> *[...] In my own development Michelson's result has not had a considerable influence. I even do not remember if I knew of it at all when I wrote my first paper on the subject (1905). The explanation is that I was, for general reasons, firmly convinced how this could be reconciled with our knowledge of electro-dynamics. One can therefore understand why in my personal struggle Michelson's experiment played no role or at least no decisive role.*

→ The Michelson Morley experiment did *not* kickstart SPECIAL RELATIVITY.

- Modern Michelson-Morley like tests of the isotropy of the speed of light achieve much higher precision than the original experiment. The authors of Refs. [32, 33], for example, report an upper bound of $\Delta c/c \sim 10^{-17}$ on potential anisotropies of the speed of light by rotating optical resonators.

↓ Lecture 4 [05.11.25]

14 | Two observations:

(1) *No* evidence that there is *no* relativity principle for electrodynamics.

(2) Why does Galilean relativity GR treat mechanics differently anyway?

Put differently: Why should mechanics, a branch of physics artificially created by human society, be different from any other branch of physics? This is not impossible, of course, but it certainly lacks simplicity! (To Galilei's defence: At his time "mechanics" was more or less identical to "physics".)

→ A. Einstein writes in §2 of Ref. [10] as his first postulate:

> *1. Die Gesetze, nach denen sich die Zustände der physikalischen Systeme ändern, sind unabhängig davon, auf welches von zwei relativ zueinander in gleichförmiger Translationsbewegung befindlichen Koordinatensystemen diese Zustandsänderungen bezogen werden.*

We reformulate this into the following postulate:

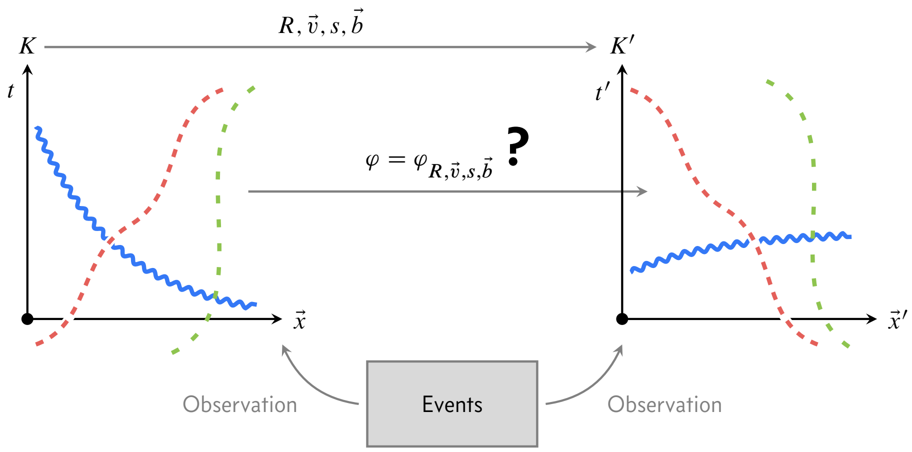
> **§ Postulate 4: (Einstein's principle of) Special Relativity SR**
>
> No ~~mechanical~~ experiment can distinguish between inertial systems.

Note the difference to Galilean relativity GR according to which no experiment *governed by classical mechanics* can distinguish between inertial systems. Einstein simply extended this idea to all of physics – no special treatment for mechanics!

¡! There are various names used in the literature to refer to SR. Here we call it the principle of *special* relativity, where the "special" refers to its restriction on *inertial systems* – as compared to the principle of *general* relativity in GENERAL RELATIVITY that refers to *all* frames (→ *later*). To emphasize its difference to Galilean relativity GR, some authors call SR the *universal* principle of relativity, where "universal" refers to its applicability on *all* laws of nature (not just the realm of classical mechanics).

15 | But now that there are more contenders (mechanics, electrodynamics, quantum mechanics) all of which must be invariant under the same transformation φ, we have to open the quest for φ again:

$$\boxed{\text{What is } \varphi?}$$



The differently colored/shaped trajectories symbolize phenomena of mechanics (red), electrodynamics (blue), and quantum mechanics (green). According to SR, *all* of them must be form-invariant under a common coordinate transformation φ.

¡! To reiterate: This is *not* a question about symmetry properties of equations or models! It is an experimentally testable fact about reality. There is only *one* correct φ and it is just as real as the three-dimensionality of space.

1.4. Transformations consistent with the relativity principle

Since this is a theory lecture, so we cannot do experiments. Let us therefore weaken the question slightly:

What is most general form of φ consistent with reasonable assumptions about reality?

§ Assumptions 1

- SR *Special Relativity*: There is no distinguished inertial system.
- IS *Isotropy:* There is no distinguished direction in space.
- HO *Homomgeneity:* There is no distinguished place in space or point in time.
- CO *Continuity:* φ is a continuous function (in the origin).

Something is "distinguished" if there exists an experiment that can be used to identify it unambiguously.

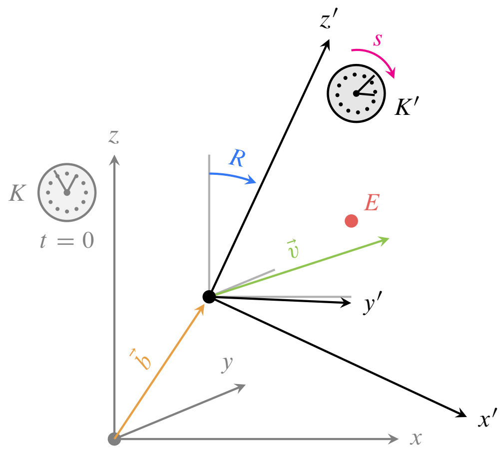
This derivation follows Straumann [9] with input from Schröder [1] and Pal [34].

Detailed calculations: ➔ Problemset 2

1 | Setup:

⊲ Two inertial systems $K \xrightarrow{R,\vec{v},s,\vec{b}} K'$.

⊲ Event $E \in \mathcal{E}$ with coordinates $x \equiv (t, \vec{x})_K \in E$ and $x' \equiv (t', \vec{x}')_{K'} \in E$:



We are interested in the transformation $\varphi \equiv \varphi_{R,\vec{v},s,\vec{b}}$ with

$$x' = \varphi(x)\,. \tag{1.32}$$

Note that SR forbids us to use the inertial system labels K or K' in the definition of φ! We can only use the relative parameters $(R, \vec{v}, s, \vec{b})$ measured in K *wrt* K'.

