Proceedings of the CAS–CERN Accelerator School: Free Electron Lasers and Energy Recovery Linacs, Hamburg, Germany, 31 May–10 June 2016, edited by R. Bailey, CERN Yellow Reports: School Proceedings, Vol. 1/2018, CERN-2018-001-SP (CERN, Geneva, 2018)

Short Overview of Special Relativity and Invariant Formulation of Electrodynamics

W. Herr
CERN, Geneva, Switzerland

Abstract
The basic concepts of special relativity are presented in this paper. Consequences for the design and operation of particle accelerators are discussed, along with applications. Although all branches of physics must fulfil the principles of special relativity, the focus of this paper is the application to electromagnetism. The formulation of physics laws in the form of four-vectors allows a fully invariant formulation of electromagnetic theory and a reformulation of Maxwell's equations. This significantly simplifies the treatment of moving charges in electromagnetic fields and can explain some open questions.

Keywords
Special relativity; electrodynamics; four-vectors.

1 Introduction and motivation

As a principle in physics, the laws of physics should take the same form in all frames of reference, i.e., they describe a symmetry, a very basic concept in modern physics. This concept of relativity was introduced by Galileo and Newton in the framework of classical mechanics. Classical electromagnetic theory as formulated by Maxwell's equations leads to asymmetries when applied to moving charges [1,2]. In this context, classical mechanics and classical electromagnetism do not fulfil the same principles of relativity. The theory of special relativity is a generalization of the Galilean and Newtonian concepts of relativity. It also paved the way to a consistent theory of quantum mechanics. It considerably simplifies the form of physics because the unity of space and time as formulated by Minkowski also applies to force and power, time and energy, and last, but not least, to electric current and charge densities. The formulation of electromagnetic theory in this framework leads to a consistent picture and explains such concepts as the Lorentz force in a natural way. Starting from basic considerations and the postulates for special relativity, we develop the necessary mathematical formalism and discuss consequences, such as length contraction, time dilation, and the relativistic Doppler effect, to mention some of the most relevant. The introduction of four-vectors automatically leads to a relativistically invariant formulation of Maxwell's equations, together with the laws of classical mechanics.

Unlike other papers on relativity, this paper concentrates on aspects of electromagnetism; other popular phenomena, such as paradoxes, are left out.

2 Concepts of relativity

The concept of relativity was introduced by Galileo and Newton and applied to classical mechanics. It was proposed by Einstein that a similar concept should be applicable when electromagnetic fields are involved. We shall move from the classical principles to electrodynamics and assess the consequences.

2.1 Relativity in classical mechanics

In the following, the terminology and definitions used are:

– co-ordinates for the formulation of physics laws:

https://doi.org/10.23730/CYRSP-2018-001.27

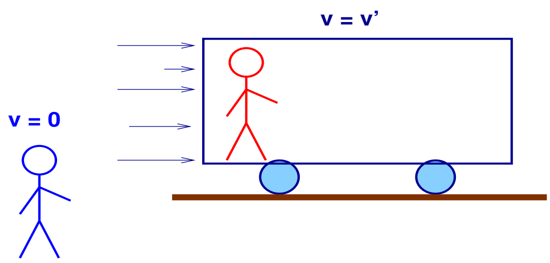

Fig. 1: Two different frames: a resting and a moving observer

  - space co-ordinates: $\vec{x} = (x, y, z)$ (not necessarily Cartesian);
  - time: t
    (side note: it might be better practice to use $\vec{r} = (x, y, z)$ instead of $\vec{x}$ as the position vector to avoid confusion with the x-component but we maintain this convention to be compatible with other textbooks and the conventions used later);
- definition of a *frame*:
  - where we observe physical phenomena and properties as functions of their position $\vec{x}$ and time t;
  - an *inertial frame* is a frame moving at a constant velocity;
  - in different frames, $\vec{x}$ and t are usually different;
- definition of an *event*:
  - something happening at $\vec{x}$ at time t is an 'event', given by four numbers: $(x, y, z), t$.

An example for two frames is shown in Fig. 1: one observer is moving at a constant relative velocity v' and another is observing from a resting frame.

2.2 Galileo transformation

How do we relate observations, e.g., the falling object in the two frames shown in Fig. 2?

- We have observed and described an event in rest frame F using co-ordinates (x, y, z) and time t, i.e., have formulated the physics laws using these co-ordinates and time.
- To describe the event in another frame F' moving at a constant velocity in the x-direction v_x, we describe it using co-ordinates (x', y', z') and t'.
- We need a transformation for:
  (x, y, z) and $t \Rightarrow (x', y', z')$ and t'.

The laws of classical mechanics are invariant, i.e., have the same form with the transformation:

$$\begin{aligned} x' &= x - v_x t\,, \\ y' &= y\,, \\ z' &= z\,, \\ t' &= t\,. \end{aligned} \tag{1}$$

The transformation (Eq. (1)) is known as the Galileo transformation. Only the position in the direction of the moving frame is transformed; time remains an absolute quantity.

Fig. 2: Observing a falling object from a moving and from a resting frame

2.3 Example of an accelerated object

An object falling with an acceleration g in the moving frame (Fig. 2, left) falls in a straight line observed within this frame.

Equation of motion in a moving frame $x'(t')$ and $y'(t')$:

$$\begin{aligned} x'(t') &= 0\,, \\ v'_y(t') &= -g \cdot t'\,, \\ y'(t') &= \int v'_y(t')\mathrm{d}t' = -\frac{1}{2}gt'^2\,. \end{aligned} \tag{2}$$

To get the equation of motion in the rest frame $x(t)$ and $y(t)$, the Galileo transform is applied:

$$\begin{aligned} y(t) &= y'(t')\,, \\ t &= t'\,, \\ x(t) &= x' + v_x \cdot t = v_x \cdot t\,, \end{aligned} \tag{3}$$

and one obtains for the trajectories $y(t)$ and $y(x)$ in the rest frame:

$$y(t) = -\frac{1}{2}gt^2\,, \qquad y(x) = -\frac{1}{2}g\frac{x^2}{v_x^2}\,. \tag{4}$$

From the resting frame, $y(x)$ describes a parabola (Fig. 2, right-hand side).

2.4 Addition of velocities

An immediate consequence of the Galileo transformation (Eq. (1)) is that the velocities of the moving object and the moving frame must be added to get the observed velocity in the rest frame:

$$v = v' + v''\,, \tag{5}$$

because (e.g., moving with the speed v_x in the x-direction):

$$\frac{\mathrm{d}x'}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t} - v_x\,. \tag{6}$$

As a very simple example (Fig. 3), the total speed of the object is 191 m/s.

2.5 Problems with Galileo transformation applied to electromagnetism

Applied to electromagnetic phenomena, the Galileo transformation exhibits some asymmetries. Assume a magnetic field and a conducting coil moving relative to the magnetic field. An induced current will be measured in the coil (Fig. 4). Depending on the frame of the observer, the interpretation of the observation is different.

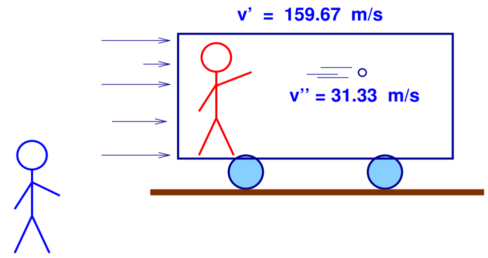

Fig. 3: Measured velocities of an object as observed from the co-moving and rest frames

Fig. 4: Effect of relative motion of a magnetic field and a conducting coil, observed from a co-moving and the rest frame.

– If you sit on the coil, you observe a changing magnetic field, leading to a circulating electric field inducing a current in the coil:

$$\frac{\mathrm{d}\vec{B}}{\mathrm{d}t} \Rightarrow \vec{\nabla} \times \vec{E} \Rightarrow \vec{F} = q \cdot \vec{E} \Rightarrow \text{current in coil}\,. \tag{7}$$

– If you sit on the magnet, you observe a moving charge in a magnetic field, leading to a force on the charges in the coil:

$$\vec{B} = \text{const., moving charge} \Rightarrow \vec{F} = q \cdot \vec{v} \times \vec{B} \Rightarrow \text{current in coil}\,. \tag{8}$$

The observed results are identical but seemingly caused by very different mechanisms! One may ask whether the physics laws are different, depending on the frame of observation.

A quantitative form can be obtained by applying the Galileo transformation to the description of an electromagnetic wave. Maxwell describes light as waves; the wave equation reads:

$$\left(\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\right)\Psi = 0\,. \tag{9}$$

Applying the Galileo transformation ($x = x' - vt, y' = y, z' = z, t' = t$), we get the wave equation in the moving frame:

$$\left(\left[1 - \frac{v^2}{c^2}\right]\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \frac{2v}{c^2}\frac{\partial^2}{\partial x \partial t} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Psi = 0\,. \tag{10}$$

The form of the transformed equation is rather different in the two frames.

