2

Special Relativity

The famous Michelson-Morley experiment discovered that the speed of light has the same value in all reference frames[1]. Albert Einstein recognized the far reaching consequences of this observation and around this curious fact of nature he built the theory of special relativity. Starting from the constant speed of light, Einstein was able to predict many interesting and strange consequences that all proved to be true. We will see how powerful this idea is, but first let's clarify what special relativity is all about. The two basic postulates are

[1] The speed of objects we observe in everyday life depend on the frame of reference. For example, when an observer standing at a train station measures that a train moves with $50\,\frac{\text{km}}{\text{h}}$, another observer running with $15\,\frac{\text{km}}{\text{h}}$ next to the same train, measures that the train moves with $35\,\frac{\text{km}}{\text{h}}$. In contrast, light always moves with $1,08 \cdot 10^9\,\frac{\text{km}}{\text{h}}$, no matter how you move relative to it.

- **The principle of relativity:** Physics is the same in all inertial frames of reference, i.e. frames moving with constant velocity relative to each other.
- **The invariance of the speed of light:** The velocity of light has the same value c in all inertial frames of reference.

In addition, we will assume that the stage our physical laws act on is homogeneous and isotropic. This means it does not matter where (=homogeneity) we perform an experiment and how it is oriented (=isotropy), the laws of physics stay the same. For example, if two physicists, one in New-York and the other one in Tokyo, perform exactly the same experiment, they would find the same[2] physical laws. Equally a physicist on planet Mars would find the same physical laws.

[2] Besides from changing constants, as, for example, the gravitational acceleration

The laws of physics, formulated correctly, shouldn't change if you look at the experiment from a different perspective or repeat it tomorrow. In addition, the first postulate tells us that a physical experiment should come up with the same result regardless of if you perform it on a wagon moving with constant speed or at rest in a laboratory. These things coincide with everyday experience. For example, if you close your eyes in a car moving with constant speed, there is no way to tell if you are really moving or if you're at rest.

J. Schwichtenberg, *Physics from Symmetry*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-319-66631-0_2

Without homogeneity and isotropy physics would be in deep trouble: If the laws of nature we deduce from experiment would hold only at one point in space, for a specific orientation of the experiment such laws would be rather useless.

The only unintuitive thing is the second postulate, which is contrary to all everyday experience. Nevertheless, all experiments until the present day show that it is correct.

2.1 The Invariant of Special Relativity

Before we dive into the details, here's a short summary of what we want to do in the following sections. We use the postulates of special relativity to derive the Minkowski metric, which tells us how to compute the "distance" between two physical events. Another name for physical events in this context is points in Minkowski space, which is how the stage the laws of special relativity act on is called. It then follows that all transformations connecting different inertial frames of reference must leave the Minkowski metric unchanged. This is how we are able to find all transformations that connect allowed frames of reference, i.e. frames with a constant speed of light. In the rest of the book we will use the knowledge of these transformations, to find equations that are unchanged by these transformations. Let's start with a thought experiment that enables us to derive one of the most fundamental consequences of the postulates of special relativity.

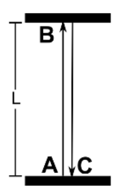

Fig. 2.1: Illustration of the thought experiment

Imagine, we have a spectator, standing at the origin of his coordinate system and sending a light pulse straight up, where it is reflected by a mirror and finally reaches again the point from where it was sent. An illustration of this can be seen in Fig. 2.1

We have three important events:

- **A** : the light leaves the starting point
- **B** : the light is reflected at a mirror
- **C** : the light returns to the starting point.

The time-interval between **A** and **C** is[3]

$$\Delta t = t_C - t_A = \frac{2L}{c}, \tag{2.1}$$

where L denotes the distance between the starting point and the mirror.

[3] For constant speed v we have $v = \frac{\Delta s}{\Delta t}$, with the distance covered Δs and the time needed Δt, and therefore $\Delta t = \frac{\Delta s}{v}$

Next imagine a second spectator, standing at t_A at the origin of his coordinate system and moving with constant velocity u to the left, relative to the first spectator[4]. For simplicity let's assume that the origin of this second spectators' coordinate system coincides at t_A with the coordinate origin of the first spectator. The second spectator sees things a little differently. In his frame of reference the point where the light ends up will not have the same coordinates as the starting point (see Fig. 2.2).