2 | Affine structure:

Our first goal is to show that φ must be an affine map.

i | $\sphericalangle$ Event $\tilde{E} \in \mathcal{E}$ with coordinates $\tilde{x} = x + a$ in K for some shift $a \in \mathbb{R}^4$.

ii | Homogeneity HO $\rightarrow$

$$\varphi(x+a) - \varphi(x) \stackrel{!}{=} a'(\varphi, a) \tag{1.33}$$

$a'(\varphi, a)$: Shift in K' *independent* of x (this reflects homogeneity in space and time)

Imagine the right-hand side $a'(\varphi, a)$ where *not* independent of x. Then there would be an interval (say, a rod of spatial extend $\vec{a}$) that has the same length $\vec{a}$ in K no matter where it is located, but *variable* length $\vec{a}(\varphi, \vec{a}, \vec{x})$ in K' as a function of $\vec{x}$. The observer in K' can then use this "magic rod" to pinpoint absolute positions in space (the same argument works in time, then with a clock instead of a rod).

iii | For $x = 0$: $a'(\varphi, a) = \varphi(a) - \varphi(0) \rightarrow$

$$\varphi(x+a) = \varphi(x) + \varphi(a) - \varphi(0)\,. \tag{1.34}$$

iv | Let $\Psi(x) := \varphi(x) - \varphi(0) \rightarrow$

$$\Psi(x+a) = \Psi(x) + \Psi(a) \quad \text{and} \quad \Psi(0) = 0\,. \tag{1.35}$$

This would be satisfied if Ψ were *linear*! But we do not know this yet ...

v | Claim: $\Psi(x)$ continuous at $x = 0$ (follows from CO) $\Rightarrow$ Ψ is *linear*.

a | Eq. (1.35) $\rightarrow$ $\Psi(nx) = n\Psi(x)$ for $n \in \mathbb{N}$ (show by induction!)

b | Eq. (1.35) $\rightarrow$ $\Psi(-x) = -\Psi(x)$ (use $\Psi(0) = 0$) $\rightarrow$ $\Psi(nx) = n\Psi(x)$ for $n \in \mathbb{Z}$

c | $\sphericalangle$ Rational number $r = \frac{m}{n}$, $m, n \in \mathbb{Z} \rightarrow$

$$r\Psi(x) = \frac{m}{n}\Psi(x) = \frac{1}{n}\Psi(mx) = \frac{1}{n}\Psi(nrx) = \frac{n}{n}\Psi(rx) = \Psi(rx)\,. \tag{1.36}$$

d | $\Psi(x)$ continuous at $x = 0 \xrightarrow{\text{Eq. (1.35)}} \Psi(x)$ continuous everywhere.
Show this using the definition of continuity, i.e., $\lim_{x\to 0} \Psi(x) = \Psi(0)$!

e | $r\Psi(x) = \Psi(rx)$ for $r \in \mathbb{Q} \xrightarrow{\Psi \text{ continuous}} r\Psi(x) = \Psi(rx)$ for $r \in \mathbb{R}$
Remember that real numbers are defined in terms of (equivalence classes of) limits of rational numbers, i.e., $\mathbb{Q}$ is dense in $\mathbb{R}$.

f | In conclusion:

$$\Psi(x + a) = \Psi(x) + \Psi(a) \quad \text{and} \quad \Psi(rx) = r\Psi(x) \tag{1.37}$$

$\rightarrow \Psi$ is *linear*.

vi | If Ψ is linear, $\varphi(x) = \Psi(x) + \varphi(0)$ is *affine*:

$$\boxed{\varphi(x) = \Lambda x + a} \tag{1.38}$$

with $\Lambda = \Lambda(R, \vec{v}, s, \vec{b})$ a 4×4 matrix and $a = a(R, \vec{v}, s, \vec{b})$ a 4-dimensional vector.

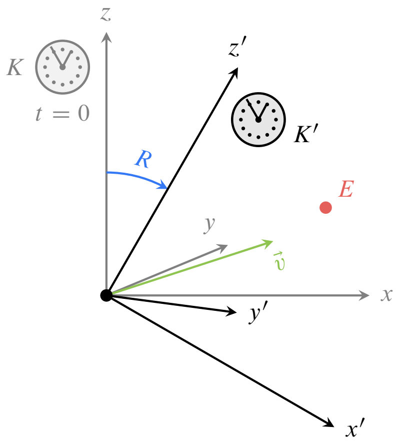
3 | a is simply a spacetime translation [recall Eqs. (1.7) and (1.9)]

$\rightarrow$ No constraints on a

If the origin O' of K' with coordinate $x'_{O'} = (0, \vec{0})$ is shifted in K by $x_{O'} = (s, \vec{b})$, we must set $a := -\Lambda \cdot (s, \vec{b})$ with the Λ to be determined → *below*. In the special case of no boost ($\vec{v} = 0$) and no rotation ($R = \mathbb{1}$), it is simply $a := (-s, -\vec{b})$.

$\rightarrow \sphericalangle$ Homogeneous transformations ($a = 0$) in the following:

$$x' = \varphi(x) = \Lambda x\,. \tag{1.39}$$



4 | We already know from our discussion of inertial systems [recall Eq. (1.11)]:

Rotation group SO(3) must be part of the transformations φ with representation

$$x' = \Lambda_{R^{-1}} x \quad \text{with} \quad \Lambda_R := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & R \end{array}\right) \quad \text{where} \quad R \in \mathrm{SO}(3)\,. \tag{1.40}$$

This is just a fancy way to rewrite Eq. (1.11).

5 | ⁂ *Pure boost* $K \xrightarrow{\mathbb{1},\vec{v},0,\vec{0}} K'$:

i | $\sphericalangle(t)_K = 0 \to \vec{x}' = \mathcal{M}\vec{x}$ for an invertible matrix $\mathcal{M} \in \mathbb{R}^{3\times 3}$:
This is the most general transformation for the position labels of the K and K'-clocks at $t = 0$. Note that we make no statements on the times t' displayed by the K'-clocks at $t = 0$.

$$\mathcal{M} = R_1 D R_2 = R_1 D R_1^T R = M R \tag{1.41}$$

with $R \in \mathrm{O}(3)$ and $M^T = M$.

