Chapter 1

Special Theory of Relativity

1.1 The Lorentz Transformation

1.1.1 Deduction from Axioms

Be there two *inertial frames* (see section 0.2) K with cartesian coordinates $\{x, y, z\}$ and K' with cartesian coordinates $\{x', y', z'\}$. We wish to establish a transformation

$$f : \{x, y, z, t\} \longrightarrow \{x', y', z', t'\} \tag{1.1}$$

of these coordinates including the respective time coordinate, t and t', according to

$$\begin{aligned} x' &= f_x(x, y, z, t) \\ y' &= f_y(x, y, z, t) \\ z' &= f_z(x, y, z, t) \\ t' &= f_t(x, y, z, t) \end{aligned} \tag{1.2}$$

and the inverse transformation. For simplification, we suppose that at $t = t' = 0$: $O = O'$, *i.e.*, the spatial origins of K and K' coincide at $t = t' = 0$. The two frames will be allowed to move at constant velocity v relative to each other. This situation is depicted in Fig. (1.1).

We will first establish the linearity of the transformation f.

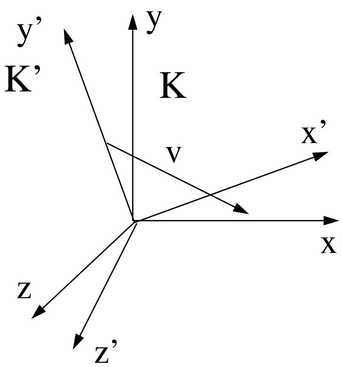


Figure 1.1:

The inertial frames at $t = t' = 0$. The origins coincide but the axes may be rotated with respect to each other. The relative velocity takes a general direction.

1.1.1.1 Homogeneity and Isotropy of Space and Time

Postulate 1: Without external influence, no point in space or time is distinguished from any other (**homogeneity of space and time**).

Postulate 2: Without external influence, no spatial direction is distinguished from any other (**isotropy of space**).

Lemma: The transformation f is linear, so $f : x \longrightarrow \eta x' + \text{const.}$

Idea of proof: Suppose the simplest form of non-linear transformation would hold[1]: $f : x \longrightarrow \zeta x'^2 + \eta x' + \text{const.}$. Then it would follow that $\zeta x'^2 + \eta x' + \text{const.} = \zeta(x' + x'_0)^2 + \text{const.'}$ by quadratic extension[2] . However, this means that $x' = -x'_0$ would be an extremum and, therefore, distinguished from other points x' (contradiction with postulate 1). The proof can be carried on to higher orders in a similar way[3].

Due to postulate 2 the inertial frames can be rotated such that $v = v_x$. So we now depart from Fig. (1.2).

Obviously, $z = 0 \Rightarrow z' = 0$ and $y = 0 \Rightarrow y' = 0$, so we can write a

[1] $\zeta, \eta, \theta \in \mathbb{R}$ are real scalar constants.

[2] In this case the equivalence means that $2\zeta x'_0 = \eta$ and $\zeta x'^2_0 + \text{const.'} = \text{const.}$.

[3] All even-order polynomials have at least one local extremal value. If an odd-order polynomial does not have a local extremal value (like $f(x) = x^3$) then it has at least one point of inflection which again represents a distinguished point.

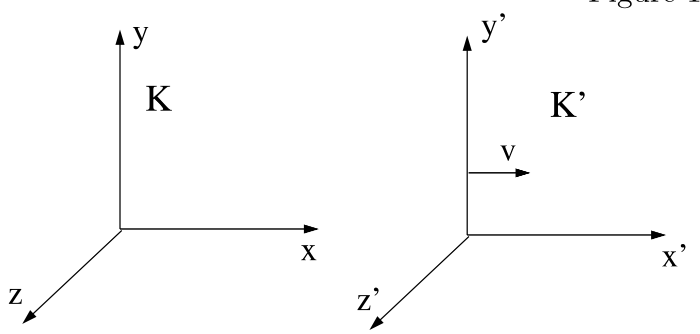


Figure 1.2: Inertial frames K and K' with axes aligned. The relative velocity can be chosen along a single axis.

transformation

$$\begin{aligned} y' &= a(v)\, y \\ z' &= \overline{a}(v)\, z \end{aligned} \tag{1.3}$$

with no constant added. Invoking postulate 2 again, we find $a(v) = \overline{a}(v)$ since the spatial coordinates y and z are on a par with respect to translation in x direction. Therefore,

$$\begin{aligned} y' &= a(v)\, y \\ z' &= a(v)\, z. \end{aligned} \tag{1.4}$$

The function $a(v)$ will be determined subsequently.