The Maxwell equations are not compatible with the Galileo transformation.

3 Special relativity

To solve this riddle, one can consider three possible options.

1. Maxwell's equations are wrong and should be modified to be invariant with Galileo's relativity (unlikely).

2. Galilean relativity applies to classical mechanics, but not to electromagnetic effects and light has a reference frame (ether). Was defended by many people, sometimes with obscure concepts…
3. A relativity principle different from Galileo for *both* classical mechanics and electrodynamics (requires modification of the laws of classical mechanics).

Against all odds and with the disbelief of his colleagues, Einstein chose the last option.

3.1 Postulate for special relativity

To arrive at the new formulation of relativity, Einstein introduced three postulates.

- All physical laws in inertial frames must have equivalent forms.
- The speed of light in a vacuum c must be the same in all frames.
- It requires a transformations (not Galilean) that makes *all* physics laws look the same.

3.2 Lorentz transformation

The transformation requires that the co-ordinates must be transformed differently, satisfying the three postulates.

Writing the equations for the front of a moving light wave in F and F':

$$F : x^2 + y^2 + z^2 - c^2t^2 = 0\,, \tag{11}$$

$$F' : x'^2 + y'^2 + z'^2 - c'^2t'^2 = 0\,. \tag{12}$$

The constant speed of light requires $c = c'$ in both equations. This leads to a set of equations known as the Lorentz transformation (Eq. (13)).

$$\begin{aligned} x' &= \frac{x - vt}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} = \gamma \cdot (x - vt)\,, \\ y' &= y\,, \\ z' &= z\,, \\ t' &= \frac{t - \frac{v \cdot x}{c^2}}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} = \gamma \cdot \left(t - \frac{v \cdot x}{c^2}\right)\,. \end{aligned} \tag{13}$$

The main difference from the Galileo transformation is that it requires a transformation of the time t. It is a direct consequence of the required constancy of the speed of light. This tightly couples the position and time and they have to be treated on equal footing.

It is common practice to introduce the relativistic variables γ and β_{r}:

$$\gamma = \frac{1}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} = \frac{1}{\sqrt{(1 - \beta_{\mathrm{r}}^2)}}\,, \tag{14}$$

where β_{r} is:

$$\beta_{\mathrm{r}} = \frac{v}{c}\,. \tag{15}$$

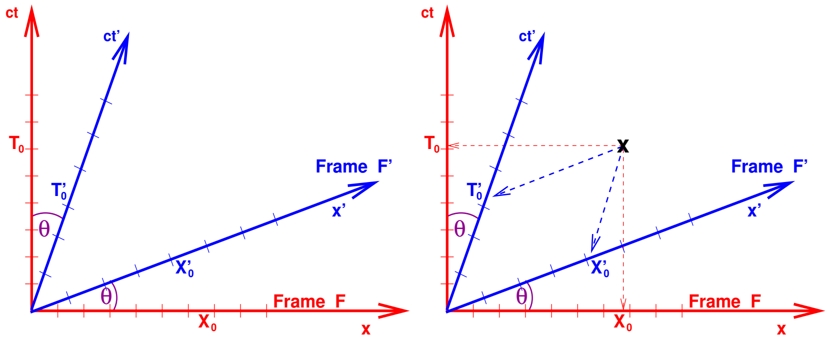

Fig. 5: The Lorentz transformation between frame F and F'. This representation is known as a Minkowski diagram.

3.3 Minkowski diagram—pictorial representation of the Lorentz transformation

An illustration of the Lorentz transformation is shown in Fig. 5. Starting from the orthogonal reference frame and using the transformation of position *and* time, both axes of the new reference system appear tilted, where the tilt angle depends on the velocity of the moving frame:

$$\tan\theta = \frac{v}{c} = \beta\,. \tag{16}$$

The position and time in the two reference frames can easily be obtained by the projection of an event onto the axes of the two frames (Fig. 5, right-hand side).

Contrary to normal (i.e., circular) rotation, where the axes remain perpendicular to each other, this type of rotation is also known as hyperbolic rotation. To quantify such a rotation, another angle ψ is introduced as:

$$\tanh\psi = \frac{v}{c} = \beta\,. \tag{17}$$

This angle ψ is also known as the rapidity. As a consequence we have:

$$\cosh\psi = \gamma \tag{18}$$

and

$$\sinh\psi = \gamma\beta\,. \tag{19}$$

Some applications become easier using this formulation.

3.4 Transformation of velocities

We assume a frame F' moving with constant speed of $\vec{v} = (v, 0, 0)$ relative to frame F. An object inside the moving frame is assumed to move with $\vec{v}' = (v'_x, v'_y, v'_z)$.

The velocity $\vec{v} = (v_x, v_y, v_z)$ of the object in the frame F is computed using the Lorentz transformation (Eq. (13))

$$v_x = \frac{v'_x + v}{1 + \frac{v'_x v}{c^2}}\,, \qquad v_y = \frac{v'_y}{\gamma\left(1 + \frac{v'_x v}{c^2}\right)}\,, \qquad v_z = \frac{v'_z}{\gamma\left(1 + \frac{v'_x v}{c^2}\right)}\,. \tag{20}$$

Adding two speeds v_1 and v_2:

$$v = v_1 + v_2 \qquad \Rightarrow \qquad v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}\,. \tag{21}$$

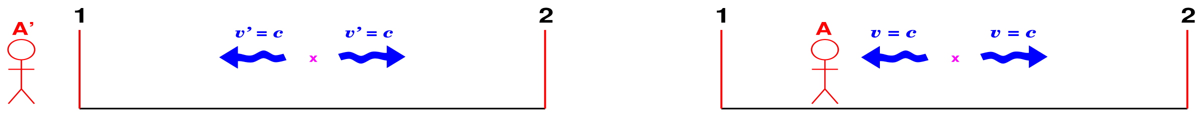

Fig. 6: Flash of light emitted in a resting frame, observed by two observers within and outside the frame

From (Eq. (21)) it can easily be seen that the speed of light can never be exceeded, in agreement with the second postulate.

An interesting result can be observed using the rapidity for this calculation. Since we have $\tanh\psi = \beta = v/c$, we can reformulate Eq. (21) as:

$$\tanh\psi = \frac{\tanh\psi_1 + \tanh\psi_2}{1 + \tanh\psi_1 \tanh\psi_2} = \tanh(\psi_1 + \psi_2)\,, \tag{22}$$

i.e., the rapidities can be added.

4 Consequences of special relativity

The use of the Lorentz transformation between two inertial frames and the required transformation of position *and* time has very significant consequences, which are rather counterintuitive.

- Space and time are *not* independent quantities.
- There is no absolute time and space, no absolute motion.
- Relativistic phenomena (with relevance for accelerators):
  - no speed of moving objects can exceed the speed of light;
  - (non-)simultaneity of events in independent frames;
  - Lorentz contraction;
  - time dilation;
  - relativistic Doppler effect;
  - Lorentz force.
- A formalism with four-vectors is introduced.
- Electrodynamics becomes very simple and consistent.

4.1 Simultaneity in special relativity

Assume that two events in frame F at (different) positions x_1 and x_2 happen simultaneously at times $t_1 = t_2$.

The times t'_1 and t'_2 in F' are obtained from the Lorentz transformation and

$$t'_1 = \gamma \cdot \left(t_1 - \frac{v \cdot x_1}{c^2}\right) \qquad \text{and} \qquad t'_2 = \gamma \cdot \left(t_2 - \frac{v \cdot x_2}{c^2}\right)\,. \tag{23}$$

One finds the surprising result that two events that are simultaneous at *different positions* x_1 and x_2 in F are not simultaneous in F': $x_1 \neq x_2$ in F implies that $t'_1 \neq t'_2$ in frame F'!

Assume the sequence of events depicted in Figs. 6–8. In a resting frame, a flash of light is emitted in the centre of the frame towards two detectors. An observer within the frame and another outside the frame observe the flash of light arriving simultaneously at detectors 1 and 2.

If the frame is moving, the detectors are reached at different times for the observer outside the frame.