[4] Transformations that allow us to transform the description of one observer into the description of a second observer, moving with constant speed relative to first observer, are called **boosts**. We derive later a formal description of such transformations.

We can express this mathematically

$$x'_A = 0 \neq x'_C = u\Delta t' \qquad \rightarrow \qquad \Delta x' = u\Delta t', \tag{2.2}$$

where the primed coordinates denote the moving spectator. For the first spectator in the rest-frame we have of course

$$x_A = x_C \qquad \rightarrow \qquad \Delta x = 0. \tag{2.3}$$

We assume movement along the x-axis, therefore

$$y'_A = y'_C \quad \text{and} \quad z'_A = z'_C \quad \rightarrow \quad \Delta y' = 0 \quad \text{and} \quad \Delta z' = 0 \tag{2.4}$$

and equally of course

$$y_A = y_C \quad \text{and} \quad z_A = z_C \quad \rightarrow \quad \Delta y = 0 \quad \text{and} \quad \Delta z = 0. \tag{2.5}$$

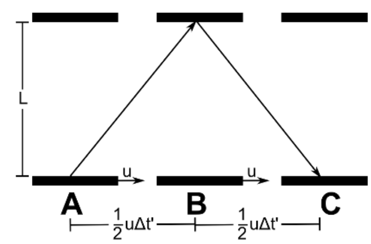

Fig. 2.2: Illustration of the thought experiment for a moving spectator. The second spectator moves to the left and therefore the first spectator (and the experiment) moves relative to him to the right.

The next question is: **What about the time interval the second spectator measures?** Because we postulate a constant velocity of light, the second spectator measures a different time interval between **A** and **C**! The time interval $\Delta t' = t'_C - t'_A$ is equal to the distance l the light travels, as the second spectator observes it, divided by the speed of light c.

$$\Delta t' = \frac{l}{c} \tag{2.6}$$

We can compute the distance traveled l using the good old Pythagorean theorem (see Fig. 2.2)

$$l = 2\sqrt{\left(\frac{1}{2}u\Delta t'\right)^2 + L^2}. \tag{2.7}$$

Using Eq. 2.6, we therefore conclude

$$c\Delta t' = 2\sqrt{\left(\frac{1}{2}u\Delta t'\right)^2 + L^2}. \tag{2.8}$$

If we now use $\Delta x' = u\Delta t'$ from Eq. 2.2, we can write

$$c\Delta t' = 2\sqrt{\left(\frac{1}{2}\Delta x'\right)^2 + L^2}$$

$$\rightarrow (c\Delta t')^2 = 4\left(\left(\frac{1}{2}\Delta x'\right)^2 + L^2\right)$$

$$\rightarrow (c\Delta t')^2 - (\Delta x')^2 = 4\left(\left(\frac{1}{2}\Delta x'\right)^2 + L^2\right) - (\Delta x')^2 = 4L^2 \qquad (2.9)$$

and now recalling from Eq. 2.1 that $\Delta t = \frac{2L}{c}$, we can write

$$(c\Delta t')^2 - (\Delta x')^2 = 4L^2 = (c\Delta t)^2 = (\Delta t c)^2 - \underbrace{(\Delta x)^2}_{=0 \text{ see Eq. 2.3}}. \qquad (2.10)$$

So finally, we arrive[5] at

$$(c\Delta t')^2 - (\Delta x')^2 - \underbrace{(\Delta y')^2}_{=0} - \underbrace{(\Delta z')^2}_{=0} = (c\Delta t)^2 - \underbrace{(\Delta x)^2}_{=0} - \underbrace{(\Delta y)^2}_{=0} - \underbrace{(\Delta z)^2}_{=0}. \qquad (2.11)$$

[5] Take note that what we are doing here is just the shortest path to the result, because we chose the origins of the two coordinate systems to coincide at t_A. Nevertheless, the same can be done, with more effort, for arbitrary choices, because physics is the same in all inertial frames. We used this freedom to choose two inertial frames where the computation is easy. In an arbitrarily moving second inertial system we do not have $\Delta y' = 0$ and $\Delta z' = 0$. Nevertheless, the equation holds, because physics is the same in all inertial frames.