For the relation between x' and x we start out from the Galilei transformation and introduce a generalization in form of a function of velocity that is $b(v) = 1$ for the Galilei transformation

$$x' = b(v)\,(x - vt) \tag{1.5}$$

Note also that a and b are in general functions of the velocity v but not of time t since the relative motion between K and K' is constant in time.

The inverse transformation can immediately be written by analogy,

$$x = b'(v')\,(x' + v't') \tag{1.6}$$

where the + sign reflects the increase of x with time, say for the origin of K'. Without loss of generality we here understand v' as the velocity of K relative to K'.

1.1.1.2 Einstein's Principle of Relativity

Homogeneity and isotropy of space and time do not lead any further than what has been presented in section 1.1.1.1. The next task is to establish a general transformation for the time coordinate. This general transformation will become evident in Eq. (1.16).

The first of Einstein's postulates of relativity reads

Postulate 3: No physical measurement can distinguish between reference frames K and K' (**Einstein's first postulate of relativity**).
In other words: "The laws of physics are the same in all inertial frames."

We will now determine general expressions for the functions a and b. We here define the **event** which is understood as a generalization of a point in real space and an instant in time. An event has three spatial and one time coordinate.

- Consider a particle at rest at time t and position z_0 in frame K. This event in K has spatial and time coordinates as

$$\text{E:} \qquad (0, 0, z_0, t) \tag{1.7}$$

  The coordinates of that same event can immediately be written for K':

$$(0, 0, z_0', t')' = (0, 0, a(v)z_0, t')' \tag{1.8}$$

  where Eq. (1.4) has been used. Comparing Eqns. (1.7) and (1.8) it follows that $a(v) = 1$, or else an objective difference between frames K and K' would exist[4], which would contradict postulate 3.

[4]Suppose that $a(v) \neq 1$, then $z_0' \neq z_0$ and the two frames would not be equivalent. Since there is no relative movement in z direction, frame K would have to be stretched (or compressed) relative to K' and the laws of physics would not be the same in both frames.

- Postulate 3 also dictates that $v = v'$, else we would again distinguish between the two frames[5]. So we have established

$$b'(v') = b'(v). \tag{1.9}$$

In order to determine the relationship between b and b' we will now carry out an elementary length measurement of a rigid object in coordinates of both frames. However, before doing so, we need to consider a further difficulty: It will be required to relate the time variables at different positions of a reference frame to each other. In other words, we have to settle the problem of **synchronizing perfect clocks**[6]. This can be achieved through the arrangement in Fig. 1.3.

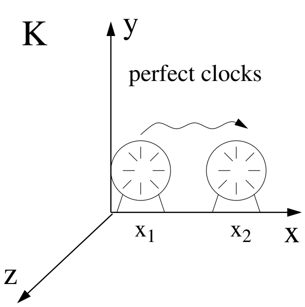


Figure 1.3:

Clocks can be synchronized by installing them in their final position and then using a signal and its duration of travel for synchronizing times on clocks 1 and 2.

It is not possible to first synchronize the clocks and then transport them to the place of measurement. The reason is that we have no guarantee that they will remain synchronized during transport (in fact, they generally don't!). So the clocks are installed in their final destinations in frame K (or K', for that matter), and **then** synchronized. This latter step can be carried out in the following way:

Be an ideal clock (1) at position x_1. A light pulse is emitted at $t(1) = 0$, so time on clock (2) is set at the time of arrival of the

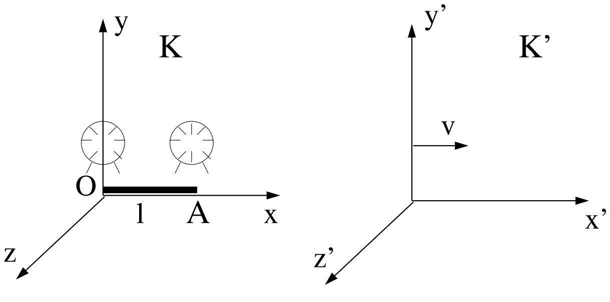
[5] v and v' are taken as positive numbers. So they designate the relative velocity between the two frames.

