles remain intact and that no further particles are created. Therefore, the four-momentum bilance reads

$$p_1^\mu + p_2^\mu = q_1^\mu + q_2^\mu , \tag{3.75}$$

where the p_i are the incoming and the q_i are the outgoing momenta, i.e. the respective particles' momentum before and after the collision. In principle, these are eight unknowns, but it will be simple to see that this number can be reduced substantially. To see this, let us start in the c.m. system of the collision. It is defined by

$$P^\mu = p_1^\mu + p_2^\mu = (E,\,\vec{0})^T , \tag{3.76}$$

i.e. by

$$\vec{p}_1 = -\vec{p}_2 . \tag{3.77}$$

Hence, in this system, also

$$P^\mu = q_1^\mu + q_2^\mu = (E,\,\vec{0})^T , \tag{3.78}$$

and

$$\vec{q}_1 = -\vec{q}_2 , \tag{3.79}$$

eliminating four unknowns directly. Adding in that both particles remain on their **mass-shell**, i.e.

$$p_{1,2}^2 = q_{1,2}^2 = m_{1,2}^2 \tag{3.80}$$

it is obvious that there are only two unknowns left. Usually, in the c.m. frame, they are taken to be the azimuthal and polar angle with respect to the axis of the incoming momenta. Thus,

$$p_{1,2}^\mu = (E_{1,2},\, \pm\vec{p})^T \quad \text{and} \quad q_{1,2}^\mu = (F_{1,2},\, \pm\vec{q})^T \tag{3.81}$$

where, due to

$$\vec{p}^2 = E_{1,2}^2 + m_{1,2}^2 \quad \text{and} \quad \vec{q}^2 = F_{1,2}^2 + m_{1,2}^2 \tag{3.82}$$

the energies before and after the collision are identical, $E_i = F_i$ and given by

$$E_{1,2} = F_{1,2} = \frac{E^2 \pm (m_1^2 - m_2^2)}{2E} \tag{3.83}$$

leading to

$$|\vec{p}| = |\vec{q}| = \lambda(E,\, m_1, m_2) =: \rho\,. \tag{3.84}$$

Orienting the incoming momenta along the z-axis therefore results in

$$\begin{aligned} p_{1,2}^\mu &= (E_{1,2},\, 0,\, 0,\, \pm\rho)^T \\ q_{1,2}^\mu &= (E_{1,2},\, \pm\rho\sin\theta\cos\phi,\, \pm\rho\sin\theta\sin\phi,\, \pm\rho\cos\theta)^T\,, \end{aligned} \tag{3.85}$$

i.e. elastic scattering results in a deflection of the trajectories of the two particles. The (primed) lab system is now connected with the c.m. system by a boost characterised by β and γ, as before. To keep things easy, and to prepare ourselves for the following examples, let us assume that one of the initial particles, say particle 2, is at rest in the lab system. Then β is along the z-axis and the total momentum in both systems is

$$\begin{aligned} P^\mu &= p_1^\mu + p_2^\mu = (E,\, \vec{0})^T \\ P'^\mu &= p_1'^\mu + p_2'^\mu = (E_1' + m_2,\, \vec{p}_1')^T = (E',\, \vec{p}_1')^T\,. \end{aligned} \tag{3.86}$$

Let us first calculate the boost parameters connecting the primed and unprimed systems. Eq. (3.30) yields

$$E' = \gamma E \quad \text{and} \quad \rho' = \beta\gamma E = \beta E'\,. \tag{3.87}$$

Therefore,

$$\begin{aligned}
\beta &= \frac{|\vec{p}_1'|}{E_1'+m_2} = \frac{\rho'}{E'} \\
\gamma &= \frac{E_1'+m_2}{E_1+E_2}\,.
\end{aligned} \tag{3.88}$$

On the other hand, since the squares of four-vectors (and of course also of four-momenta) are invariant under Lorentz-transformations, one has

$$P^2 = E^2 = P'^2 = E'^2 - \rho'^2 \quad\Longrightarrow\quad \rho' = \sqrt{E'^2 - E^2}\,. \tag{3.89}$$

Compton scattering

The findings above are now applied to the case of Compton scattering. There, a photon collides with an electron at rest and is deflected. The four-momentum bilance of this reaction reads

$$p_\gamma^\mu + p_e^\mu \;=\; q_\gamma^\mu + q_e^\mu\,, \tag{3.90}$$

where the quantum mechanical finding has been employed that photons indeed have energy and momentum, given by

$$E_\gamma = h\nu = \hbar\omega = |\vec{p}_\gamma|\,, \tag{3.91}$$

where again factors of c have been neglected and where explicit use has been made of the fact that photons are massless, implying that $p^2 = 0$. In the lab system, where the electron is at rest, the bilance above then reads

$$\begin{aligned}
\hbar\omega + m_e &= \hbar\bar{\omega} + E_e \\
\hbar\omega\vec{e}_z &= \hbar\bar{\omega}\vec{e}_r + \vec{q}_e\,,
\end{aligned} \tag{3.92}$$

where ω and $\bar{\omega}$ are the circular frequencies before and after the collision. Using that

$$E_e = \sqrt{m_e^2 + \vec{q}_e^{\,2}} \geq m_e \tag{3.93}$$

it becomes apparent that in this process the photon looses energy if the electron is not at rest after the collision. From the reasoning so far, it should be clear that there is thus a correlation between the scattering angle (the angle of $\vec{e}_r$ w.r.t. $\vec{e}_z$) of the photon and its energy loss.

