Special Relativity: Notation and Lorentz transformations

Mathis Habich
August 21, 2016

Prologue

This is an introduction to 4-vectors in special relativity (SR), and a derivation of Lorentz transformations which was held in the wake of SSP 2016 at the New Mexico Institute of Technology and Mining, Socorro, NM. Most of the content is based on lecture notes by Prof. Dr. Dirk H. Rischke from the Goethe University, Frankfurt, Germany [1]. I had the pleasure to attend his formidable lectures on inter alia analytical mechanics [1], quantum mechanics, and statistical mechanics. All of his notes can be found at [2]. I would also like to mention Alberto “Albi” “What ...a legend” Mosconi, who provided insightful comments and feedback for this document.

1 Position vectors in two coordinate systems

Given two inertial systems Σ and Σ', with Σ' moving relatively to Σ with a velocity $\vec{v}$. Furthermore, both inertial systems' origins coincide at $t = 0$. Now, one conducts the following experiment: At $t = 0$, a light ray is sent out at the origin of Σ (which is also the origin of Σ' at $t = 0$); thus, the position and squared position in both coordinate systems is

$$ct = |\vec{r}|\,, \tag{1}$$
$$ct' = |\vec{r}'|\,; \tag{2}$$
$$c^2t^2 = \vec{r}^2 = x^2 + y^2 + z^2\,, \tag{3}$$
$$c^2t'^2 = \vec{r}'^2 = x'^2 + y'^2 + z'^2\,. \tag{4}$$

Rearranging the squared equations and moving the spatial terms to the same side yields

$$0 = c^2t^2 - x^2 + y^2 + z^2\,, \tag{5}$$
$$0 = c^2t'^2 - x'^2 + y'^2 + z'^2\,, \tag{6}$$

or

$$0 = c^2t^2 - x^2 + y^2 + z^2 = c^2t'^2 - x'^2 + y'^2 + z'^2\,, \tag{7}$$

which resembles the invariance of the absolute square of a vector under a transformation from unprimed to primed coordinates, such as a rotation. In order to better understand this, one can think of the absolute square of a vector in Euclidean space

$$\vec{r}_E^{\,2} = \sum_{\mu=1}^{4} r_\mu^2 = r_1^2 + r_2^2 + r_3^2 + r_4^2\,. \tag{8}$$

The only difference between the absolute square in Euclidean space and for the light wave example is a relative minus sign between the first and all the other components[1]. In Euclidean space, the

[1]N.b., that this is the so-called ’minus convention’, there is also the notation with flipped signs. Ultimately, the relative minus sign is important.

absolute square of a vector can be expressed via

$$\vec{r}_E^{\,2} = \vec{r}^T\vec{r} = \vec{r}^T \mathcal{I}d\,\vec{r}\,, \tag{9}$$

which is the formal way of multiplying vectors (or also matrices), that is, that the first object's columns must match the second object's rows. Additionally, we introduced the identity matrix

$$\mathcal{I}d \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv \mathrm{diag}(1,1,1,1)\,, \tag{10}$$

which does not change the result of the scalar product; however, one can now generalize the scalar product of vectors in arbitrary spaces. For the case with the light wave, one can now write

$$c^2t^2 - x^2 + y^2 + z^2 = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{11}$$

$$= (ct,\, x,\, y,\, z) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \sum_{\mu,\nu=0}^{3} x^\mu g_{\mu\nu} x^\nu\,, \tag{12}$$

where the matrix $(g_{\mu\nu})$ was introduced that deals with the relative minus sign between the spatial and temporal components. This matrix is also called *metric tensor* and besides the minus signs, it also transposes a 4-vector[2] based on the position of the indices.

1.1 Invariance of the absolute square of a 4-vector

The traveling and expanding light wave was used as an example, yet the position vector of a light wave is *not special or distinct* from any other object; therefore, it is correct to state that the absolute square of 4-vector in different inertial systems is *always invariant.*

$$c^2t^2 - x^2 + y^2 + z^2 = c^2t'^2 - x'^2 + y'^2 + z'^2 \tag{13}$$

N.b. that this does *only* equate to zero in case of a *light wave*[3].