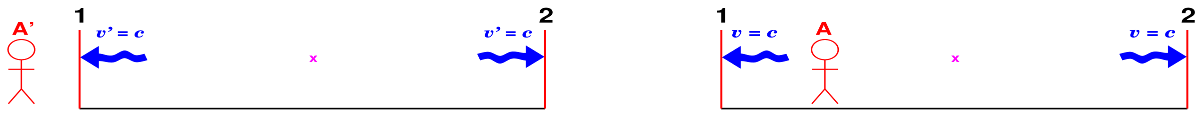

Fig. 7: For both observers, the flash of light reaches detectors 1 and 2 simultaneously

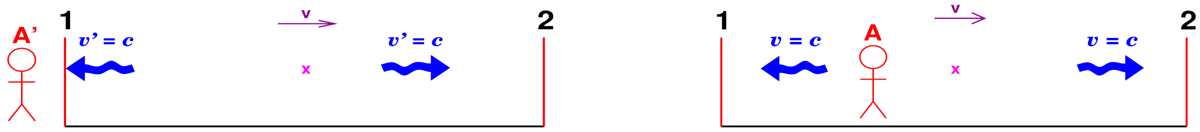

Fig. 8: Emitted in a moving frame, the flash reaches the detectors at different times for the outside observer

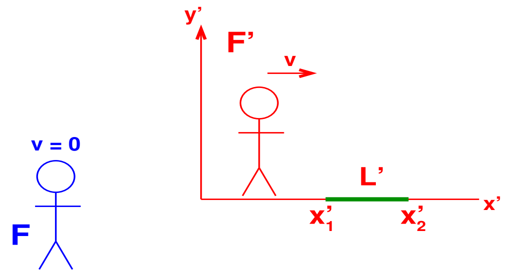

Fig. 9: Measuring the length of an object in a moving and a rest frame

Why should we bother about simultaneity?

- Simultaneity is not frame independent.
- It plays a pivotal role in special relativity.
- Almost all paradoxes are explained by it!
- Different observers see a different reality; in particular, the sequence of events can change!

For $t_1 < t_2$, we may find (not always!) a frame where $t_1 > t_2$ (the concept of before and after depends on the observer).

4.2 Length contraction

To measure the length of an object (Fig. 9), the procedure is to measure the position at *both* ends *simultaneously!*

The measured length of a rod in F' is the difference between the two positions $L' = x'_2 - x'_1$, measured simultaneously at a fixed time t' in frame F'.

From the frame F, we follow the same procedure, i.e., we have to measure simultaneously (!) the ends of the rod at a fixed time t in frame F, i.e.,:

$$L = x_2 - x_1 \,.$$

We therefore make a Lorentz transformation of the 'rod co-ordinates' into the rest frame:

$$x'_1 = \gamma \cdot (x_1 - vt) \qquad \text{and} \qquad x'_2 = \gamma \cdot (x_2 - vt)\,, \tag{24}$$

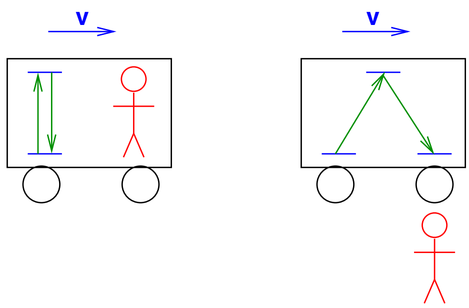

Fig. 10: Path of light between two mirrors in (left) a moving frame and (right) a rest frame

$$L' = x'_2 - x'_1 = \gamma \cdot (x_2 - x_1) = \gamma \cdot L\,. \tag{25}$$

We obtain

$$L = L'/\gamma\,. \tag{26}$$

The length appears contracted in the rest frame (Eq. (26)), depending on the relative velocity, i.e., the relativistic γ.

In accelerators, this has consequences, as objects in the frame of the particle and the frame of the accelerator appear to have different lengths. This is of particular importance for bunch length, electromagnetic fields, magnets, and the distances between magnets and other objects in the accelerator.

4.3 Time dilation

Applying the Lorentz transformation, we have to transform the time t as well as the position.

We assume a moving frame where a flash of light is moving upwards and reflected downwards by a mirror (Fig. 10).

We assume that the frame moves with velocity v. Seen from outside, the flash arrives at the mirror, but at a different position. This means that the apparent total path is longer, but c must be the same. The geometry of this process is shown in Fig. 11.

In frame F': light travels distance L in time $\Delta t'$.
In frame F: light travels distance D in time $\Delta t'$;
the entire system moves distance d in time Δt.

We look at the trajectories in the two frames and simple calculation leads to the result that the time needed by the flash is longer by the factor γ when it is observed from the outside. After a round trip, the full distance observed in the moving frame is $2 \cdot L$; measured from the outside, it is $2 \cdot D$.

$$L = c \cdot \Delta t'\,, \qquad D = c \cdot \Delta t\,, \qquad d = v \cdot \Delta t\,,$$
$$(c \cdot \Delta t)^2 = (c \cdot \Delta t')^2 + (v \cdot \Delta t)^2\,.$$

We obtain:

$$\rightarrow \Delta t = \gamma \cdot \Delta t'\,. \tag{27}$$

4.4 Proper time and proper length

This derivation can lead to some confusion.

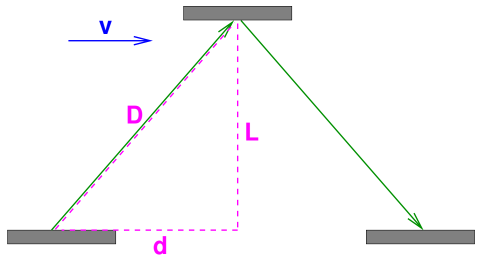

Fig. 11: Path of light between two mirrors, as observed from a rest frame

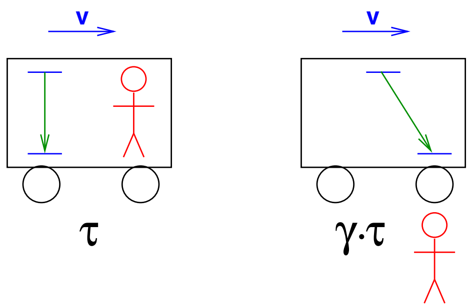

Fig. 12: Time measured within the moving frame (left-hand side) and from the rest frame (right-hand side). In the moving frame, the measured time is always the proper time τ, independent of the velocity of the moving frame.

- The car is moving: $\Delta t = \gamma \cdot \Delta t'$.
- The observer is moving: $\Delta t' = \gamma \cdot \Delta t$.

This seems like a contradiction. This paradox is solved by introducing the concept of proper time τ. The proper time τ is the time measured by the observer at rest relative to the process.

Or: the proper time for a given observer is measured by the clock that travels with the observer:

$$c^2 \Delta\tau^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \,. \tag{28}$$

Equation (28) defines the proper time τ.

A similar argument holds for the Lorentz contraction and we define the 'proper length' in the same way.

We illustrate this in Fig. 12, where the time is measured within the frame as the proper time τ. From outside, the measured time is $\gamma \cdot \tau$.

To summarize, it is found that time and distances are relative.

- τ is a fundamental time: the proper time τ.
- The proper time is the time measured by an observer in its own frame.
- From frames moving *relative* to it, the time appears longer.
- $\mathcal{L}$ is a fundamental length: the proper length $\mathcal{L}$.
- The proper length is the length measured by an observer in its own frame.
- From frames moving *relative* to it, the length appears shorter.

4.5 Proper acceleration

Although not very much used in accelerator physics, one has a proper acceleration. It can be computed as the derivative of the rapidity with respect to the proper time:

$$\alpha = \frac{\mathrm{d}\psi}{\mathrm{d}\tau} \, .$$

4.6 Relativistic space travel

The formula for time dilation also holds for an accelerated system! Assuming a spaceship starting from Earth and moving with a constant acceleration, g = 9.81 m/ s^2. We denote the time on Earth t and the proper time in the spaceship t_p. For β, we then have:

$$\beta = \tanh\left(\frac{v}{c}\right) = \tanh\left(\frac{g \cdot t_\mathrm{p}}{c}\right) \, , \tag{29}$$

and therefore for γ:

$$\gamma = \cosh\left(\frac{g \cdot t_\mathrm{p}}{c}\right) = \sqrt{1 + \left(\frac{g \cdot t_\mathrm{p}}{c}\right)^2} \, . \tag{30}$$

Finally, for the distance to the start (e.g., Earth), we obtain:

$$d = \left(\frac{c^2}{g}\right) \cdot \left[\cosh\left(\frac{g \cdot t_\mathrm{p}}{c}\right) - 1\right] \, . \tag{31}$$

When 12 years have passed for the passenger of the spaceship, the distance to Earth is $d \approx 120\,000$ light years! This is approximately the diameter of the Milky Way!

4.7 Muon lifetime

A popular example is the lifetime of a μ particle moving at high speed. We can compute the (observed) lifetime of the muon.

In the frame of the muon, the lifetime is $\tau \approx 2 \cdot 10^{-6}$ s. Measured from the laboratory frame, the time $\gamma \cdot \tau$ is observed. The muon appears to have a longer lifetime, in contradiction with the postulates.