[6] We are not really doing this in practice. This is a so-called "Gedankenexperiment" (yes, they use a German word in English for this, it means in French "expérience à pensées").

pulse at x_2 to $t(2) := t(1) + \frac{x_2 - x_1}{c} = \frac{x_2 - x_1}{c}$. From then onward, clocks (1) and (2) are synchronized.

Now back to the elementary length measurement.

Figure 1.4:



Length measurement of the same physical object in the rest frame and in the moving frame.

– We first measure the left-hand situation in Fig. (1.4). The left tip and the right tip of the bar in coordinates of K are

$$\begin{aligned} O: &\quad (0,0,0,t) \\ A: &\quad (l,0,0,t) \end{aligned} \tag{1.10}$$

where the replacement $x = l$ is permissive since the bar rests in frame K. We carry out the same measurement in coordinates of K' **at the same time**[7] (using synchronized clocks), say at $t' = 0$. With the above discussion and Eq. (1.9), Eq. (1.6) becomes

$$\begin{aligned} x &= b'(v)\,(x' + vt') \\ x' &= \frac{x}{b'(v)} - vt' \qquad \text{and with } t' = 0 \text{ we have} \\ x' &= \frac{x}{b'(v)} \end{aligned}$$

So the coordinates of the events in K' can be written as

$$\begin{aligned} O: &\quad (0,0,0,0)' \\ A: &\quad (x',0,0,0)' = (l/b'(v),0,0,0)' \end{aligned} \tag{1.11}$$

[7]This is equivalent to supposing that there exists an inertial frame K' in which the two events occur simultaneously. From the Galilean point of view this is a trivial assumption. We will later see that this assumption is compatible with the general structure of SpaceTime in Einsteinian relativity.

– Carrying out the same type of measurement, but now with the bar at rest in frame K', Fig.

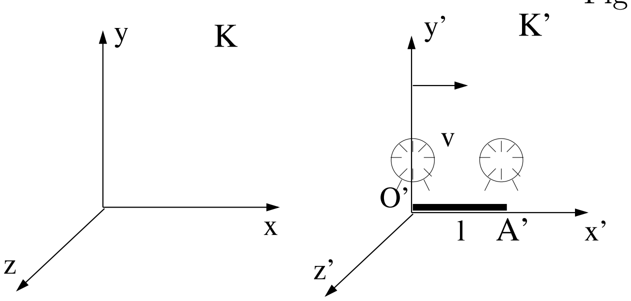


Figure 1.5:

Length measurement of the same physical object in the rest frame and in a frame in relative movement with respect to the rest frame.

(1.5), we measure in K' coordinates

$$\begin{aligned} O' : &\quad (0,0,0,t')' \\ A' : &\quad (l,0,0,t')' \end{aligned} \tag{1.12}$$

Expressed in coordinates of K at the same time, say at $t = 0$, we obtain

$$\begin{aligned} O' : &\quad (0,0,0,0) \\ A' : &\quad (x,0,0,0) = (l/b(v),0,0,0) \end{aligned} \tag{1.13}$$

where now (1.5) has been used for the equality. In detail:

$$\begin{aligned} x' &= b(v)\,(x - vt) \\ x &= \frac{x'}{b(v)} + vt \\ x &= \frac{x'}{b(v)} \end{aligned}$$

Invoking postulate 3 once again, it follows necessarily that

$$b(v) = b'(v) \tag{1.14}$$

since otherwise we would have an objective distinction between frames K and K' (inverting the measurement should not change the measured length of the bar).

