

Special Theory of Relativity

for

Electrical Engineers

skimu@me.com

October 18, 2026

Abstract

This is my study note on special relativity. I started this note with the hope that it will eventually become part of my training course on the foundations of electronics for analog design engineers. To that end, we follow the notation used in *The Feynman Lectures on Physics* (FLP) [1], as it serves as a lifelong companion for electrical engineers. For example, we use $(a_t, \mathbf{a}) \cdot (b_t, \mathbf{b}) = a_t b_t - \mathbf{a} \cdot \mathbf{b}$ as the definition of the inner product of two four-vectors. We chose the SI unit system, but use ε_0 and c instead of ε_0 and μ_0 for electromagnetic constants. [2] Additionally, we refrain from using tensor component expressions as much as possible, opting instead for ordinary vectors, as they are more intuitive and convenient for solving our electrical engineering problems.

This note is not for first-time learners of relativity, but it can be used as a quick review or recap for those who have some exposure to the subject, or as a supplementary material for things you may have missed at school.

Acknowledgment I acknowledge the use of Google AI mode (80%) and ChatGPT (20%) to check the grammar of the text.

Contents

1	Special Theory of Relativity	4
1.1	Spacetime	4
1.1.1	Interval	4
1.1.2	Proper time	5
1.1.3	Four-vectors	6
1.2	Relativistic Mechanics	10
1.2.1	Four-velocity	10
1.2.2	Relativistic energy and momentum	10
1.2.3	Relativistic equation of motion and four-force	14
1.2.4	Force tensor	15
1.2.5	Angular momentum	16
1.2.6	Summary	18
1.3	Lorentz Transformation	19
1.3.1	Lorentz group	19
1.3.2	Inner product	20
1.3.3	Velocity addition law	21
1.3.4	Velocity	22
1.3.5	Force	22
1.3.6	Force tensor	24
1.3.7	Four-gradient	24
1.3.8	Thomas precession	25
	Appendices	31
	Appendix A Derivation of (1.13)	31
	Appendix B Lorentz Transformation	32
B.1	Wave equation	32
B.2	Constancy of the speed of light	33
B.3	Rotation analogy	34
B.4	Symmetry and transformation group	35
B.4.1	Introduction	35
B.4.2	Derivation	35
B.5	Mermin's derivation of velocity addition law	40
B.6	Vector notation	40
	Appendix C Details of (1.50)	42

Appendix D Vectors and tensors	43
D.1 Vectors	43
D.2 Tensors	43
D.3 Anti-symmetric tensor	45
D.3.1 Two ordinary vectors	45
D.3.2 Linear map	46
D.3.3 Lorentz transformation	47
D.3.4 Dual tensor	48
D.3.5 Scalar	48

1 Special Theory of Relativity

1.1 Spacetime

In this section, we introduce basic concepts such as spacetime, interval, proper time, and four-vector. Readers are expected to be familiar with the concept of inertial frame.

1.1.1 Interval

Experience tells us that the speed of light is the same regardless of the inertial frame it is measured. Consider a light ray emitted from a point P_1 and is arriving at another point P_2 . Suppose that we observe it from two inertial frames K and K' . Let's say the origin of K' is moving with velocity $\mathbf{V}$ relative to K . The constancy of the speed of light can then be expressed as follows:

$$(c \Delta t)^2 - (\Delta \mathbf{r})^2 = (c \Delta t')^2 - (\Delta \mathbf{r}')^2 = 0 \quad \text{for any } \mathbf{V}, \quad (1.1)$$

where Δt and $\Delta \mathbf{r}$ are, respectively, the elapsed time and displacement vector from P_1 to P_2 measured in K . $\Delta t'$ and $\Delta \mathbf{r}'$ are the corresponding quantities measured in K' .

Now, we define spacetime interval Δs between two points and times as

$$(\Delta s)^2 = (c \Delta t)^2 - (\Delta \mathbf{r})^2.$$

Let us show that this interval $(\Delta s)^2$ is independent of inertial frame in which it is measured, in other words, $(\Delta s)^2$ is an invariant. According to (1.1), if $(\Delta s)^2 = 0$ in one inertial frame, then $(\Delta s')^2 = 0$ in another. This suggests that the interval transforms according to

$$(\Delta s')^2 = A (\Delta s)^2,$$

where factor A is a function of the magnitude of the relative velocity $\mathbf{V}$ only, because the isotropy of space requires that A cannot depend on the direction of $\mathbf{V}$. Similarly, the homogeneity of space and time requires that A cannot depend on the locations of the two points or the times. Therefore, we can write

$$(\Delta s')^2 = A(V^2) (\Delta s)^2.$$

When $V = 0$, the two inertial frames coincide. Therefore

$$A(0) = 1. \quad (1.2)$$

Now consider a third inertial frame K'' is moving at $-\mathbf{V}$ relative to K' , i.e., K'' is identical to K . Applying the transformation twice, we get

$$(\Delta s'')^2 = A(V^2)(\Delta s')^2 = A(V^2)^2 (\Delta s)^2.$$

Since K'' and K are identical, $A(V^2)^2$ must be unity. Together with (1.2), we find $A(V^2) = 1$. Therefore spacetime interval $(\Delta s)^2$ is an invariant:

$$(\Delta s)^2 = (c \Delta t)^2 - (\Delta \mathbf{r})^2 = (c \Delta t')^2 - (\Delta \mathbf{r}')^2 \quad \text{for any } \mathbf{V}.$$

1.1.2 Proper time

Let us consider the motion of a mass-point. Because its velocity may vary with time, we examine an infinitesimal displacement $d\mathbf{r}$ during an infinitesimal time interval dt . The differential interval ds at a given instant t is

$$ds = \sqrt{(c dt)^2 - (d\mathbf{r})^2} = c dt \sqrt{1 - (v/c)^2},$$

where

$$v^2 = |\mathbf{v}|^2, \quad \mathbf{v} = \frac{d\mathbf{r}}{dt}.$$

Suppose an inertial frame K' moves at the same instantaneous velocity $\mathbf{v}$ as the mass-point relative to K at an instant t . In K' , the velocity of the mass-point is zero, $v' = 0$. Therefore, recalling the invariance of the interval, we find

$$ds = c dt \sqrt{1 - (v/c)^2} = c dt' \sqrt{1 - (v'/c)^2} = c dt'. \quad (1.3)$$

For a mass-point traveling along a path p from P_1 at t_1 to P_2 at t_2 , the elapsed time measured by a clock traveling with the mass-point is

$$t'_2 - t'_1 = \int_p ds/c,$$

This elapsed time is called the proper time of the mass-point.

From (1.3), we find

$$\frac{dt'}{dt} = \sqrt{1 - (v/c)^2}. \quad (1.4)$$

Since the right hand side is less than unity, the proper time “felt” by a moving mass-point is less than the elapsed time measured in K , or by the observer. This is known as time dilation. In the non-relativistic limit of $c \rightarrow \infty$, we get $dt' = dt$, as expected.

Example Let's say $T(0)$ is the lifetime of a particle at rest ($v = 0$), the apparent lifetime $T(v)$ of the particle moving at speed v will be

$$T(v) = dt = \frac{dt'}{\sqrt{1 - (v/c)^2}} = \frac{T(0)}{\sqrt{1 - (v/c)^2}}.$$

A muon¹ at rest has a mean life of $T(0) = 2.2 \times 10^{-6}$ seconds. Cosmic rays hitting nitrogen and oxygen nuclei make pions high in the Earth's atmosphere. The pions rapidly decay into muons in 2.6×10^{-8} sec. A muon moving at the speed of light from

¹This example is taken from “Physical Mathematics” by Kevin Cahill [3]

10 km takes at least $t = 10 \text{ km} / 300,000 \text{ km/sec} = 3.3 \times 10^{-5} \text{ sec}$ to hit the ground. Were it not for time dilation, the probability P of such a muon reaching the ground as a muon would be

$$P = e^{t/T(0)} = \exp(-33/2.2) = 2.6 \times 10^{-7}.$$

A muon moving at a speed of $v = 0.99c$ has an apparent mean life $T(v)$

$$T(v) = \frac{T(0)}{\sqrt{1 - (v/c)^2}} = \frac{2.2 \times 10^{-6}}{\sqrt{1 - (0.99)^2}} = 1.6 \times 10^{-5} \text{ sec.}$$

The probability of survival with time dilation is

$$P = e^{t/T(v)} = \exp(-33/16) = 0.12$$

so that 12% survive. Time dilation increases the chance of survival.

1.1.3 Four-vectors

Just as the magnitude of an ordinary vector is invariant under spatial coordinate transformations such as rotation or translation, we can think of four-dimensional vector $(a_t, \mathbf{a})$ of which magnitude $(a_t)^2 - (\mathbf{a})^2$ is invariant under certain coordinate transformations. We call such coordinate transformations Lorentz transformations, the four-dimensional vector space the spacetime, and any quantities invariant under Lorentz transformations a Lorentz invariant. When physical laws are expressed in terms of ordinary vectors, their covariance under rotation or translation is apparent. Our goal is to express our physical laws in terms of four-vectors so that their covariance under Lorentz transformations is likewise apparent. We will take deeper look at Lorentz transformations later in Section 1.3 and Appendix B.

The first example of such four-vector is spacetime displacement $d\mathbf{r} = (c dt, d\mathbf{r})$. The magnitude of $d\mathbf{r}$ is the interval: $(ds)^2 \equiv d\mathbf{r}^2 = (dt)^2 - (d\mathbf{r})^2$.

As with ordinary vectors, the inner product of two four-vectors is also a Lorentz invariant, or a scalar. To see this, let $\mathbf{a}$ and $\mathbf{b}$ be four-vectors, the sum $(\mathbf{a} + \mathbf{b})$ is a four-vector and hence its magnitude $(\mathbf{a} + \mathbf{b})^2$ is a scalar:

$$(\mathbf{a} + \mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2 + 2\mathbf{a} \cdot \mathbf{b}. \quad (\text{Scalar})$$

Since $\mathbf{a}^2$ and $\mathbf{b}^2$ are also scalars, $\mathbf{a} \cdot \mathbf{b}$ must be a scalar. Therefore

$$\mathbf{a} \cdot \mathbf{b} = (a_t, \mathbf{a}) \cdot (b_t, \mathbf{b}) = a_t b_t - \mathbf{a} \cdot \mathbf{b}. \quad (\text{Scalar})$$

We will derive this in a different way later in Section 1.3.

Consider a scalar function $f(t, \mathbf{r})$. Recalling that the total differential df of a scalar function $f(t, \mathbf{r})$ is also a scalar. That is

$$df = \frac{\partial f}{\partial t} dt + \nabla f \cdot d\mathbf{r}. \quad (\text{Scalar})$$

This can be viewed as the inner product of a displacement four-vector $d\mathbf{r} = (c dt, d\mathbf{r})$ and the four-gradient ∂ :

$$df = \partial f \cdot d\mathbf{r} \quad \text{where} \quad \partial = \left(\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right).$$

Similarly, the continuity of charge,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0,$$

can be viewed as inner product of the four-gradient ∂ and the four-current $\mathbf{j}$:

$$\partial \cdot \mathbf{j} = 0 \quad \text{where} \quad \mathbf{j} = (c\rho, \mathbf{j}).$$

The magnitude of the four-gradient is d'Alembertian:

$$\square^2 = \partial \cdot \partial = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2.$$

Just as ordinary vector space, there are spacetime version of tensors, four-tensors. Appendix D gives some more descriptions about four-vectors and four-tensors.