A clock in the muon frame shows the *proper time* and the muon decays in $\approx 2 \cdot 10^{-6}$ s, independent of the muon's speed.

Seen from the lab frame, the muon lives γ times longer. In the muon storage ring at CERN, the lifetime of muons circulating with γ = 29.327 was found to be dilated to 64.378 µs, confirming time dilation.

4.8 Moving electron

Important effects are experienced by a fast-moving electron in an accelerator. We assume: $v \approx c$.

Bunch length:
In lab frame: σ_z,   In frame of electron: $\gamma \cdot \sigma_z$.

Length of an object (e.g., magnet, *distance* between magnets!):
In lab frame: L,   In frame of electron: L/γ.

4.9 Relativistic Doppler effect

A rather important relativistic effect is the Doppler shift of the frequency of a fast-moving particle. In light sources, this effect is most significant. Unlike sound, light does not have a medium of propagation and the nature of this effect is purely relativistic.

The frequency observed ν depends on the velocity γ and the observation angle θ. The frequency of the light to an observer looking against the direction of motion is increased by a factor γ; at high speeds, this is a very significant effect:

$$\nu = \nu_0 \cdot \gamma \cdot (1 - \beta_r \cdot \cos(\theta)) \,. \tag{32}$$

4.10 Everyday effects

Although the speeds we experience in everyday life are small compared with the speed of light, time dilation has to be taken into account for some applications.

4.10.1 Intercontinental flights

As the time for moving objects is passing more slowly with respect to a reference on Earth, one expects that during an eastward intercontinental flight the passengers age more slowly. This was tested and confirmed experimentally with atomic clocks in 1971 and 1977 [3].

For a 6 hour flight eastwards with a regular aeroplane cruising at ≈900 km/h, one computes a difference of 25–30 ns. This can easily be measured [3] but has little effect in everyday life. Taking into account the effects of general relativity, the effect is very different (in the opposite direction, but still small). General relativity is not the topic of this lecture, but the relevant equations are given in Appendix A for the gifted reader.

4.10.2 GPS

Contrary to flights with an aeroplane, other everyday effects experience very strong relativistic effects. A prominent example is the global positioning system (GPS). The flight parameters of the satellites of the system and the effects are:

– Orbital speed, i.e., relative to a reference on Earth is 14000 km/h ≈ 3.9 km/s

$$\Rightarrow \beta \approx 1.3 \cdot 10^{-5} \,, \qquad \gamma \approx 1.000000000084 \,.$$

– This is very small, but it accumulates 7 μs during one day compared with the reference time on Earth!
– After one day, the position is wrong by ≈2 km!
– Including general relativity, the error is as large as 10 km per day; the computation is given in Appendix A.

Without corrections for the effects of special and general relativity, the satellite navigation cannot work. Countermeasures:

– A minimum of four satellites is used (avoid reference time on Earth);
– The data transmission frequency is reduced from 1.023 MHz to 1.022999999543 MHz prior to launch.

5 Relativistic mass and momentum

After one has established the relativistic corrections for position and time, it is necessary to evaluate the consequences for momentum, energy, and mass of a moving particle.

5.1 Consequences of momentum conservation

To simplify the computation, we assume that an object moves in frame F' with $\vec{u}' = (0, u'_y, 0)$. To have an invariant form, one requires that the expression

$$\vec{F} = m \cdot \vec{a} = m \cdot \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} \tag{33}$$

has the same form in all frames.

The transverse momentum must be conserved; this demands:

$$\begin{aligned} p_y &= p'_y \,, \\ m u_y &= m' u'_y \,, \\ m u'_y/\gamma &= m' u'_y \,. \end{aligned} \tag{34}$$

This implies that

$$m = \gamma m' \,. \tag{35}$$

As a consequence of momentum conservation, *mass must also be transformed*! Using the expression for the mass m (using $m' = m_0$):

$$m = m_0 \sqrt{1 - \left(\frac{v}{c}\right)^2} = \gamma \cdot m_0 \,, \tag{36}$$

and expand it for small speeds:

$$m \cong m_0 + \frac{1}{2} m_0 v^2 \left(\frac{1}{c^2}\right) \;; \tag{37}$$

multiplied by c^2:

$$mc^2 \cong m_0 c^2 + \frac{1}{2} m_0 v^2 = m_0 c^2 + T \,. \tag{38}$$

The second term is the kinetic energy T :

$$E = mc^2 = m_0 c^2 + T \,. \tag{39}$$

5.2 Interpretation of relativistic mass and energy

- The total energy E is $E = mc^2$.
- It is the sum of the kinetic energy and the rest energy.
- The energy of a particle in its rest frame is $E_0 = m_0 c^2$.

Using the definition of the relativistic mass $m = \gamma m_0$, we can write:

$$E = m \cdot c^2 = \gamma m_0 \cdot c^2 \,. \tag{40}$$

For any object, $m \cdot c^2$ is the total energy, and this follows directly from momentum conservation. We can say that m is the mass (energy) of the object 'in motion', while m_0 is the mass (energy) of the object 'at rest'. The mass m is not the same in all inertial systems; the *rest mass* m_0 is. Following previous arguments, m_0 is the 'proper mass'.

5.3 Relativistic relations

Starting from

$$p = mv\,, \tag{41}$$

with $m = \gamma m_0$:

$$p = \gamma \cdot m_0 v = \gamma \cdot \beta \cdot c \cdot m_0\,. \tag{42}$$

We rewrite:

$$E^2 = m^2c^4 = \gamma^2 m_0^2 c^4 = (1 + \gamma^2\beta^2) m_0^2 c^4\,, \tag{43}$$

and finally get:

$$E^2 = (m_0c^2)^2 + (pc)^2 \qquad \Rightarrow \qquad \frac{E}{c} = \sqrt{(m_0c)^2 + p^2}\,. \tag{44}$$

This is a rather important formula in practice, e.g., for kinematics in accelerators.

5.4 Units used in special relativity and particle physics

Standard SI units are not very convenient; other units are much easier to use and have become a standard in accelerator and particle physics:

$$[E] = \text{ eV}\,, \qquad [p] = \text{ eV}/c\,, \qquad [m] = \text{ eV}/c^2\,.$$

Then one can use a very convenient form for Eq. (44):

$$E^2 = m_0^2 + p^2 m\,. \tag{45}$$

Examples for elementary particles (proton):
Mass of a proton: $m_\mathrm{p} = 1.672 \cdot 10^{-27}$ kg,
Energy(at rest): $m_\mathrm{p}c^2 = 938$ MeV $= 0.15$ nJ.

We can take an everyday object like a ping-pong ball:
Ping-pong ball: $m_\mathrm{pp} = 2.7 \cdot 10^{-3}$ kg ($\approx 1.6 \cdot 10^{24}$ protons),
Energy (at rest): $m_\mathrm{pp}c^2 = 1.5 \cdot 10^{27}$ MeV $= 2.4 \cdot 10^{14}$ J.
This corresponds to an energy of:

- $\approx$750 000 times the full LHC beam;
- $\approx$60 kilotons of TNT.

The typical kinetic energy of a ping-pong ball is completely negligible.

5.5 Mass of a moving particle in a particle accelerator

The mass of a fast-moving particle increases as:

$$m = \gamma m_0 = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}\,. \tag{46}$$

When we accelerate:

- for $v \ll c$:
  - E, m, p, v increase…
- for $v \approx c$:
  - E, m, p increase, but v (almost) does not!