In order to calculate this correlation, the momentum relations above could be employed in a somewhat tedious but straightforward calculation. Here, another approach is chosen. For this to work, it is worthwhile to realise that

$$(p_1^\mu - q_1^\mu)u_\mu = (p_2^\mu - q_2^\mu)u_\mu = 0\,, \tag{3.94}$$

where u_μ is the four-velocity of the c.m. frame in the frame under consideration. Since this equation is Lorentz-invariant, therefore it suffices to show that it is valid in one system, for simplicity the c.m. frame of the collision itself is chosen. There (remember, $c = 1$!),

$$u_\mu = (1,\, \vec{0})\,, \tag{3.95}$$

and the above relations become

$$E_1 - F_1 = E_2 - F_2 = 0\,, \tag{3.96}$$

which is trivially fulfilled with the reasoning of the above section, cf. Eq. (3.83). In the lab system,

$$(p'^\mu_1 - q'^\mu_1)u_\mu = \gamma(E_1' - F_1') - \gamma(\vec{p}_1' - \vec{q}_1')\cdot\vec{\beta} = 0 \tag{3.97}$$

leading to

$$E_1' - F_1' = |\vec{p}_1'||\vec{\beta}| - |\vec{q}_1'||\vec{\beta}|\cos\theta_\gamma\,, \tag{3.98}$$

where θ_γ is the scattering angle of the photon in the lab frame, and where the fact has been used that $\vec{u} \parallel \vec{p}_1'$. Using that according to Eq. (3.88)

$$|\vec{\beta}| = |\vec{p}_1'|/(E_1' + m_2)\,, \tag{3.99}$$

θ_γ can be fixed:

$$\begin{aligned}\cos\theta_\gamma &= -\frac{E_1' - F_1' - \frac{\vec{p}'^2_1}{E_1'+m_2}}{\frac{|\vec{p}_1'||\vec{q}_1'|}{E_1'+m_2}}\\ &= \frac{\vec{p}'^2_1 - (E_1'+m_2)(E_1'-F_1')}{|\vec{p}_1'||\vec{q}_1'|} = \frac{(E_1'+m_2)F_1' - m_2E_1'}{|\vec{p}_1'||\vec{q}_1'|}\,,\end{aligned} \tag{3.100}$$

where in the last line

$$E_1'^2 - \vec{p}'^2_1 = m_1^2 = 0 \tag{3.101}$$

has been employed. In full analogy the scattering angle for the electron in the lab frame, θ_e can be fixed. Starting with

$$E_2' - F_2' = |\vec{p}_2'||\vec{\beta}| - |\vec{q}_2'||\vec{\beta}|\cos\theta_e\,, \tag{3.102}$$

one finds

$$\cos\theta_e = \frac{(E_1' + m_2)(F_2' - m_2)}{|\vec{p}_1'||\vec{q}_2'|}\,. \tag{3.103}$$

In principle, this solves the problem; the only thing that remains to be done is to plug in numbers. To this end, remember that for the photons

$$|\vec{p}_1'| = E_1' = \hbar\omega \quad \text{and} \quad |\vec{q}_1'| = F_1' = \hbar\bar{\omega} \tag{3.104}$$

and solve the first equation for $\bar{\omega}$:

$$\begin{aligned} \cos\theta_\gamma &= \frac{(\hbar\omega + m_2c^2)\hbar\bar{\omega} - m_2c^2\hbar\omega}{\hbar\omega\hbar\bar{\omega}} \\ \Longrightarrow \hbar\bar{\omega} &= \frac{\hbar\omega m_2c^2}{\hbar\omega(1-\cos\theta_\gamma) + m_2c^2}\,, \end{aligned} \tag{3.105}$$

where the factors of c have been supplemented. Identifying angular frequencies and wavelengths, e.g.

$$\hbar\omega = \frac{hc}{\lambda} \tag{3.106}$$

then yields

$$\bar{\lambda} = \lambda + \lambda_C(1-\cos\theta_\gamma)\,, \tag{3.107}$$

where

$$\lambda_C = \frac{h}{m_2c} \tag{3.108}$$

is the Compton wavelength. This shows that the energy loss of the photon is maximal (i.e. the resulting wavelength is minimal), if the photon “bounces” back, and it is minimal if the photon does not experience any deflection.

Inelastic scattering

So far, only elastic scattering, leaving the colliding particles intact and their number constant, has been considered. In many particle and nuclear physics experiments, however, the **aim is to produce new particles**, i.e. to investigate inelastic scattering. Let us consider now a few examples for such inelastic processes.

1. $\pi^- p \to \Lambda K^0$:
   Let us start with the reaction above, in which the number of particles remains constant, but change their identity and, thus, their masses. The respective masses are

$$\begin{aligned} m_\pi = 139.6\,\mathrm{MeV}/c^2\,, &\quad m_p = 938.3\,\mathrm{MeV}/c^2\,, \\ m_\Lambda = 1115.6\,\mathrm{MeV}/c^2\,, &\text{ and } m_K = 497.7\,\mathrm{MeV}/c^2\,. \end{aligned} \tag{3.109}$$

   Obviously, the total energy of the incoming pion and proton must be large enough to allow for the transition to the heavier Kaon and Lambda-particle.