Quick demo Let me show you four-vector alone will take you to something resembles Maxwell's equations. First, we know that following Poisson's equation is a direct consequence of the Coulomb's law.

$$-\nabla^2 \phi = \rho / \varepsilon_0$$

To make this Lorentz invariant, we need to replace Laplacian with d'Alembertian:

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = \frac{\rho}{\varepsilon_0}. \quad (1.5)$$

Since $c\rho$ is the temporal component of the four-current, we should have the same equation for the spatial component. That means we need a spatial component of the potential. Let's say $c\mathbf{A}$ is such a component, i.e., $c\mathbf{A} = (\phi, c\mathbf{A})$ is a four-vector. Then the equation for the spatial component will be

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) c\mathbf{A} = \frac{\mathbf{j}}{\varepsilon_0 c}. \quad (1.6)$$

The four-gradient of the four-potential, $\partial \cdot c\mathbf{A}$, is a scalar function. Let's say it is f :

$$\partial \cdot c\mathbf{A} = \frac{1}{c} \frac{\partial \phi}{\partial t} + \nabla \cdot c\mathbf{A} = f. \quad (1.7)$$

Plugging

$$\nabla^2 \mathbf{A} = \nabla \nabla \cdot \mathbf{A} - \nabla \times \nabla \times \mathbf{A}$$

into (1.6) and using (1.7), we find

$$\frac{1}{c} \frac{\partial}{\partial t} \left(\nabla \phi + \frac{1}{c} \frac{\partial c\mathbf{A}}{\partial t} \right) + \nabla \times \nabla \times c\mathbf{A} - \nabla f = \frac{\mathbf{j}}{\varepsilon_0 c}. \quad (1.8)$$

From (1.5) and (1.7), we find

$$\nabla \cdot \left(-\nabla \phi - \frac{1}{c} \frac{\partial c\mathbf{A}}{\partial t} \right) + \frac{1}{c} \frac{\partial f}{\partial t} = \frac{\rho}{\varepsilon_0}. \quad (1.9)$$

If we define $\mathbf{E}$ and $c\mathbf{B}$ as

$$\mathbf{E} = -\nabla \phi - \frac{1}{c} \frac{\partial c\mathbf{A}}{\partial t}, \quad c\mathbf{B} = \nabla \times c\mathbf{A}, \quad (1.10)$$

(1.8) and (1.9) becomes

$$\nabla \times c\mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{j}}{\varepsilon_0 c} + \nabla f, \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} - \frac{1}{c} \frac{\partial f}{\partial t}. \quad (1.11)$$

From (1.10), together with the fact that rotation of a gradient vanishes and that the divergence of a rotation vanishes, we find

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \nabla \times c\mathbf{A}}{\partial t} = -\frac{1}{c} \frac{\partial c\mathbf{B}}{\partial t}, \quad \nabla \cdot c\mathbf{B} = 0. \quad (1.12)$$

If f is zero or at least constant, (1.11) and (1.12) are nothing but Maxwell's equations, except we don't know what $\mathbf{B}$ or $\mathbf{A}$ does to charged particles. To figure out, we need to learn about force, i.e., relativistic mechanics. See Ref. [4] and the references therein for complete derivations.

Some words on my notation I am using the SI unit system's convention on the definition of the standard quantities such as $\mathbf{E}$, $\mathbf{B}$, $\mathbf{A}$ and ϕ . However, some reader may wonder why I write (1.10) like that instead of

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}.$$

Or why I don't write the left hand side of (1.7) as below?

$$\frac{1}{c^2} \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A}$$

The answer is that I am writing formulae in a way I can see dimensions of quantities easily. [2] For example, $\partial/c\partial t$ has the same dimension as ∇ . Therefore, when you encounter

$$\frac{1}{c} \frac{\partial \phi}{\partial t} + \nabla \cdot c\mathbf{A}$$

you will immediately see the dimensions of ϕ and $c\mathbf{A}$ are the same. When you encounter

$$\nabla \times c\mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{j}}{\epsilon_0 c}$$

you will immediately see the dimensions of $\mathbf{E}$ and $c\mathbf{B}$ are the same.

There is another popular unit system called Gaussian units, where centimeters, grams and seconds (cgs) are used instead meters, kilograms and seconds (mks). It sounds they are different only in those units they use. But in fact definitions of quantities are different. For example, $\mathbf{B}$ in Gaussian is $c\mathbf{B}$ in SI, $\mathbf{A}$ in Gaussian is $c\mathbf{A}$ in SI, etc. Gaussian units are popular among physicists at some point because mathematical formulae expressed in Gaussian units look cleaner. A lot of excellent literature uses Gaussian units or their derivatives. My notation is an attempt to make it easier to go back and forth between those units and SI units.

1.2 Relativistic Mechanics

In this section, we derive the relativistic equation of motion by introducing the relativistic energy-momentum four-vector and four-force. Relativistic angular momentum and torque are also introduced.

1.2.1 Four-velocity

We would like to find the relativistic version of the equation of motion. To do so, we need to find the four-vector version of velocity and acceleration. If we simply divide the displacement $(c dt, d\mathbf{r})$ by dt , the resulting vector will not be a four-vector since the divisor dt is not a Lorentz invariant. As we saw in Section 1.1.2 that ds is a Lorentz invariant and is proportional to the differential of the proper time of the mass-point, ds is something act like differential time for the mass-point. Therefore, we consider following quantity $\mathbf{u}$.

$$\mathbf{u} = \frac{d(ct, \mathbf{r})}{ds} = \frac{(c dt, d\mathbf{r})}{c dt \sqrt{1 - (v/c)^2}} = \left(\frac{1}{\sqrt{1 - (v/c)^2}}, \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} \right),$$

where

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}.$$

We call $\mathbf{u}$ four-velocity. Note that this is a dimensionless quantity and its magnitude is always unity:

$$\mathbf{u} \cdot \mathbf{u} = \left(\frac{1}{\sqrt{1 - (v/c)^2}} \right)^2 - \left(\frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} \right)^2 = 1.$$

1.2.2 Relativistic energy and momentum

Let's find the relativistic energy and momentum in a rather heuristic way first, so that we can more easily identify them in the four-acceleration formula in the next section.

We begin with this identity.²

$$\frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} = \frac{\mathbf{v}}{c} \cdot \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} \quad (1.13)$$

Multiplying mc^2 to both sides yields

$$\frac{d}{dt} \frac{mc^2}{\sqrt{1 - (v/c)^2}} = \mathbf{v} \cdot \frac{d}{dt} \frac{m\mathbf{v}}{\sqrt{1 - (v/c)^2}}. \quad (1.14)$$

²See Appendix A for derivation.

This is telling us something. Let us define $\mathbf{p}$ and $\mathbf{f}$ as

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - (v/c)^2}} \quad \text{and} \quad \mathbf{f} = \frac{d\mathbf{p}}{dt}. \quad (1.15)$$

We find that $\mathbf{p}$ becomes Newtonian momentum, $\mathbf{p} = m\mathbf{v}$, if we bring c to infinity, the non-relativistic limit. Therefore, we guess $\mathbf{p}$ in (1.15) can be the relativistic version of momentum. If so, $\mathbf{f}$ is equal to the force acting on the mass-point and the right hand side of (1.14) becomes

$$\mathbf{v} \cdot \frac{d\mathbf{p}}{dt} = \mathbf{v} \cdot \mathbf{f} = \frac{\mathbf{f} \cdot d\mathbf{r}}{dt}.$$

This can be interpreted as the work done to the mass-point per unit time. Therefore the left hand side of (1.14) is increase of energy (ϵ) of the mass-point per unit time.

$$\frac{d}{dt} \frac{mc^2}{\sqrt{1 - (v/c)^2}} = \frac{d\epsilon}{dt}$$

Let us ignore the constant of integration and say

$$\epsilon = \frac{mc^2}{\sqrt{1 - (v/c)^2}}, \quad (1.16)$$

we find

$$\mathbf{p} = (\epsilon/c, \mathbf{p}) = \left(\frac{mc^2/c}{\sqrt{1 - (v/c)^2}}, \frac{m\mathbf{v}}{\sqrt{1 - (v/c)^2}} \right)$$

forms a four-vector, since its magnitude is invariant. That is

$$\mathbf{p}^2 = \epsilon^2/c^2 - p^2 = m^2c^2. \quad (\text{invariant})$$

We would like to express the energy ϵ in terms of the momentum p . That is

$$\epsilon = \sqrt{m^2c^4 + (cp)^2}. \quad (1.17)$$

Note that the ratio of the energy and the momentum does not depend on the mass.

$$c\mathbf{p}/\epsilon = \mathbf{v}/c \quad \text{or} \quad c\mathbf{p} = \epsilon \mathbf{v}/c. \quad (1.18)$$

It is known that the energy and the momentum of the electromagnetic field has relationship of $\epsilon = cp$, i.e., $v = c$. Do definitions (1.15) and (1.16) describe the nature correctly? That is subject to verification. We need to ask nature. So far, all the results from this theory is consistent with observations.

Kinetic energy When $v \ll c$, by performing polynomial expansion with respect to (v/c) of (1.16), we find the Newtonian kinetic energy as the first term after mc^2 .

$$\epsilon = \frac{mc^2}{\sqrt{1 - (v/c)^2}} = mc^2 \left(1 + \frac{1}{2}(v/c)^2 + \dots \right) = mc^2 + \frac{1}{2}mv^2 + \dots \quad (1.19)$$

Therefore (1.16) includes kinetic energy.

The same can be said from (1.17),

$$\epsilon = \sqrt{m^2c^4 + (cp)^2} = mc^2 \sqrt{1 + \frac{1}{c^2} \left(\frac{p}{m} \right)^2} = mc^2 + \frac{p^2}{2m} + \dots \quad (1.20)$$

Example The rest mass of muon is 105.66 MeV. So the speed v of a muon of energy 749 MeV can be calculated as follows.

$$\frac{1}{\sqrt{1 - (v/c)^2}} = \frac{\epsilon}{mc^2} = \frac{749}{105.66} = 7.089 \quad \rightarrow \quad v/c = 0.99$$

Example Let us consider an ideal gas confined in a cubic box of which side length is L . Gas particles bounce back and forth at the walls elastically. When a particle of momentum $\mathbf{p}$ velocity $\mathbf{v}$ bounces at one of the walls normal to x -axis, its momentum changes by $2p_x$ and it will travel $2L$ before coming back to the same wall. The average force f on the wall would be

$$\begin{aligned} f &= \text{momentum change per impact} \times \text{number of impact per unit time}, \\ &= 2p_x \cdot \frac{v_x}{2L} = p_x \cdot v_x / L. \end{aligned}$$

The pressure P is

$$P = f/L^2 = p_x \cdot v_x / V,$$

where $V = L^3$ is the volume of the box. Using (1.18), we find

$$P = \frac{\epsilon}{c^2} \cdot v_x^2 / V.$$

Now summing up for all N particles and bringing V to the other side,

$$PV = \sum_{i=1}^N \frac{\epsilon_i}{c^2} \cdot \frac{1}{3} v_i^2,$$

where ϵ_i and v_i is the energy and the velocity of particle i , respectively. The factor $1/3$ is came from the fact that velocity is uniformly distributed over x , y and z component. For non-relativistic gas, we find by using (1.19),

$$PV = \frac{2}{3} \sum_i \frac{1}{2} m_i v_i^2, \quad (v \ll c)$$

The summation is the sum of the kinetic energy of each particle, we call it internal energy U . Using U above becomes

$$PV = \frac{2}{3} U. \quad (v \ll c) \quad (1.21)$$

In case $v_i = c$, the internal energy U is the sum of each particle's energy ϵ_i .

$$PV = \frac{1}{3} \sum_i \epsilon_i = \frac{1}{3} U. \quad (v = c) \quad (1.22)$$

Example The invariance of $\mathbf{p}^2$ can be used to facilitate converting laboratory energies to center-of-mass energies in the following way (consider identical particles of mass m , for simplicity):³

In K' (laboratory), particle 1 has energy ϵ_1 and particle 2 is at rest.

$$c\mathbf{p}_1' \cdot c\mathbf{p}_2' = (\epsilon_1, c\mathbf{p}_1') \cdot (mc^2, 0) = \epsilon_1 mc^2$$

In K (center of mass), both has the same energy ϵ .

$$c\mathbf{p}_1 \cdot c\mathbf{p}_2 = (\epsilon, c\mathbf{p}) \cdot (\epsilon, -c\mathbf{p}) = \epsilon^2 + c^2 p^2$$

Inner product of four-vector is invariant, i.e., $c\mathbf{p}_1 \cdot c\mathbf{p}_2 = c\mathbf{p}_1' \cdot c\mathbf{p}_2'$. Therefore

$$\epsilon^2 + c^2 p^2 = \epsilon_1 mc^2.$$

And both has the same momentum magnitude:

$$c^2 \mathbf{p}_1^2 = c^2 \mathbf{p}_2^2 = \epsilon^2 - mc^4 = m^2 c^4.$$

Therefore, we find by adding above two together,

$$2\epsilon^2 = mc^2(\epsilon_1 + mc^2) \quad \rightarrow \quad \epsilon = \sqrt{\frac{mc^2}{2} (\epsilon_1 + mc^2)}$$

Example How much mass is used to run 1000MW generator for a year?