1.2 Minkowski space and notation

The Minkowski space has three types of objects:

1. *Contravariant* objects have only upper indices. In case of vectors, they are written as row vectors

$$(x^\mu) = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{14}$$

2. *Covariant* objects have only lower indices. In case of vectors, they are written as column

[2]Vectors in Minkowski space are called 4-vectors.

[3]Vectors with vanishing absolute square are also called *light vectors* or *null vectors.*

vectors

$$(x_\mu) \quad = \quad (ct,\, -x,\, -y,\, -z) \tag{15}$$

N.b. the minus sign is an artefact of being a covariant vector. Those vectors exist in the dual space to the covariant vector space.

3. Objects with mixed indices.

1.3 Einstein's summation convention

As one can see in eq. (12), a simple scalar product already requires two sums; therefore, one suppresses the sum signs if the same index appears twice as one covariant (lower) and one contravariant (upper) index. That is,

$$\sum_{\mu,\nu=0}^{3} x^\mu g_{\mu\nu} x^\nu \quad \equiv \quad x^\mu g_{\mu\nu} x^\nu \,. \tag{16}$$

This is Einstein's summation convention. Note that the position of the indices is crucial for this to work.

1.4 Matrix and index picture

It is often times confusing whether one is talking about a matrix or a component. Components are numbers, so their relative position to each other is completely commutative; whereas, matrices are, in general, non-commutative, that is, their order matters!
To avoid confusion between these two pictures, *superfluous parentheses around an object with indices* indicate matrix-like objects, e.g., $(g_{\mu\nu})$ is a 4×4 matrix; objects with just indices are components, thus scalars, e.g., $g_{\mu\nu}$ is a scalar.
Lastly, transitioning from one picture to another can become quite involved.

1. From components picture to matrix picture

$$a_{\mu\nu} b^\nu{}_\sigma \quad = \quad b^\nu{}_\sigma a_{\mu\nu} \tag{17}$$

where we interchanged simple components/scalars. The left-hand side of this eq. (17) allows for easy conversion into the matrix picture since the summed indices are adjacent to each other:

$$a_{\mu\nu} b^\nu{}_\sigma \quad \to \quad (a_{\mu\nu})\,(b^\nu{}_\sigma) = \begin{pmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b^0{}_0 & b^0{}_1 & b^0{}_2 & b^0{}_3 \\ b^1{}_0 & b^1{}_1 & b^1{}_2 & b^1{}_3 \\ b^2{}_0 & b^2{}_1 & b^2{}_2 & b^2{}_3 \\ b^3{}_0 & b^3{}_1 & b^3{}_2 & b^3{}_3 \end{pmatrix} \tag{18}$$

However, the right-hand side of eq. (17) does not allow for a conversion into matrices since matrices do not commute in general.

2. Transposing

$$d_{\mu\nu} e_\sigma{}^\nu \quad = \quad e_\sigma{}^\nu d_{\mu\nu} \tag{19}$$

Here, the indices are not adjacent to each other, and even rearranging does not help, but they *still imply summation*! In order to convert this expression into matrix form, one needs to swap the indices of one of the scalars, e.g., $\mu \leftrightarrow \nu$, and, thus, denote this scalar as transposed.

$$d_{\mu\nu}e_{\sigma}{}^{\nu} = d^{T}{}_{\nu\mu}e_{\sigma}{}^{\nu} = e_{\sigma}{}^{\nu}d^{T}{}_{\nu\mu} \rightarrow \left(e_{\sigma}{}^{\nu}\right)\left(d^{T}{}_{\nu\mu}\right) , \quad (20)$$

$$d_{\mu\nu}e_{\sigma}{}^{\nu} = d_{\mu\nu}e^{T\nu}{}_{\sigma} \rightarrow \left(d_{\mu\nu}\right)\left(e^{T\nu}{}_{\sigma}\right) \quad (21)$$

2 Lorentz transformations

One of the most important aspects of SR is the transformation between two inertial systems that are moving with respect to each other (as above). It was stipulated that their coordinates are different, but it is hitherto unclear what the exact form of the new primed coordinates is. However, the previous sections provided the foundation for the coming calculations. It is safe to assume that such a coordinate transformation can be expressed by a matrix $(\Lambda^{\mu}{}_{\nu})$; it is, thus, a linear transformation:

$$x'^{\mu} \equiv \Lambda^{\mu}{}_{\nu}x^{\nu} . \quad (22)$$

In order to determine the exact form of this matrix, one requires additional constraints:

1. The boost occurs only in one principal direction, e.g., $\vec{z} = v\hat{z}$;
2. The other *spatial* components should not be affected by such a boost, e.g., $x = x'$ and $y = y'$;
3. The other *spatial* components should not be mixed with the component along the boost direction and the temporal component.
4. The absolute square of a 4-vector in two different inertial systems is the same.

