Introduction to

Covariant Formulation

by

Gerhard Kristensson

April 1981 (typed and with additions June 2013)

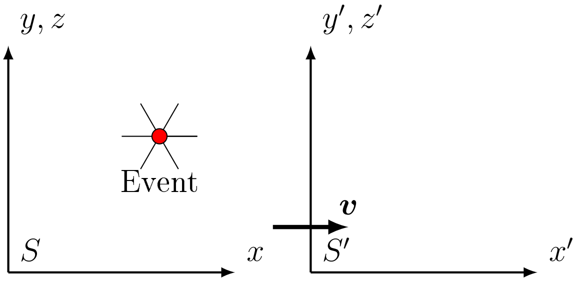

Figure 1: The two coordinate systems S and S'.

1 Introduction and the Lorentz' transformation

An event in spacetime is characterized by the coordinates (t, x, y, z) and (t', x', y', z') as seen in two different frames of reference S and S', respectively. These reference frames are also named **inertial reference frames**.[1]

The coordinates are related by the Lorentz' transformation [1], see Figure 1.[2]

$$\begin{cases} t' = (t - vx/c^2)\gamma \\ x' = (x - vt)\gamma \\ y' = y \\ z' = z \end{cases} \tag{1.1}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

Introduce the **contravariant coordinate vector** $x^\mu = (x^0, x^1, x^2, x^3)$, where

$$\begin{cases} x^0 = ct \\ x^1 = x \\ x^2 = y \\ x^3 = z \end{cases}$$

Define the **metric tensor** (matrix)

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{matrix} \mu = 0 \\ \mu = 1 \\ \mu = 2 \\ \mu = 3 \end{matrix}$$

[1] In an inertial reference frames (or inertial frame of reference) a particle moves along a straight line in the absence of acting forces.

[2] This transformation is based on two fundamental postulates, *i.e.*,

1. All laws of nature should have the same form in all frames of inertia
2. The speed of light in vacuum is finite and independent of the frame of inertia

and the **covariant coordinate vector**, defined by

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu$$

or $x_\mu = (x^0, -x^1, -x^2, -x^3) = (ct, -x, -y, -z)$.

It is convenient to introduce and use the Einstein's summation rule or convention.

$$\sum_{\nu=0}^{3} g_{\mu\nu} x^\nu = g_{\mu\nu} x^\nu$$

This means that repeated indices indicate a summation. Note that one index in the summation is a superscript and one is a subscript.

In particular,

$$\begin{aligned} x^\mu x_\mu &= x^\mu g_{\mu\nu} x^\nu = g_{\mu\nu} x^\mu x^\nu \\ &= (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (ct)^2 - x^2 - y^2 - z^2 = (ct)^2 - r^2 \end{aligned}$$

We are now in a position to write the Lorentz' transformation in a very compact form.

$$x'^\mu = L^\mu{}_\nu x^\nu \tag{1.2}$$

where, for a transformation along the x axis ($\boldsymbol{v} = v\hat{\boldsymbol{e}}_x$), $L^\mu{}_\nu$ is

$$L^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.3}$$

since in components the transformation (1.2) becomes

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

or

$$\begin{cases} ct' = x'^0 = \gamma x^0 - \gamma x^1 v/c = \gamma(ct - xv/c) \\ x' = x'^1 = -\gamma x^0 v/c + \gamma x^1 = \gamma(x - ct) \\ y' = x'^2 = x^2 = y \\ z' = x'^3 = x^3 = z \end{cases}$$

which is identical to (1.1).

Note: If $\boldsymbol{v}$ is not parallel to the x axis, the relation (1.2) still holds, but the tensor (matrix) (1.3) is more complex. For a general direction of the velocity $\boldsymbol{v}$, the tensor $L^\mu{}_\nu$ is "filled".