Energy generated is

$$1 \times 10^9 \text{ W} \times 365 \times 24 \times 60 \times 60 \text{ s} = 3.15 \times 10^{16} \text{ J}$$

Let's say overall efficiency is 35%. Using $c = 299792458 \text{ m/s}$,

$$\delta m c^2 \times 0.35 = 3.15 \times 10^{16} \quad \rightarrow \quad \delta m \simeq 1 \text{ kg}.$$

³This example is taken from R.P. Feynman, "Quantum Electrodynamics" [5]

1.2.3 Relativistic equation of motion and four-force

Differentiating $\mathbf{u}$ with respect to s yields four-acceleration:

$$\frac{d\mathbf{u}}{ds} = \frac{d}{ds} \frac{d(c\,dt, \mathbf{r})}{ds} = \frac{d}{ds} \left(\frac{1}{\sqrt{1-(v/c)^2}}, \frac{\mathbf{v}/c}{\sqrt{1-(v/c)^2}} \right).$$

Note that the four-acceleration is orthogonal to $\mathbf{u}$ since $\mathbf{u} \cdot \mathbf{u}$ is constant:

$$\mathbf{u} \cdot \frac{d\mathbf{u}}{ds} = \frac{1}{2} \frac{d\mathbf{u} \cdot \mathbf{u}}{ds} = 0.$$

Noting (1.13) and

$$\frac{d}{ds} = \frac{dt}{ds} \frac{d}{dt} = \frac{1}{c\sqrt{1-(v/c)^2}} \frac{d}{dt},$$

we find

$$\frac{d\mathbf{u}}{ds} = \frac{1}{c\sqrt{1-(v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1-(v/c)^2}}, \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1-(v/c)^2}} \right).$$

Notice that the four-momentum $\mathbf{p}$ is mc times the four-velocity $\mathbf{u}$:

$$c\mathbf{p} = mc^2\mathbf{u} = (\epsilon, c\mathbf{p}). \quad (1.23)$$

Recalling (1.15), we find the equation of motion in four-vector notation.

$$\begin{aligned} \frac{d c\mathbf{p}}{ds} &= mc^2 \frac{d\mathbf{u}}{ds} = \frac{1}{\sqrt{1-(v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \frac{d\mathbf{p}}{dt}, \frac{d\mathbf{p}}{dt} \right), \\ &= \frac{1}{\sqrt{1-(v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{f}, \mathbf{f} \right) = \mathbf{K}. \end{aligned} \quad (1.24)$$

Here, we defined the four-force $\mathbf{K} = K_\mu = (K_t, \mathbf{K})$:

$$K_t = \frac{(\mathbf{v}/c) \cdot \mathbf{f}}{\sqrt{1-(v/c)^2}}, \quad \mathbf{K} = \frac{\mathbf{f}}{\sqrt{1-(v/c)^2}}. \quad (1.25)$$

The magnitude of a four-force is negative magnitude of the ordinary force f^2 :

$$K_\mu K_\mu = -\mathbf{f} \cdot \mathbf{f} = -f^2.$$

Note that the four-force $\mathbf{K}$ is composed not only of the ordinary force $\mathbf{f}$ but also of the velocity of the mass-point $\mathbf{v}$.

In tensor component notation, the equation of motion (1.24) becomes

$$\frac{d c p_\mu}{ds} = mc^2 \frac{d u_\mu}{ds} = K_\mu. \quad (1.26)$$

1.2.4 Force tensor

We have just learned that the four-force $\mathbf{K}$ is composed not only of the ordinary force $\mathbf{f}$ but also of the velocity of the mass-point $\mathbf{v}$. However, we would like to consider something independent of the motion of the mass-point as a source of force, so that we can study it by itself. Let's see if we can bring the velocity out from the four-force K_μ by factorizing it into the four-velocity and a tensor $R_{\mu\nu}$, like this:

$$mc^2 \frac{du_\mu}{ds} = K_\mu = u_\nu R_{\mu\nu}.$$

If such $R_{\mu\nu}$ exists, it must be anti-symmetric. The explanation follows.

Taking inner product with u_μ yields

$$mc^2 u_\mu \frac{du_\mu}{ds} = u_\mu u_\nu R_{\mu\nu}.$$

The left hand side is zero, because $u_\mu u_\mu = 1$.

$$mc^2 u_\mu \frac{du_\mu}{ds} = \frac{mc^2}{2} \frac{du_\mu u_\mu}{ds} = 0$$

Therefore, the right hand side, $u_\mu u_\nu R_{\mu\nu}$, must be zero as well, which means $R_{\mu\nu}$ must be anti-symmetric.

An anti-symmetric tensor has six independent components and they are represented by two ordinary vectors, say $\mathbf{p}$ and $\mathbf{q}$. We use $(\mathbf{p}, \mathbf{q})$ to represent an anti-symmetric tensor. Note that, in our notation, when it is a pair of a scalar and an ordinary vector such as $(a_t, \mathbf{a})$, it is a four-vector and that when it is a pair of two ordinary vectors such as $(\mathbf{p}, \mathbf{q})$, it is an anti-symmetric tensor. See Appendix D.3 for more about anti-symmetric tensor.

We get (1.25) with $\mathbf{R} = R_{\mu\nu} = (\mathbf{f}, 0)$:

$$\begin{aligned} \mathbf{uR} &= (u_t, \mathbf{u}) (\mathbf{f}, 0), \\ &= \frac{1}{\sqrt{1 - (v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{f}, \mathbf{f} \right) = (K_t, \mathbf{K}) = \mathbf{K}, \end{aligned}$$

where we used (D.3) and $\mathbf{u} = (1, \mathbf{v}/c)/\sqrt{1 - (v/c)^2}$.

In case $\mathbf{R} = (\mathbf{P}, \mathbf{Q})$,

$$\begin{aligned} \mathbf{K} = \mathbf{uR} &= (u_t, \mathbf{u}) (\mathbf{P}, \mathbf{Q}) = (\mathbf{u} \cdot \mathbf{P}, u_t \mathbf{P} + \mathbf{u} \times \mathbf{Q}), \\ &= \frac{1}{\sqrt{1 - (v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{P}, \mathbf{P} + \frac{\mathbf{v}}{c} \times \mathbf{Q} \right). \end{aligned}$$

Therefore,

$$\mathbf{f} = \mathbf{P} + \frac{\mathbf{v}}{c} \times \mathbf{Q}.$$

If we use $(\mathbf{E}, c\mathbf{B})$, this will become Lorentz force.

1.2.5 Angular momentum

Let's take a look at following anti-symmetric tensor $\mathbf{L} \equiv L_{\mu\nu}$:

$$L_{\mu\nu} = r_\mu p_\nu - r_\nu p_\mu = \mathbf{L} = -(ct\mathbf{p} - \epsilon/c\mathbf{r}, \mathbf{r} \times \mathbf{p}).$$

Taking derivative with respect to ds , we find

$$\begin{aligned} \frac{d\mathbf{L}}{ds} &= \frac{1}{c\sqrt{1-(v/c)^2}} \cdot \frac{d\mathbf{L}}{dt}, \\ &= -\frac{1}{c\sqrt{1-(v/c)^2}} \cdot \frac{d}{dt}(ct\mathbf{p} - \epsilon/c\mathbf{r}, \mathbf{r} \times \mathbf{p}) \\ &= -\frac{1}{c\sqrt{1-(v/c)^2}} \left(c\mathbf{p} + ct\frac{d\mathbf{p}}{dt} - \frac{d\epsilon/c}{dt}\mathbf{r} - \epsilon/c\mathbf{v}, \mathbf{v} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt} \right), \\ &= -\frac{1}{c\sqrt{1-(v/c)^2}} \left(ct\mathbf{f} - \frac{d\epsilon/c}{dt}\mathbf{r}, \mathbf{r} \times \mathbf{f} \right), \end{aligned} \tag{1.27}$$

where we used

$$\frac{d}{ds} = \frac{1}{c\sqrt{1-(v/c)^2}} \frac{d}{dt}, \quad c\mathbf{p} = \epsilon/c\mathbf{v}, \quad \mathbf{v} \times \mathbf{p} = 0, \quad \frac{d\mathbf{p}}{dt} = \mathbf{f}.$$

Now, we define another anti-symmetric tensor $\mathbf{N} \equiv N_{\mu\nu}$:

$$N_{\mu\nu} = r_\mu K_\nu - r_\nu K_\mu = \mathbf{N} = -\frac{1}{\sqrt{1-(v/c)^2}}(ct\mathbf{f} - (\mathbf{v}/c \cdot \mathbf{f})\mathbf{r}, \mathbf{r} \times \mathbf{f}). \tag{1.28}$$

Noting that $(d\epsilon/dt = \mathbf{v} \cdot \mathbf{f})$, we see the equation of motion for $\mathbf{L}$:

$$\frac{d\mathbf{L}}{ds} = \mathbf{N} \quad \text{or} \quad \frac{dL_{\mu\nu}}{ds} = N_{\mu\nu}. \tag{1.29}$$

In ordinary vector, with $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ and $\mathbf{N} = \mathbf{r} \times \mathbf{f}$, above becomes

$$\frac{d\mathbf{L}}{dt} = \mathbf{N}. \tag{1.30}$$

The ordinary vectors $\mathbf{L}$ and $\mathbf{N}$ are the angular momentum and torque, respectively, which are axial vectors. The anti-symmetric tensors $\mathbf{L}$ and $\mathbf{N}$ are their spacetime versions.

Isolated interacting particles Let us consider interacting particles those positions and momentums are $\mathbf{r}_i = (ct, \mathbf{r}_i)$, $\mathbf{p} = (\epsilon_i/c, \mathbf{p}_i)$, respectively. Since the system is isolated, total angular momentum as well as total energy and momentum are conserved.

$$\mathbf{p} = \sum_i (\epsilon_i/c, \mathbf{p}_i) = \text{constant}$$

$$\mathbf{L} = \sum_i - (ct \mathbf{p}_i - \epsilon_i/c \mathbf{r}_i, \mathbf{r}_i \times \mathbf{p}_i) = \text{constant}$$

From the temporal component of $\mathbf{L}$,

$$\sum_i ct \mathbf{p}_i - \sum_i \frac{\epsilon_i}{c} \mathbf{r}_i = \text{constant}.$$

Since the total energy ($\sum \epsilon_i$) is conserved, we can rewrite above

$$\frac{\sum \epsilon_i \mathbf{r}_i}{\sum \epsilon_i} = \frac{c^2 \sum \mathbf{p}_i}{\sum \epsilon_i} t + \text{constant},$$

or

$$\mathbf{R} = \mathbf{V}t + \text{constant},$$

where

$$\mathbf{R} = \frac{\sum \epsilon_i \mathbf{r}_i}{\sum \epsilon_i}, \quad \mathbf{V} = \frac{c^2 \sum \mathbf{p}_i}{\sum \epsilon_i},$$

is the center of mass and its velocity, respectively.