Table 1: Relations between relativistic parameters

	cp	T	E	γ
β	$\frac{1}{\sqrt{\left(\frac{E_0}{cp}\right)^2+1}}$	$\sqrt{1-\frac{1}{\left(1+\frac{T}{E_0}\right)^2}}$	$\sqrt{1-\left(\frac{E_0}{E}\right)^2}$	$\sqrt{1-\gamma^{-2}}$
cp	cp	$\sqrt{T(2E_0+T)}$	$\sqrt{E^2-E_0^2}$	$E_0\sqrt{\gamma^2-1}$
E_0	$\frac{cp}{\sqrt{\gamma^2-1}}$	$T/(\gamma-1)$	$\sqrt{E^2-c^2p^2}$	E/γ
T	$cp\sqrt{\frac{\gamma-1}{\gamma+1}}$	T	$E-E_0$	$E_0(\gamma-1)$
γ	$cp/E_0\beta$	$1+T/E_0$	E/E_0	γ

Table 2: Logarithmic relations between relativistic parameters

	$\frac{\mathrm{d}\beta}{\beta}$	$\frac{\mathrm{d}p}{p}$	$\frac{\mathrm{d}T}{T}$	$\frac{\mathrm{d}E}{E}=\frac{\mathrm{d}\gamma}{\gamma}$
$\frac{\mathrm{d}\beta}{\beta}$	$\frac{\mathrm{d}\beta}{\beta}$	$\frac{1}{\gamma^2}\frac{\mathrm{d}p}{p}$	$\frac{1}{\gamma(\gamma+1)}\frac{\mathrm{d}T}{T}$	$\frac{1}{(\beta\gamma)^2}\frac{\mathrm{d}\gamma}{\gamma}$
$\frac{\mathrm{d}p}{p}$	$\gamma^2\frac{\mathrm{d}\beta}{\beta}$	$\frac{\mathrm{d}p}{p}$	$[\gamma/(\gamma+1)]\frac{\mathrm{d}T}{T}$	$\frac{1}{\beta^2}\frac{\mathrm{d}\gamma}{\gamma}$
$\frac{\mathrm{d}T}{T}$	$\gamma(\gamma+1)\frac{\mathrm{d}\beta}{\beta}$	$\left(1+\frac{1}{\gamma}\right)\frac{\mathrm{d}p}{p}$	$\frac{\mathrm{d}T}{T}$	$\frac{\gamma}{(\gamma-1)}\frac{\mathrm{d}\gamma}{\gamma}$
$\frac{\mathrm{d}E}{E}$	$(\beta\gamma)^2\frac{\mathrm{d}\beta}{\beta}$	$\beta^2\frac{\mathrm{d}p}{p}$	$\left(1-\frac{1}{\gamma}\right)\frac{\mathrm{d}T}{T}$	$\frac{\mathrm{d}\gamma}{\gamma}$
$\frac{\mathrm{d}\gamma}{\gamma}$	$(\gamma^2-1)\frac{\mathrm{d}\beta}{\beta}$	$\frac{\mathrm{d}p}{p}-\frac{\mathrm{d}\beta}{\beta}$	$\left(1-\frac{1}{\gamma}\right)\frac{\mathrm{d}T}{T}$	$\frac{\mathrm{d}\gamma}{\gamma}$

$$\beta=\frac{v}{c}\approx\sqrt{1-\frac{m_0^2c^4}{T^2}}\,. \tag{47}$$

A consequence of this effect is that, in a synchrotron, a particle with higher speed takes longer to complete a turn than a particle with lower speed. This leads to the effect of transition in synchrotrons.

For small rest masses (e.g., electrons), this effect is very small; the effect is only relevant for hadrons.

5.6 Kinematic relations

A summary of kinematic relations between the relativistic variables is shown in Tables 1 and 2.

As an example, one can compute the relative spread of particle velocities from the momentum spread:

$$\text{LHC (7 TeV):}\qquad \frac{\Delta p}{p}=10^{-4}\,,\qquad \frac{\Delta\beta}{\beta}=10^{-12}\,,$$

$$\text{LEP (100 GeV):}\qquad \frac{\Delta p}{p}=10^{-4}\,,\qquad \frac{\Delta\beta}{\beta}=10^{-15}\,.$$

The reason for the much smaller velocity spread in the LEP is, of course, the larger γ factor.

6 First summary

The theory of special relativity is derived from two postulates.

- Physics laws are the same in all inertial frames.
- The speed of light in vacuum c is the same in all frames and requires Lorentz transformation.

Consequences of special relativity are:

- simultaneity is not independent of the frame of observation;
- moving objects *appear* shorter;
- moving clocks *appear* to go slower;
- the mass is not independent of motion ($m = \gamma \cdot m_0$) and the total energy is $E = m \cdot c^2$;
- absolute space or time do not exist: where it happens and when it happens are not independent.

In the following, it is demonstrated how to simplify calculations using the concept of four-vectors.

7 Space–time and four-vectors

Since space and time are not independent, we must reformulate the physics laws, taking both into account on an equal footing:

$$t, \qquad \vec{x} = (x, y, z) \qquad \Rightarrow \qquad \text{Replace by one vector including the time.} \tag{48}$$

All four-vectors are constructed from a temporal and a spatial part, where the temporal part (e.g., time t) is multiplied by c to get the same units.

We need two types of four-vector (here *position four-vector*):

$$X^{\mu} = (ct, x, y, z) \qquad \text{and} \qquad X_{\mu} = (ct, -x, -y, -z)\,. \tag{49}$$

This is due to the 'skewed' reference system; for details see the bibliography.

Four-vectors of the type X^{μ} are called *contravariant* vectors; X_{μ} are called *covariant* vectors. It is common practice to use capital letters for four-vectors to distinguish from the 3D vectors. Using four-vectors, the Lorentz transformation can easily be written in matrix form:

$$X'^{\mu} = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = X^{\mu}\,, \tag{50}$$

$$X'^{\mu} = \Lambda \circ X^{\mu} \qquad (\Lambda \text{ for 'Lorentz'})\,, \tag{51}$$

but note:

$$X'_{\mu} = \begin{pmatrix} ct' \\ -x' \\ -y' \\ -z' \end{pmatrix} = \begin{pmatrix} \gamma & +\gamma\beta & 0 & 0 \\ +\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix} = X_{\mu}\,. \tag{52}$$

It can be verified that the transformation for covariant vectors is the inverse of Λ and:

$$X'_{\mu} = \Lambda^{-1} \circ X_{\mu}\,. \tag{53}$$

7.1 Hyperbolic transformation

Using the definitions from Eq. (17), the transformation becomes:

$$X'^{\mu} = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh(\psi) & -\sinh(\psi) & 0 & 0 \\ -\sinh(\psi) & \cosh(\psi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = X^{\mu} \,. \tag{54}$$

This takes the form of a rotation.

7.2 Vector operations

Having introduced four-vectors, we have to look at operations on and with four-vectors, in particular scalar products, because they play a key role in special relativity.

7.3 Scalar product

The well-known scalar product using Cartesian co-ordinates and the Euclidean metric is:

$$\vec{x} \cdot \vec{y} = (x_a, y_a, z_a) \cdot (x_b, y_b, z_b) = (x_a \cdot x_b + y_a \cdot y_b + z_a \cdot z_b) \,. \tag{55}$$

Space–time four-vectors are of the form:

$$A^{\mu} = (ct_a, x_a, y_a, z_a) \,, \qquad B_{\mu} = (ct_b, -x_b, -y_b, -z_b) \,.$$

The four-vector scalar product is:

$$A^{\mu} B_{\mu} = \underbrace{\sum_{\mu=0}^{3} A^{\mu} B_{\mu}}_{\text{Einstein convention}} = (ct_a \cdot ct_b - x_a \cdot x_b - y_a \cdot y_b - z_a \cdot z_b) \,. \tag{56}$$

In Eq. (56), we have used the so-called *Einstein convention* to write the equations in a more compact form: when an index appears more than once in an equation (here: μ) it implies summation over this index.

For many applications, a simplified rule can be used:

$$AB = (ct_a \cdot ct_b - x_a \cdot x_b - y_a \cdot y_b - z_a \cdot z_b) \,. \tag{57}$$

7.4 Four-vectors

We have important four-vectors:

Co-ordinates:	$X^{\mu} = (ct, x, y, z) = (ct, \vec{x})\,,$
Velocities:	$U^{\mu} = \frac{\mathrm{d}X^{\mu}}{\mathrm{d}\tau} = \gamma(c, \dot{\vec{x}}) = \gamma(c, \vec{u})\,,$
Momenta:	$P^{\mu} = mU^{\mu} = m\gamma(c, \vec{u}) = \gamma(mc, \vec{p})\,,$
Force:	$F^{\mu} = \frac{\mathrm{d}P^{\mu}}{\mathrm{d}\tau} = \gamma \frac{\mathrm{d}}{\mathrm{d}\tau}(mc, \vec{p})\,,$
Wave propagation vector:	$K^{\mu} = \left(\frac{\omega}{c}, \vec{k}\right)\,,$
We define the gradient operator as a four-vector:	$\partial^{\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right)\,.$

It is a key element in this analysis that *all* four-vectors A^{μ} transform as:

$$A'^{\mu} = \Lambda \circ A^{\mu} \,. \tag{58}$$

7.5 Invariant forms

The main objective of the principle of special relativity is to arrive at invariant laws of physics in different frames.

- The solution: write the laws of physics in terms of four-vectors and use Lorentz transformations between the frames.
- Without proof (it is rather straightforward using Eqs. (51) and (52)): *any* four-vector (scalar) product $Z^\mu Z_\mu$ has the same value in all inertial frames:

$$Z^\mu Z_\mu = Z'^\mu Z'_\mu \, . \tag{59}$$

All scalar products of any four-vectors are invariant!

but: $Z^\mu Z^\mu$ and $Z'_\mu Z'_\mu$ are not!