   In the c.m. frame, the momenta of the proton and the pion compensate each other,

$$\vec{p}_\pi = -\vec{p}_p = \vec{p}\,, \tag{3.110}$$

   and the total energy is thus given by

$$\begin{aligned} E &= E_\pi + E_p = \sqrt{m_\pi^2 + \vec{p}^2} + \sqrt{m_p^2 + \vec{p}^2} \\ &= E_K + E_\Lambda = \sqrt{m_K^2 + \vec{p}_K^2} + \sqrt{m_\Lambda^2 + \vec{p}_\Lambda^2} \\ &\geq m_K + m_\Lambda\,. \end{aligned} \tag{3.111}$$

   Again, factors of c have been ignored, i.e. $c = 1$. The "="-sign in the equation above characterises the minimal energy neccessaary for this reaction to happen, it thus implies a minimal value of $|\vec{p}|$, namely

$$|\vec{p}_{\mathrm{min}}| = \lambda(m_K + m_\Lambda,\, m_\pi,\, m_p) \approx 520\,\mathrm{MeV}/c\,. \tag{3.112}$$

2. $e^+e^- \to e^+e^- + n\pi$:
   This is another example, where an electron-positron pair produces additionally n pions in its reaction. Since the lectron mass is tiny w.r.t.

the pion mass ($m_e = 511$ keV$/c^2$), it will be neglected here. Clearly, in the c.m. system the incoming energies must be larger than the mass of the n pions,

$$E_{\text{c.m.}} \geq n m_\pi \,. \tag{3.113}$$

3.2 Newton's mechanics

3.2.1 Force, relativistic version

Force and momentum

In order to move on towards relativistic dynamics rather than mere kinematics, it is neccessary to introduce the concept of force in relativistic notation. In principle, this is quite a simple task. So far, such quantities have been introduced by analogies, for instance four-velocity and –acelaration have been defined by differentiating with respect to the eigentime,

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \quad \text{and} \quad b^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} \,. \tag{3.114}$$

It seems then straightforward to introduce force by an analogy with the non-relativistic version

$$\vec{F} = \dot{\vec{p}} = \frac{\mathrm{d}\vec{p}}{\mathrm{d}t} \tag{3.115}$$

as

$$F^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} \,. \tag{3.116}$$

Assuming the mass to be constant, this implies

$$F^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = m\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = m b^\mu \,, \tag{3.117}$$

where b^μ is the four-acceleration.

Just to convince ourselves that this is reasonable, let us check the spatial components,

$$\begin{aligned}
\vec{F} &= m\frac{\mathrm{d}\vec{u}}{\mathrm{d}\tau} = m\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\vec{v}}{\sqrt{1-v^2/c^2}}\right) = \frac{m}{\sqrt{1-v^2/c^2}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\vec{v}}{\sqrt{1-v^2/c^2}}\right)\\
&= \frac{1}{\sqrt{1-v^2/c^2}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\vec{v}}{\sqrt{1-v^2/c^2}}\right) = \frac{1}{\sqrt{1-v^2/c^2}}\frac{\mathrm{d}}{\mathrm{d}t}\left[m(v)\vec{v}\right]\\
&= \frac{1}{\sqrt{1-v^2/c^2}}\frac{\mathrm{d}\vec{p}}{\mathrm{d}t}\,.
\end{aligned} \tag{3.118}$$

Expanding this in terms of v^2/c^2 and keeping only the first term immediately yields the non-relativistic version, cf. Eq. (3.115). Of course, also dour-accelaration could be used. Then, emplyoing Eq. (3.50),

$$b^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = \left(\frac{\frac{\vec{v}\cdot\vec{a}}{c}}{\left(1-\frac{v^2}{c^2}\right)^2},\ \frac{\vec{a}}{1-\frac{v^2}{c^2}} + \frac{\vec{v}}{c}\frac{\frac{\vec{v}\cdot\vec{a}}{c}}{\left(1-\frac{v^2}{c^2}\right)^2}\right)^T\,, \tag{3.119}$$

the spatial components of force read

$$\begin{aligned}
\vec{F} &= m\vec{b} = m\left(\frac{\vec{a}}{1-\frac{v^2}{c^2}} + \frac{\vec{v}}{c}\frac{\frac{\vec{v}\cdot\vec{a}}{c}}{\left(1-\frac{v^2}{c^2}\right)^2}\right)\\
&= m\left[\vec{a} + \frac{v^2}{c^2}\vec{a} + \left(\frac{\vec{v}\cdot\vec{a}}{c}\right)\frac{\vec{v}}{c} + \dots\right]\,,
\end{aligned} \tag{3.120}$$

where terms of order v^3/c^3 or higher are neglected. Taking a closer look at the relativistic correction terms, it is interesting to note that the second of them does not point into the direction of three-acceleration $\vec{a}$, but rather in the direction of the velocity.
Phrased the other way around, this implies that the acceleration of a body has a component which does not point into the direction of the force it experiences. This implies that the motion will be more complicated than in the ordinary, non-relativistic case. The difference of both is due to the relativistic dependence of the energy and, hence, mass on the velocity. Only if the force acting on a body is parallel or orthogonal w.r.t. its direction of motion, the force and the resulting acceleration are parallel. The former of the two cases will be discussed now.

A simple example

In this section, a simple cases for the relativistic force equation will be discussed, namely a particle in a gravitational field. It will be assumed that it is just falling down, with no lateral motion. In order to start the investigation, it will be borrowed from general relativity that gravity acts on the total energy of the particle, such that

$$\vec{F} = \frac{gE(v)}{c}\vec{e}_z\,, \tag{3.121}$$

where g is the constant acceleration in the gravitational field. It should be stressed here that this *ansatz* is not Lorentz-invariant. However, this is not important here, because the notion of constant gravitational acceleration implies the existence of a preferred reference frame - the planet (or so), the particle is falling on. So, the following discussion is not fully correct, but it serves the purpose.
Inserting this force into the equation of motion (E.o.M.), then

$$\frac{m}{\sqrt{1-v^2/c^2}}g\vec{e}_z = \frac{m}{\sqrt{1-v^2/c^2}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\vec{v}}{\sqrt{1-v^2/c^2}}\right)\,. \tag{3.122}$$