Considering a third observer, moving with a different velocity relative to the first observer, we can use the same reasoning to arrive at

$$(c\Delta t'')^2 - (\Delta x'')^2 - (\Delta y'')^2 - (\Delta z'')^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \qquad (2.12)$$

Therefore, we have found something which is the same for all observers: the quadratic form

$$(\Delta s)^2 \equiv (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2. \qquad (2.13)$$

In addition, we learned in this section that $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$ or $(c\Delta t)^2$ aren't the same for different observers. We will talk about the implications of this curious property in the next section.

2.2 Proper Time

We derived in the last section the invariant of special relativity Δs^2, i.e. a quantity that has the same value for all observers. Now, we want to think about the physical meaning of this quantity.

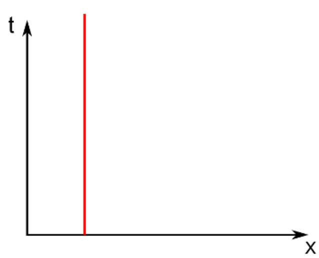

Fig. 2.3: World line of an object at rest. The position of the object stays the same as time goes on.

For brevity, let's restrict ourselves to one spatial dimension. An object at rest, relative to some observer, has a **spacetime diagram** as drawn in Fig. 2.3. In contrast, an object moving with constant velocity, relative to the same observer, has a spacetime diagram as drawn in Fig. 2.4.

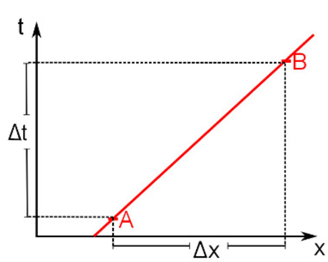

Fig. 2.4: World line of a moving object with two events A and B. The distance travelled between A and B is Δx and the time that passed the events is Δt.

The lines we draw to specify the position of objects in spacetime are called **world lines**. World lines are always observer dependent. Two different observers may draw completely different world lines for the same object. The moving object with world line drawn in

Fig. 2.4, looks for a second observer who moves with the same constant speed as the object, as drawn in Fig. 2.5. For this second observer the object is at rest. Take note, to account for the two different descriptions we introduce primed coordinates for the second observer: x' and t'.

We can see that both observers do not agree on the distance the object travels between some events A and B in spacetime. For the first observer we have $\Delta x \neq 0$, but for the second observer $\Delta x' = 0$. For both observers the time interval between A and B is non-zero: $\Delta t \neq 0$ and $\Delta t' \neq 0$. Both observers agree on the value of the quantity $(\Delta s)^2$, because as we derived in the last section, this invariant of special relativity has the same value for all observers. A surprising consequence is that both observers do not agree on the time elapsed between the events A and B

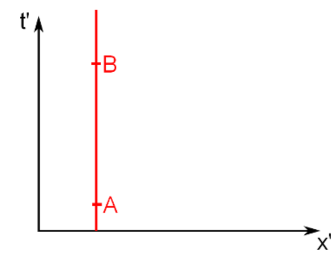

Fig. 2.5: World line of the same moving object, as observed from someone moving with the same constant speed as the object. The distance travelled between A and B is for this observer $\Delta x' = 0$.

$$(\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 \tag{2.14}$$

$$(\Delta s')^2 = (c\Delta t')^2 - \underbrace{(\Delta x')^2}_{=0} = (c\Delta t')^2 \tag{2.15}$$

$$(\Delta s)^2 = (\Delta s')^2 \rightarrow (\Delta t')^2 \neq (\Delta t)^2 \quad \text{because } (\Delta x)^2 \neq 0\,. \tag{2.16}$$

This is one of the most famous phenomena of special relativity and commonly called **time-dilation**. Time-intervals and spatial distances are observer dependent. The clocks tick differently for different observers and therefore they observe a different number of ticks between two events.