In the non-relativistic limit, where $\epsilon_i = m_i c^2$, we find Newtonian center-of-mass and its velocity:

$$\mathbf{R} = \frac{\sum m_i \mathbf{r}_i}{\sum m_i}, \quad \mathbf{V} = \frac{\sum \mathbf{p}_i}{\sum m_i}.$$

1.2.6 Summary

The equation of motion in ordinary vector:

$$\frac{d\mathbf{p}}{dt} = \mathbf{f} \quad \text{where} \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - (v/c)^2}}. \quad (1.15)$$

The energy and momentum form a four-vector:

$$c\mathbf{p} = mc^2\mathbf{u} = (\epsilon, c\mathbf{p}). \quad (1.23)$$

The energy ϵ is expressed using the velocity v ,

$$\epsilon = \frac{mc^2}{\sqrt{1 - (v/c)^2}}, \quad (1.16)$$

or using the momentum p ,

$$\epsilon = \sqrt{m^2c^4 + (cp)^2}. \quad (1.17)$$

The ratio of the energy and the momentum does not depend on the mass:

$$c\mathbf{p}/\epsilon = \mathbf{v}/c \quad \text{or} \quad c\mathbf{p} = \epsilon\mathbf{v}/c. \quad (1.18)$$

The equation of motion in four-vector:

$$\frac{d c\mathbf{p}}{ds} = mc^2 \frac{d\mathbf{u}}{ds} = \mathbf{K}, \quad (1.24)$$

where the interval ds is a Lorentz invariant and it is similar to the proper time:

$$ds = \sqrt{(cdt)^2 - d\mathbf{r} \cdot d\mathbf{r}} = cdt\sqrt{1 - (v/c)^2}.$$

The four-force $\mathbf{K} = (K_t, \mathbf{K})$ is composed of the ordinary force $\mathbf{f}$ and the velocity of the mass-point $\mathbf{v}$ as

$$K_t = \frac{(\mathbf{v}/c) \cdot \mathbf{f}}{\sqrt{1 - (v/c)^2}}, \quad \mathbf{K} = \frac{\mathbf{f}}{\sqrt{1 - (v/c)^2}}. \quad (1.25)$$

In case $\mathbf{K}$ can be factorized into $\mathbf{u}$ and a tensor $\mathbf{R}$, $\mathbf{R}$ must be anti-symmetric:

$$\text{If } K_\mu = u_\nu R_{\mu\nu} \text{ then } R_{\mu\nu} = -R_{\nu\mu}, \quad \mathbf{R} = (\mathbf{P}, \mathbf{Q}), \quad (1.31)$$

and $\mathbf{K}$ is written with $\mathbf{P}$ and $\mathbf{Q}$ as

$$\mathbf{K} = \mathbf{u}\mathbf{R} = (u_t, \mathbf{u}) (\mathbf{P}, \mathbf{Q}) = \frac{1}{\sqrt{1 - (v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{P}, \mathbf{P} + \frac{\mathbf{v}}{c} \times \mathbf{Q} \right). \quad (1.32)$$

The equation of motion for angular momentum in ordinary vector is unchanged:

$$\frac{d\mathbf{L}}{dt} = \mathbf{N}, \quad \mathbf{L} = \mathbf{r} \times \mathbf{p}, \quad \mathbf{N} = \mathbf{r} \times \mathbf{f}. \quad (1.30)$$

1.3 Lorentz Transformation

We consider two inertial frames K and K' such that K' is moving at a velocity $\mathbf{V}$ relative to K . In this section, we will use γ and β defined below for shorthand and we use them only for the relative velocity between inertial frames, not for the velocity of mass-point to avoid confusion. We call β and γ normalized velocity and Lorentz factor, respectively.

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}}, \quad \beta = |\beta|, \quad \beta = \mathbf{V}/c. \quad (1.33)$$

The Lorentz transformation of a four-vector $(x_t, \mathbf{x})$ from K to K' is

$$\mathbf{x}' = \mathbf{x}_\perp + \gamma (\mathbf{x}_\parallel - \beta x_t), \quad x'_t = \gamma (x_t - \beta \cdot \mathbf{x}), \quad (1.34)$$

where $\mathbf{x}_\perp$ and $\mathbf{x}_\parallel$ is normal and parallel component of $\mathbf{x}$ with respect to β , namely,

$$\mathbf{x} = \mathbf{x}_\perp + \mathbf{x}_\parallel, \quad \mathbf{x}_\parallel = \frac{\beta (\beta \cdot \mathbf{x})}{\beta^2}. \quad (1.35)$$

The inverse transformation is obtained by inverting $\mathbf{V}$, thus β .

$$\mathbf{x} = \mathbf{x}'_\perp + \gamma (\mathbf{x}'_\parallel + \beta x'_t), \quad x_t = \gamma (x'_t + \beta \cdot \mathbf{x}'). \quad (1.36)$$

See Appendix B for derivation. Here are some useful identities.

$$\mathbf{a}_\parallel = \frac{\beta (\beta \cdot \mathbf{a})}{\beta^2}, \quad \mathbf{a}_\perp = -\frac{\beta \times (\beta \times \mathbf{a})}{\beta^2} = \mathbf{a} - \mathbf{a}_\parallel, \quad (1.37)$$

$$(\mathbf{a} \times \mathbf{b})_\parallel = (\mathbf{a}_\perp \times \mathbf{b}_\perp)_\parallel, \quad (\mathbf{a} \times \mathbf{b})_\perp = (\mathbf{a}_\parallel \times \mathbf{b}_\perp) + (\mathbf{a}_\perp \times \mathbf{b}_\parallel). \quad (1.38)$$

1.3.1 Lorentz group

Here, we describe a few facts about Lorentz transformation and Lorentz group without proof, just to introduce terminologies such as proper Lorentz transformation and Lorentz boost.

Lorentz transformation is linear transformation. Let $\mathbf{B}(\mathbf{v})$ be linear operator which transforms vectors in a inertial frame K to another inertial frame K' .

$$\mathbf{r}' = \mathbf{B}(\mathbf{v}) \mathbf{r}.$$

We have already seen that the inverse transformation is $\mathbf{B}(-\mathbf{v})$.

$$\mathbf{r} = \mathbf{B}(-\mathbf{v}) \mathbf{r}' \quad \rightarrow \quad \mathbf{B}^{-1}(\mathbf{v}) = \mathbf{B}(-\mathbf{v}).$$

However, as we will see in Section 1.3.8, two consecutive transformations can not be a single Lorentz transformation in general. It must be a combination of a spatial rotation $R(\boldsymbol{\omega})$ and a Lorentz transformation $B(\mathbf{v})$. For example, for given $\mathbf{v}_1$ and $\mathbf{v}_2$, there exist $\mathbf{v}$ and $\boldsymbol{\omega}$ such that

$$B(\mathbf{v}_1) B(\mathbf{v}_2) = R(\boldsymbol{\omega}) B(\mathbf{v}).$$

$R(\boldsymbol{\omega}) B(\mathbf{v})$ is not equal to $B(\mathbf{v}) R(\boldsymbol{\omega})$ in general, but there exist $\mathbf{v}'$ and $\boldsymbol{\omega}'$ such that

$$B(\mathbf{v}) R(\boldsymbol{\omega}) = R(\boldsymbol{\omega}') B(\mathbf{v}').$$

Therefore, combination of Lorentz transformation and spatial rotation $L = R(\boldsymbol{\omega}) B(\mathbf{v})$ forms a group. We call it proper Lorentz group and L is called proper Lorentz transformation and $B(\mathbf{v})$ is called Lorentz boost. Yet, time reversal T and spatial inversion P also conserves magnitude of four-vector. When people say Lorentz group, it usually refers the one T and P included.

Here in this article, when we simply say Lorentz transformation, it refers Lorentz boost $B(\mathbf{v})$.

1.3.2 Inner product

Let us show that the inner product of four-vectors is a Lorentz invariant. We did this in another way back in Section 1.1.3. Here, we do it by applying Lorentz transformation.

Suppose two four-vector $\mathbf{a} = (a_t, \mathbf{a})$ and $\mathbf{b} = (b_t, \mathbf{b})$. Their Lorentz transformations are

$$\begin{aligned} \mathbf{a}' &= \mathbf{a}_\perp + \gamma (\mathbf{a}_\parallel - \boldsymbol{\beta} a_t), & a'_t &= \gamma (a_t - \boldsymbol{\beta} \cdot \mathbf{a}), \\ \mathbf{b}' &= \mathbf{b}_\perp + \gamma (\mathbf{b}_\parallel - \boldsymbol{\beta} b_t), & b'_t &= \gamma (b_t - \boldsymbol{\beta} \cdot \mathbf{b}). \end{aligned}$$

The inner product is calculated to be

$$\mathbf{a}' \cdot \mathbf{b}' = a'_t b'_t - \mathbf{a}' \cdot \mathbf{b}'.$$

The first term is

$$a'_t b'_t = \gamma^2 (a_t b_t - a_t (\boldsymbol{\beta} \cdot \mathbf{b}) - b_t (\boldsymbol{\beta} \cdot \mathbf{a}) + (\boldsymbol{\beta} \cdot \mathbf{a}) (\boldsymbol{\beta} \cdot \mathbf{b})). \quad (1.39)$$

The second term is

$$\mathbf{a}' \cdot \mathbf{b}' = \mathbf{a}_\perp \cdot \mathbf{b}_\perp + \gamma^2 (\mathbf{a}_\parallel \cdot \mathbf{b}_\parallel - (\mathbf{a}_\parallel \cdot \boldsymbol{\beta}) b_t - (\mathbf{b}_\parallel \cdot \boldsymbol{\beta}) a_t + \beta^2 a_t b_t). \quad (1.40)$$

Recalling that

$$\mathbf{a}_\parallel = \frac{\boldsymbol{\beta} (\boldsymbol{\beta} \cdot \mathbf{a})}{\beta^2}, \quad \mathbf{b}_\parallel = \frac{\boldsymbol{\beta} (\boldsymbol{\beta} \cdot \mathbf{b})}{\beta^2},$$

we find

$$\mathbf{a}_{\parallel} \cdot \mathbf{b}_{\parallel} = (\boldsymbol{\beta} \cdot \mathbf{a})(\boldsymbol{\beta} \cdot \mathbf{b}) / \beta^2, \quad (1.41)$$

and

$$\boldsymbol{\beta} \cdot \mathbf{a}_{\parallel} = \boldsymbol{\beta} \cdot \mathbf{a}, \quad \boldsymbol{\beta} \cdot \mathbf{b}_{\parallel} = \boldsymbol{\beta} \cdot \mathbf{b}. \quad (1.42)$$

Plugging (1.41) and (1.42) into (1.39) minus (1.40), we find

$$\mathbf{a}' \cdot \mathbf{b}' = a'_t b'_t - \mathbf{a}' \cdot \mathbf{b}' = a_t b_t - \mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b}.$$

The inner product of two four-vectors is indeed Lorentz invariant.

1.3.3 Velocity addition law

We want to find the velocity $\mathbf{v}$ of an object launched from a moving station. Suppose that the station is moving at $\mathbf{V}$ relative to the observer, the inertial frame attached to the station is K' , and the inertial frame attached to the observer is K , therefore K is moving at $-\mathbf{V}$ relative to K' . The object is launched at velocity $\mathbf{v}'$ relative to the station. The object's position displacement vector $d\mathbf{r}' = (c dt', d\mathbf{r}')$ is transformed to K as follows,

$$d\mathbf{r} = d\mathbf{r}'_{\perp} + \gamma (d\mathbf{r}'_{\parallel} + \boldsymbol{\beta} c dt'), \quad c dt = \gamma (c dt' + \boldsymbol{\beta} \cdot d\mathbf{r}'). \quad (1.43)$$

Therefore $\mathbf{v}$ is calculated as,

$$\begin{aligned} \frac{\mathbf{v}}{c} &= \frac{1}{c} \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}'_{\perp} + \gamma (d\mathbf{r}'_{\parallel} + \boldsymbol{\beta} c dt')}{\gamma (c dt' + \boldsymbol{\beta} \cdot d\mathbf{r}')}, \\ &= \frac{d\mathbf{r}'_{\perp}/\gamma + d\mathbf{r}'_{\parallel} + \boldsymbol{\beta} c dt'}{c dt' + \boldsymbol{\beta} \cdot d\mathbf{r}'}, \\ &= \frac{\mathbf{v}'_{\perp}/c\gamma + \mathbf{v}'_{\parallel}/c + \boldsymbol{\beta}}{1 + \boldsymbol{\beta} \cdot \mathbf{v}'/c}, \\ &= \frac{\mathbf{v}'_{\perp}/c\gamma + \mathbf{v}'_{\parallel}/c + \mathbf{V}/c}{1 + \mathbf{V} \cdot \mathbf{v}'/c^2}. \end{aligned}$$

Finally, we find the velocity addition law:

$$\mathbf{v} = \frac{\mathbf{v}'_{\perp}/\gamma + \mathbf{v}'_{\parallel} + \mathbf{V}}{1 + \mathbf{V} \cdot \mathbf{v}'/c^2}. \quad (1.44)$$

1.3.4 Velocity

The velocity addition law can also be derived from the Lorentz transformation of the four-velocity:

$$\mathbf{u}' = \frac{d(ct', \mathbf{r}')}{ds} = \left(\frac{1}{\sqrt{1 - (v'/c)^2}}, \frac{\mathbf{v}'/c}{\sqrt{1 - (v'/c)^2}} \right).$$