The situation is symmetric: The length of the bar fixed in K' appears changed in K, and the bar fixed in K also appears changed in K' in the same manner. This is not surpising, since in both situations an observer has a bar moving away from him. Having obtained this result, we can rewrite Eqns. (1.5) and (1.6) as

$$\begin{aligned} x' &= b(v)\,(x - vt) \\ x &= b(v)\,(x' + vt'). \end{aligned} \tag{1.15}$$

We can now summarize the findings up to this point. Eliminating x' from the second equation in Eq. (1.15) by inserting the first (with the goal of obtaining t' as a function of the unprimed variables) yields

$$\begin{aligned} x &= b(v)\,\big(b(v)\,(x - vt) + vt'\big) \\ vt' &= \frac{x}{b(v)} - b(v)\,(x - vt) \\ t' &= \frac{x}{vb(v)} + tb(v) - \frac{xb(v)}{v} \end{aligned}$$

and the transformation from coordinates of frame K to K' can be written as

$$\boxed{\begin{aligned} x' &= b(v)\,(x - vt) \\ y' &= y \\ z' &= z \\ t' &= b(v)\,\left[t + \frac{x}{v}\left(\frac{1}{b(v)^2} - 1\right)\right] \qquad (1.16) \end{aligned}}$$

It is not possible to determine $b(v)$ without experimental information. Setting $b(v) = 1$ yields the Galilei transformation, which was the accepted form until measurements at very high relative velocity and very high precision became available. Then, deviations from $b(v) = 1$ could be determined[8].

1.1.1.3 The Constancy of the Speed of Light

The result of various measurements has been elevated by Einstein to become a postulate:

Postulate 4: A measurement of the speed of light in any direction in a reference frame always yields a constant value, $c \approx 2.99792458 \times 10^8 \left[\frac{m}{s}\right]$[a] (**Einstein's second postulate of relativity**).

[a]National Institute of Standards (NIST), values of physical constants

Corollary: It is sufficient to formulate Postulate 4 for one reference frame, because it follows from Postulate 3 that $c =$const. in all reference frames, since they have to be equivalent.

We are now in the position to determine the function $b(v)$ unambiguously. It follows from a simple thought experiment:

Fig. (1.6) defines the arrangement. The coordinates of the emission event are

$$\begin{aligned} E &: \ (0,0,0,0) \\ E &: \ (0,0,0,0)' \end{aligned} \tag{1.17}$$

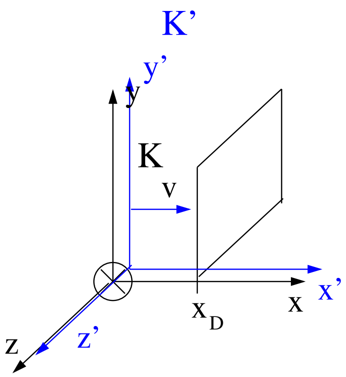
An observer at rest in K measures the detection of the pulse in x at time t. An observer at rest in K' measures the detection at x' at time t'.

Let us first see what non-relativistic transformation laws would tell us. The propagation of the detector in coordinates of K' follows

$$x'_D = x_D - vt \tag{1.18}$$

[8]Fizeau, Michelson-Morley, Joos

Figure 1.6:



A light pulse is emitted from the origin of K at time $t = 0$. At this instant the origins of K and K' coincide, so $t = t' = 0$. The detector is fixed in frame K at position x_D. K' moves with constant velocity v relative to the light source.

for any time instant t. So at the specific instant $t = t_D$ we can write this as

$$x'_D = ct_D - vt_D \tag{1.19}$$

since $x_D = ct_D$. Slightly reformulated we have

$$x'_D = (c - v)t_D \tag{1.20}$$

Since time is absolute (independent of the inertial frame) in Newtonian/Galilean physics, meaning $t_D = t'_D$, this can be written as

$$x'_D = (c - v)t'_D \tag{1.21}$$

The last expression means nothing else than that the light propagates with the velocity $c - v$ seen from the frame K' which is in accord with our classical (non-relativistic) view of physics. However, postulate 4 and the corrolary enforce that the speed of light is also c in frame K'! Therefore, for general positions and instances in time, we write

$$\begin{aligned} x &= ct \\ x' &= ct' \end{aligned} \tag{1.22}$$

Obviously, at this point there is the clear departure from the non-relativistic notion: According to the Galilei transformation we would expect $x' = (c - v)t'$. Eq. (1.22) also tells us that since after some propagation of the pulse $x_D \neq x'_D$ it follows that $t_D \neq t'_D$ which means that the detection event does **not** happen simultaneously in frames K and K'.