2 Invariance of interval

We study two different events in the two reference frames S and S'. Orient the reference frames such that the first event occurs when the two frames of reference coincide (the origins in the two frames coincide). The first event then is $x^\mu = (0,0,0,0)$ and $x'^\mu = (0,0,0,0)$, respectively. The second event, we denote $x^\mu = (x^0, x^1, x^2, x^3)$ and $x'^\mu = (x'^0, x'^1, x'^2, x'^3)$ in the two frames, respectively. The Lorentz' transformation (1.2) relates the coordinates of the two events, *i.e.*, $x'^\mu = L^\mu{}_\nu x^\nu$. Moreover, we notice that[3]

$$x'^\mu x'_\mu = g_{\mu\nu} x'^\mu x'^\nu = g_{\mu\nu} L^\mu{}_\alpha x^\alpha L^\nu{}_\beta x^\beta = g_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta x^\alpha x^\beta$$

Compute $g_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta = L^\mu{}_\alpha g_{\mu\nu} L^\nu{}_\beta$. For a translation along the x axis, we get

$$\begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^t \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where ^{t} stands for the transpose of the matrix (rows and columns change places). Multiplication of the matrices implies:

$$\begin{aligned} &\begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ \gamma v/c & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ = &\begin{pmatrix} \gamma^2 - \gamma^2 v^2/c^2 & -\gamma^2 v/c + \gamma^2 v/c & 0 & 0 \\ -\gamma^2 v/c + \gamma^2 v/c & \gamma^2 v^2/c^2 - \gamma^2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

We get

$$x'^\mu x'_\mu = \underbrace{g_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta}_{g_{\alpha\beta}} x^\alpha x^\beta = g_{\alpha\beta} x^\alpha x^\beta = x^\alpha x_\alpha = x^\mu x_\mu$$

Definition: The interval ("distance" in space-time) between two events is defines as $s^2 = x^\mu x_\mu$.

The interval s^2 between two events is independent of which frame of reference it is evaluated in. The interval is **invariant**.

Note: Compare the distance between two points in three dimensions. This distance is independent of which frame of reference it is evaluated in, *i.e.*, $\boldsymbol{r} \cdot \boldsymbol{r}$, where $\boldsymbol{r}$ is the relative vector between the points.

[3]It is really the difference between the two events that is used, but since the first event is at the origin, the difference coincide with x^μ and x'^μ, respectively.

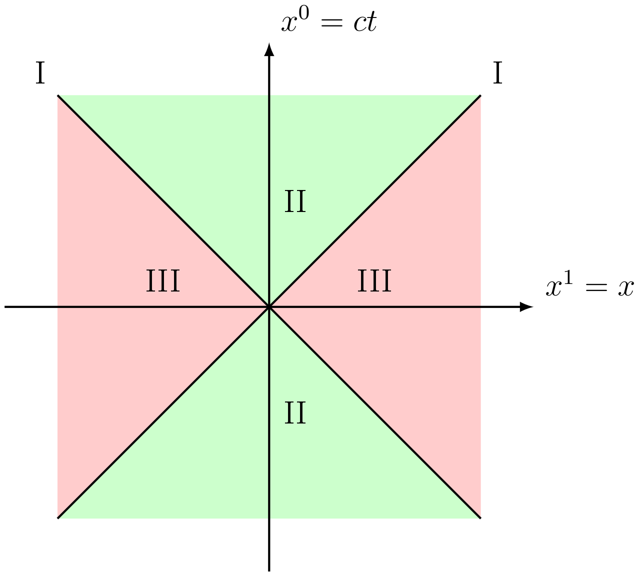

Figure 2: The light cone.

3 Light cone

The interval $s^2 = x^\mu x_\mu$ acts like a scalar product in the four dimensional spacetime. We can use $s^2 = x^\mu x_\mu$ to classify two events separated by the four-vector x^μ.

I) $s^2 = 0$, the interval in a **lightlike** separation

II) $s^2 > 0$, the interval in a **timelike** separation

III) $s^2 < 0$, the interval in a **spacelike** separation

In general

$$x^\mu x_\mu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (ct)^2 - x^2 - y^2 - z^2 = (ct)^2 - r^2$$

Space-time is divided in three different areas, see Figure 2. A lightlike interval (I) defines the **light cone**.

1. Two events that are separated by a lightlike interval can be connected by a light signal

2. Two events that are separated by a timelike interval can be connected by a signal with a speed $v < c$

3. Two events that are separated by a spacelike interval cannot be connected by a signal with a speed $v < c$

Events in the last category are not causal (the effect precedes the cause i some frame of reference S). Events in the backward cone, see Figure 3, affect an event at the origin causally. Similarly, an event at the origin affects an event in the forward cone causally.