Assuming the same scenario as in Section 1.3.3. For the spatial component, we have

$$\begin{aligned} \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} &= \frac{\mathbf{v}'_{\perp}/c}{\sqrt{1 - (v'/c)^2}} + \gamma \left(\frac{\mathbf{v}'_{\parallel}/c}{\sqrt{1 - (v'/c)^2}} + \frac{\boldsymbol{\beta}}{\sqrt{1 - (v'/c)^2}} \right), \\ &\rightarrow \frac{1}{\gamma} \sqrt{\frac{1 - (v'/c)^2}{1 - (v/c)^2}} \mathbf{v} = \mathbf{v}'_{\perp}/\gamma + \mathbf{v}'_{\parallel} + \mathbf{V}, \end{aligned} \quad (1.45)$$

and for the temporal component, we have

$$\begin{aligned} \frac{1}{\sqrt{1 - (v/c)^2}} &= \gamma \left(\frac{1}{\sqrt{1 - (v'/c)^2}} + \boldsymbol{\beta} \cdot \frac{\mathbf{v}'/c}{\sqrt{1 - (v'/c)^2}} \right), \\ &\rightarrow \frac{1}{\gamma} \sqrt{\frac{1 - (v'/c)^2}{1 - (v/c)^2}} = 1 + \mathbf{V} \cdot \mathbf{v}'/c^2. \end{aligned} \quad (1.46)$$

Plugging (1.46) into (1.45) yields the velocity addition law:

$$\mathbf{v} = \frac{\mathbf{v}'_{\perp}/\gamma + \mathbf{v}'_{\parallel} + \mathbf{V}}{1 + \mathbf{V} \cdot \mathbf{v}'/c^2}. \quad (1.44)$$

Similarly, starting from the inverse transformation, we get

$$\frac{1}{\gamma} \sqrt{\frac{1 - (v/c)^2}{1 - (v'/c)^2}} = 1 - \mathbf{V} \cdot \mathbf{v}/c^2, \quad (1.47)$$

and

$$\mathbf{v}' = \frac{\mathbf{v}_{\perp}/\gamma + \mathbf{v}_{\parallel} - \mathbf{V}}{1 - \mathbf{V} \cdot \mathbf{v}/c^2}. \quad (1.48)$$

1.3.5 Force

Plugging the four-force,

$$\mathbf{K}' = (K'_t, \mathbf{K}') = \frac{1}{\sqrt{1 - (v'/c)^2}} \left(\frac{\mathbf{v}'}{c} \cdot \mathbf{f}', \mathbf{f}' \right),$$

into the Lorentz transformation,

$$\mathbf{K} = \mathbf{K}'_{\perp} + \gamma \left(\mathbf{K}'_{\parallel} + \beta K'_t \right),$$

yields

$$\frac{\mathbf{f}}{\sqrt{1 - (v/c)^2}} = \frac{\mathbf{f}'_{\perp}}{\sqrt{1 - (v'/c)^2}} + \gamma \left(\frac{\mathbf{f}'_{\parallel}}{\sqrt{1 - (v'/c)^2}} + \beta \frac{\mathbf{v}'/c \cdot \mathbf{f}'}{\sqrt{1 - (v'/c)^2}} \right).$$

Using (1.46), we find

$$\mathbf{f} = \frac{\mathbf{f}'_{\perp}/\gamma + \mathbf{f}'_{\parallel} + \beta (\mathbf{v}'/c \cdot \mathbf{f}')}{1 + \beta \cdot \mathbf{v}'/c}. \quad (1.49)$$

It will be more convenient to express $\mathbf{f}$ with the velocity measured in the same frame. We would like to use $\mathbf{v}$ instead of $\mathbf{v}'$. Multiplying (1.46) and (1.47), we find

$$\frac{1}{\gamma^2} = \left(1 + \beta \cdot \frac{\mathbf{v}'}{c} \right) \left(1 - \beta \cdot \frac{\mathbf{v}}{c} \right).$$

Plugging this into (1.49) yields

$$\mathbf{f} = \left(\gamma \mathbf{f}'_{\perp} + \gamma^2 \mathbf{f}'_{\parallel} + \gamma^2 \beta \left(\frac{\mathbf{v}'}{c} \cdot \mathbf{f}' \right) \right) \left(1 - \left(\beta \cdot \frac{\mathbf{v}}{c} \right) \right).$$

Using (1.48), above becomes

$$\mathbf{f} = \left(\gamma \mathbf{f}'_{\perp} + \gamma^2 \mathbf{f}'_{\parallel} \right) \left(1 - \left(\beta \cdot \frac{\mathbf{v}}{c} \right) \right) + \gamma^2 \beta \left(\left(\frac{\mathbf{v}_{\perp}/\gamma}{c} + \frac{\mathbf{v}_{\parallel}}{c} - \beta \right) \cdot \mathbf{f}' \right).$$

After a few lines of calculation,⁴ we get

$$\mathbf{f} = \mathbf{f}'_{\parallel} + \gamma \mathbf{f}'_{\perp} + \left(\frac{\mathbf{v}}{c} \cdot \gamma \mathbf{f}' \right) \beta - \left(\frac{\mathbf{v}}{c} \cdot \beta \right) \gamma \mathbf{f}'.$$

Recalling that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}$, we find

$$\mathbf{f} = \mathbf{f}'_{\parallel} + \gamma \mathbf{f}'_{\perp} + \frac{\mathbf{v}}{c} \times \left(\beta \times \gamma \mathbf{f}' \right). \quad (1.50)$$

We see that force is composed of two components, $\mathbf{P}$ and $\mathbf{Q}$ as below.

$$\mathbf{f} = \mathbf{P} + \frac{\mathbf{v}}{c} \times \mathbf{Q}$$

If we redo the calculation with

$$\mathbf{f}' = \mathbf{P}' + \frac{\mathbf{v}'}{c} \times \mathbf{Q}',$$

we should be able to show that if $\mathbf{P}'$ and $\mathbf{Q}'$ does not depend on $\mathbf{v}'$, $\mathbf{P}$ and $\mathbf{Q}$ will not depend on $\mathbf{v}$. And we should be able to find the transformation for $\mathbf{P}$ and $\mathbf{Q}$. However, it is a bit tedious to proceed. But it is exactly the case when four-force K_{μ} can be expressed as $u_{\nu} R_{\mu\nu}$, as we saw in Section 1.2.4.

⁴See Appendix C for details.

1.3.6 Force tensor

As we saw in Section 1.2.4, if four-force $\mathbf{K}$ is factorized into the four-velocity of the mass-point $\mathbf{u}$ and a tensor $\mathbf{R}$, $\mathbf{R}$ must be an anti-symmetric four-tensor, $\mathbf{R} = (\mathbf{P}, \mathbf{Q})$.

$$\begin{aligned} \mathbf{uR} &= (u_t, \mathbf{u}) (\mathbf{P}, \mathbf{Q}) = (\mathbf{u} \cdot \mathbf{P}, u_t \mathbf{P} + \mathbf{u} \times \mathbf{Q}), \\ &= \frac{1}{\sqrt{1 - (v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{P}, \mathbf{P} + \frac{\mathbf{v}}{c} \times \mathbf{Q} \right), \\ &= \frac{1}{\sqrt{1 - (v/c)^2}} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{f}, \mathbf{f} \right) = \mathbf{K}, \end{aligned}$$

Therefore

$$\mathbf{f} = \mathbf{P} + \frac{\mathbf{v}}{c} \times \mathbf{Q}. \quad (1.51)$$

In case both $\mathbf{P}$ and $\mathbf{Q}$ does not depend on $\mathbf{v}$, the Lorentz transformation of $\mathbf{R} = (\mathbf{P}, \mathbf{Q})$ will be, from (D.4) and (D.5),

$$\mathbf{P}' = \mathbf{P}_{\parallel} + \gamma(\mathbf{P} + \boldsymbol{\beta} \times \mathbf{Q})_{\perp} \quad (1.52)$$

$$\mathbf{Q}' = \mathbf{Q}_{\parallel} + \gamma(\mathbf{Q} - \boldsymbol{\beta} \times \mathbf{P})_{\perp} \quad (1.53)$$

The inverse transformation of above is obtained by changing the sign of $\boldsymbol{\beta}$.

$$\mathbf{P} = \mathbf{P}'_{\parallel} + \gamma(\mathbf{P}' - \boldsymbol{\beta} \times \mathbf{Q}')_{\perp}, \quad (1.54)$$

$$\mathbf{Q} = \mathbf{Q}'_{\parallel} + \gamma(\mathbf{Q}' + \boldsymbol{\beta} \times \mathbf{P}')_{\perp}. \quad (1.55)$$

Some reader may have noticed that (1.51), (1.52) and (1.53) are basically the Lorentz force and the Lorentz transformation of electric and magnetic fields. We got these result without using anything specific to electricity nor magnetism.

1.3.7 Four-gradient

Let's use a Cartesian coordinate whose x-axis coincides with $\boldsymbol{\beta}$, i.e., $\boldsymbol{\beta} = (\beta_x \ 0 \ 0)$. The Lorentz transformation will be

$$ct = \gamma(ct' + \beta_x x'), \quad x = \gamma(x' + \beta_x ct'), \quad y = y', \quad z = z'$$

Lorentz transformation of four-gradient can be found by the chain rule as follows.

$$\begin{aligned} \frac{\partial}{\partial x'} &= \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} + \frac{\partial ct}{\partial x'} \frac{\partial}{\partial ct} = \gamma \left(\frac{\partial}{\partial x} + \beta_x \frac{\partial}{\partial ct} \right), \\ \frac{\partial}{\partial y'} &= \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial z'} = \frac{\partial}{\partial z}, \end{aligned}$$

$$\frac{\partial}{\partial ct'} = \frac{\partial x}{\partial ct'} \frac{\partial}{\partial x} + \frac{\partial ct}{\partial ct'} \frac{\partial}{\partial ct} = \gamma \left(\frac{\partial}{\partial ct} + \beta_x \frac{\partial}{\partial x} \right).$$

Bringing these back to vector notation yields

$$\nabla' = \nabla_{\perp} + \gamma \left(\nabla_{\parallel} + \beta \frac{1}{c} \frac{\partial}{\partial t} \right), \quad \frac{1}{c} \frac{\partial}{\partial t'} = \gamma \left(\frac{1}{c} \frac{\partial}{\partial t} + \beta \cdot \nabla \right).$$

1.3.8 Thomas precession

We would like to find the rest frame for an accelerating particle, namely an electron orbiting around the nucleus of an atom. Suppose that the particle's velocity is $\mathbf{v}$ at time t and is $\mathbf{v}_2 = \mathbf{v} + \delta\mathbf{v}$ at time $t + \delta t$, all measured in the observer's inertial frame K . Let K' be the instantaneous rest frame of the particle at t and K'' be that of at $t + \delta t$. We will find the infinitesimal transformation from K' to K'' , which is two consecutive Lorentz transformations, is composed of not only a Lorentz transformation but also a spatial rotation of $\delta\mathbf{\Omega}$. The main purpose of this section is to find the angular velocity of the rotation, $\boldsymbol{\omega}_T = \delta\mathbf{\Omega}/\delta t$.