- Note: the four-vectors in the scalar product do not have to be of the same type.
- $P^\mu X_\mu$ is also an invariant four-vector product.

One should look at two particularly important invariants.

7.5.1 A special invariant—invariant velocity

From the velocity four-vector U^μ:

$$U^\mu = \gamma(c, \vec{u}) \, , \tag{60}$$

we get the scalar product:

$$U^\mu U_\mu = \gamma^2(c^2 - \vec{u}^2) = c^2 \, . \tag{61}$$

We find that the invariant of the velocity has the same value in all inertial frames and is the speed of light. The speed of light c is the same in all frames.

7.5.2 Invariant momentum

Starting from the four-momentum P^μ:

$$P^\mu = m_0 U^\mu = m_0\gamma(c, \vec{u}) = (mc, \vec{p}) = \left(\frac{E}{c}, \vec{p}\right) \, , \tag{62}$$

$$P'^\mu = m_0 U'^\mu = m_0\gamma(c, \vec{u'}) = (mc, \vec{p'}) = \left(\frac{E'}{c}, \vec{p'}\right) \, . \tag{63}$$

We can get another important invariant:

$$P^\mu P_\mu = P'^\mu P'_\mu = m_0^2 c^2 \, . \tag{64}$$

The invariant of the four-momentum vector is the mass m_0. It follows that the rest mass is the same in all frames (it has to be, otherwise we could not tell whether we are moving or not!)

7.6 Four-vectors and kinematics

The use of four-vectors enables a very straightforward and simple procedure to compute the kinematics of moving particles.

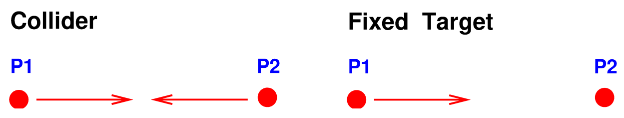

Fig. 13: Schematic illustration of particle collisions. Left: for the case of a colliding beam facility; right: when one particle is at rest (fixed-target collisions).

Table 3: Centre of mass energies for fixed-target collisions and colliding beams for different energies and particle types.

Collision	E **beam energy**	$E_{\rm cm}$ **(collider)**	$E_{\rm cm}$ **(fixed target)**
pp	315 (GeV)	630 (GeV)	24.3 (GeV)
pp	6500 (GeV)	13000 (GeV)	110.4 (GeV)
pp	90 (PeV)	180 (PeV)	13000 (GeV)
e^+e^-	100 (GeV)	200 (GeV)	0.320 (GeV)

7.6.1 Particle collisions

For collisions of particles in high energy physics experiments, the available centre of mass energy $E_{\rm cm}$ is the relevant parameter.

We distinguish two types of collision: when a particle hits a particle at rest and when two high energy particles collide, either head on or at an angle. This is shown in Fig. 13.

The computation of $E_{\rm cm}$ is rather straightforward when four-vectors are used. The centre of mass energy must be an invariant of the collision process and is therefore the scalar product of two four-vectors.

It was shown before that the relevant four-vector is the momentum four-vector. Taking the scalar product of the sum of the four-momentum of the colliding particles gives the centre of mass energy.

$$
\begin{aligned}
P_1^\mu &= (E, \vec{p})\,, \quad P_2^\mu = (E, -\vec{p})\,, & P_1^\mu &= (E, \vec{p})\,, \quad P_2^\mu = (m_0, 0)\,,\\
P^\mu &= P_1^\mu + P_2^\mu = (2E, 0)\,, & P^\mu &= P_1^\mu + P_2^\mu = (E + m_0, \vec{p})\,,\\
E_{\rm cm} &= \sqrt{P^\mu P_\mu} = 2 \cdot E\,, & E_{\rm cm} &= \sqrt{P^\mu P_\mu} = \sqrt{2 m_0 E}\,.
\end{aligned} \tag{65}
$$

This procedure also works for more than two colliding particles: $P^\mu = P_1^\mu + P_2^\mu + P_3^\mu + \ldots$ It works for any configuration; also, for particle decay.

In Table 3, a comparison is made of the centre of mass energies for fixed-target collisions and colliding beams for different energies and particle types. The differences are very significant; centre of mass energies as provided by the LHC cannot be reached by fixed-target collisions.

7.6.2 Particle decay

We consider a particle P_0 decaying into two (or more) particles (Fig. 14): $P_0 \Rightarrow P_1 + P_2$.

We can measure the properties of the decay products, (i.e., $\vec{p_2}$, $\vec{p_1}$, m_1, m_2, E_1, E_2). The parameters (in particular the mass m_0) of the original particle are unknown.

For every decay product, we construct the corresponding four-momentum from the measured values:

$$
\begin{aligned}
P_1^\mu &= (E_1\,, \vec{p_1}), & E_1 &= \sqrt{m_1^2 + \vec{p_1}^2}\,,\\
P_2^\mu &= (E_2\,, \vec{p_2}), & E_2 &= \sqrt{m_2^2 + \vec{p_2}^2}\,.
\end{aligned} \tag{66}
$$

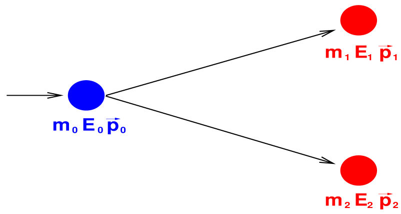

Fig. 14: A particle decaying into two particles

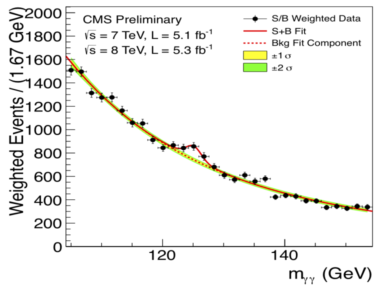

Fig. 15: Invariant mass of a $\gamma\gamma$ decay, showing an enhancement at the Higgs mass

The centre of mass of the decaying particle is an invariant: the rest mass m_0 of the particle.

We obtain the *sum* of the four-momenta, which is the four-momentum of the original particle:

$$P_0^\mu = (P_1^\mu + P_2^\mu) = (E_1 + E_2, \vec{p_1} + \vec{p_2})\,. \tag{67}$$

The particle mass we get from the scalar product and, since we know $P_0^2 = m_0^2 c^4$:

$$m_0 c^2 = \sqrt{(E_1 + E_2)^2 - (\vec{p_1}c + \vec{p_2}c)^2}\,. \tag{68}$$

When we draw m_0 for every observed decay, we obtain a histogram with the mass of the original particle (Fig. 15).

As an example, Fig. 15 shows the invariant mass of two photons, an expected channel for the decay of the Higgs particle. At the mass of the Higgs particle, we see an enhancement above the background. This procedure works for any number of decay products.

7.6.3 Particle scattering and four-vectors

In theoretical physics, the Mandelstam variables are numerical quantities that encode the energy, momentum, and angles of particles in a scattering process in a Lorentz-invariant fashion. They are used for scattering processes of two incoming particles and two outgoing particles.

Assume that there are two particles (P_1 and P_2) in the initial state of an interaction and two particles (P_3 and P_4) after the interaction. The two most important relations are described by the variables s and t, defined as:

$$s = (P_1 + P_2)^2 = (P_3 + P_4)^2\,, \tag{69}$$

$$t = (P_1 - P_3)^2 = (P_2 - P_4)^2\,. \tag{70}$$

The physical interpretation of these variables is rather straightforward. Compared with earlier results, the variable $s = (P_1 + P_2)^2$ is the square of the centre of mass energy in the collision process (Fig. 16).

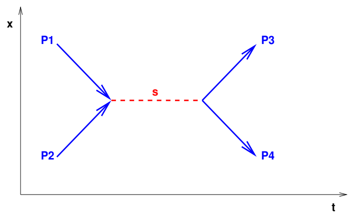

Fig. 16: Invariant variable s for colliding particles

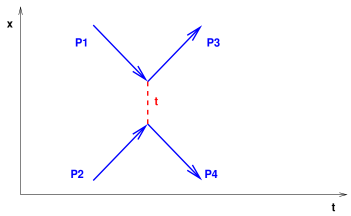

Fig. 17: Invariant variable t for colliding particles

The variable $t = (P_1 - P_3)^2$ describes an interaction where an incoming particle is scattered from another particle (Fig. 17), exchanging momentum t, typically through the exchange of an intermediate particle (vector boson).

7.6.4 Cross-sections and luminosity

The probability for a collision process between particles is characterized by the corresponding cross-section for this process.