This follows from the ↓ *singular value decomposition* of real matrices with $R_1, R_2 \in \mathrm{O}(3)$ and D a diagonal matrix.

ii | With spatial rotations Eq. (1.40) we can always transform the K-coordinates by $\vec{x} \mapsto R^{-1}\vec{x}$ such that $\vec{x}' = \mathcal{M}\vec{x} = M\vec{x}$ at $t = 0 \to$

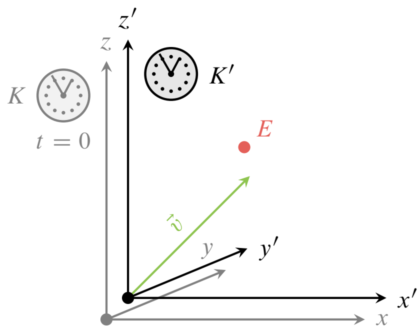
⁂ *Pure boost* $K \xrightarrow{\mathbb{1},\vec{v},0,\vec{0}} K'$:

$$x' = \Lambda_{\vec{v}} x \quad \Leftrightarrow \quad \begin{cases} t' = a(\vec{v})\, t + \vec{b}(\vec{v}) \cdot \vec{x} \\ \vec{x}' = M(\vec{v})\, \vec{x} + \vec{e}(\vec{v})\, t \end{cases} \tag{1.42}$$

- a: $\vec{v}$-dependent scalar
- $\vec{b}, \vec{e}$: $\vec{v}$-dependent vectors
- $M^T = M$: $\vec{v}$-dependent 3×3-matrix

Pure boosts are therefore characterized by a symmetric transformation of the spatial coordinates at $t = 0$ in K. Geometrically, this implies that there are *three* (orthogonal) lines through the origin of K which are mapped onto themselves under the boost [spanned by the real eigenvectors of $M(\vec{v})$]. One can show that the only other possibility is that there is a *single* invariant line. This follows because a real 3×3 matrix $\mathcal{M}$ always has at least one real eigenvector; that the possibility of *two* real eigenvectors is excluded is a special property of Lorentz transformations and cannot be derived at this point. Note that rotation matrices also have only one real eigenvector (the rotation axis). Thus *pure* boosts (with three real eigenvectors) are those boosts *without* a rotation mixed in.

→ We focus on pure boosts in the remainder of this derivation:



¡! Our characterization of a pure boost does *not* imply that at $t = 0$ the axes of the two systems K and K' align (as suggested by the sketch and naïvely expected). If this were the case, the eigenbasis of $M(\vec{v})$ would be given by the basis vectors $\hat{e}_i$ in K. Since we do not know the form of $M(\vec{v})$ (yet), we cannot make this assumption! So do not take this sketch literally, it only illustrates *symbolically* the situation of a pure boost in an arbitrary direction.

6 | Isotropy:

Here are two lines of arguments that use isotropy **IS** to restrict the form of Eq. (1.42) further:

- Argument A:

  i | We claim that isotropy **IS** requires the following multiplicative structure for pure boosts and rotations:

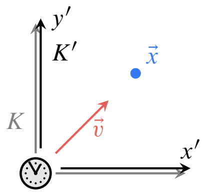
$$\Lambda_R \Lambda_{\vec{v}} \Lambda_{R^{-1}} \stackrel{!}{=} \Lambda_{R\vec{v}} \quad \Leftrightarrow \quad \forall_x : \Lambda_R \Lambda_{\vec{v}} x = \Lambda_R x' \stackrel{!}{=} \Lambda_{R\vec{v}} \Lambda_R x \,. \tag{1.43a}$$

$$\Leftrightarrow \quad \forall_x : \Lambda_{\vec{v}} x \stackrel{!}{=} \Lambda_{R^{-1}} \Lambda_{R\vec{v}} (\Lambda_R x) \,. \tag{1.43b}$$

  The reasoning goes as follows:

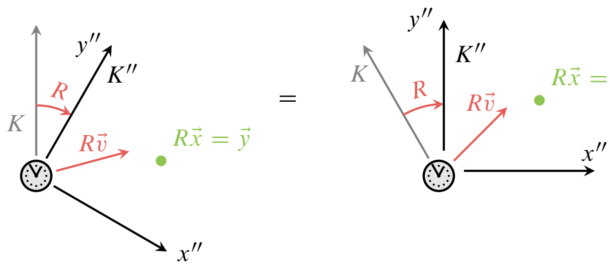
  1. $\triangleleft$ Left-hand side of Eq. (1.43b):

     $x = (t, \vec{x})$ are the coordinates of some event in K and $\Lambda_{\vec{v}} x$ of the same event in K':



  2. $\triangleleft$ Right-hand side of Eq. (1.43b):

     We consider $y = (t, \vec{y}) := \Lambda_R x = (t, R\vec{x})$ as an *active* transformation, i.e., y denotes a different event that is spatially rotated from x by R. To state our isotropy claim **IS**, we now rotate the coordinate system K'' in which we want to express this event *in the same way*. This implies a rotated boost $\Lambda_{R\vec{v}}$ and a subsequent rotation of the coordinate axes by R via $\Lambda_{R^{-1}}$. (Remember that when rotating the coordinate *axes* by R, the *coordinates* of an event transform by $\Lambda_{R^{-1}}$.):



  3. Spatial isotropy **IS** is the property that the event x as seen from K' cannot be distinguished from the rotated event y as seen from the rotated system K''; this is Eq. (1.43b).

  ii | Now we can use Eq. (1.42) to rewrite Eq. (1.43a) as

$$t' \stackrel{!}{=} a(R\vec{v})\, t + \vec{b}(R\vec{v}) \cdot R\vec{x} \tag{1.44a}$$

$$R\vec{x}' \stackrel{!}{=} M(R\vec{v})\, R\vec{x} + \vec{e}(R\vec{v})\, t \tag{1.44b}$$

  iii | A comparison with Eq. (1.42) (for all t and $\vec{x}$ and arbitrary $\vec{v}$ and R) leads to constraints on the unknown functions:

  - $a(\vec{v}) \stackrel{!}{=} a(R\vec{v}) \to a(\vec{v}) = a_v$ with $v = |\vec{v}|$

    Functions invariant under arbitrary rotations can only depend on the norm $|\vec{v}|$.