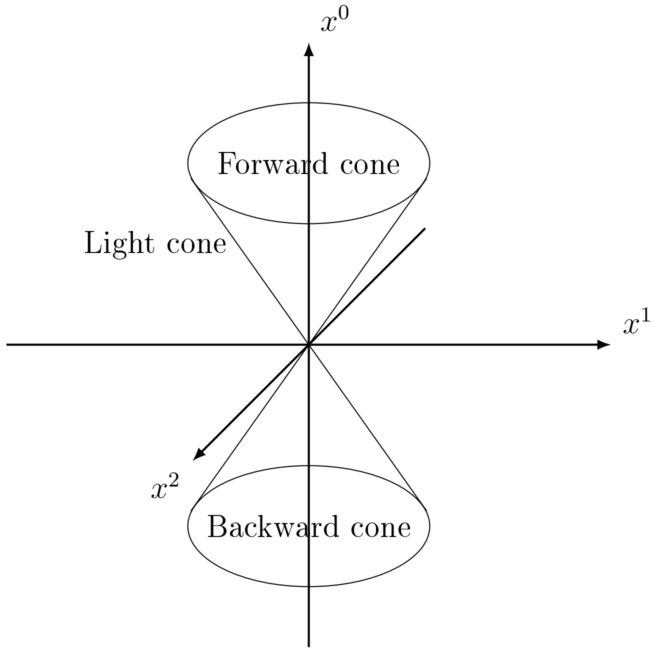

Figure 3: The light cone.

A final question in this section: What is the speed u of the signal that connects two events that are related causally? The interval is

$$\begin{aligned} s^2 &= (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (ct)^2 - r^2 \\ &= (ct)^2(1 - r^2/(ct)^2) = (ct)^2(1 - u^2/c^2) > 0 \end{aligned}$$

since $u = r/t$.

Define the **proper time** $\tau = t\sqrt{1 - u^2/c^2}$ for a particle. The proper time is the time an observer in the rest system (frame of reference where the particle is at rest) measures $s^2 = c^2\tau^2$.

4 Covariant formulation of linear momentum and energy

The relativistic momentum $\boldsymbol{p}$ and the energy E are defined as

$$\begin{cases} \boldsymbol{p} = m\boldsymbol{u} = \dfrac{m_0\boldsymbol{u}}{\sqrt{1 - u^2/c^2}} \\ E = mc^2 = \dfrac{m_0c^2}{\sqrt{1 - u^2/c^2}} \end{cases}$$

where $\boldsymbol{u}$ is the velocity of the particle and its **rest mass** is m_0. We have also have

$$p^2 - E^2/c^2 = -m_0c^2 = \text{constant, independent of the frame of reference} \qquad (4.1)$$

We define the contravariant momentum (4-vector) $p^\mu = (p^0, p^1, p^2, p^3)$ where

$$\begin{cases} p^0 = m/c = \dfrac{m_0 c}{\sqrt{1-u^2/c^2}} \\ p^1 = m u_x = \dfrac{m_0 u_x}{\sqrt{1-u^2/c^2}} \\ p^2 = m u_y = \dfrac{m_0 u_y}{\sqrt{1-u^2/c^2}} \\ p^3 = m u_z = \dfrac{m_0 u_z}{\sqrt{1-u^2/c^2}} \end{cases}$$

In analogy to $s^2 = x^\mu x_\mu$, we calculate

$$p^\mu p_\mu = g_{\mu\nu} p^\nu p^\mu = (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = E^2/c^2 - p^2 = m_0^2 c^2$$

where we also used (4.1). Again, we see that $p^\mu p_\mu$ acts as a scalar product in the four dimensional space-time, in complete analogy with what we found for the vector $x^\mu x_\mu = s^2$.

This indicates that the 4-vector, p^μ, transforms between two frames of reference S and S', the same way as x^μ, *i.e.*,

$$\begin{cases} x'^\mu = L^\mu{}_\nu x^\nu \\ p'^\mu = L^\mu{}_\nu p^\nu \end{cases}$$

Explicitly, with $\boldsymbol{v} = v\hat{\boldsymbol{e}}_x$

$$\begin{cases} p'^0 = \gamma p^0 - \gamma p^1 v/c \\ p'^1 = -\gamma p^0 v/c + \gamma p^1 \\ p'^2 = p^2 \\ p'^3 = p^3 \end{cases}$$

or

$$\begin{cases} E'/c = \gamma(E/c - p_x v/c) \\ p'_x = \gamma(p_x - vE/c^2) \\ p'_y = p_2 \\ p'_3 = p_3 \end{cases}$$

Note: The momentum $\boldsymbol{p}$ and the energy E are now covariantly connected in the 4-vector p^μ in the same way as space-time is connected in the vector x^μ.