Relation between these three inertial frames are illustrated as below.

$$K' \xrightarrow{-\mathbf{v}} K \xrightarrow{\mathbf{v} + \delta\mathbf{v}} K''$$

Let us use following normalized velocities and Lorentz factors:

$$\beta = \mathbf{v}/c, \quad \beta_2 = (\mathbf{v} + \delta\mathbf{v})/c, \quad \delta\beta = \beta_2 - \beta = \delta\mathbf{v}/c,$$

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}}, \quad \gamma_2 = \frac{1}{\sqrt{1 - \beta_2^2}},$$

as well as these:

$$\mathbf{r}'_{\perp} = \mathbf{r}' - \mathbf{r}'_{\parallel}, \quad \mathbf{r}'_{\parallel} = \frac{(\boldsymbol{\beta} \cdot \mathbf{r}')}{\beta^2}, \quad \mathbf{r}_{\perp} = \mathbf{r} - \mathbf{r}_{\parallel}, \quad \mathbf{r}_{\parallel} = \frac{(\boldsymbol{\beta}_2 \cdot \mathbf{r})}{\beta_2^2}.$$

And we will use followings later.

$$\begin{aligned} \gamma_2 - \gamma &= \delta\gamma = \frac{\partial\gamma}{\partial\boldsymbol{\beta}} \cdot \delta\boldsymbol{\beta} = \gamma^3 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \\ &\rightarrow \gamma_2 = \gamma + \gamma^3 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \end{aligned} \tag{1.56}$$

$$\begin{aligned} \frac{1}{\beta_2^2} - \frac{1}{\beta^2} &= \delta \left(\frac{1}{\beta^2} \right) = \frac{\partial}{\partial\boldsymbol{\beta}} \frac{1}{\beta^2} \cdot \delta\boldsymbol{\beta} = -\frac{2(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^4} \\ &\rightarrow \frac{1}{\beta_2^2} = \frac{1}{\beta^2} - \frac{2(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^4} \end{aligned} \tag{1.57}$$

The transformation from K' to K is

$$\mathbf{r} = \mathbf{r}'_{\perp} + \gamma \left(\mathbf{r}'_{\parallel} + \boldsymbol{\beta} ct' \right), \quad (1.58)$$

$$ct = \gamma \left(ct' + \boldsymbol{\beta} \cdot \mathbf{r}' \right). \quad (1.59)$$

The transformation from K to K'' is

$$\begin{aligned} \mathbf{r}'' &= \mathbf{r}_{\perp} + \gamma_2 \left(\mathbf{r}_{\parallel} - \boldsymbol{\beta}_2 ct \right), \\ &= \mathbf{r} + (\gamma_2 - 1) \frac{(\boldsymbol{\beta}_2 \cdot \mathbf{r})}{\beta_2^2} \boldsymbol{\beta}_2 - \gamma_2 \boldsymbol{\beta}_2 ct \end{aligned} \quad (1.60)$$

$$ct'' = \gamma_2 (ct - \boldsymbol{\beta}_2 \cdot \mathbf{r}). \quad (1.61)$$

Let the normalized velocity of K'' measured in K' be $\delta\boldsymbol{\beta}'$. Since $\delta\boldsymbol{\beta}'$ is small, we can set the Lorentz factor associated with $\delta\boldsymbol{\beta}'$ unity because it is an even function of $\delta\boldsymbol{\beta}'$, i.e., there must be no first (odd) order term (terms) in the polynomial expansion. You can verify this by writing it down explicitly,

$$\frac{1}{\sqrt{1 - \delta\beta'^2}} = 1 + \frac{1}{2} \delta\beta'^2 + \dots \simeq 1.$$

The transformation from K' to K'' should include the Lorentz transformation of $\delta\boldsymbol{\beta}'$, we guess the transformation in the form of

$$ct'' = ct' - \delta\boldsymbol{\beta}' \cdot \mathbf{r}', \quad (1.62)$$

$$\mathbf{r}'' = \mathbf{r}' - \delta\boldsymbol{\beta}' ct' + (\text{remainder}). \quad (1.63)$$

We want to know the term, (*remainder*). The calculation from here is straightforward but it is a bit tedious. Readers who wants do it by yourself can skip to (1.67).

Now we want to find the coefficients of ct' and $\mathbf{r}'$ in (1.62) to first order in $\delta\boldsymbol{\beta}$. Plugging (1.58) and (1.59) into (1.61) yields

$$\begin{aligned} ct'' &= \gamma_2 \gamma \left(ct' + \boldsymbol{\beta} \cdot \mathbf{r}'_{\parallel} \right) - \gamma_2 \boldsymbol{\beta}_2 \cdot \left(\mathbf{r}'_{\perp} + \gamma \left(\mathbf{r}'_{\parallel} + \boldsymbol{\beta} ct' \right) \right), \\ &= \gamma_2 \gamma (1 - (\boldsymbol{\beta}_2 \cdot \boldsymbol{\beta})) ct' + \gamma_2 \gamma \boldsymbol{\beta} \cdot \mathbf{r}'_{\parallel} - \gamma_2 \boldsymbol{\beta}_2 \cdot \left(\mathbf{r}'_{\perp} + \gamma \mathbf{r}'_{\parallel} \right), \\ &= \gamma_2 \gamma (1 - (\boldsymbol{\beta}_2 \cdot \boldsymbol{\beta})) ct' + \gamma_2 \gamma \boldsymbol{\beta} \cdot \mathbf{r}'_{\parallel} \\ &\quad - \gamma_2 \underbrace{\boldsymbol{\beta} \cdot \mathbf{r}'_{\perp}}_0 - \gamma_2 \gamma \boldsymbol{\beta} \cdot \mathbf{r}'_{\parallel} - \delta\boldsymbol{\beta} \cdot \gamma_2 \left(\mathbf{r}'_{\perp} + \gamma \mathbf{r}'_{\parallel} \right), \\ &= \gamma_2 \gamma (1 - (\boldsymbol{\beta}_2 \cdot \boldsymbol{\beta})) ct' - \delta\boldsymbol{\beta} \cdot \gamma_2 \left(\mathbf{r}'_{\perp} + \gamma \mathbf{r}'_{\parallel} \right). \end{aligned}$$

The coefficient of ct' is,

$$\begin{aligned}\gamma_2\gamma(1 - (\boldsymbol{\beta}_2 \cdot \boldsymbol{\beta})) &= (\gamma^2 + \gamma^4(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}))(1 - \beta^2 - (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})), \\ &= \left(\frac{1}{1 - \beta^2} + \frac{1}{(1 - \beta^2)^2}(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \right) (1 - \beta^2 - (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})), \\ &= 1 + \frac{1}{1 - \beta^2}(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) - \frac{1}{1 - \beta^2}(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) + \mathcal{O}(\delta\beta^2)\end{aligned}$$

This is indeed unity.

The coefficient of $\mathbf{r}'$ is rather straightforward. That is

$$\delta\boldsymbol{\beta} \cdot \gamma_2 \left(\mathbf{r}'_{\perp} + \gamma \mathbf{r}'_{\parallel} \right) = \gamma_2 \left(\delta\boldsymbol{\beta} + (\gamma - 1) \frac{(\delta\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \right) \cdot \mathbf{r}'.$$

Recalling (1.56), above becomes by taking first order terms only,

$$\delta\boldsymbol{\beta} \cdot \gamma_2 \left(\mathbf{r}'_{\perp} + \gamma \mathbf{r}'_{\parallel} \right) = \gamma \left(\delta\boldsymbol{\beta} + (\gamma - 1) \frac{(\delta\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \right) \cdot \mathbf{r}'.$$

Finally, we get

$$ct'' = ct' + \gamma \left(\delta\boldsymbol{\beta} + (\gamma - 1) \frac{(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \right) \cdot \mathbf{r}'. \quad (1.64)$$

By comparing this with (1.62), we find

$$\delta\boldsymbol{\beta}' = \gamma \left(\delta\boldsymbol{\beta} + (\gamma - 1) \frac{(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \right) = \gamma \delta\boldsymbol{\beta}_{\perp} + \gamma^2 \delta\boldsymbol{\beta}_{\parallel}, \quad (1.65)$$

where

$$\delta\boldsymbol{\beta}_{\perp} = \delta\boldsymbol{\beta} - \delta\boldsymbol{\beta}_{\parallel}, \quad \delta\boldsymbol{\beta}_{\parallel} = \frac{(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta}.$$

Now we want to find the coefficient of ct' and $\mathbf{r}'$ in (1.63) to first order in $\delta\boldsymbol{\beta}$. Plugging (1.58) and (1.59) into (1.60) yields

$$\begin{aligned}\mathbf{r}'' &= \mathbf{r}' + (\gamma - 1) \frac{(\boldsymbol{\beta} \cdot \mathbf{r}')}{\beta^2} \boldsymbol{\beta} - \gamma \boldsymbol{\beta} ct' \\ &\quad + \left(\gamma - 1 + \gamma^3(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \right) \left(\frac{1}{\beta^2} - \frac{2(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^4} \right) (\boldsymbol{\beta} + \delta\boldsymbol{\beta}) \\ &\quad \times \left(\gamma(\boldsymbol{\beta} \cdot \mathbf{r}') + \gamma\beta^2 ct' + \delta\boldsymbol{\beta} \cdot (\mathbf{r}'_{\perp} + \gamma(\mathbf{r}'_{\parallel} - \boldsymbol{\beta} ct')) \right) \\ &\quad - \left(\gamma + \gamma^3(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \right) (\boldsymbol{\beta} + \delta\boldsymbol{\beta}) \gamma (ct' + \boldsymbol{\beta} \cdot \mathbf{r}').\end{aligned} \quad (1.66)$$

The coefficient of ct' in zeroth order in $\delta\boldsymbol{\beta}$ must be zero and it is:

$$\gamma\boldsymbol{\beta} + (\gamma - 1) \frac{1}{\beta^2} \boldsymbol{\beta} \gamma \beta^2 - \gamma^2 \boldsymbol{\beta} = (\gamma + \gamma(\gamma - 1) - \gamma^2) \boldsymbol{\beta} = 0.$$

The coefficient of ct' in first order in $\delta\boldsymbol{\beta}$ must be $-\delta\boldsymbol{\beta}'$ and it is:

$$\begin{aligned} & \gamma^2 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \frac{1}{\beta^2} \boldsymbol{\beta} \gamma \beta^2 - (\gamma - 1) \frac{2(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^4} \boldsymbol{\beta} \gamma \beta^2 \\ & + (\gamma - 1) \frac{\delta\boldsymbol{\beta}}{\beta^2} \gamma \beta^2 + (\gamma - 1) \frac{\boldsymbol{\beta}}{\beta^2} (\delta\boldsymbol{\beta} \cdot \gamma\boldsymbol{\beta}) \\ & - \gamma^3 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \boldsymbol{\beta} \gamma - \gamma \delta\boldsymbol{\beta} \gamma, \\ & = \gamma^4 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \boldsymbol{\beta} - \gamma(\gamma - 1) \frac{2(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \\ & + \gamma(\gamma - 1) \delta\boldsymbol{\beta} + \gamma(\gamma - 1) \frac{(\delta\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \\ & - \gamma^4 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \boldsymbol{\beta} - \gamma^2 \delta\boldsymbol{\beta}, \\ & = -\gamma \left(\delta\boldsymbol{\beta} + (\gamma - 1) \frac{(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^2} \boldsymbol{\beta} \right) = -\delta\boldsymbol{\beta}'. \end{aligned}$$

Terms including $\boldsymbol{r}'$ to zeroth order in $\delta\boldsymbol{\beta}$ must be $\boldsymbol{r}'$ and it is:

$$\begin{aligned} & \boldsymbol{r}' + (\gamma - 1) \frac{(\boldsymbol{\beta} \cdot \boldsymbol{r}')}{\beta^2} \boldsymbol{\beta} + (\gamma - 1) \frac{1}{\beta^2} \boldsymbol{\beta} \gamma (\boldsymbol{\beta} \cdot \boldsymbol{r}') - \gamma^2 \boldsymbol{\beta} (\boldsymbol{\beta} \cdot \boldsymbol{r}'), \\ & = \boldsymbol{r}' + (\gamma - 1 + \gamma^2 - \gamma) \frac{(\boldsymbol{\beta} \cdot \boldsymbol{r}')}{\beta^2} \boldsymbol{\beta} - \gamma^2 \boldsymbol{\beta} (\boldsymbol{\beta} \cdot \boldsymbol{r}'), \\ & = \boldsymbol{r}' + \left((\gamma^2 - 1) \frac{1}{\beta^2} - \gamma^2 \right) (\boldsymbol{\beta} \cdot \boldsymbol{r}') \boldsymbol{\beta}, \\ & = \boldsymbol{r}' + \left(\frac{1 - 1 + \beta^2}{1 - \beta^2} \cdot \frac{1}{\beta^2} - \gamma^2 \right) (\boldsymbol{\beta} \cdot \boldsymbol{r}') \boldsymbol{\beta} = \boldsymbol{r}'. \end{aligned}$$

Terms including $\mathbf{r}'$ to first order in $\delta\boldsymbol{\beta}$ should be (*remainder*):

$$\begin{aligned}
& \gamma^3(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \frac{1}{\beta^2} \boldsymbol{\beta} \gamma(\boldsymbol{\beta} \cdot \mathbf{r}') - (\gamma - 1) \frac{2(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta})}{\beta^4} \boldsymbol{\beta} \gamma(\boldsymbol{\beta} \cdot \mathbf{r}') + (\gamma - 1) \frac{1}{\beta^2} \delta\boldsymbol{\beta} \gamma(\boldsymbol{\beta} \cdot \mathbf{r}') \\
& + (\gamma - 1) \frac{1}{\beta^2} \boldsymbol{\beta} \left(\delta\boldsymbol{\beta} \cdot \left(\mathbf{r}' + (\gamma - 1) \frac{(\boldsymbol{\beta} \cdot \mathbf{r}')}{\beta^2} \boldsymbol{\beta} \right) \right) \\
& - \gamma^3(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) \boldsymbol{\beta} \gamma(\boldsymbol{\beta} \cdot \mathbf{r}') - \gamma \delta\boldsymbol{\beta} \gamma(\boldsymbol{\beta} \cdot \mathbf{r}') \\
& = \frac{\gamma^4}{\beta^2} (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) (\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta} - \frac{2\gamma(\gamma - 1)}{\beta^4} (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) (\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta} + \frac{\gamma(\gamma - 1)}{\beta^2} (\boldsymbol{\beta} \cdot \mathbf{r}') \delta\boldsymbol{\beta} \\
& + \frac{\gamma - 1}{\beta^2} (\delta\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta} + \frac{(\gamma - 1)^2}{\beta^4} (\delta\boldsymbol{\beta} \cdot \boldsymbol{\beta}) (\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta} \\
& - \gamma^4 (\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) (\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta} - \gamma^2 (\boldsymbol{\beta} \cdot \mathbf{r}') \delta\boldsymbol{\beta}.
\end{aligned}$$