- $\vec{b}(\vec{v}) \stackrel{!}{=} R^T\vec{b}(R\vec{v}) \to \vec{b}(\vec{v}) = b_v\vec{v}$

  Note that $\vec{b}(R\vec{v})\cdot R\vec{x} = [R^T\vec{b}(R\vec{v})]\cdot\vec{x}$. Let $R_{\hat{v}}$ be some rotation with axis $\hat{v} = \vec{v}/v$ such that $R_{\hat{v}}\vec{v} = \vec{v}$; then $\vec{b}(\vec{v}) \stackrel{!}{=} R_{\hat{v}}^T\vec{b}(\vec{v})$ and therefore $\vec{b}(\vec{v}) \propto \vec{v}$ since rotation matrices have only a single eigenvector.

- $RM(\vec{v}) \stackrel{!}{=} M(R\vec{v})R \to M(\vec{v}) = c_v\,\mathbb{1} + d_v\,\hat{v}\hat{v}^T$

  First recall that $M^T(\vec{v}) = M(\vec{v})$ such that $M(\vec{v})$ can be written as sum of orthogonal projectors (projecting onto its eigenspaces). It is in particular $R_{\hat{v}}M(\vec{v})R_{\hat{v}}^T \stackrel{!}{=} M(\vec{v})$ such that one of the eigenvectors must be $\hat{v} \propto \vec{v}$. The remaining two eigenvectors are orthogonal to $\hat{v}$ and can therefore be mapped onto each other by $R_{\hat{v}}$. Since $R_{\hat{v}}$ commutes with $M(\vec{v})$, their eigenvalues must be degenerate such that the two-dimensional subspace orthogonal to $\hat{v}$ is a degenerate eigenspace. The most general spectral decomposition of $M(\vec{v})$ is then the one given above.

- $R\vec{e}(\vec{v}) \stackrel{!}{=} \vec{e}(R\vec{v}) \to \vec{e}(\vec{v}) = e_v\vec{v}$

  This is the same argument as for $\vec{b}(\vec{v})$.

- Argument B:

  A shorter (but less rigorous) line of arguments goes as follows:

  i | To define the unknown functions algebraically, we are only allowed to use the vector $\vec{v}$ and constant scalars. We *cannot* use $\vec{x}$ or t due to linearity, and any other constant vector (like $\hat{e}_x = (1,0,0)^T$) would pick out some direction and therefore violate isotropy **IS**.

  ii | Since the only *scalar* one can construct from a single vector is its norm, $|\vec{v}|^2 = \vec{v}\cdot\vec{v}$, it must be $a(\vec{v}) = a_v$.

  iii | Similarly, since the only *vector* one can construct from a single vector is a scalar multiplied by the vector itself, it must be $\vec{b}(\vec{v}) = b_v\vec{v}$ and $\vec{e}(\vec{v}) = e_v\vec{v}$.

  iv | Lastly, since $M^T(\vec{v}) = M(\vec{v})$, we can decompose the matrix into orthogonal projectors: $M(\vec{v}) = \sum_i \lambda_i(v)P_i(\vec{v})$. The only projectors that can be defined by a single vector are $P_0 = \hat{v}\hat{v}^T$ and $P_1 = \mathbb{1} - P_0 = \mathbb{1} - \hat{v}\hat{v}^T$ which leads to the most general form $M(\vec{v}) = c_v\,\mathbb{1} + d_v\,\hat{v}\hat{v}^T$.

Both arguments lead to the same form for pure boosts $\Lambda_{\vec{v}}$ consistent with isotropy **IS**:

$$t' = a_v\,t + b_v\,(\vec{v}\cdot\vec{x}) \tag{1.45a}$$

$$\vec{x}' = c_v\,\vec{x} + \tfrac{d_v}{v^2}\,\vec{v}(\vec{v}\cdot\vec{x}) + e_v\,\vec{v}\,t \tag{1.45b}$$

with $v = |\vec{v}| = |R\vec{v}|$ and $(R\vec{v}\cdot R\vec{x}) = (\vec{v}\cdot\vec{x})$.

7 | $\triangleleft$ Trajectory of origin O' of K':

- In K': $\vec{x}'_{O'} = 0$ (This is the operational *definition* of the origin O'.)
- In K: $\vec{x}_{O'} = \vec{v}t$ (This is the operational *definition* of $\vec{v}$ in $K \xrightarrow{\mathbb{1},\vec{v},0,\vec{0}} K'$.)

In Eq. (1.45b):

$$\vec{0} = c_v\,\vec{v}t + \tfrac{d_v}{v^2}\,\vec{v}(\vec{v}\cdot\vec{v})t + e_v\,\vec{v}\,t \tag{1.46a}$$

$$\vec{v} \neq \vec{0}\ \&\ \forall t \quad \Rightarrow \quad 0 = c_v + d_v + e_v \tag{1.46b}$$

8 | Reciprocity:

i | $\sphericalangle$ Inverse transformation $K' \xrightarrow{\mathbb{1},\vec{v}',0,\vec{0}} K$ from K' to K:

$$\Lambda_{\vec{v}'}\Lambda_{\vec{v}} = \mathbb{1} \quad \Leftrightarrow \quad \Lambda_{\vec{v}'} = \Lambda_{\vec{v}}^{-1} . \tag{1.47}$$

Note that $\vec{v}'$ is the velocity of the origin O of K as measured in K'.

In general: $\vec{v}' = \vec{V}(\vec{v})$ with unknown function $\vec{V}$.

We assume *reciprocity*: $\vec{v}' = -\vec{v}$ such that

$$\Lambda_{\vec{v}}^{-1} = \Lambda_{-\vec{v}} . \tag{1.48}$$

While this is clearly the most reasonable/intuitive assumption, it is not trivial! Recall that $\vec{v}$ is the speed of the origin O' of K' measured with the clocks in K, whereas $\vec{v}'$ is the speed of the origin O of K measured with *different* clocks in K'. So without additional assumptions we cannot conclude that the results of these measurements yield reciprocal results.

However, the assumption of reciprocity can be rigorously derived from relativity **SR**, isotropy **IS** and homogeneity **HO**, see Ref. [35]. Reciprocity is therefore not an independent assumption.

ii | $\sphericalangle$ Inverse transformation in Eq. (1.45):

$$t = a_v\, t' - b_v\, (\vec{v}\cdot\vec{x}') \tag{1.49a}$$

$$\vec{x} = c_v\, \vec{x}' + \frac{d_v}{v^2}\, \vec{v}(\vec{v}\cdot\vec{x}') - e_v\, \vec{v}\, t' \tag{1.49b}$$

iii | Eq. (1.49) in Eq. (1.45) & Eq. (1.46b) $\overset{\circ}{\rightarrow}$ (we suppress the v dependence)

$$c^2 = 1 , \tag{1.50a}$$

$$a^2 - ebv^2 = 1 , \tag{1.50b}$$

$$e^2 - ebv^2 = 1 , \tag{1.50c}$$

$$e(a+e) = 0 , \tag{1.50d}$$

$$b(a+e) = 0 . \tag{1.50e}$$

To show this, use $\vec{v} = (v_x, 0, 0)^T$ with $v_x \neq 0$ and remember that the equations you obtain from plugging Eq. (1.49) into Eq. (1.45) must be valid for all t' and $\vec{x}'$. Use Eq. (1.46b) to replace $c_v + d_v$ by $-e_v$.