With the first three constraints, one can construct the following matrix for a boost in the z-direction

$$(\Lambda^{\mu}{}_{\nu}) \equiv \begin{pmatrix} \Lambda^{0}{}_{0} & 0 & 0 & \Lambda^{0}{}_{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda^{3}{}_{0} & 0 & 0 & \Lambda^{3}{}_{3} \end{pmatrix} , \quad (23)$$

where the zeros guarantee that the other two spatial entries are not mixed with the boost direction or the temporal coordinate. The identity in the center just projects the old, unprimed coordinates onto the new, primed coordinates without any changes as required by constraint 2.
The remaining constraint yields

$$x^{\sigma}g_{\sigma\mu}x^{\mu} = x_{\mu}x^{\mu} \equiv x'_{\mu}x'^{\mu} = x'^{\nu}g_{\nu\mu}x'^{\mu} = \Lambda^{\nu}{}_{\sigma}x^{\sigma}g_{\nu\mu}\Lambda^{\mu}{}_{\tau}x^{\tau} = x^{\sigma}\Lambda^{\nu}{}_{\sigma}g_{\nu\mu}\Lambda^{\mu}{}_{\tau}x^{\tau} \quad (24)$$

$$= x^{\sigma}\left(\Lambda^{T}\right)_{\sigma}{}^{\nu}g_{\nu\mu}\Lambda^{\mu}{}_{\tau}x^{\tau} , \quad (25)$$

where the invariance of the absolute square was used at the second equal sign, and then the vectors were successively expressed in terms of the old variables via the Lorentz transformations. In the last line, the known transposing allows for a conversion into matrix notation. In other words, the invariance of the absolute square requires

$$g_{\sigma\tau} = \left(\Lambda^{T}\right)_{\sigma}{}^{\nu}g_{\nu\mu}\Lambda^{\mu}{}_{\tau} . \quad (26)$$

Therefore, this multiplication in matrix notation with (23) becomes

$$g = \Lambda^{T}g\Lambda , \quad (27)$$

or with the components written out

$$\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} = \begin{pmatrix}\Lambda^0{}_0&0&0&\Lambda^3{}_0\\0&1&0&0\\0&0&1&0\\\Lambda^0{}_3&0&0&\Lambda^3{}_3\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}\Lambda^0{}_0&0&0&\Lambda^0{}_3\\0&1&0&0\\0&0&1&0\\\Lambda^3{}_0&0&0&\Lambda^3{}_3\end{pmatrix} \tag{28}$$

$$= \begin{pmatrix}\Lambda^0{}_0&0&0&-\Lambda^3{}_0\\0&-1&0&0\\0&0&-1&0\\\Lambda^0{}_3&0&0&-\Lambda^3{}_3\end{pmatrix}\begin{pmatrix}\Lambda^0{}_0&0&0&\Lambda^0{}_3\\0&1&0&0\\0&0&1&0\\\Lambda^3{}_0&0&0&\Lambda^3{}_3\end{pmatrix} \tag{29}$$

$$= \begin{pmatrix}\left(\Lambda^0{}_0\right)^2-\left(\Lambda^3{}_0\right)^2&0&0&\Lambda^0{}_0\Lambda^0{}_3-\Lambda^3{}_0\Lambda^3{}_3\\0&-1&0&0\\0&0&-1&0\\\Lambda^0{}_3\Lambda^0{}_0-\Lambda^3{}_3\Lambda^3{}_0&0&0&\left(\Lambda^0{}_3\right)^2-\left(\Lambda^3{}_3\right)^2\end{pmatrix}. \tag{30}$$

This calculation results in three more equations for the four components of the ansatz of the boost matrix Λ.