The coefficient of $(\boldsymbol{\beta} \cdot \delta\boldsymbol{\beta}) (\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta}$:

$$\begin{aligned}
& \frac{\gamma^4}{\beta^2} - \frac{2\gamma(\gamma - 1)}{\beta^4} + \frac{(\gamma - 1)^2}{\beta^4} - \gamma^4 \\
& = \gamma^4 \left(\frac{1}{\beta^2} - 1 \right) + \frac{-2\gamma^2 + 2\gamma + \gamma^2 - 2\gamma + 1}{\beta^4} \\
& = \frac{\gamma^2}{\beta^2} + \frac{1 - \gamma^2}{\beta^4} = \frac{\gamma^2}{\beta^2} \left(1 - \frac{1}{\beta^2} \right) + \frac{1}{\beta^4} = -\frac{1}{\beta^4} + \frac{1}{\beta^4} = 0.
\end{aligned}$$

The coefficient of $(\boldsymbol{\beta} \cdot \mathbf{r}') \delta\boldsymbol{\beta}$:

$$\frac{\gamma(\gamma - 1)}{\beta^2} - \gamma^2 = \frac{\gamma^2 - \gamma}{\beta^2} - \gamma^2 = \gamma^2 \left(\frac{1}{\beta^2} - 1 \right) - \frac{\gamma}{\beta^2} = -\frac{\gamma - 1}{\beta^2}.$$

The coefficient of $(\delta\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta}$ is $(\gamma - 1)/\beta^2$ and this can be written using γ only:

$$\frac{\gamma - 1}{\beta^2} = \frac{\gamma^2 - 1}{\beta^2(\gamma + 1)} = \left(\frac{1}{1 - \beta^2} - 1 \right) \frac{1}{\beta^2(\gamma + 1)} = \frac{\gamma^2}{\gamma + 1}.$$

Therefore (*remainder*) is, noting that $\mathbf{a} \times \mathbf{b} \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}$,

$$\frac{\gamma^2}{\gamma + 1} \left((\mathbf{r}' \cdot \delta\boldsymbol{\beta}) \boldsymbol{\beta} - (\mathbf{r}' \cdot \boldsymbol{\beta}) \delta\boldsymbol{\beta} \right) = \frac{\gamma^2}{\gamma + 1} \mathbf{r}' \times \boldsymbol{\beta} \times \delta\boldsymbol{\beta} = -\frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} \times \delta\boldsymbol{\beta} \times \mathbf{r}'. \quad (1.67)$$

This represents the displacement of an infinitesimal spatial rotation, $-\delta\boldsymbol{\Omega} \times \mathbf{r}'$, where

$$\delta\boldsymbol{\Omega} = \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} \times \delta\boldsymbol{\beta} = \frac{\gamma^2}{\gamma + 1} \frac{\mathbf{v} \times \delta\mathbf{v}}{c^2}.$$

Recalling that δt is the time for the particle to accelerate from $\mathbf{v}$ to $\mathbf{v} + \delta \mathbf{v}$ in K . The angular velocity $\boldsymbol{\omega}_T$ is

$$\begin{aligned}\boldsymbol{\omega}_T &= \lim_{\delta t \rightarrow 0} \frac{\delta \boldsymbol{\Omega}}{\delta t}, \\ &= \lim_{\delta t \rightarrow 0} \frac{\gamma^2}{\gamma + 1} \frac{\mathbf{v}}{c} \times \frac{\delta \mathbf{v}}{c \delta t} = \frac{\gamma^2}{\gamma + 1} \frac{\mathbf{v} \times \mathbf{a}}{c^2}.\end{aligned}\quad (1.68)$$

Note that γ , $\mathbf{v}$, and $\mathbf{a}$ are all measured in the observer's frame K while $\delta \boldsymbol{\Omega}$ is a vector in K' and that this rotation is relativistic effect since $\boldsymbol{\omega}_T$ vanishes as $c \rightarrow \infty$. When the particle's velocity $\mathbf{v}$ is small compared to c , i.e., $v \ll c$, $\gamma \sim 1$,

$$\boldsymbol{\omega}_T \simeq \frac{\mathbf{v} \times \mathbf{a}}{2c^2}. \quad (v \ll c)$$

The particle's rest frame rotates at the instantaneous angular velocity $\boldsymbol{\omega}_T$. The precession associated with this rotation is called Thomas precession. This is essential for full understanding of Zeeman effect. [6, 7]

Finally, (1.62) and (1.63) becomes

$$ct'' = ct' - \delta \boldsymbol{\beta}' \cdot \mathbf{r}', \quad (1.69)$$

$$\mathbf{r}'' = \mathbf{r}' - \delta \boldsymbol{\beta}' ct' - \delta \boldsymbol{\Omega} \times \mathbf{r}'. \quad (1.70)$$

We can rewrite above using $\mathbf{B}$ and $\mathbf{R}$ introduced in Appendix 1.3.1.

$$\mathbf{r}'' = \mathbf{B}(\delta \boldsymbol{\beta}') \mathbf{R}(\delta \boldsymbol{\Omega}) \mathbf{r}' + \mathcal{O}(\delta^2).$$

where

$$\begin{aligned}\mathbf{B}(\delta \boldsymbol{\beta}') \mathbf{r}' &= \mathbf{B}(\delta \boldsymbol{\beta}')(ct', \mathbf{r}') = (ct' - \delta \boldsymbol{\beta}' \cdot \mathbf{r}', \mathbf{r}' - \delta \boldsymbol{\beta}' ct'), \\ \mathbf{R}(\delta \boldsymbol{\Omega}) \mathbf{r}' &= \mathbf{R}(\delta \boldsymbol{\Omega})(ct', \mathbf{r}') = (ct', \mathbf{r}' - \delta \boldsymbol{\Omega} \times \mathbf{r}').\end{aligned}$$

$\mathbf{B}$ and $\mathbf{R}$ commutes for infinitesimal transformations, although they does not in general:

$$\mathbf{B}(\delta \boldsymbol{\beta}') \mathbf{R}(\delta \boldsymbol{\Omega}) = \mathbf{R}(\delta \boldsymbol{\Omega}) \mathbf{B}(\delta \boldsymbol{\beta}') + \mathcal{O}(\delta^2).$$

Recalling that transformation from $\mathbf{r}'$ to $\mathbf{r}''$ is also $\mathbf{B}(\boldsymbol{\beta} + \delta \boldsymbol{\beta}) \mathbf{B}(-\boldsymbol{\beta})$, we find that two consecutive Lorentz transformations is equal to a spatial rotation followed by a Lorentz transformation or a Lorentz transformation followed by a spatial rotation:

$$\mathbf{B}(\boldsymbol{\beta} + \delta \boldsymbol{\beta}) \mathbf{B}(-\boldsymbol{\beta}) = \mathbf{B}(\delta \boldsymbol{\beta}') \mathbf{R}(\delta \boldsymbol{\Omega}) = \mathbf{R}(\delta \boldsymbol{\Omega}) \mathbf{B}(\delta \boldsymbol{\beta}').$$

Appendix A Derivation of (1.13)

(1.13):

$$\frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} = \frac{\mathbf{v}}{c} \cdot \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}}, \quad (1.13)$$

The right hand side is

$$\frac{\mathbf{v}}{c} \cdot \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} = \frac{\mathbf{v}}{c} \cdot \frac{d\mathbf{v}/c}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} + \left(\frac{v}{c}\right)^2 \frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} \quad (A.1)$$

Noting that

$$\frac{dv^2}{dt} = \frac{d\mathbf{v} \cdot \mathbf{v}}{dt} = 2\mathbf{v} \cdot \frac{d\mathbf{v}}{dt},$$

(A.1) becomes

$$\frac{\mathbf{v}}{c} \cdot \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} = \frac{1}{2} \frac{1}{\sqrt{1 - (v/c)^2}} \frac{d(v/c)^2}{dt} + \left(\frac{v}{c}\right)^2 \frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}}. \quad (A.2)$$

Now

$$\frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} = \frac{1}{2} \left(1 - \left(\frac{v}{c}\right)^2\right)^{-3/2} \frac{d(v/c)^2}{dt}.$$

By multiplying $(1 - (v/c)^2)$ to both sides and exchange left side and right side, we find

$$\frac{1}{2} \frac{1}{\sqrt{1 - (v/c)^2}} \frac{d(v/c)^2}{dt} = \left(1 - \left(\frac{v}{c}\right)^2\right) \frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}}.$$

Inserting this to (A.2) yields (1.13).

$$\begin{aligned} \frac{\mathbf{v}}{c} \cdot \frac{d}{dt} \frac{\mathbf{v}/c}{\sqrt{1 - (v/c)^2}} &= \left(1 - \left(\frac{v}{c}\right)^2\right) \frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} + \left(\frac{v}{c}\right)^2 \frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}}, \\ &= \frac{d}{dt} \frac{1}{\sqrt{1 - (v/c)^2}} \end{aligned}$$

Appendix B Lorentz Transformation

In this appendix, we will derive the Lorentz transformation in 5 different ways, 1) wave equation, 2) constancy of the speed of light, 3) rotation analogy, 4) symmetry and transformation group, 5) vector notation. In the first four derivations, we consider only linear transformations of time t and a one dimensional spatial coordinate x . Justification for this will be provided in the forth derivation, Appendix B.4. For the final derivation, we will use three dimensional vectors for spatial coordinates.

Let's consider an inertial frame K with coordinates (ct, x) , and another inertial frame K' with coordinates (ct', x') . Suppose that K' is moving at velocity V relative to K and the origin of K' coincides with the origin of K at $t = 0$. A trivial condition is that a point $x = Vt$ is transformed to $x' = 0$. Therefore, the transformation for x can be written as

$$x' = \alpha (x - (V/c) ct).$$

Here, we will treat ct as a single symbol to ensure it has the same dimension as x (length). Now, we have only one parameter α for the spatial coordinate. For ct , we introduce two parameters β and γ , as follows:

$$ct' = \gamma (ct - \beta x)$$

Our task is to find three parameters, α , β and γ , under the given conditions.