The importance of a covariant formulation is now obvious:

If energy and linear momentum are conserved quantities for one observer in S, then another observer in S' — connected by a Lorentz's transformation — also finds these quantities conserved.

5 The 4-force F^μ

We also have the possibility to define a relativistic force concept by defining a 4-vector F^μ by

$$F^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau}$$

Notice that in this definition the derivative is taken w.r.t. the proper time τ and not the local time t. The reason for this is that the proper time $\tau = t\sqrt{1 - u^2/c^2}$ is an invariant quantity (same value $\tau = s/c$ in all inertial systems). In component form the force is

$$\begin{cases} F^0 = \dfrac{\mathrm{d}p^0}{\mathrm{d}\tau} = \dfrac{\mathrm{d}}{\mathrm{d}\tau}(E/c) = \dfrac{\mathrm{d}E}{\mathrm{d}t}\dfrac{1}{c\sqrt{1-u^2/c^2}} \\ F^1 = \dfrac{\mathrm{d}p^1}{\mathrm{d}\tau} = \dfrac{\mathrm{d}p_x}{\mathrm{d}t}\dfrac{1}{\sqrt{1-u^2/c^2}} \\ F^2 = \dfrac{\mathrm{d}p^2}{\mathrm{d}\tau} = \dfrac{\mathrm{d}p_y}{\mathrm{d}t}\dfrac{1}{\sqrt{1-u^2/c^2}} \\ F^3 = \dfrac{\mathrm{d}p^3}{\mathrm{d}\tau} = \dfrac{\mathrm{d}p_z}{\mathrm{d}t}\dfrac{1}{\sqrt{1-u^2/c^2}} \end{cases}$$

If we introduce the force $\boldsymbol{f} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}$ and $\frac{\mathrm{d}E}{\mathrm{d}t} = \boldsymbol{f} \cdot \boldsymbol{u}$, we get

$$F^\mu = (\boldsymbol{f} \cdot \boldsymbol{u}/c, f_x, f_y, f_z)\frac{1}{\sqrt{1-u^2/c^2}}$$

and the 4-vector transforms in the same way as p^μ and x^μ under a Lorentz transformation, *i.e.*,

$$F'^\mu = L^\mu{}_\nu F^\nu \tag{5.1}$$

5.1 Action reaction

In non-relativistic mechanics we have learned that two bodies attract each other with equally large and oppositely directed forces. This the Newton's third law and it is a consequence of the conservation of momentum. We now investigate this case in a relativistic setting.

Assume two particles 1 and 2 in the unprimed frame of reference, with force and velocities, see also Figure 4

$$\begin{cases} \boldsymbol{f}_1 = (0, 0, f) \\ \boldsymbol{f}_2 = (0, 0, -f) \end{cases} \qquad \begin{cases} \boldsymbol{u}_1 = (0, 0, 0f) \\ \boldsymbol{u}_2 = (0, 0, u_2) \end{cases}$$

The 4-vectors in S are

$$\begin{cases} F_1^\mu = (0, 0, 0, f) \\ F_2^\mu = (-fu_2/c, 0, 0, -f)\dfrac{1}{\sqrt{1-u_2^2/c^2}} \end{cases}$$

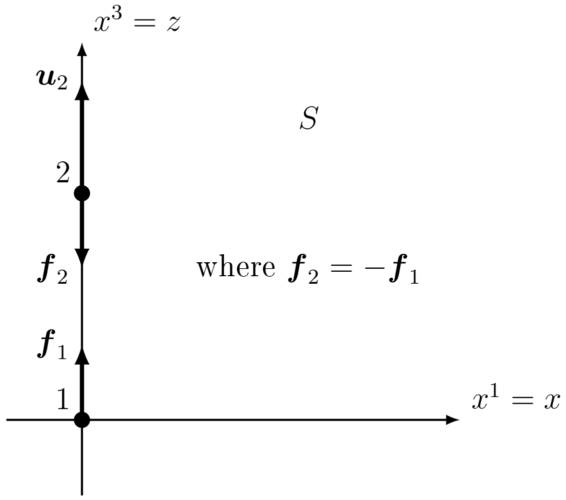

Figure 4: The forces and velocities of the two particles.