$$1 = \left(\Lambda^0{}_0\right)^2 - \left(\Lambda^3{}_0\right)\,, \tag{31}$$

$$0 = \Lambda^0{}_3\Lambda^0{}_0 - \Lambda^3{}_3\Lambda^3{}_0\,, \tag{32}$$

$$-1 = \left(\Lambda^0{}_3\right)^2 - \left(\Lambda^3{}_3\right)^2 \tag{33}$$

Ergo, the system is underdetermined. Again, by guessing an ansatz[4], one can write

$$\Lambda^0{}_0 = \Lambda^3{}_3 = \cosh\chi\,, \tag{34}$$

$$\Lambda^0{}_3 = \Lambda^3{}_0 = -\sinh\chi\,, \tag{35}$$

which obviously fulfills the system of equations; therefore, the Lorentz matrix depending on χ takes on the form

$$\Lambda = \begin{pmatrix}\cosh\chi&0&0&-\sinh\chi\\0&1&0&0\\0&0&1&0\\-\sinh\chi&0&0&\cosh\chi\end{pmatrix}. \tag{36}$$

However, the free parameter χ exists now[5]. How does the free parameter χ relate to other physical quantities, such as the velocity between both inertial systems?
To this end, writing down the transformed 4-vector in the old coordinate allows to shed some light on the connection between χ and the boost velocity v.

$$\left(x'^\mu\right) = \begin{pmatrix}ct'\\x'\\y'\\z'\end{pmatrix} = \begin{pmatrix}ct\cosh\chi - z\sinh\chi\\x\\y\\-ct\sinh\chi + z\cosh\chi\end{pmatrix} \tag{37}$$

[4] An ansatz is a mere guess which is normally informed by experience. The hyperbolic functions are chosen because of the hyperbolic equivalent of the trigonometric Pythagorean theorem : $\cosh^2\alpha - \sinh^2\alpha = 1$
More about hyperbolic functions can be found on Wikipedia [3]. In particular, the rise and run of the hyperbola are interesting comparisons to the trigonometric functions.

[5] This parameter χ is also called rapidity

With respect to the coordinate system Σ, the origin of Σ' is at $z = vt$. If the transformed vector (x'^μ) describes the origin of the moving coordinate system $z' = 0$, then

$$\begin{aligned} 0 &= z' = -ct \sinh\chi + z\cosh\chi\,, &(38)\\ \Rightarrow \tanh\chi &= \frac{z}{ct} = \frac{v}{c} \equiv \beta\,, &(39)\end{aligned}$$

where we introduced the relative velocity $\beta \equiv v/c$. Eq. (39) is the desired relationship between $v = \beta c$ and χ! Furthermore, the identity $\cosh^2\chi - \sinh^2\chi = 1$ can be employed to rewrite the hyperbolic functions as

$$\begin{aligned} \cosh\chi &= \frac{1}{\sqrt{1-\tanh^2\chi}} \equiv \frac{1}{\sqrt{1-\beta^2}} \equiv \gamma(\beta)\,, &(40)\\ \sinh\chi &= \frac{\tanh^2\chi}{1-\tanh^2\chi} = \frac{\beta}{\sqrt{1-\beta^2}} \equiv \beta\gamma(\beta)\,; &(41)\end{aligned}$$

hereby, the Lorentz gamma factor $\gamma(\beta)$ was defined. These factors are omnipresent in all of SR. Finally, the Lorentz transformation can be solely written in terms of the boost velocity $v = \beta c$, i.e.,

$$\Lambda = \begin{pmatrix} \cosh\chi & 0 & 0 & -\sinh\chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\chi & 0 & 0 & \cosh\chi \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}. \qquad (42)$$

Summary

To recapitulate the previous steps:

1. 4-vectors transform between relatively moving coordinate systems via Lorentz transformations (22):

$$x'^\mu \equiv \Lambda^\mu{}_\nu x^\nu\,. \qquad (43)$$

2. Four additional constraints pertaining the boosts' direction and the invariance of the absolute square of 4-vectors allowed for determining the boost matrix Λ in terms of χ which is called the *rapidity.*

3. By looking at the origin of the boosted coordinate system Σ', one can establish a mathematical relationship between the *rapditity* χ and the *boost velocity* $v = \beta c$.

4. Since the expressions for the boost matrix occur very often, the following short-hand notations were introduced:

$$\begin{aligned} \beta &\equiv \frac{v}{c}\,, &(44)\\ \gamma &\equiv \frac{1}{\sqrt{1-\beta^2}}\,. &(45)\end{aligned}$$

5. Note that γ diverges as a v approaches the speed of light c! Later you will see that massive objects (non-zero mass) cannot travel at or faster than the speed of light because of this γ factor in the framework of SR.