Now that the concept of time has become relative, a new notion of time that all observers agree on may be useful. In the example above we can see that for the second observer, moving with the same speed as the object, we have

$$(\Delta s)^2 = (c\Delta t')^2\,. \tag{2.17}$$

This means the invariant of special relativity is equivalent, up to a constant c, to the time interval measured by this observer. With this in mind, we can interpret $(\Delta s)^2$ and define a notion of time that all observers agree on. We define

$$(\Delta s)^2 = (c\Delta\tau)^2, \tag{2.18}$$

where τ is called the **proper time**. The proper time is the time measured by an observer in the special frame of reference where the object in question is at rest.

Of course objects in the real world aren't restricted to motion with constant speed, but if the time interval is short enough, in the extremal case infinitesimal, any motion is linear and the notion of proper time is sensible. In mathematical terms this requires we make the transition to infinitesimal intervals $\Delta \to d$:

$$(ds)^2 = (cd\tau)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2. \tag{2.19}$$

Therefore, even if an object moves wildly, we can still imagine some observer with a clock travelling with the object and therefore observing the object at rest. The time interval this special observer measures is the proper time and all observers agree on its value, because $(ds)^2 = (cd\tau)^2$ has the same value for all observers. Again, this does not mean that all observers measure the same time interval! They just agree on the value of the time interval measured by someone who travels with the object in question.

2.3 Upper Speed Limit

Now that we have an interpretation for the invariant of special relativity, we can go a step further and explore one of the most stunning consequences of the postulates of special relativity.

It follows from the minus sign in the definition of Δs^2that it can be zero for two events that are separated in space and time. It even can be negative, but then we would get a complex value for the proper time[6], which is commonly discarded as unphysical. We conclude, we have a minimal proper time $\tau = 0$ for two events if $\Delta s^2 = 0$. Then we can write

[6] Recall $(ds)^2 = (cd\tau)^2$ and therefore if $(ds)^2 < 0 \to d\tau$ is complex.

$$\Delta s^2_{\text{min}} = 0 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

$$\to (c\Delta t)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

$$\to c^2 = \frac{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}{(\Delta t)^2}. \tag{2.20}$$

On the right-hand side we have a squared velocity v^2, i.e. distance divided by time. We can rewrite this in the infinitesimal limit

$$\to c^2 = \frac{(dx)^2 + (dy)^2 + (dz)^2}{(dt)^2}. \tag{2.21}$$

The functions $x(t), y(t), z(t)$ describe the path between the two events. Therefore, we have on the right-hand side the velocity between the events.

We conclude the lowest value for the proper time is measured by someone travelling with speed

$$\rightarrow c^2 = v^2. \tag{2.22}$$

This means nothing can move with a velocity larger than c! **We have an upper speed limit for everything in physics**. Two events in spacetime can't be connected by anything faster than c.

From this observation follows the **principle of locality**, which means that everything in physics can only be influenced by its immediate surroundings. Every interaction must be local and there can be no action at a distance, because everything in physics needs time to travel from some point to another.

2.4 The Minkowski Notation

> Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.
>
> **- Hermann Minkowski**[7]

[7] In a speech at the 80th Assembly of German Natural Scientists and Physicians (21 September 1908)

We can rewrite the invariant of special relativity

$$ds^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2 \tag{2.23}$$

by using a new notation, which looks quite complicated at first sight, but will prove to be invaluable:

$$ds^2 = \eta^{\mu\nu} dx_\mu dx_\nu = \eta^{00}(dx_0)^2 + \eta^{11}(dx_1)^2 + \eta^{22}(dx_2)^2 + \eta^{33}(dx_3)^2$$

$$= (dx_0)^2 - (dx_1)^2 - (dx_2)^2 - (dx_3)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2. \tag{2.24}$$

Here we use several new notations and conventions one needs to become familiar with, because they are used everywhere in modern physics:

- Einsteins summation convention: If an index occurs twice, a sum is implicitly assumed : $\sum_{i=1}^3 a_i b_i = a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3$, but $\sum_{i=1}^3 a_i b_j = a_1 b_j + a_2 b_j + a_3 b_j \neq a_i b_j$
- Greek indices[8], like μ, ν or σ, are always summed from 0 to 3: $x_\mu y_\mu = \sum_{\mu=0}^3 x_\mu y_\mu$.
- Renaming of the variables $x_0 \equiv ct$, $x_1 \equiv x$, $x_2 \equiv y$ and $x_3 \equiv z$, to make it obvious that time and space are now treated equally and to be able to use the rules introduced above

[8] In contrast, Roman indices like i, j, k are always summed: $x_i x_i \equiv \sum_i^3 x_i x_i$ from 1 to 3. Much later in the book we will use capital Roman letters like A, B, C that are summed from 1 to 8.

- Introduction of the Minkowski metric $\eta^{00} = 1$, $\eta^{11} = -1$, $\eta^{22} = -1$, $\eta^{33} = -1$ and $\eta^{\mu\nu} = 0$ for $\mu \neq \nu$ (an equal way of writing this is[9] $\eta = \text{diag}(1, -1, -1, -1)$)

[9] $\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$

In addition, it's conventional to introduce the notion of a **four-vector**

$$dx_\mu = \begin{pmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{pmatrix}, \tag{2.25}$$

because the equation above can be written equally using four-vectors and the Minkowski metric in matrix form

$$(ds)^2 = dx_\mu \eta^{\mu\nu} dx_\nu = \begin{pmatrix} dx_0 & dx_1 & dx_2 & dx_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{pmatrix}$$
$$= dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2 \tag{2.26}$$

This is really just a clever way of writing things. A physical interpretation of ds is that it is the "distance" between two events in spacetime. Take note that we don't mean here only the spatial distance, but also have to consider a separation in time. If we consider 3-dimensional Euclidean[10] space the squared (shortest) distance between two points is given by[11]

[10] 3-dimensional Euclidean space is just the space of classical physics, where time was treated differently from space and therefore it was not included into the geometric considerations. The notion of spacetime, with time as a fourth coordinate was introduced with special relativity, which enables mixing of time and space coordinates as we will see.

[11] The Kronecker delta δ_{ij}, which is the identity matrix in index notation, is defined in Appendix B.5.5.

$$(ds)^2 = dx_i \delta_{ij} dx_j = \begin{pmatrix} dx_1 & dx_2 & dx_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \end{pmatrix}$$
$$= (ds)^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 \tag{2.27}$$

The mathematical tool that tells us the distance between two infinitesimal separated points is called **metric**. In boring Euclidean space the metric is just the identity matrix δ. In the curved spacetime of general relativity much more complicated metrics can occur. The geometry of the spacetime of special relativity is encoded in the relatively simple Minkowski metric η. Because the metric is the tool to compute length, we need it to define the **length of a four-vector**, which is given by the scalar product of the vector with itself[12]

[12] The same is true in Euclidean space: $\text{length}^2(v) = \vec{v} \cdot \vec{v} = v_1^2 + v_2^2 + v_3^2$, because the metric is here simply $\delta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

$$x^2 = x \cdot x \equiv x_\mu x_\nu \eta^{\mu\nu} .$$

Analogously, the **scalar product** of two arbitrary four-vectors is defined by

$$x \cdot y \equiv x_\mu y_\nu \eta^{\mu\nu}. \tag{2.28}$$

There is another, notational convention to make computations more streamlined. We define a four-vector with upper index as[13]

[13] Four-vectors with a lower index are often called covariant and four-vectors with an upper index contravariant.

$$x^\mu \equiv \eta^{\mu\nu} x_\nu \tag{2.29}$$

or equally

$$y^\nu \equiv \eta^{\mu\nu} y_\mu \underbrace{=}_{\text{The Minkowski metric is symmetric } \eta^{\mu\nu}=\eta^{\nu\mu}} \eta^{\nu\mu} y_\mu \,. \tag{2.30}$$

Therefore, we can write the scalar product of two four-vectors as[14]

[14] The name of the index makes no difference. For more information about this have a look at Appendix B.5.1.

$$x \cdot y \equiv x_\mu y_\nu \eta^{\mu\nu} = x_\mu y^\mu = x^\nu y_\nu. \tag{2.31}$$

It doesn't matter which index we transform to an upper index. This is just a way of avoiding writing the Minkowski metric all the time, just as Einstein's summation convention is introduced to avoid writing the summation sign.