We can conclude:

- $\xrightarrow{\text{Eq. (1.50a)}} c = 1$ ($c = -1$ contradicts $\lim_{v\to 0}\Lambda_{\vec{v}} \overset{!}{=} \mathbb{1}$)
- $\xrightarrow{\text{Eq. (1.50c)}} e \neq 0 \xrightarrow{\text{Eq. (1.50d)}} a + e = 0$
  $\rightarrow$ Eq. (1.50b) $\equiv$ Eq. (1.50c) & Eq. (1.50e) satisfied

9 | Collecting results from Eq. (1.50) & Eq. (1.46b):

$$c = 1 , \quad e = -a , \quad d = a - 1 , \quad b = \frac{1-a^2}{av^2} . \tag{1.51}$$

$d = a - 1$ follows from Eq. (1.46b) and the first two equations.

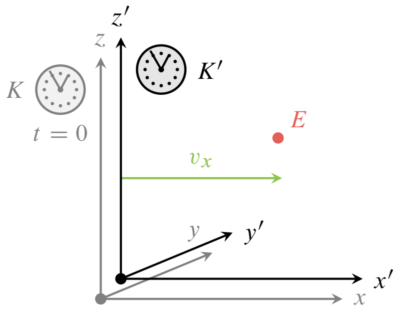
Eq. (1.45) $\xrightarrow{\text{Eq. (1.51)}}$

$$t' = a_v\, t + \frac{1-a_v^2}{va_v}\, (\hat{v}\cdot\vec{x}) \tag{1.52a}$$

$$\vec{x}' = \vec{x} + [a_v - 1]\, \hat{v}(\hat{v}\cdot\vec{x}) - va_v\, \hat{v}\, t \tag{1.52b}$$

with $\hat{v} := \vec{v}/|\vec{v}|$.

10 | ⊲ Special boost $\vec{v} = (v_x, 0, 0)^T$ in x-direction:



$$t' = a_v\, t + \frac{1-a_v^2}{v_x a_v}\, x \tag{1.53a}$$

$$x' = a_v\, x - v_x a_v\, t \tag{1.53b}$$

$$y' = y \tag{1.53c}$$

$$z' = z \tag{1.53d}$$

Note that $v = |v_x|$ with $v_x \in \mathbb{R}$.

Matrix form:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \underbrace{\left(\begin{array}{cc|cc} a_v & \frac{1-a_v^2}{v_x a_v} & & \\ -v_x a_v & a_v & & \\ \hline & & 1 & 0 \\ & & 0 & 1 \end{array}\right)}_{=:\Lambda_{v_x}} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{1.54}$$

In the following, we refer to the upper 2×2-block as $A(v_x)$.

11 | Group structure:

i | Relativity principle SR →

$$\varphi(K' \xrightarrow{R_2,\vec{v}_2,s_2,\vec{b}_2} K'') \circ \varphi(K \xrightarrow{R_1,\vec{v}_1,s_1,\vec{b}_1} K') \stackrel{!}{=} \varphi(K \xrightarrow{R_3,\vec{v}_3,s_3,\vec{b}_3} K'') \tag{1.55}$$

for some parameters $(R_3, \vec{v}_3, s_3, \vec{b}_3)$ that are a function of $(R_i, \vec{v}_i, s_i, \vec{b}_i)_{i=1,2}$.

In words:

The concatenation of a coordinate transformations from K to K' and from K' to K'' must be another coordinate transformation *that is parametrized by data that relates the reference systems K with K'' directly (without referring to K' in any way).*

You may ask why Eq. (1.55) is a constraint on φ in the first place. After all, we could just *define* that

$$\varphi(K \xrightarrow{R_3,\vec{v}_3,s_3,\vec{b}_3} K'') := \varphi(K' \xrightarrow{R_2,\vec{v}_2,s_2,\vec{b}_2} K'') \circ \varphi(K \xrightarrow{R_1,\vec{v}_1,s_1,\vec{b}_1} K')\,. \tag{1.56}$$

The problem is that the function defined such generically depends on 8 (!) parameters $R_1, \vec{v}_1, s_1, \vec{b}_1, R_2, \vec{v}_2, s_2, \vec{b}_2$ – it is a non-trivial functional constraint on φ that these can be compressed to four parameters $R_3, \vec{v}_3, s_3, \vec{b}_3$. This "compression" is mandated by the relativity principle SR according to which all inertial systems must be treated equally. In particular, the transformation between two systems K and K'' can only depend on parameters that can be experimentally determined from within these two systems. (The existence of) a third frame K' cannot be of relevance for this transformation as this would make K' special.

Combined with the existence of an inverse transformation (← *above*):

→ The set of all transformations forms a ↓ (multiplicative) group.

Note that *associativity* is implicit since we talk about the concatenation of linear/affine maps.

ii | In particular:

$$\Lambda_{v_x}\Lambda_{u_x} \stackrel{!}{=} \Lambda_{w_x} \quad \Leftrightarrow \quad A(v_x)A(u_x) \stackrel{!}{=} A(w_x) \tag{1.57}$$

where $w_x = W(v_x, u_x)$ has to be determined.

- ¡! Using the restricted form of the boost Eq. (1.54) that followed from previous arguments, it follows indeed that the concatenation of two pure boosts *in the same direction* has again the form of a pure boost (in the same direction). For the arguments that follow, this is sufficient.

  However, in general, the multiplicative group structure Eq. (1.55) allows for two boosts to concatenate to a *combination* of boosts and rotations. As we will see → *later*, this is indeed what happens: The concatenation of two pure boosts (in different directions) produces a boost with a rotation mixed in (↑ *Thomas-Wigner rotation*).