6. With this relationship, one can express the boost matrix in terms of β and γ. All boost matrices, including boosts in x- and y-direction, are

$$\Lambda_x(\beta) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{46}$$

$$\Lambda_y(\beta) = \begin{pmatrix} \gamma & 0 & -\beta\gamma & 0 \\ 0 & 1 & 0 & 0 \\ -\beta\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{47}$$

$$\Lambda_z(\beta) = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{48}$$

7. For example, the primed 4-vector for a boost along the z-axis expressed in the old coordinates is

$$X' = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma ct - \beta\gamma z \\ x \\ y \\ -\beta\gamma ct + \gamma z \end{pmatrix} \tag{49}$$

3 Exercises

1. Time dilation: a time interval is measured in two coordinate systems that are moving relatively with velocity v along the y-axis of Σ. The measurement is performed at the same spatial location.
   - In Σ, the initial time is t_1 and the final time is t_2. What is the (trivial) time interval?
   - Perform a Lorentz transformation on a 4-vector at two different times, but same positions. What is the time interval now? Is it smaller or larger than the time interval in Σ?

2. Length contraction: now a distance is measured in two relatively moving coordinate systems at the same time in each coordinate system. The end points of the measurement are at rest w.r.t. Σ.
   - What is the trivial measurement of length in Σ?
   - Now, enforce the condition that the measurement in Σ' occurs at the same time, and use this information to find the difference between the transformed spatial coordinates.

4 Solutions

1. Time dilation: In Σ, the measurement of a time interval[6] is simply

$$t_2 - t_1 \equiv \Delta t \,. \tag{50}$$

In Σ', both components transform to

$$t_2' - t_1' = \left(\gamma t_2 - \gamma\beta\frac{y}{c}\right) - \left(\gamma t_1 - \gamma\beta\frac{y}{c}\right) = \gamma(t_2 - t_1) = \gamma\Delta t > \Delta t \,. \tag{51}$$

Thus, time passes on more slowly in the moving coordinate system Σ'.

2. Length contraction: In Σ, the measurement of a distance (which is at rest) is simply

$$y_2 - y_1 \equiv \Delta y \,. \tag{52}$$

In Σ', the measurement shall also occur at the same time[7], $t_1' = t_2'$.

$$t_2' = \gamma t_2 - \beta\gamma\frac{y_2}{c} \overset{!}{=} t_1' = \gamma t_1 - \beta\gamma\frac{y_1}{c} \,, \tag{53}$$

$$\Rightarrow \quad t_2 - t_1 = \frac{\beta}{c}\left(y_2 - y_1\right) \,. \tag{54}$$

Interestingly, the measurement in Σ' occurs at different times $t_1 \neq t_2$ in Σ. Therefore, the length measurement with this constraint yields

$$y_2' - y_1' = \left(-\beta\gamma c t_2 + \gamma y_2\right) - \left(-\beta\gamma c t_1 + \gamma y_1\right) = -\beta\gamma\frac{\beta}{c}\left(y_2 - y_1\right) + \gamma\left(y_2 - y_1\right) \tag{55}$$

$$= \Delta y\gamma\left(1 - \beta^2\right) = \Delta y\gamma\gamma^{-2} = \frac{\Delta y}{\gamma} < \Delta y \tag{56}$$

Thus, the length in the moving coordinate system Σ' is less than the length in the resting coordiante system Σ.

References

[1] Dirk H. Rischke. *Theoretische Physik II: Analytische Mechanik und Spezielle Relativitätstheorie*. http://th.physik.uni-frankfurt.de/~drischke/. 2010. URL: http://th.physik.uni-frankfurt.de/~drischke/Skript_Mechanik_II.pdf.

[2] Dirk H. Rischke. *Lecture notes*. 2010-2012. URL: http://th.physik.uni-frankfurt.de/~drischke/.

[3] Wikipedia. *Hyperbolic function*. URL: https://en.wikipedia.org/wiki/Hyperbolic_function.

[6] E.g., a clock which is at rest w.r.t. Σ.

[7] This is a crucial constraint; otherwise, one of the end points travels farther while the measurement is taking place.