If the particle was at rest in the beginning, $\vec{v}(0) = 0$ then at each point $\vec{v} = v\vec{e}_z$ and therefore

$$g = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{v}{\sqrt{1-v^2/c^2}}\right) \tag{3.123}$$

or, for each time $t \geq 0$,

$$gt = \left(\frac{v}{\sqrt{1-v^2/c^2}}\right)\,. \tag{3.124}$$

Solving this equation for v yields

$$v = \frac{gt}{\sqrt{1+(gt/c)^2}} \tag{3.125}$$

and thus the trajectory reads, with $z(t=0) = z_0$,

$$z(t) = z_0 + \int\limits_0^t \mathrm{d}t'v(t') = z_0 + \frac{c^2}{g}\left(\sqrt{1+\frac{g^2t^2}{c^2}} - 1\right)\,. \tag{3.126}$$

Zeroth component of force

So far the discussion focused on the spatial components of the force, but of course, there is also the 0th component. From Eq. (),

$$F^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} \tag{3.127}$$

it can be read of as

$$F^0 = \frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}\tau} = \frac{1}{c\sqrt{1-v^2/c^2}}\frac{\mathrm{d}E}{\mathrm{d}t}\,, \tag{3.128}$$

i.e. it is connected with a change of energy. The term $\mathrm{d}E/\mathrm{d}t$ describes the energy that is absorbed or released by the body per unit time while accelerating, hence it is the power. Of course, this can also be expressed through the 0th component of the four-acceleration,

$$F^0 = m\frac{\frac{\vec{v}\cdot\vec{a}}{c}}{\left(1-v^2/c^2\right)^2}\,. \tag{3.129}$$

Having completed the four-vector describing force it is natural to start a deeper investigation of its properties. To do so, remember that

$$b^\mu u_\mu = 0\,, \tag{3.130}$$

i.e. that four-acceleration and -velocity are orthogonal to each other. Since $F^\mu \parallel b^\mu$, this immediately implies that

$$F^\mu u_\mu = F^0\frac{c\mathrm{d}t}{\mathrm{d}\tau} - \vec{F}\cdot\frac{\mathrm{d}\vec{x}}{\mathrm{d}\tau} = 0\,. \tag{3.131}$$

Inserting F^0 from above, this translates into

$$\frac{\mathrm{d}E}{\mathrm{d}t} - \sqrt{1-v^2/c^2}\vec{F}\cdot\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = 0\,. \tag{3.132}$$

In a conservative force field,

$$\vec{F}\cdot\mathrm{d}\vec{x} = -\mathrm{d}V\,, \tag{3.133}$$

and absorbing the constant factor $\sqrt{1-v^2/c^2}$ into the potential V this finally yields

$$\frac{\mathrm{d}}{\mathrm{d}t}(E+V) = 0\,. \tag{3.134}$$

Here E is the kinetic energy of the particle, and V is its potential energy. It is connected with the acting force through

$$\vec{F} = -\frac{\vec{\nabla} V}{\sqrt{1 - v^2/c^2}}\,, \tag{3.135}$$

and therefore

$$\dot{\vec{p}} = -\vec{\nabla} V\,. \tag{3.136}$$

In the equation above, describing the conservation of energy, V shows up in connection with E. Now, E transforms like the 0th component of a four-vector – the four-momentum – and therefore V must also behave like the 0th component of a four-vector. This requirement ensures Lorentz-invariance of our equations. Therefore, a four-vector A^μ must be introduced, whose 0th component is V. This vector is known as the four-potential. In the rest-frame,

$$A^\mu = (V,\, \vec{0})^T \tag{3.137}$$

and therefore, in all frames,

$$A^2 = A^\mu A_\mu = V^2\,. \tag{3.138}$$

Then, energy conservation can be written as four-momentum conservartion, namely

$$\frac{\mathrm{d}}{\mathrm{d}t}(p^\mu + A^\mu) = 0\,. \tag{3.139}$$

3.2.2 Lorentz force

Four-potential in electromagnetism

Having introduced the four-potential A^μ in quite an abstract way, let us now consider a well-known system which has such a potential, namely electromagnetism. There, two potentials, a scalar potential Φ and a vector-potential $\vec{A}$ have been introduced. They are connected to the measurable electric and magnetic fields by

$$\begin{aligned} \vec{E} &= -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\Phi \\ \vec{B} &= \vec{\nabla} \times \vec{A}\,. \end{aligned} \tag{3.140}$$

Here (and already in the previos section) the $\vec{\nabla}$-operator has been employed,

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)^T . \tag{3.141}$$

It can be shown that it deserves the vector-arrow very well, because it transforms like a vector.

Lorentz force

Particles in electromagnetic fields experience a force, known as the Lorentz-force. Its spatial components read

$$\vec{F}' = q\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right) , \tag{3.142}$$

where q is the charge of the particle. This equation does not exhibit any clear sign of being the spatial components of a four-vector, so let us try to recover this. To this end, the electric and magnetic fields in the Lorentz force are replaced by the respective expressions emplyoing the potential(s) Φ and $\vec{A}$. Later, they will be subsumed in the four-potential

$$A^\mu = (\Phi,\, \vec{A})^T \tag{3.143}$$

of electromagnetism. However, expressed by them, the non-relativistic Lorentz force reads

$$\begin{aligned} \vec{F}' &= \frac{q}{c}\left(-\frac{\partial\vec{A}}{\partial t} - c\vec{\nabla}\Phi + \vec{v}\times\vec{\nabla}\times\vec{A}\right) \\ &= \frac{q}{c}\left[-c\vec{\nabla}\Phi - \frac{\partial\vec{A}}{\partial t} + \vec{\nabla}(\vec{v}\cdot\vec{A}) - (\vec{v}\cdot\vec{\nabla})\vec{A}\right] , \end{aligned} \tag{3.144}$$