According to the postulates, the cross-section must be an invariant in all frames. This requires the correction term from Eq. (71):

$$K = \sqrt{(\vec{v_1} - \vec{v_2})^2 - (\vec{v_1} \times \vec{v_2})^2/c^2}\,. \tag{71}$$

It is easily verified that in the case of two colliding beams with the same energy, this factor becomes $K = 2$.

Since the luminosity [4] must also be an invariant, the same correction factor must be used in the calculation of the luminosity.

8 Special relativity and electrodynamics

We come back to the original starting point and derive a relativistic formulation of Maxwell's equations.

8.1 Four-vector formulation of electromagnetic quantities

According to the rules for four-vectors, one can write the potentials and currents as four-vectors:

$$\phi, \vec{A} \Rightarrow A^{\mu} = \left(\frac{\phi}{c}, \vec{A}\right), \tag{72}$$

$$\rho, \vec{j} \Rightarrow J^{\mu} = (\rho \cdot c, \vec{j})\,. \tag{73}$$

What about the transformation of current and potentials? Since we have formulated the potentials and currents as four-vectors, we transform the four-current as:

$$\begin{pmatrix} \rho' c \\ j'_x \\ j'_y \\ j'_z \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho c \\ j_x \\ j_y \\ j_z \end{pmatrix} .$$

It transforms via: $J'^{\mu} = \Lambda J^{\mu}$ and the potential is transformed correspondingly: $A'^{\mu} = \Lambda A^{\mu}$. Since scalar products of four-vectors are invariant, one writes:

$$\partial_{\mu} J^{\mu} = \frac{\partial \rho}{\partial t} + \vec{\nabla}\vec{j} = 0 . \qquad (74)$$

Equation (74) implies the conservation (invariance) of charge.

8.2 Electromagnetic quantities and field tensor

Having written the currents and potentials as four-vectors [1], we derive a formulation for the fields. The *Magnetic field is derived from the potential*:

$$\vec{B} = \nabla \times \vec{A} ,$$

e.g., written explicitly for the x-component:

$$B_x = \frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} .$$

The scalar potential provides the *electric field:*

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} ,$$

e.g. written explicitly for the x-component:

$$E_x = -\frac{\partial A_0}{\partial x} - \frac{\partial A_1}{\partial t} = -\frac{\partial A_t}{\partial x} - \frac{\partial A_x}{\partial t} .$$

These operations can be performed for all components of the fields. Then the electromagnetic fields can be condensed to form a field tensor $F^{\mu\nu}$:

$$F^{\mu\nu} := \partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu} = \begin{pmatrix} 0 & \frac{-E_x}{c} & \frac{-E_y}{c} & \frac{-E_z}{c} \\ \frac{E_x}{c} & 0 & -B_z & B_y \\ \frac{E_y}{c} & B_z & 0 & -B_x \\ \frac{E_z}{c} & -B_y & B_x & 0 \end{pmatrix} . \qquad (75)$$

It transforms via: $F'^{\mu\nu} = \Lambda F^{\mu\nu} \Lambda^T$ (using the same transformation Λ as before). This corresponds to the transformation of the components of the fields into a moving frame.

Using the Lorentz transformation for $F^{\mu\nu}$, and writing the components explicitly, one gets easily:

$$\begin{aligned} E'_x &= E_x , & B'_x &= B_x , \\ E'_y &= \gamma(E_y - v \cdot B_z) , & B'_y &= \gamma\left(B_y + \frac{v}{c^2} \cdot E_z\right) , \\ E'_z &= \gamma(E_z + v \cdot B_y) z , & B'_z &= \gamma\left(B_z - \frac{v}{c^2} \cdot E_y\right) z . \end{aligned} \qquad (76)$$

For the example of a Coulomb field (a charge moving with constant speed, Fig. 18):

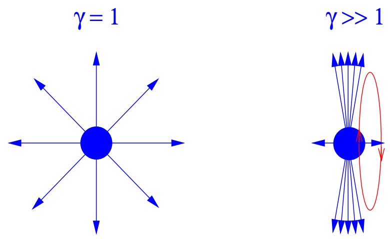

Fig. 18: Transformation of a Coulomb field to a moving frame

– in the rest frame, one has only electrostatic forces;
– in the moving frame, $\vec{E}$ is transformed and a magnetic field $\vec{B}$ appears.

When we rewrite the second component of Eq. (76):

$$B'_y = \gamma \cdot \frac{v}{c^2} \cdot E_z \,. \tag{77}$$

For small velocities v, this corresponds to the component of the Biot–Savart law. The Lorentz transformation automatically yields this law, including relativistic corrections for larger v.

8.3 An important consequence for moving charges in an accelerator

$$\begin{aligned} E'_x &= E_x \,, & B'_x &= B_x \,, \\ E'_y &= \gamma(E_y - v \cdot B_z) \,, & B'_y &= \gamma \left(B_y + \frac{v}{c^2} \cdot E_z \right) \,, \\ E'_z &= \gamma(E_z + v \cdot B_y) \,, & B'_z &= \gamma \left(B_z - \frac{v}{c^2} \cdot E_y \right) \,. \end{aligned} \tag{78}$$

Assuming that $\vec{B}' = 0$, we get for the *transverse* forces:

$$\vec{F}_{\text{mag}} = -\beta^2 \cdot \vec{F}_{\text{el}} \,. \tag{79}$$

For particles close to the speed of light, β = 1, and electric and magnetic *forces* cancel.

This has many important consequences, (e.g., space charge effects) because transverse fields generated by the charges vanish as the beam is accelerated.

This is most important for the stability of beams and (for $\beta \ll 1$) it cannot be ignored.

8.4 Retarded potentials

To compute these fields of moving charges, one can start with the four-potential of the charges at rest and apply the transformation.

For the static charge, we have the Coulomb potential [1] and $\vec{A} = 0$. The transformation into the new frame (moving in the x-direction) gives, for the new potentials:

$$\frac{\phi'}{c} = \gamma \left(\frac{\phi}{c} - A_x \right) = \gamma \frac{\phi}{c} \,, \tag{80}$$

$$A'_x = \gamma \left(A_x - \frac{v\phi}{c^2} \right) = -\gamma \frac{v}{c^2} \phi = -\frac{v}{c^2} \phi' \,. \tag{81}$$

i.e., all that is needed to compute the fields is the new scalar potential ϕ':

$$\phi'(\vec{x}) = \gamma \phi(\vec{x}) = \gamma \cdot \frac{1}{4\pi\epsilon_0} \cdot \frac{q}{|\vec{x} - \vec{x}_q|} \,. \tag{82}$$

After transformation of the co-ordinates, e.g., $x = \gamma(x' - vt')$, the resulting potentials can be used to compute the fields, as observed in the system at rest. However, one has to take care of causality.

The field observed at a position $\vec{x}$ at time t was caused at an *earlier* time $t_x < t$ at the location $\vec{x}_0(t_\mathrm{r})$ and the potentials have to be written as:

$$\phi(\vec{x}, t) = \frac{qc}{|\vec{X}|c - \vec{X}\vec{v}}\,, \qquad \vec{A}(\vec{x}, t) = \frac{q\vec{v}}{|\vec{X}|c - \vec{X}\vec{v}}\,. \tag{83}$$

The potentials $\phi(\vec{x}, t)$ and $\vec{A}(\vec{x}, t)$ depend on the state at the *retarded* time t_r, not t. Here ,$\vec{v}$ is the velocity at time t_r and $\vec{X} = \vec{x} - \vec{x}_0(t_\mathrm{r})$ relates the retarded position to the observation point.

8.5 Invariant formulation of Maxwell's equations

One can now rewrite Maxwell's equations using four-vectors and $F^{\mu\nu}$:

$$\nabla\vec{E} = \frac{\rho}{\epsilon_0} \qquad \text{and} \qquad \nabla \times \vec{B} - \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} = \mu_0 \vec{J}\,,$$

$$\stackrel{1+3}{\Rightarrow} \partial_\mu F^{\mu\nu} = \mu_0 J^\nu \qquad \text{(inhomogeneous Maxwell equation)}\,; \tag{84}$$

$$\nabla\vec{B} = 0 \qquad \text{and} \qquad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0\,,$$

$$\stackrel{1+3}{\Rightarrow} \partial_\gamma F^{\mu\nu} + \partial_\mu F^{\nu\lambda} + \partial_\nu F^{\lambda\mu} = 0 \qquad \text{(homogeneous Maxwell equation)}\,. \tag{85}$$

We have Maxwell's equations in a very compact form; transformation between moving systems is now very easy. The equivalent formulation in matter (macroscopic Maxwell's equation) is shown in Appendix B.