In another frame of reference (frame of inertia) S' ($\boldsymbol{v} = v\hat{\boldsymbol{e}}_x$) the 4-force is F'^{μ}. Use (1.3) and (5.1)

$$\begin{cases} F'^0 = \gamma F^0 - \gamma F^1 v/c \\ F'^1 = -\gamma F^0 v/c + \gamma F^1 \\ F'^2 = p^2 \\ F'^3 = p^3 \end{cases}$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$. We get

$$\begin{cases} F'^{\mu}_1 = (0, 0, 0, f) \\ F'^{\mu}_2 = (-\gamma f u_2/c, \gamma v u_2 f/c^2, 0, -f)\dfrac{1}{\sqrt{1 - u_2^2/c^2}} \end{cases}$$

The velocities $\boldsymbol{u}_1 = \boldsymbol{0}$ and $\boldsymbol{u}_2$ transform to the primed frame of reference [1, (11.31)]

$$\begin{cases} \boldsymbol{u}'_1 = (-v, 0, 0) \\ \boldsymbol{u}'_2 = (-v, 0, u_2\sqrt{1 - v^2/c^2}) \end{cases}$$

which implies the magnitudes

$$\begin{cases} u'_1 = v \\ u'_2 = \sqrt{(v^2 + u_2^2(1 - v^2/c^2))} \end{cases}$$

and

$$\begin{aligned} 1 - {u'_2}^2/c^2 &= 1 - (v^2 + u_2^2(1 - v^2/c^2))/c^2 = 1 - u_2^2/c^2 - v^2/c^2(1 - u_2^2/c^2) \\ &= (1 - u_2^2/c^2)(1 - v^2/c^2) \end{aligned}$$

The force in S' is obtained from F'^{μ}

$$F'^{\mu} = (\boldsymbol{f}' \cdot \boldsymbol{u}'/c, f'_x, f'_y, f'_z)\frac{1}{\sqrt{1-u'^2/c^2}}$$

Explicitly, we get

$$\begin{cases} \boldsymbol{f}'_1 = (0,0,f)\sqrt{1-{u'_1}^2/c^2} = (0,0,f\sqrt{1-v^2/c^2}) \\ \boldsymbol{f}'_2 = (\gamma v u_2 f/c^2, 0, -f)\dfrac{\sqrt{1-{u'_2}^2/c^2}}{\sqrt{1-u_2^2/c^2}} = (v u_2 f/c^2, 0, -f\sqrt{1-v^2/c^2}) \end{cases}$$

An additional component of the force in the x direction appears in the primed frame of reference. The forces $\boldsymbol{f}'_1$ and $\boldsymbol{f}'_2$ are no longer equal and oppositely directed in S'. We conclude that the law of conservation of momentum is more fundamental than the Newton"s third law.

6 Electromagnetic field

The electric field $\boldsymbol{E}$ and the magnetic flux density $\boldsymbol{B}$ do not transform as 4-vectors. Instead they transform as a **second-rank, antisymmetric tensor** $F^{\mu\nu}$ — the field-strength tensor, see [1]. This tensor is defined as

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}$$

and the fields transform as, *cf.*, the transformation of x^{μ} in (1.2)

$$F'^{\mu\nu} = L^{\mu}{}_{\alpha} L^{\nu}{}_{\beta} F^{\alpha\beta}$$

In terms of their components ($\boldsymbol{v} = v\hat{\boldsymbol{e}}_x$) we get

$$\begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^t$$

$$= \begin{pmatrix} 0 & -\gamma^2 E_x(1-v^2/c^2) & -\gamma E_y + \gamma B_z v/c & -\gamma E_z - \gamma B_y v/c \\ \gamma^2 E_x(1-v^2/c^2) & 0 & -\gamma B_z + \gamma E_y v/c & \gamma B_y + \gamma E_z v/c \\ \gamma E_y - \gamma B_z v/c & \gamma B_z - \gamma E_y v/c & 0 & -B_x \\ \gamma E_z + \gamma B_y v/c & -\gamma B_y - \gamma E_z v/c & B_x & 0 \end{pmatrix}$$

from which we identity the components

$$\begin{cases} \boldsymbol{E}' = (E_x, \gamma E_y - \gamma B_z v/c, \gamma E_z + \gamma B_y v/c) \\ \boldsymbol{B}' = (B_x, \gamma B_y + \gamma E_z v/c, \gamma B_z - \gamma E_y v/c) \end{cases}$$

References

[1] J. D. Jackson. *Classical Electrodynamics.* John Wiley & Sons, New York, third edition, 1999.