where

$$\vec{a}\times\vec{\nabla}\times\vec{b} = \vec{\nabla}(\vec{a}\cdot\vec{b}) - (\vec{a}\cdot\vec{\nabla})\vec{b} - (\vec{b}\cdot\vec{\nabla})\vec{a} - \vec{b}\times\vec{\nabla}\times\vec{a} \tag{3.145}$$

has been used together with the fact that $\vec{v}$ does not explicitly depend on position and that hence all its spatial derivatives vanish. The partial derivative of $\vec{A}$ w.r.t. time can be reepxressed in a suitable fashion by realising that the total derivative w.r.t. time can be written as

$$\frac{\mathrm{d}\vec{A}}{\mathrm{d}t} = \frac{\partial\vec{A}}{\partial t} + \sum_{i=1}^{3}\frac{\partial\vec{A}}{\partial x_i}\frac{\partial x_i}{\partial t} = \frac{\partial\vec{A}}{\partial t} + (\vec{v}\cdot\vec{\nabla})\vec{A}\,. \tag{3.146}$$

Inserting this into the Lorentz force yields

$$\vec{F}' = \frac{q}{c}\left[-c\vec{\nabla}\Phi - \frac{\mathrm{d}\vec{A}}{\mathrm{d}t} + \vec{\nabla}(\vec{v}\cdot\vec{A})\right], \tag{3.147}$$

It can be shown that, written in a covariant fashion and up to a multiplicative constant, this can be cast into

$$F^{\mu} = \frac{q}{c}\left[\frac{\partial}{\partial x_{\mu}}(u_{\nu}A^{\nu}) - \frac{\mathrm{d}}{\mathrm{d}\tau}A^{\mu}\right], \tag{3.148}$$

where the contra-variant four-vector

$$\frac{\partial}{\partial x_{\mu}} = \partial^{\mu} = (\frac{\partial}{c\partial t}, -\vec{\nabla})^{T} \tag{3.149}$$

has been introduced. The minus sign is due to the fact that the derivative is taken w.r.t. the covariant components. In order to check that this is indeed the case, consider the spatial components and

$$A^{\mu}u_{\mu} = A^{\mu}\frac{\mathrm{d}x_{\mu}}{\mathrm{d}\tau} = \gamma(c\Phi - \vec{v}\cdot\vec{A})\,. \tag{3.150}$$

Then,

$$\begin{aligned}\vec{F} &= \frac{q}{c}\left[-\vec{\nabla}(\gamma c\Phi - \gamma\vec{v}\cdot\vec{A}) - \frac{\mathrm{d}}{\mathrm{d}\tau}\vec{A}\right] = \frac{q\gamma}{c}\left[-\vec{\nabla}(c\Phi - \vec{v}\cdot\vec{A}) - \frac{\mathrm{d}}{\mathrm{d}t}\vec{A}\right]\\ &= \frac{q\gamma}{c}\left[-\vec{\nabla}(c\Phi) - \frac{\partial}{\partial t}\vec{A}\right],\end{aligned} \tag{3.151}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{A} = \frac{\partial\vec{A}}{\partial t} + (\vec{v}\cdot\vec{\nabla})\vec{A} \tag{3.152}$$

has been used. The second term of the equation above, $(\vec{v}\cdot\vec{\nabla})\vec{A}$, cancels the term $\vec{\nabla}(\vec{v}\cdot\vec{A})$ in the Lorentz force, because

$$\vec{\nabla}(\vec{v}\cdot\vec{A}) - (\vec{v}\cdot\vec{\nabla})\vec{A} = \vec{A}(\vec{\nabla}\cdot\vec{v}) = 0\,. \tag{3.153}$$

From the first line of the expression of the force vector above, the multiplicative constant is readily identified with the (constant) γ, which does not depende explicitly on space or time. Expanding $\vec{F}$ in terms of v^2/c^2 gives $\vec{F}'$ at lowest order. Let us also scruitinize the 0-component of this relativistic Lorentz force,

$$F^0 = \frac{q}{c}\left[-\frac{\partial}{c\partial t}(\gamma c\Phi - \gamma\vec{v}\cdot\vec{A}) - \gamma\frac{\mathrm{d}}{\mathrm{d}t}\Phi\right] = \frac{q\gamma}{c}\left[-\frac{\vec{v}}{c}\cdot\frac{\partial\vec{A}}{\partial t} - \vec{v}\cdot\vec{\nabla}\Phi\right], \tag{3.154}$$

where again

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi = \frac{\partial}{\partial t}\Phi + \vec{v}\cdot\vec{\nabla}\Phi \tag{3.155}$$

has been used. Together this yields

$$F^\mu = (F^0,\, \vec{F})^T = \left(\frac{\vec{v}\cdot\vec{F}}{c},\, \vec{F}\right)^T \tag{3.156}$$

However, from the discussion in the previous section, it is known that

$$F^0 = \frac{\gamma}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, \tag{3.157}$$

connecting the 0th component of the force with the temporal derivative of the energy. Identifying the right-hand side of the force equation above with the electric field then yields

$$\frac{\mathrm{d}E}{\mathrm{d}t} = q\vec{v}\cdot\vec{E}, \tag{3.158}$$

making manifest that the particle's change in energy does depend on the electric field only. At this point it should be stressed that E is the energy, and $\vec{E}$ is the electric field, both of which are not components of the same four-vector.