- Note that due to Eq. (1.43a) all that follows holds for any pair of *collinear* velocities $\vec{v}$ and $\vec{u}$ (there is nothing special about the x-direction). Indeed, let R be a rotation that maps $\vec{v}$ and $\vec{u}$ to vectors on the x-axis, $\vec{v}_x := R\vec{v}$ and $\vec{u}_x := R\vec{u}$. Then

$$\Lambda_{\vec{v}}\Lambda_{\vec{u}} \stackrel{1.43a}{=} \Lambda_{R^{-1}}\Lambda_{\vec{v}_x}\Lambda_{\vec{u}_x}\Lambda_R \stackrel{!}{=} \Lambda_{R^{-1}}\Lambda_{\vec{w}_x}\Lambda_R \stackrel{1.43a}{=} \Lambda_{\vec{w}} \tag{1.58}$$

  where $\vec{w}$ is again collinear with $\vec{v}$ and $\vec{u}$.

$\stackrel{\circ}{\rightarrow}$ (use that the diagonal elements of $A(w_x)$ must be equal)

$$\forall_{v_x,u_x}: \quad \frac{1-a_v^2}{v_x^2 a_v^2} \stackrel{!}{=} \frac{1-a_u^2}{u_x^2 a_u^2} \tag{1.59}$$

→ Universal constant:

$$\kappa := \frac{a_v^2 - 1}{v_x^2 a_v^2} = \text{const} \tag{1.60}$$

Note: $[\kappa] = \text{Velocity}^{-2}$

→

$$a_v = \frac{1}{\sqrt{1-\kappa v_x^2}}\,. \tag{1.61}$$

We use the positive solution for a_v since $\lim_{v\to 0} A(v) \stackrel{!}{=} \mathbb{1}$, i.e., $\lim_{v\to 0} a_v \stackrel{!}{=} 1$.

iii | With this we check: $A(v_x)A(u_x) \stackrel{\circ}{=} A(w_x)$ with

$$w_x = W(v_x, u_x) \stackrel{\circ}{=} \frac{v_x + u_x}{1 + u_x v_x \kappa}\,. \tag{1.62}$$

Eq. (1.62) becomes important later: it tells us how to add velocities in SPECIAL RELATIVITY.

12 | Preliminary result:

Eq. (1.52) & Eq. (1.60) → Boost $\Lambda_{\vec{v}}$ in direction $\hat{v}$ with velocity $\vec{v} = v\hat{v}$:

$$\boxed{\begin{aligned} t' &= a_v\left[t - \kappa\,(\vec{v}\cdot\vec{x})\right] \\ \vec{x}' &= \vec{x} + [a_v - 1]\,\hat{v}(\hat{v}\cdot\vec{x}) - a_v\,\vec{v}\,t \end{aligned}} \tag{1.63a, 1.63b}$$

with

$$\boxed{a_v = \frac{1}{\sqrt{1-\kappa v^2}}\,.} \tag{1.64}$$

This is the most general transformation between two inertial coordinate systems that move with relative velocity $\vec{v}$ [with coinciding origins $x'_{O'} = x'_O = x_O = (0,\vec{0})$ and without spatial rotations] that is consistent with our basic assumptions stated at the beginning of this section: SR, HO, and IS.

The only undetermined parameter left is κ!

1.5. The Lorentz transformation

The purpose of this section is to select the value for κ that describes our reality.

13 | Since $[\kappa] = \text{Velocity}^{-2}$ define formally: $\kappa \equiv 1/v_{\text{max}}^2$.

Why we subscribe the velocity v_{max} with "max" will become clear below.

14 | Three cases:

- $\kappa = 0 \;\Leftrightarrow\; v_{\text{max}} = \infty$:

$$\boxed{\text{Eq. (1.63)} \;\Rightarrow\; \left.\begin{aligned} t' &= t \\ \vec{x}' &= \vec{x} - \vec{v}\,t \end{aligned}\right\} \;⁂\; \textit{Galilei boost}} \tag{1.65a}$$

  → Maxwell equations are *not* form-invariant under φ.

  → Maxwell equations cannot be correct and must be modified.

  → Experiment that shows the invalidity of Maxwell equations?

  Note that we cannot conclude the *validity* of classical mechanics from this; Newton's equations may still require modifications (without spoiling the Galilean symmetry, of course).

- $\kappa > 0 \;\Leftrightarrow\; v_{\text{max}} < \infty$:

$$\boxed{\text{Eq. (1.63)} \;\Rightarrow\; \left.\begin{aligned} t' &= \gamma\left(t - \frac{\vec{v}\cdot\vec{x}}{v_{\text{max}}^2}\right) \\ \vec{x}' &= \vec{x} + (\gamma - 1)\,\hat{v}(\hat{v}\cdot\vec{x}) - \gamma\,\vec{v}\,t \end{aligned}\right\} \;⁂\; \textit{Lorentz boost}} \tag{1.66a}$$

with the ⁂ *Lorentz factor* (previously denoted a_v)

$$\boxed{\gamma_v \equiv \gamma := \frac{1}{\sqrt{1-\beta^2}} \quad \text{and} \quad \beta := v/v_{\max}\,.} \tag{1.67}$$

→ Newton's equations are *not* form-invariant under φ.

→ Classical mechanics cannot be correct and must be modified.

→ Experiment that shows the invalidity of Newton's equations?

Similarly, we cannot conclude the *validity* of electrodynamics from this; Maxwell equations may still require modifications (without spoiling the Lorentz symmetry).

- $\kappa < 0$: Physically not relevant. (➔ Problemset 2; we ignore this solution in the following.)

  This solution is not self-consistent (see e.g. Ref. [34]) and immediately leads to implications that are not observed in nature.

  For example, the rule Eq. (1.62) to compute the velocity w_x between K/K'' from the velocities v_x and u_x between K/K' and K'/K'' reads for $\kappa < 0$

  $$w_x = \frac{v_x + u_x}{1 - u_x v_x |\kappa|}\,. \tag{1.68}$$

  Let $u_x, v_x > 0$ be positive, i.e., K' moves in *positive* x-direction *wrt* K and K'' moves also in *positive* x-direction *wrt* K'. But for large enough velocities $u_x v_x > 1/|\kappa|$ we find $w_x < 0$ such that K'' moves in *negative* x-direction *wrt* K.