From all of this we get

$$\begin{aligned} D:\ (x, 0, 0, t) &= (ct, 0, 0, t) \\ D:\ (x', 0, 0, t')' &= (ct', 0, 0, t')' \end{aligned} \tag{1.23}$$

and the rest is calculation. Introducing this result into Eq. (1.15) we obtain

$$\begin{aligned} ct' &= b(v)\,(ct - vt) = b(v)\,(c - v)t \\ ct &= b(v)\,(ct' + vt') = b(v)\,(c + v)t'. \end{aligned} \tag{1.24}$$

Multiplication of the two equations eliminates the time coordinates and yields

$$\begin{aligned} c^2 tt' &= b(v)^2\,(c^2 - v^2)tt' \\ b(v)^2 &= \frac{c^2}{c^2 - v^2} \end{aligned} \tag{1.25}$$

$$b(v) = \pm\sqrt{\frac{1}{1 - \frac{v^2}{c^2}}} \tag{1.26}$$

The sign ambiguity is resolved through an additional consideration. Taking $v = 0$ in the first line of Eq. (1.16) gives $x' = b(0)x$. But in this case $x' = x \Rightarrow b(0) = +1$, so we must use the positive sign for consistency. We can write

$$b(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}} =: \gamma(v) \tag{1.27}$$

with $\beta := \frac{v}{c}$ and γ the **Lorentz factor**. Introduction of the Lorentz factor into the preliminary transformation Eq. (1.16) results in

$$t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left[t + \frac{x}{v}\left(1 - \frac{v^2}{c^2} - 1\right)\right]$$

and so

$$\boxed{\begin{aligned} x' &= \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t - x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \qquad (1.28) \end{aligned}}$$

as the final form of the Lorentz transformation of coordinates. This type of transformation is called a "**Lorentz boost**" (rotation-free Lorentz transformation).

At this point the verification of the Lemma from subsection 1.1.1.1 is in place. In y and z the Lorentz transformation is trivially linear. The transformations in x and t

$$\begin{aligned} x' &= \gamma x - \gamma v t \\ t' &= \gamma t - \gamma x \frac{v}{c^2} \end{aligned}$$

are linear, too, since γ, v, t are constant in x (first line) and γ, v, x are constant in t (second line)[9]. Therefore, the Lorentz transformation is manifestly a linear transformation.

It is also immediately evident that for relative velocities very small compared to the speed of light, $v << c$, the Galilei transformation

[9] x and t are of course independent coordinates in frame K.

from non-relativistic physics is approximately obtained:

$$\left(\frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}}\right)_{v<<c} \approx x-vt$$
$$\left(\frac{t-x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}\right)_{v<<c} \approx t \tag{1.29}$$

However, since in non-relativistic physics no limit for relative velocities exists, this can only be considered an approximation for special cases. In order to obtain a non-relativistic limit of the theory as such, we have to allow for arbitrary relative velocities v.

Formally, we take the speed of light to infinity in expressions involving velocity ratios[10]

$$\lim_{c\longrightarrow\infty} \frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}} = x-vt$$
$$\lim_{c\longrightarrow\infty} \frac{t-x\frac{v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} = t. \tag{1.30}$$

This **defines** the non-relativistic limit of a physical theory.