A simple example

To highlight some of the aspects discusse so far, consider as an example the motion of an electron (charge e) in a constant homogeneous magnetic field $\vec{B}$. In order to determine its trajectory, employ

$$\vec{F} = \gamma\vec{F}' = \gamma\left[-e\vec{\nabla}\Phi + \frac{e}{c}\vec{\nabla}(\vec{v}\cdot\vec{A}) - \frac{e}{c}\frac{\mathrm{d}\vec{A}}{\mathrm{d}t}\right]. \tag{3.159}$$

Setting $\vec{E} = 0$ yields

$$\vec{F} = \frac{e\gamma}{c}\vec{v} \times \vec{B}\,, \tag{3.160}$$

where again the vector identity of Eq. (3.145) has been used. Dividing the velocity into a component parallel and one component perpendicular w.r.t. the magnetic field,

$$\vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp} \tag{3.161}$$

then results in

$$\vec{F} = \gamma\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \frac{e\gamma}{c}\vec{v}_{\perp} \times \vec{B}\,, \tag{3.162}$$

hence

$$\frac{\mathrm{d}[m(v)\vec{v}]}{\mathrm{d}t} = \frac{\mathrm{d}[m\gamma\vec{v}]}{\mathrm{d}t} = \frac{e}{c}\vec{v}_{\perp} \times \vec{B}\,. \tag{3.163}$$

The rest mass m is independent of time. Thus, $\mathrm{d}(\gamma\vec{v})/\mathrm{d}t$ must be calculated,

$$\frac{\mathrm{d}(\gamma\vec{v})}{\mathrm{d}t} = \gamma^3\left(\frac{\vec{v}\cdot\vec{a}}{c^2}\right)\vec{v} + \gamma\vec{a}\,, \tag{3.164}$$

and therefore

$$\gamma^3\left(\frac{\vec{v}\cdot\vec{a}}{c^2}\right)\vec{v} + \gamma\vec{a} = \frac{e}{mc}\vec{v}_{\perp} \times \vec{B}\,. \tag{3.165}$$

Since the force is perpendicular to $\vec{v}$ and because $\vec{a} \parallel \vec{F}$, $\vec{a}\cdot\vec{v} = 0$ and thus

$$\vec{a} = \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \frac{e}{mc\gamma}\vec{v}_{\perp} \times \vec{B}\,. \tag{3.166}$$

From the findings so far, two things should be obvious:

1. The particle is not accelerated in the direction of the field but perpendicular to it. Therefore, the parallel velocity component $\vec{v}_{\parallel}$ is constant.

2. Because the force and hence the change in velocity is also orthogonal w.r.t. the perpendicular component, it only changes its direction, but not its size. Taken together, this implies that $v = |\vec{v}|$ is constant. The trajectory is thus a helix.

Projecting this helix onto the transversal plane, it is possible to determine its radius r. It is given by the condition that the force through the magnetic field equals the centripetal force,

$$\frac{e}{c}|\vec{v}_\perp||\vec{B}| = \frac{m\gamma}{r}v_\perp^2 \tag{3.167}$$

and therefore

$$r = \frac{mc\gamma}{e}\frac{|\vec{v}_\perp|}{|\vec{B}|}\,. \tag{3.168}$$

In addition, the circular frequency is

$$\omega = \frac{v_\perp}{r} = \frac{eB}{cm\gamma}\,. \tag{3.169}$$

3.3 Lagrange formulation

will be supplemented

Chapter 4

Relativistic electrodynamics

In this chapter, electrodynamics will be investigated in a covariant formulation. The interesting point here is that, while **the laws of mechanics had to be modified in order to accomodate Lorentz invariance**, the laws governing electrodynamics, i.e. **Maxwell's equations, actually are Lorentz invariant**. The only tricky point here, is that this is obscured by their original formulation in terms of electric and magentic fields. Therefore, the first task in this chapter is to unveil the true, covariant nature of electrodynamics.

4.1 Maxwell's equations

4.1.1 The field strength tensor

Conservation of charges

The starting point for the relativistic invariant reconstruction of the Maxwell equations is the principle of conserved charges. The idea here is clear: Consider charges in a finite volume V. If the total charge in this volume changes, some charge carriers must have left the volume through its surface S. Thus, assuming the volume to be constant with time,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int\limits_V \mathrm{d}^3V\,\frac{\partial\varrho}{\partial t} = \int\limits_{S=\delta V} \mathrm{d}^2S\,\vec{n}_{\mathrm{d}^2S}\cdot\vec{j}\,, \tag{4.1}$$

where ϱ is the charge density, $\vec{n}_S$ is a unit vector normal to the surface element d^2S, and $\vec{j}$ is the charge current. Now, Gauss' law of integration can be used[1], which connects the integration over a surface with an integration over the enclosed volume by

$$\int\limits_{S=\delta V} \mathrm{d}^2 S\, \vec{n}_{\mathrm{d}^2 S} \cdot \vec{j} = \int\limits_{V} \mathrm{d}^3 V\, \vec{\nabla} \cdot \vec{j}\,. \tag{4.2}$$

Then, charge conservation can be reformulated as

$$\int\limits_{V} \mathrm{d}^3 V \left[\frac{\partial \varrho}{\partial t} - \vec{\nabla} \cdot \vec{j}\right] = 0\,. \tag{4.3}$$

In this way, the integral version of charge conservation has been prepared to yield the differential version, namely

$$\frac{\partial \varrho}{\partial t} - \vec{\nabla} \cdot \vec{j} = 0 \tag{4.4}$$

and, emplyoing our knowledge of relativistic invariance, it is straightforward to introduce a charge current four-vector j^μ such that

$$j^\mu = (c\varrho, \vec{j})^T \tag{4.5}$$

and

$$\partial_\mu j^\mu = 0\,. \tag{4.6}$$

Relativistic properties the charge current

- Lorentz-transformations of the charge current:
  It is quite simple to transform the charge current from one reference frame to another one: Just apply a Lorentz-transformation to it with a velocity $\vec{v}$, say along the x-axis. Then

$$j'^\mu = \Lambda^\mu{}_\nu(\vec{v}) j^\nu \tag{4.7}$$

[1] Don't become nervous here, in principle, Gauss' law of integration is simple and straightforward to understand; if you do not know it by now, just consult e.g. the first chapters of Feynman lectures, part II.