2.5 Lorentz Transformations

Next, we try to figure out in what ways we can transform our description in a given frame of reference without violating the postulates of special-relativity. We learned above that it follows directly from the two postulates that $ds^2 = \eta^{\mu\nu} dx_\mu dx_\nu$ is the same in all inertial frames of reference:

$$ds'^2 = dx'_\mu dx'_\nu \eta^{\mu\nu} = ds^2 = dx_\mu dx_\nu \eta^{\mu\nu}. \tag{2.32}$$

Therefore, allowed transformations are those which leave this quadratic form or equally the scalar product of Minkowski spacetime invariant. Denoting a generic transformation that transforms the description in one frame of reference into the description in another frame with Λ, the transformed coordinates dx'_μ can be written as:

$$dx_\mu \to dx'_\mu = \Lambda_\mu^{\ \sigma} dx_\sigma. \tag{2.33}$$

Then we can write the invariance condition as

$$(ds)^2 = (ds')^2$$
$$\rightarrow dx \cdot dx \stackrel{!}{=} dx' \cdot dx'$$
$$\rightarrow dx_\mu dx_\nu \eta^{\mu\nu} \stackrel{!}{=} dx'_\mu dx'_\nu \eta^{\mu\nu} \underbrace{=}_{\text{Eq. 2.33}} \Lambda_\mu{}^\sigma dx_\sigma \Lambda_\nu{}^\gamma dx_\gamma \eta^{\mu\nu}$$
$$\underbrace{\rightarrow}_{\text{Renaming dummy indices}} dx_\mu dx_\nu \eta^{\mu\nu} \stackrel{!}{=} \Lambda_\sigma{}^\mu dx_\mu \Lambda_\gamma{}^\nu dx_\nu \eta^{\sigma\gamma}$$
$$\underbrace{\rightarrow}_{\text{Because the equation holds for arbitrary } dx_\mu} \eta^{\mu\nu} \stackrel{!}{=} \Lambda_\sigma{}^\mu \eta^{\sigma\gamma} \Lambda_\gamma{}^\nu . \tag{2.34}$$

Or written in matrix notation[15]

[15] If you wonder about the transpose here have a look at Appendix C.1.

$$\eta = \Lambda^T \eta \Lambda \tag{2.35}$$

This is the condition that transformations Λ between allowed frames of reference must fulfil.

If this seems strange at this point don't worry, because we will see that such a condition is a quite natural thing. In the next chapter we will learn that, for example, rotations in ordinary Euclidean space are defined as those transformations O that leave the scalar product of Euclidean space invariant[16]

[16] The $\cdot$ is used for the scalar product of vectors, which corresponds to $\vec{a} \cdot \vec{b} = \vec{a}^T \vec{b}$ for ordinary matrix multiplication, where a vector is an 1×3 matrix. The fact $(Oa)^T = a^T O^T$ is explained in Appendix C.1, specifically Eq. C.3.

$$\vec{a} \cdot \vec{b} \stackrel{!}{=} \vec{a}' \cdot \vec{b}' \underbrace{=}_{\text{Take note that } (Oa)^T = a^T O^T} \vec{a}^T O^T O \vec{b}. \tag{2.36}$$

Therefore[17] $O^T 1 O \stackrel{!}{=} 1$ and we can see that the metric of Euclidean space, which is just the unit matrix 1, plays the same role as the Minkowski metric in Eq. 2.35. This is one part of the definition for rotations, because the defining feature of rotations is that they leave the length of a vector unchanged, which corresponds mathematically to the invariance of the scalar product[18]. Additionally we must include that rotations do not change the orientation[19] of our coordinate system, which means mathematically $\det O \stackrel{!}{=} 1$, because there are other transformations which leave the length of any vector invariant: spatial inversions[20]

[17] This condition is often called **orthogonality**, hence the symbol O. A matrix satisfying $O^T O = 1$ is called orthogonal, because its columns are orthogonal to each other. In other words: Each column of a matrix can be thought of as a vector and the orthogonality condition for matrices means that each such vector is orthogonal to all other column vectors.