8.5.1 Derivation of Gauss' law

Starting from Eq. (84):

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu\,. \tag{86}$$

Written explicitly (Einstein convention, sum over μ):

$$\partial_\mu F^{\mu\nu} = \sum_{\mu=0}^{3} \partial_\mu F^{\mu\nu} = \partial_0 F^{0\nu} + \partial_1 F^{1\nu} + \partial_2 F^{2\nu} + \partial_3 F^{3\nu} = \mu_0 J^\nu\,. \tag{87}$$

As an example, one can choose ν = 0 and replace $F^{\mu\nu}$ with the corresponding elements:

$$\partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} = \mu_0 J^0\,,$$

$$0 + \partial_x \frac{E_x}{c} + \partial_y \frac{E_y}{c} + \partial_z \frac{E_z}{c} = \mu_0 J^0 = \mu_0 c\rho\,. \tag{88}$$

This corresponds exactly to:

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad (c^2 = \epsilon_0\mu_0)\,. \tag{89}$$

8.5.2 Derivation of Ampère's law

For $\nu = 1, 2, 3$, one obtains Ampère's law.

For example in the x-plane (ν = 1) and the F frame:

$$\partial_y B_z - \partial_z B_y - \partial_t \frac{E_x}{c} = \mu_0 J^x\,. \tag{90}$$

After transforming ∂^γ and $F^{\mu\nu}$ to the F' frame:

$$\partial'_y B'_z - \partial'_z B'_y - \partial'_t \frac{E'_x}{c} = \mu_0 J'^x \,. \tag{91}$$

It should be mentioned that now Maxwell's equations have the identical form in F and F'.

8.5.3 Combining the equations

Finally: since we have $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$,

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \,, \tag{92}$$

$$\partial_\gamma F^{\mu\nu} + \partial_\mu F^{\nu\lambda} + \partial_\nu F^{\lambda\mu} = 0 \,. \tag{93}$$

We can rewrite them two-in-one in a new form:

$$\frac{\partial^2 A^\mu}{\partial x_\nu \partial x^\nu} = \mu_0 J^\mu \,. \tag{94}$$

This expression contains all four Maxwell's equations, and it is the only one that stays the same in all frames!

One can conclude that there are *no* separate electric and magnetic fields; they are just a frame-dependent manifestation of a single electromagnetic field.

8.6 Electromagnetic forces in the framework of relativity

Starting with the four-force as the time derivative of the four-momentum:

$$\mathcal{F}^\mu_\mathcal{L} = \frac{\partial \mathcal{P}^\mu}{\partial \tau} \,. \tag{95}$$

We get the four-vector for the Lorentz force, where the spatial part is the well-known expression:

$$\mathcal{F}^\mu_L = \gamma q \left(\frac{\vec{E} \cdot \vec{u}}{c}, \vec{E} + \vec{u} \times \vec{B} \right) = q \cdot F^{\mu\nu} U_\nu \,. \tag{96}$$

8.7 Interpretation

To quote Einstein [2]:

> For a charge moving in an electromagnetic field, the force experienced by the charge is equal to the electric force, transformed into the rest frame of the charge.

Therefore, the Lorentz force is not an add-on to Maxwell's equations but just a consequence of two reference frames.

9 Summary

9.1 Summary—relativity basics

- Special relativity is very simple; there are a few basic principles.
  - Physics laws are the same in all inertial systems.
  - The speed of light in vacuum is the same in all inertial systems.
- Everyday phenomena lose their meaning (do not ask what is 'real').
  - Only the union of space and time preserve an independent reality: *space–time*.

– Electric and magnetic fields do not exist!
– They are simply different aspects of a *single electromagnetic field*.
– The manifestation of the electromagnetic field, i.e., division into electric $\vec{E}$ and magnetic $\vec{B}$ components, depends on the chosen reference frame.

9.2 Summary—consequences for particle accelerators

Write everything as four-vectors, blindly follow the rules and you get it all easily, in particular, transformation of fields, etc.

- Relativistic effects in accelerators (used in later lectures):
  - Lorentz contraction and time dilation (e.g., free-electron laser, . . .);
  - relativistic Doppler effect (e.g., free-electron laser, . . .);
  - invariants!
  - relativistic mass effects and dynamics;
  - new interpretation of electric and magnetic fields, in particular 'Lorentz force'.
- If you do not take relativity into account, you are sunk. . .

References

[1] W. Herr, Classical electrodynamics and applications to particle accelerators, these proceedings.
[2] A. Einstein, *Ann. Phys.* **322** (1905) 891. http://dx.doi.org/10.1002/andp.19053221004
[3] J. Hafele, *Science* **177**(4044) (1972) 166. http://dx.doi.org/10.1126/science.177.4044.166
[4] W. Herr, in Proc. of the CAS-CERN Accelerator School: Intermediate Course on Accelerator Physics, Zeuthen, Germany, 16–26 September 2003, edited by D. Brandt, CERN-2006-002 (CERN, Geneva, 2006), pp. 361-378. http://dx.doi.org/10.5170/CERN-2006-002.361

Bibliography

A. Chao and M. Tigner, *Handbook of Accelerator Physics and Engineering* (World Scientific Publishing, New Jersey, 1998).

R.P. Feynman, *Feynman Lectures on Physics* (Basic Books, New York, 2001), Vols. 1 + 2.

J. Freund, *Special Relativity for Beginners: A Textbook for Undergraduates* (World Scientific, New Jersey, 2008), http://dx.doi.org/10.1142/6601

J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1998).

L. Landau and E. Lifschitz, *The Classical Theory of Fields* (Butterworth-Heinemann, Oxford, 1975), Vol. 2.

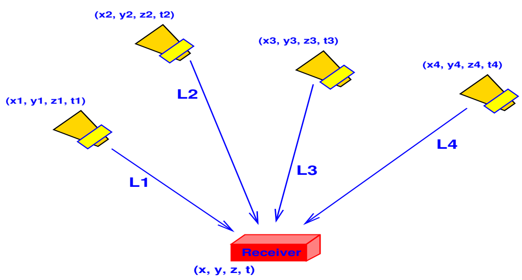

Fig. A.1: Schematic view of the GPS

Appendices

A Time dilation and general relativity

The time dilation due to the difference of height between two systems can be calculated as:

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \sqrt{1 - \frac{2GM}{Rc^2}}\,, \tag{A.1}$$

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} \approx 1 - \frac{GM}{Rc^2}\,. \tag{A.2}$$

where G is the gravitational constant and R the distance from the centre of the Earth, and

$$\Delta\tau = \frac{GM}{c^2} \cdot \left(\frac{1}{R_{\mathrm{Earth}}} - \frac{1}{R_{\mathrm{GPS}}} \right)\,. \tag{A.3}$$

With the following parameters for the GPS (Fig. A.1):

$$R_{\mathrm{Earth}} = 6357000\ \mathrm{m}\,, \qquad R_{\mathrm{GPS}} = 26541000\ \mathrm{m}\,,$$

$$G = 6.674 \cdot 10^{-11}\ \mathrm{N\,m^2/\,kg^2}\,, \qquad M = 5.974 \cdot 10^{24}\ \mathrm{kg}\,,$$

we have:

$$\Delta\tau \approx 5.3 \cdot 10^{-10}\ \mathrm{s}\,. \tag{A.4}$$

B Tensors and macroscopic Maxwell's equations

It was mentioned in a previous lecture [1] that electric displacement $\vec{D}$ and magnetic field $\vec{H}$ are linked to the electric field $\vec{E}$ and induction $\vec{B}$ as:

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P}\,,$$

$$\vec{H} = \frac{1}{\mu_0}\vec{B}\vec{M}\,, \tag{B.1}$$

where $\vec{P}$ and $\vec{M}$ are the polarization and magnetization, respectively.

Each can be represented as a covariant tensor:

$$M^{\mu\nu} := \begin{pmatrix} 0 & P_x c & P_y c & P_z c - P_x c & 0 & -M_z & M_y \\ -P_y c & M_z & 0 & -M_x \\ -P_z c & -M_y & M_x & 0 \end{pmatrix}\,, \tag{B.2}$$

$$D^{\mu\nu} := \begin{pmatrix} 0 & -D_x c & -D_y c & -D_z c \\ D_x c & 0 & -H_z & H_y \\ D_y c & H_z & 0 & -H_x \\ D_z c & -H_y & H_x & 0 \end{pmatrix}\,. \tag{B.3}$$

The tensors are linked as:

$$D_{\mu\nu} = \frac{1}{\mu_0} F_{\mu\nu} - M^{\mu\nu} \,. \tag{B.4}$$

It can easily be verified that Eq. (B.4) is equivalent to Eq. (B.1).

The Gauss–Ampère law (Eq. (84)) becomes:

$$\partial_\mu D_{\mu\nu} = J^\nu \,. \tag{B.5}$$