and in components,

$$\begin{aligned}
\varrho' &= \frac{\varrho - \frac{v}{c^2} j_1}{\sqrt{1-\frac{v^2}{c^2}}} \\
j_1' &= \frac{j_1 - v\varrho}{\sqrt{1-\frac{v^2}{c^2}}} \\
j_{2,3}' &= j_{2,3}\,.
\end{aligned} \tag{4.8}$$

- Relativistic invariants:
  Apart from the four-divergence $\partial_\mu j^\mu$ of the charge current there is another, obviously Lorentz-invariant quantity that can be built, namely

$$j_\mu j^\mu = (c\varrho)^2 - \vec{j}^{\,2}\,. \tag{4.9}$$

- Conserved charge:
  Consider now a fixed charge distribution given by ϱ in its rest frame S. In another frame S' the charges move with a velocity $\vec{v} = v\vec{e}_x$. The differential charge in S reads

$$\mathrm{d}Q = \varrho \mathrm{d}^3 V \tag{4.10}$$

  and in S'

$$\begin{aligned}
\mathrm{d}Q' &= \varrho' \mathrm{d}^3 V' = \varrho' \mathrm{d}x' \mathrm{d}y' \mathrm{d}z' \\
&= \frac{\varrho}{\sqrt{1-\frac{v^2}{c^2}}} \left(\sqrt{1-\frac{v^2}{c^2}}\mathrm{d}x\right) \mathrm{d}y\mathrm{d}z = \mathrm{d}Q\,.
\end{aligned} \tag{4.11}$$

Lorentz force, once again

Let us now appraoch the Lorentz-force already encountered in the previous chapter in a more structural context. From a more "practical point" of view, the electric and magnetic fields are measurable quantities, which can be measured by subjecting a test particle to them. This test particle with charge q will then experience a force, the non-relativistic Lorentz-force,

$$\vec{F}' = \frac{q}{c}(c\vec{E} + \vec{v} \times \vec{B})\,, \tag{4.12}$$

cf. Eq. (3.142). Then, acceleration of the particle in different directions can be connected to the electric and magnetic field, $\vec{E}$ and $\vec{B}$, respectively. In the previous chapter this non-relativistic version has been translated into an Lorentz-invariant version. Here, let us construct this relativistic version more directly. To do so, first note that the Lorentz-force is linear in the charge and the charge times its velocity. It seems natural then to demand that the four-vector Lorentz-force is linear in the four-velocity of the charge, qu^μ. However, a mere proportionality is out of the question, since the force has a component orthogonal to the current $q\vec{v}$. Thus there must be a tensor, connecting the force and the charge four velocity

$$F^\mu = qF^{\mu\nu}u_\nu \,, \tag{4.13}$$

where $F^{\mu\nu}$ is known as the field strength tensor. This is a natural, obvious name. Just comparing the two equations reveals that it must be a linear function of the fields $\vec{E}$ and $\vec{B}$. In the previous chapter, F^μ has been found to be

$$F^\mu = \frac{q}{\sqrt{1-v^2/c^2}} \left(\frac{\vec{v}\cdot\vec{F}'}{c}, \vec{F}'\right)^T \,, \tag{4.14}$$

cf. Eq. (3.156).

Explicit construction of the field strength tensor

This can now be employed to construct the field strength tensor by explicit comparison:

- $\mu = k \in [1,\, 3]$:
  The spatial components of the Lorentz force read

$$\begin{aligned} F^k &= \frac{q}{\sqrt{1-v^2/c^2}}(E^k + \epsilon^{klm}v_l B_m) \\ &= F^{k\nu}u_\nu = \frac{q}{\sqrt{1-v^2/c^2}}(F^{k0}c - F^{kl}v_l) \,, \end{aligned} \tag{4.15}$$

  implying that

$$F^{k0} = E^k/c \quad \text{and} \quad F^{kl} = -\epsilon^{klm}B_m = -F^{lk} \,. \tag{4.16}$$

- $\mu = 0$:
  The temporal component of the Lorent force reads

$$\begin{aligned} F^0 &= \frac{q}{\sqrt{1-v^2/c^2}}\frac{\vec{v}\cdot\vec{F}'}{c} = \frac{q}{\sqrt{1-v^2/c^2}}\frac{(v_k E^k + \epsilon^{klm} v_k v_l B_m)}{c} \\ &= \frac{q}{\sqrt{1-v^2/c^2}}\frac{(v_k E^k}{c} \\ &= F^{0\nu} u_\nu = \frac{q}{\sqrt{1-v^2/c^2}}(F^{00}c - F^{0l}v_l)\,, \end{aligned} \tag{4.17}$$

  where the second line is a result of the antisymmetric ϵ^{klm} being contracted with the symmetric $v_k v_l$-piece, yielding a vanishing term. This however implies that

$$F^{00} = 0 \quad \text{and} \quad F^{0l} = -E^l/c\,. \tag{4.18}$$

Hence, the field strength tensor reads

$$F^{\mu\nu} = -F^{\nu\mu} = \begin{pmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & -B_3 & B_2 \\ E_2/c & B_3 & 0 & -B_1 \\ E_3/c & -B_2 & B_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\vec{E}/c \\ \vec{E}/c & \epsilon^{ijk}B_k \end{pmatrix}. \tag{4.19}$$