[18] Recall that the length of a vector is given by the scalar product of a vector with itself.

[19] This is explained in Appendix A.5.

[20] A spatial inversion is simply a map $\vec{x} \rightarrow -\vec{x}$. Mathematically such transformations are characterized by the conditions $\det O \stackrel{!}{=} -1$ and $O^T O = 1$. Therefore, if we only want to talk about rotations we have the extra condition $\det O \stackrel{!}{=} 1$. Another name for spatial inversions are parity transformations.

We define the **Lorentz transformations** as those transformations that leave the scalar product of Minkowski spacetime invariant. In physical terms this means that Lorentz transformations describe changes between frames of references that respect the postulates of special relativity. In turn this does mean, of course, that everytime we want to get a term that does not change under Lorentz transformations, we must combine an upper with a lower index:

$x_\mu y^\mu = x_\mu y_\nu \eta^{\mu\nu}$. We will construct explicit matrices for the allowed transformations in the next chapter, after we have learned some very elegant techniques for dealing with conditions like this.

2.6 Invariance, Symmetry and Covariance

Before we move on, we have to talk about some very important notions. Firstly, we call something **invariant**, if it does not change under transformations. For instance, let's consider something arbitrary like $F = F(A, B, C, ...)$ that depends on different quantities $A, B, C, ...$ If we transform $A, B, C, ... \to A', B', C', ...$ and we have

$$F(A', B', C', ...) = F(A, B, C, ...) \tag{2.37}$$

F is called invariant under this transformation. We can express this differently using the word symmetry. **Symmetry** is defined as invariance under a transformation or class of transformations. For example, some physical system is symmetric under rotations if we can rotate it arbitrarily and it always stays **exactly** the same. Another example would be a room with constant temperature. The quantity temperature does not depend on the position of measurement. In other words, the quantity temperature is invariant under translations. A translation means that we move every point a given distance in a specified direction. Therefore, we have translational symmetry within this room.

Covariance means something similar, but may not be confused with invariance. An equation is called covariant, if it takes the same **form** when the objects in it are transformed. For instance, if we have an equation

$$E_1 = aA^2 + bBA + cC^4$$

and after the transformation this equation reads

$$E_1' = aA'^2 + bB'A' + cC'^4$$

the equation is called covariant, because the form stayed the same. Another equation

$$E_2 = x^2 + 4axy + z$$

that after a transformation looks like

$$E_2' = y'^3 + 4az'y' + y'^2 + 8z'x'$$

is not covariant, because it changed its form completely.

All physical laws must be covariant under Lorentz transformations, because only such laws are valid in all reference frames. Formulating the laws of physics in a non-covariant way would be a very bad idea, because such laws would only hold in **one** frame of reference. The laws of physics would look differently in Tokyo and New York. There is no preferred frame of reference and we therefore want our laws to hold in all reference frames. We will learn later how we can formulate the laws of physics in a covariant manner.

Further Reading Tips

- **E. Taylor and J. Wheeler - Spacetime Physics: Introduction to Special Relativity**[21] is a very good book to start with.
- **D. Fleisch - A Student's Guide to Vectors and Tensors**[22] has very creative explanations for the tensor formalism used in special relativity, for example, for the differences between covariant and contravariant components.
- **N. Jeevanjee - An Introduction to Tensors and Group Theory for Physicists**[23] is another good source for the mathematics needed in special relativity.
- **A. Zee - Einstein Gravity in a nutshell**[24] is a book about general relativity, but has many great explanations regarding special relativity, too.

[21] Edwin F. Taylor and John Archibald Wheeler. *Spacetime Physics*. W. H. Freeman, 2nd edition, 3 1992. ISBN 9780716723271

[22] Daniel Fleisch. *A Student's Guide to Vectors and Tensors*. Cambridge University Press, 1st edition, 11 2011. ISBN 9780521171908

[23] Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

[24] Anthony Zee. *Einstein Gravity in a Nutshell*. Princeton University Press, 1st edition, 5 2013. ISBN 9780691145587