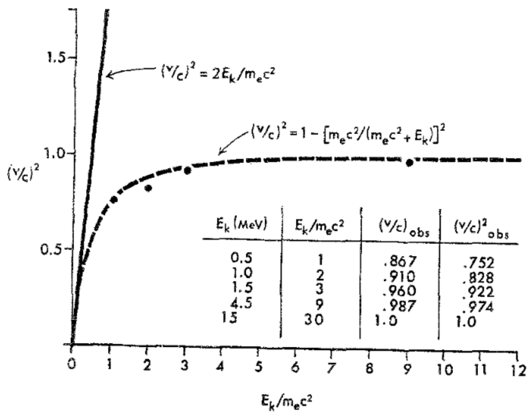
  No such effect has ever been observed; if you do, let us know!

Note that at no point we used or claimed that $v_{\max}$ is the speed of light!

$$\boxed{\text{Which transformation describes reality: } v_{\max} < \infty \text{ or } v_{\max} = \infty?}$$

15 | Evidence:

- Maximum velocity $v_{\max} \approx c < \infty$ for electrons:
  [Plot from Ref. [36]: ↑ *Bertozzi experiment* (1964)]



→ Newton's equations are clearly invalid for high velocities!

See Refs. [36, 37] for more technical details. Note that these results were obtained decades after Einstein published his seminal paper in 1905.

- *By contrast:*

  No evidence for the invalidity of Maxwell equations (on the macroscopic level).

  Electrodynamics, as encoded by the Maxwell equations, is of course not a truly fundamental theory as it is the classical limit of a quantum theory: Quantum electrodynamics (QED). For example, the linearity of the Maxwell equations (= EM waves cannot scatter off each other) is an *approximation*; in QED photons *can* (weakly) scatter off each other! This is why I emphasize that Maxwell theory is experimentally valid only on the *macroscopic level*. Note, however, that QED has the same spacetime symmetry group as electrodynamics, namely Lorentz transformations.

16 | Hence it is reasonable stipulate $v_{\max} < \infty$ and postulate:

> The transformations φ between inertial systems are given by *Lorentz transformations*.
> These transformations must be (part of) the spacetime symmetries of *all* physical theories.

The last statement is often rephrased as follows:

> All (fundamental) theories must be *form-invariant* (covariant) under Lorentz transformations.

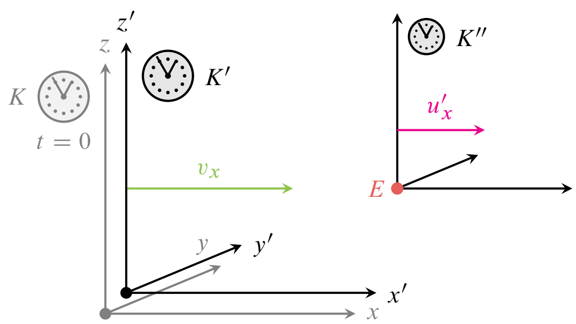
This is just SR all over again: The equations of models that describe reality must "look the same" (more precisely: be functionally equivalent) in all inertial systems. Since the transformations between inertial systems are given by Lorentz transformations (and not Galilean transformations, as historically anticipated), this requires their form-invariance under Lorentz transformations.

→ SPECIAL RELATIVITY restricts the structure of *all fundamental theories of physics*!

This is what is meant by the statement that SPECIAL RELATIVITY is a *theoretical framework* (German: *Rahmentheorie*) or "meta theory": It provides a "recipe" (*ordering principle*) of how to construct consistent theories of physics. The Standard Model of particle physics, for example, is form-invariant under Lorentz transformations, and if you propose an extension thereof (for example to give neutrinos a mass) you better make sure that the terms you write down are also form-invariant under Lorentz transformations (otherwise you will not be taken seriously!). Note, however, that this perspective prevents an important insight: What we really study is an entity called *spacetime*, and this entity has a property: Lorentz symmetry. Since all our (fundamental) physical theories are formulated *on* spacetime, it should not come as a surprise that the Lorentz symmetry of spacetime shows up all over the place.

17 | Interpretation of $v_{\max}$:

i | $\sphericalangle$ Systems $K \xrightarrow{v_x} K'$ and signal with velocity $\frac{\mathrm{d}x'}{\mathrm{d}t'} = u'_x$:



Question: What is the velocity $u_x = \frac{dx}{dt}$ of this signal in K?

ii | Remember (Group structure!):

$$\varphi(K' \xrightarrow{v_2} K'') \circ \varphi(K \xrightarrow{v_1} K') = \varphi(K \xrightarrow{v_3} K'') \quad \text{with} \quad v_3 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{v_{\max}^2}} . \tag{1.69}$$

Let $v_1 = v_x$ and $v_2 = u'_x$ so that $v_3 = u_x$ (i.e., the signal is at rest in the origin of K'').

You can also derive this by computing the time derivative of the position of the signal in K using a Lorentz transformation; you will do this properly when you derive a more general addition of velocities (➔ Problemset 2).

iii | Addition formula for collinear velocities:

$$\boxed{u_x = \frac{v_x + u'_x}{1 + \frac{v_x u'_x}{v_{\max}^2}}} \tag{1.70}$$

Because of isotropy **IS** this formula must be true in all directions (not just in x-direction) *as long as the two velocities to be added are parallel*. We still keep the index x to signify that these are not absolute values of velocities.

- Note that for $v_{\max} \to \infty$ we get back the "conventional" (= Galilean) additivity of velocities:

$$u_x = (v_x + u'_x)\left[1 - \frac{v_x u'_x}{v_{\max}^2} + \dots\right] \overset{v_{\max}\to\infty}{=} v_x + u'_x \tag{1.71}$$

  From this expansion and the validity of classical mechanics for small velocities (in particular its law for adding velocities), we can also conclude that $v_{\max}$ must be *large compared to everyday experience*.

- A historically influential experiment that (in hindsight) can be explained by the relativistic addition of velocities Eq. (1.70) is the ↑ *Fizeau experiment* [38, 39] (see also ↑ *Fresnel drag coefficient*). The Fizeau experiment was one of the crucial hints that led Einstein to SPECIAL RELATIVITY.

iv | $\sphericalangle$ $0 \leq v_x, u'_x \leq v_{\max}$: ($\tilde{v}_x := v_x/v_{\max}$ so that $0 \leq \tilde{v}_x, \tilde{u}_x \leq 1$)

$$u_x = v_{\max} \frac{\tilde{v}_x + \tilde{u}'_x}{1 + \tilde{v}_x \tilde{u}'_x} \leq v_{\max} \tag{1.72}$$

Here we used that $a + b \leq 1 + ab$ for numbers $0 \leq a, b \leq 1$ since then $(1-a)(1-b) \geq 0$.