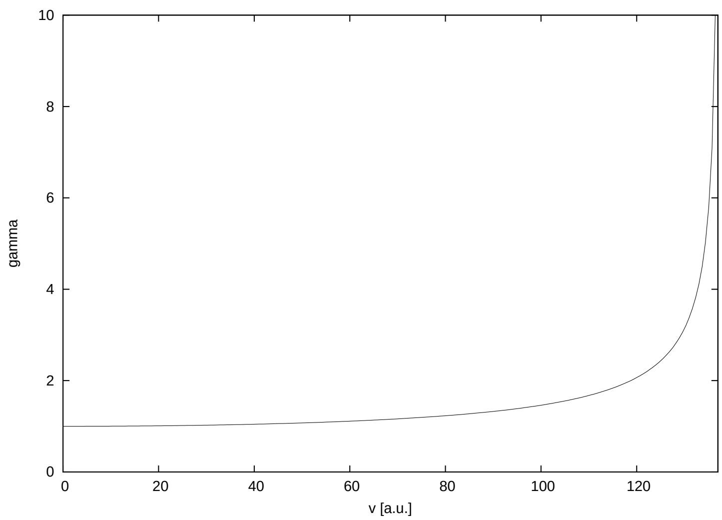
Before continuing, we will here take a look at the important Lorentz factor.

1.1.2 Lorentz Factor γ

Fig. (1.7) displays the function $\gamma(v) = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$.

[10]In electromagnetism, this limit can also be taken, although the dependency on factors $\frac{v}{c}$ is more subtle. For example, $E_m = -\boldsymbol{\mu}\cdot\mathbf{B} \propto \frac{\mathbf{v}\cdot\mathbf{v}'}{c^2}$. The factors $\frac{1}{c}$ become "visible" in the Gaussian unit system, and magnetic moment as well as magnetic field are proportional to velocities of charged particles. This means that magnetism does not exist in the non-relativistic limit of electromagnetism!

Figure 1.7:



Lorentz factor as a function of relative velocity (in atomic units, $c = \frac{1}{\alpha} \approx 137.136$ where α is Sommerfeld's fine-structure constant)

- $\lim_{v \to 0} \gamma(v) = 1$
- $\lim_{v \to c} \gamma(v) = +\infty$
- $\gamma(v > c) \in \mathbb{C}$

It is to be noted that for a relative velocity $v > c$ between inertial frames the Lorentz factor becomes a complex number. From Eq. (1.28) it would then be directly inferred that $x' \in \mathbb{C}$, too. But this would be in contradiction with the basic notion that observables (such as lengths) should be real numbers. We, therefore, take this finding as a first indication that relative velocities greater than the speed of light are not possible.

1.1.2.1 Corollaries

It is interesting to inspect the direct consequences of the Lorentz boost in Eq. (1.28) in terms of two aspects:

1. Time is absolute in non-relativistic physics. However, if one considers a time interval in frame K' (at fixed position x') and transforms the time interval as measured in coordinates of K, then (transformation from K' to K):

$$\Delta t = t_2 - t_1 = \gamma\, t_2' + \gamma\, x' \frac{v}{c^2} - \gamma\, t_1' - \gamma\, x' \frac{v}{c^2} = \gamma\, \Delta t' \qquad (1.31)$$

This means that generally $\Delta t' \neq \Delta t$. In other words, the passage of time depends on the relative motion between inertial frames, and the phenomenon is generally called **time dilation**.

2. We know that length is conserved under Galilei transformations. Under Lorentz boosts, let us consider a spatial interval in K', $\Delta x' = x'_2 - x'_1$, measured at fixed time t' in K'. An observer in K measures

$$\Delta x = x_2 - x_1 = \gamma\, x'_2 + \gamma\, vt' - \gamma\, x'_1 - \gamma\, vt = \gamma\, \Delta x' \qquad (1.32)$$

Since for the relative velocity $v \neq 0$ we have $\gamma > 1$, it follows that the observed length (in K') appears shorter than the measured length in K. Fixed lengths in the direction of relative movement are, therefore, **not conserved** and this phenomenon is generally called **length contraction**.

Various situations can be investigated under these two aspects and interpreted in terms of length contraction and time dilation in special relativity, but the details depend crucially on the specific setup.

1.1.3 Addition Theorem for Velocities

1.1.3.1 Derivation

Newtonian mechanics is invariant under Galilei transformations, and the addition theorem of velocities in the non-relativistic context is well known:

$$\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2 \qquad (1.33)$$

We will now deduce the corresponding law in the framework of special relativity based on the Lorentz transformation. Fig. (1.1.3.1) shows the setup of the thought experiment.

We wish to deduce the relative velocity between frames K and K". The starting point is Eq. (1.28) with the Lorentz transformation in its