Here, the right-down square F^{ij} is given by the total antisymmetric tensor ϵ^{ijk} times the kth component of the magnetic field vector. The tensor is defined by

$$\epsilon^{123} = \epsilon_{123} = 1 \tag{4.20}$$

and a change of sign for anticyclic permutations. Summarising this: The field strength tensor is an antisymmetric tensor of rank two (it has two Lorentz indices), where the spatial components are given by the components of the magentic field and the mixed spatial-temporal components are given by the electric field components. This property can be sean quite easily. If six independent quantities are to be represented by an object in four-dimensional space-time, there remains antisymmetric tensors of rank two as a first guess,

since they have six independent components (start with sixteen, six are fixed by $F^{\mu\nu} = -F^{\nu\mu}$ and four are zero because of $F^{\mu\mu} = -F^{\mu\mu}$).
Having constructed the contravariant field strength tensor let us now generate its covariant counterpart. To do so, the metric tensor will be employed, yielding

$$F_{\mu\nu} = g_{\mu\kappa} g_{\nu\lambda} F^{\kappa\lambda} = \begin{pmatrix} 0 & E_1/c & E_2/c & E_3/c \\ -E_1/c & 0 & -B_3 & B_2 \\ -E_2/c & B_3 & 0 & -B_1 \\ -E_3/c & -B_2 & B_1 & 0 \end{pmatrix} . \tag{4.21}$$

Here, only the mixed spatial-temporal components change their sign (the purely spatial ones change the sign twice, once for each metric tensor!). From this, it is straightforward to construct an invariant quantity, namely

$$\begin{aligned} F^{\mu\nu} F_{\mu\nu} &= -F^{\mu\nu} F_{\nu\mu} \\ &= \mathrm{Tr}\left[\begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{E_1}{c} & \frac{E_2}{c} & \frac{E_3}{c} \\ -\frac{E_1}{c} & 0 & -B_3 & B_2 \\ -\frac{E_2}{c} & B_3 & 0 & -B_1 \\ -\frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \right] \\ &= 2\left(\vec{B}^2 - \frac{\vec{E}^2}{c^2} \right) , \end{aligned} \tag{4.22}$$

where "Tr" denotes the trace operation in the equation above. This quantity is the same in all reference frames.

Lorentz transformations of the fields

The field strength tensor can now be used to determine the behaviour of the electric and magnetic fields under Lorentz transformations. In order to transform a tensor of rank n, however, more than one transformation tensor Λ is used. In fact, **to perform a Lorentz transformation, each Lorentz index μ must be transformed with one transformation tensor $\Lambda^{\mu'}_{\ \mu}$.** Hence, the field strength tensor in the primed system reads

$$F'^{\mu'\nu'} = \Lambda^{\mu'}_{\ \mu} \Lambda^{\nu'}_{\ \nu} F^{\mu\nu} , \tag{4.23}$$

if the Λ connect the two systems. Specifying for the case, where the two systems move with constant relative velocity $\vec{v} = v\vec{e}_x$ results in

$$F'^{\mu\nu} = \begin{pmatrix} 0 & -E_1'/c & -E_2'/c & -E_3'/c \\ E_1'/c & 0 & -B_3' & B_2' \\ E_2'/c & B_3' & 0 & -B_1' \\ E_3'/c & -B_2' & B_1' & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma & -\frac{v\gamma}{c} & 0 & 0 \\ -\frac{v\gamma}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v\gamma}{c} & 0 & 0 \\ -\frac{v\gamma}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -\frac{E_1}{c} & -\gamma\frac{E_2}{c} + \frac{v}{c}\gamma B_3 & -\gamma\frac{E_3}{c} - \frac{v}{c}\gamma B_2 \\ \frac{E_1}{c} & 0 & \gamma\frac{v}{c}\frac{E_2}{c} + \gamma B_3 & -\gamma\frac{v}{c}\frac{E_3}{c} - \gamma B_2 \\ \gamma\frac{E_2}{c} - \frac{v}{c}\gamma B_3 & -\gamma\frac{v}{c}\frac{E_2}{c} + \gamma B_3 & 0 & -B_1 \\ \gamma\frac{E_3}{c} + \frac{v}{c}\gamma B_2 & -\gamma\frac{v}{c}\frac{E_3}{c} - \gamma B_3 & B_1 & 0 \end{pmatrix} \tag{4.24}$$

The first obvious statement is that under Lorentz transformations, electric and magnetic fields start to mix. In other words, components of electric fields in one reference frame become magnetic in another frame and vice versa. And seemingly, one may find situations, where in the original frame there are only electrical fields giving rise to electric and magnetic fields in the primed system. Closer inspection now allows to identify

$$E_1' = E_1 \quad \text{and} \quad B_1' = B_1\,, \tag{4.25}$$

and

$$E_{2,3}' = \gamma\left(E_{2,3} \mp vB_{3,2}\right) \quad \text{and} \quad B_{2,3}' = \gamma\left(B_{2,3} \pm vE_{3,2}\right)\,. \tag{4.26}$$

This identification gives rise to the more general set of transformation laws

$$\begin{aligned} \vec{E}_\parallel' &= \vec{E}_\parallel & \vec{B}_\parallel' &= \vec{B}_\parallel \\ \vec{E}_\perp' &= \frac{\vec{E}_\perp + \vec{v}\times\vec{B}_\perp}{\sqrt{1-v^2/c^2}} & \vec{B}_\perp' &= \frac{\vec{B}_\perp - \vec{v}\times\vec{E}_\perp}{\sqrt{1-v^2/c^2}}\,, \end{aligned}$$

where the suvscripts $\|$ and $\perp$ denote the vectors parallel with $\vec{v}$ or perpendicular to it.

4.2 Electromagnetic waves