→ "Addition" of velocities Eq. (1.70) never exceeds $v_{\max}$.

→ $v_{\max}$ plays the role of a *maximum velocity*.

v | $\sphericalangle$ Signal with maximum velocity in K': $u'_x = v_{\max}$:

$$u_x = \frac{v_{\max} + v_x}{1 + \frac{v_{\max} v_x}{v_{\max}^2}} = v_{\max} \frac{v_{\max} + v_x}{v_{\max} + v_x} = v_{\max} \tag{1.73}$$

Note that the result is completely independent of the velocity v_x of K'!

→ Whatever moves with the maximum velocity $v_{\max}$ does so *in all inertial systems*!

Please appreciate how counterintuitive this effect is *from the perspective of everyday experience*! But also notice that we didn't have to postulate it: The relativity principle SR together with the *existence* of a (finite) maximum velocity is sufficient.

If you think about it: Assuming a maximum velocity (in the absence of a preferred reference frame) automatically invalidates the simple Galilean law of additive velocities. So it is actually not surprising at all that the maximum velocity must be independent of the reference system.

18 | Experiments (in particular: the validity of Maxwell equations) show:

$$\boxed{v_{\max} = c = 299\,792\,458\,\mathrm{m\,s^{-1}}} \tag{1.74}$$

Note that since 1983 the value of c in the international system of units (SI) is *exact* by definition.

A. Einstein incorporated this insight in §2 of Ref. [10] as his second postulate:

> *2. Jeder Lichtstrahl bewegt sich im "ruhenden" Koordinatensystem mit der bestimmten Geschwindigkeit V, unabhängig davon, ob dieser Lichtstrahl von einem ruhenden oder bewegten Körper emittiert ist.*

Note that at the time it was conventional to denote the speed of light with a capital V. The convention switched to our now standard lower-case c just a few years later. For more historical background:

➔ https://math.ucr.edu/home/baez/physics/Relativity/SpeedOfLight/c.html

We can condense this into:

> **§ Postulate 5: Constancy of the speed of light** SL
>
> The speed of light is independent of the inertial system in which it is measured.

Comments:

- *If* you take the validity of the Maxwell equations for granted, then $v_{\max} = c < \infty$ (and thereby SL) follows immediately from the relativity principle SR because then the Maxwell equations must be valid in all inertial systems. But you've learned in your course on electrodynamics that the wavelike solutions of these equations always propagate with group velocity c in vacuum. This is only possible if the speed of light plays the role of the limiting velocity: $v_{\max} = c$.

  Einstein acknowledges as much at the beginning of Ref. [11]. However, SL is empirically weaker than claiming the validity of Maxwell's equations (after all, there could be alternative equations that also predict the velocity c of wavelike solutions). At the time when Einstein formulated SL in [10], he also worked on the photoelectric effect (another of his *annus mirabilis* papers [40]). The postulation of "quanta of light" is the foundation of quantum mechanics, but cannot be explained by Maxwell's equations. It is therefore reasonable to assume that Einstein didn't want to rely on the validity of this specific theory when formulating his SPECIAL RELATIVITY. He therefore opted for the empirically weaker (but still sufficient) assumption SL.

- If you derive the transformation φ using *both* postulates SR and SL the derivation is shorter (see e.g. [1] or [5]); one then of course doesn't find the Galilei transformations as an option. Note, however, that the relativity principle SR is a reasonable and intuitive starting point that doesn't need much convincing (after all, we witness the relativity of Newtonian mechanics in our everyday life). By contrast, the speed of light postulate SL clashes directly with our everyday experience (how velocities add up, that is). Through our elaborate derivation we learned how much is already implied by the simple, reasonable assumption of relativity. We only had to check whether there is any evidence of a finite maximum velocity $v_{\max}$. The counterintuitive feature that this velocity is the same viewed from all inertial systems was then a necessary conclusion from our derivation.

† Note: Finite speed of causality (Locality)

Another insight from our SR-based derivation of the Lorentz transformation is that the formulation of the speed-of-light postulate SL is conceptually misleading:

- The constant $v_{\max}$ and its role as a maximum velocity followed *without* referring to light (or electrodynamics) in any way!

  Put bluntly: SPECIAL RELATIVITY is *not* about the "strange behavior" of light!

- The relevant speed for SPECIAL RELATIVITY is the *speed of causality*: How fast can information travel, i.e., one event affect another. $v_{\max}$ is the maximum speed of *causal interactions*, irrespective of the mediator of these interactions.

  In our world, the fastest and most salient information carrier just happens to be the electromagnetic field ("light"). For example, to synchronize our clocks with light signals, it wasn't the light *per se* we were interested in; we just used it as carrier of information to correlate the clocks.

- Given the relativity principle SR and our derivation in Section 1.4, we showed that there are only two possibilities: (1) There is *no* upper bound on velocities (Galilean symmetry) or (2) there *is* such an upper bound $v_{\max}$ (Lorentz symmetry). In the latter case, every signal that propagates with $v_{\max}$ in some frame automatically does so in all inertial systems. (Which immediately leads to the counterintuitive conclusion, akin to SL, that there are signals the velocity of which does not depend on the velocity of the observer.)

- We could replace SL therefore by the (empirically weaker) postulate that there are *no instantaneous actions* at a distance (this is essentially a statement about *locality*). This modified postulate implies the existence of a maximal velocity $v_{\max} < \infty$ which, in turn, selects the Lorentz transformation as the correct symmetry. That $v_{\max} = c$ is then a fact to be discovered by experiments.

- It turns out that *everything with vanishing rest mass* travels at this maximum speed $v_{\max} = c$. Since photons are the only elementary particles that are massless and can be easily detected, we just happen to refer to this maximum velocity as "speed of light."

  For example: Without Higgs symmetry breaking, the $W^{\pm}$ and Z bosons of the weak interaction are massless and would propagated with light velocity, just as the photon (the weak interactions would then be no longer "weak"). For a long time it was believed that neutrinos are massless as well, and therefore would also propagate with the speed of light (today we know that they have a very tiny mass).

↓ **Lecture 5** [12.11.25]

19 | Special Lorentz transformations = Lorentz boosts:

Now that everything is settled, let us write down our final result in their conventional form.

¡! These are not the most general (homogeneous) Lorentz transformations since we omit rotations,