B.1 Wave equation

We require that the condition that the wave equation

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} = 0.$$

is invariant under a coordinate transformation

$$x' = \alpha (x - (V/c) ct), \quad ct' = \gamma (ct - \beta x), \quad (\text{B.1})$$

where V is the velocity of K' relative to K . The first term of the wave equation transforms as

$$\begin{aligned} \frac{\partial^2}{\partial ct^2} &= \frac{\partial}{\partial ct} \left(\frac{\partial ct'}{\partial ct} \frac{\partial}{\partial ct'} + \frac{\partial x'}{\partial ct} \frac{\partial}{\partial x'} \right), \\ &= \left(\frac{\partial ct'}{\partial ct} \right)^2 \frac{\partial^2}{\partial ct'^2} + \left(\frac{\partial x'}{\partial ct} \right)^2 \frac{\partial^2}{\partial x'^2} + 2 \left(\frac{\partial ct'}{\partial ct} \right) \left(\frac{\partial x'}{\partial ct} \right) \frac{\partial^2}{\partial ct' \partial x'} \end{aligned} \quad (\text{B.2})$$

The second term of the wave equation transforms as

$$\begin{aligned}\frac{\partial^2}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial ct'}{\partial x} \frac{\partial}{\partial ct'} + \frac{\partial x'}{\partial x} \frac{\partial}{\partial x'} \right), \\ &= \left(\frac{\partial ct'}{\partial x} \right)^2 \frac{\partial^2}{\partial ct'^2} + \left(\frac{\partial x'}{\partial x} \right)^2 \frac{\partial^2}{\partial x'^2} + 2 \left(\frac{\partial ct'}{\partial x} \right) \left(\frac{\partial x'}{\partial x} \right) \frac{\partial^2}{\partial ct' \partial x'}\end{aligned}\quad (\text{B.3})$$

Here from Eqs. (B.1),

$$\left(\frac{\partial ct'}{\partial ct} \right) = \gamma, \quad \left(\frac{\partial x'}{\partial ct} \right) = -\alpha V/c, \quad \left(\frac{\partial ct'}{\partial x} \right) = -\gamma\beta, \quad \left(\frac{\partial x'}{\partial x} \right) = \alpha.$$

For the wave equation to be Lorentz invariant, we require Eq. (B.2) minus Eq. (B.3) becomes

$$\frac{\partial^2}{\partial ct^2} - \frac{\partial^2}{\partial x^2} = \frac{\partial^2}{\partial ct'^2} - \frac{\partial^2}{\partial x'^2}.$$

In other words, we require the coefficient of $\partial^2/\partial ct'^2$, $\partial^2/\partial x'^2$ and $\partial^2/\partial ct' \partial x'$ to be 1, -1 and 0, respectively. That is

$$\gamma^2 - \gamma^2 \beta^2 = 1, \quad \alpha^2 (V/c)^2 - \alpha^2 = -1, \quad -\gamma\alpha V/c + \gamma\beta\alpha = 0.$$

The solution is

$$\beta = V/c, \quad \gamma^2 = \alpha^2 = \frac{1}{1 - \beta^2}.$$

Because we want the transformation to be the identity when $V = 0$, we take positive values for both α and γ .

$$\gamma = \alpha = \frac{1}{\sqrt{1 - (V/c)^2}}$$

B.2 Constancy of the speed of light

We begin with transformation of (cdt, dx) .

$$dx' = \alpha (dx - (V/c) cdt), \quad cdt' = \gamma (cdt - \beta dx).$$

We want to find α , β , γ which keeps the interval $(cdt)^2 - (dx)^2$ constant. They are calculated as follows:

$$\begin{aligned}(cdt')^2 &= \gamma^2 ((cdt)^2 + \beta^2 (dx)^2 - 2\beta cdt dx), \\ (dx')^2 &= \alpha^2 ((dx)^2 + (V/c)^2 (cdt)^2 - 2(V/c) cdt dx).\end{aligned}$$

The interval in K' is

$$\begin{aligned}(cdt')^2 - (dx')^2 &= (\gamma^2 - \alpha^2 (V/c)^2) (cdt)^2 - (\alpha^2 - \gamma^2 \beta^2) (dx)^2 \\ &\quad + 2(-\gamma^2 \beta + \alpha^2 V/c) cdt dx.\end{aligned}$$

This must be equal to the interval in K , $(cdt)^2 - (dx)^2$. Therefore

$$\gamma^2 - \alpha^2(V/c)^2 = 1, \quad \alpha^2 - \gamma^2\beta^2 = 1, \quad -\gamma^2\beta + \alpha^2 V/c = 0.$$

The solution is

$$\beta = V/c, \quad \gamma^2 = \alpha^2 = \frac{1}{1 - \beta^2}.$$

Because we want the transformation to be the identity when $V = 0$, we take positive values for both α and γ .

$$\gamma = \alpha = \frac{1}{\sqrt{1 - (V/c)^2}}$$

B.3 Rotation analogy

In a 2-dimensional vector space, it is rotation that keeps $x^2 + y^2$ constant.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Here, we want to find a transformation that keeps $(ct)^2 - x^2$ constant. Let's consider the following transformation.

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh \theta & \sinh \theta \\ \sinh \theta & \cosh \theta \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \tag{B.4}$$

One can easily verify that this transformation indeed gives $(ct)^2 - x^2 = (ct')^2 - x'^2$.

Let (ct, x) be the coordinate of the origin of K' in K . Since the coordinate of the origin of K' in K' is $(ct', 0)$,

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh \theta & \sinh \theta \\ \sinh \theta & \cosh \theta \end{pmatrix} \begin{pmatrix} ct' \\ 0 \end{pmatrix}.$$

Therefore

$$ct = ct' \cosh \theta, \quad x = ct' \sinh \theta$$

Since O' is moving at V in K , i.e., $V = x/t$,

$$\frac{V}{c} = \frac{x}{ct} = \frac{\sinh \theta}{\cosh \theta} = \tanh \theta.$$

Recalling that $\cosh^2 \theta - \sinh^2 \theta = 1$, we find

$$\cosh \theta = \frac{1}{\sqrt{1 - \tanh^2 \theta}}, \quad \sinh \theta = \frac{\tanh \theta}{\sqrt{1 - \tanh^2 \theta}}.$$

Therefore

$$\cosh \theta = \frac{1}{\sqrt{1 - (V/c)^2}}, \quad \sinh \theta = \frac{V/c}{\sqrt{1 - (V/c)^2}}.$$

Inserting these to Eq. (B.4),

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{1}{\sqrt{1 - (V/c)^2}} \begin{pmatrix} 1 & V/c \\ V/c & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}.$$

The inverse transformation will be

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \frac{1}{\sqrt{1 - (V/c)^2}} \begin{pmatrix} 1 & -V/c \\ -V/c & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}.$$

B.4 Symmetry and transformation group

B.4.1 Introduction

In this appendix, we derive the Lorentz transformation by considering the symmetry of spacetime and transformation group. It was constancy of the speed of light which brought the idea that time must be transformed along with the spatial coordinates. Once we accept this idea, we can demonstrate that the transformation from one inertial frame to another is either a Lorentz transformation or a Galilean transformation, based on a few assumption of space, time and the transformation itself. Those are 1) isotropy of space, 2) homogeneity of space and time, 3) the transformation forms a transformation group. [8, 9]

The outline of the derivation is as follows: We first show that the transformation is linear and separates into (t, x) and (y, z) coordinates from the isotropy of space and the homogeneity of space and time, both of which were implicitly assumed in the previous derivations in this Appendix. Then, the reciprocity of the transformation and combined transformation are applied to determine the matrix element of the transformation. The resulting transformation includes an arbitrary parameter K , which is identified as $1/c^2$ for the transformation to be the Lorentz transformation.

B.4.2 Derivation

Consider a Cartesian coordinate system S' moving at a velocity v relative to another Cartesian coordinate system S . Suppose that the x, y, z axes of S and the x', y', z' axes of S' coincides at $t = 0$ and the transformation of spatial and time coordinate

(t, x, y, z) of S to spatial and time coordinate (t', x', y', z') of S' is written as follows:

$$\begin{aligned} t' &= T(t, x, y, z; v), \\ x' &= X(t, x, y, z; v), \\ y' &= Y(t, x, y, z; v), \\ z' &= Z(t, x, y, z; v). \end{aligned}$$

The transformation should be the identity when $v = 0$.

$$\begin{aligned} T(t, x, y, z; 0) &= t, \\ X(t, x, y, z; 0) &= x, \\ Y(t, x, y, z; 0) &= y, \\ Z(t, x, y, z; 0) &= z. \end{aligned} \tag{B.5}$$

From the isotropy of space, the directions of axes are arbitrary as long as they are orthogonal each other. However, since we set the direction of relative velocity v to x -axis, we only consider rotations that will not alter the direction of x -axis. If we rotate the y and z axes of S around the x -axis by θ , the coordinates of a point (x, y, z) will become $(x, y \cos \theta + z \sin \theta, -y \sin \theta + z \cos \theta)$. The same thing can be said for S' . Therefore

$$\begin{aligned} T(t, x, y \cos \theta + z \sin \theta, -y \sin \theta + z \cos \theta; v) &= t', \\ X(t, x, y \cos \theta + z \sin \theta, -y \sin \theta + z \cos \theta; v) &= x', \\ Y(t, x, y \cos \theta + z \sin \theta, -y \sin \theta + z \cos \theta; v) &= y' \cos \theta + z' \sin \theta, \\ Z(t, x, y \cos \theta + z \sin \theta, -y \sin \theta + z \cos \theta; v) &= -y' \sin \theta + z' \cos \theta. \end{aligned}$$

In case $\theta = 180^\circ$,

$$\begin{aligned} T(t, x, -y, -z; v) &= t', \\ X(t, x, -y, -z; v) &= x', \\ Y(t, x, -y, -z; v) &= -y', \\ Z(t, x, -y, -z; v) &= -z'. \end{aligned} \tag{B.6}$$

Now, let's consider a rotation around the z -axis. This time, we can only perform a 180° rotation since we don't want to alter the direction of the x -axis. If we rotate the x and y axes of S around the z -axis by 180° , the coordinates of a point (x, y, z) will become $(-x, -y, z)$. The same can be said for S' . This time, the x -axis got inverted, the velocity v must also be inverted.

$$\begin{aligned} T(t, -x, -y, z; -v) &= t', \\ X(t, -x, -y, z; -v) &= -x', \\ Y(t, -x, -y, z; -v) &= -y', \\ Z(t, -x, -y, z; -v) &= z'. \end{aligned} \tag{B.7}$$

From the homogeneity of space and time, all functions have translation symmetry. For example, for two time points t_2 and t_1 ,

$$T(t_2, x, y, z) - T(t_1, x, y, z) = T(t_2 + t, x, y, z) - T(t_1 + t, x, y, z),$$

for any t . Rearranging above yields

$$T(t_1 + t, x, y, z) - T(t_1, x, y, z) = T(t_2 + t, x, y, z) - T(t_2, x, y, z).$$

Since this is true for any t_1 and t_2 , T must be a linear function of t .

$$nT(t) = at + \text{const.}$$

The same applies to all coordinates (t, x, y, z) and functions (T, X, Y, Z) . Therefore, the transformation is linear and it can be written in a matrix form as follows:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{tt}(v) & a_{tx}(v) & a_{ty}(v) & a_{tz}(v) \\ a_{xt}(v) & a_{xx}(v) & a_{xy}(v) & a_{xz}(v) \\ a_{yt}(v) & a_{yx}(v) & a_{yy}(v) & a_{yz}(v) \\ a_{zt}(v) & a_{zx}(v) & a_{zy}(v) & a_{zz}(v) \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

From Eq. (B.6), we have

$$\begin{pmatrix} t' \\ x' \\ -y' \\ -z' \end{pmatrix} = \begin{pmatrix} a_{tt}(v) & a_{tx}(v) & a_{ty}(v) & a_{tz}(v) \\ a_{xt}(v) & a_{xx}(v) & a_{xy}(v) & a_{xz}(v) \\ a_{yt}(v) & a_{yx}(v) & a_{yy}(v) & a_{yz}(v) \\ a_{zt}(v) & a_{zx}(v) & a_{zy}(v) & a_{zz}(v) \end{pmatrix} \begin{pmatrix} t \\ x \\ -y \\ -z \end{pmatrix}$$

By comparing the above two expressions, we see some of coefficients must be zero.

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{tt}(v) & a_{tx}(v) & 0 & 0 \\ a_{xt}(v) & a_{xx}(v) & 0 & 0 \\ 0 & 0 & a_{yy}(v) & a_{yz}(v) \\ 0 & 0 & a_{zy}(v) & a_{zz}(v) \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

From Eq. (B.7), we have

$$\begin{pmatrix} t' \\ -x' \\ -y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{tt}(-v) & a_{tx}(-v) & 0 & 0 \\ a_{xt}(-v) & a_{xx}(-v) & 0 & 0 \\ 0 & 0 & a_{yy}(-v) & a_{yz}(-v) \\ 0 & 0 & a_{zy}(-v) & a_{zz}(-v) \end{pmatrix} \begin{pmatrix} t \\ -x \\ -y \\ z \end{pmatrix}$$

By comparing the above two expressions, we see that

$$\begin{aligned} a_{yy}(v) &= a_{zz}(v) = a_{yy}(-v) = a_{zz}(-v), \\ a_{yz}(v) &= a_{zy}(v) = a_{yz}(-v) = a_{zy}(-v) = 0, \\ a_{tt}(-v) &= a_{tt}(v), \quad a_{tx}(-v) = -a_{tx}(v), \\ a_{xt}(-v) &= -a_{tx}(v), \quad a_{xx}(-v) = a_{xx}(v). \end{aligned}$$

Recalling that S' is moving at v along the x -axis of S and the origin of S' coincides with the origin of S at $t = 0$. That is, when $x = vt$, then $x' = 0$, or

$$x' = \alpha(v) (x - vt),$$

where

$$\alpha(v) = \alpha(-v) = a_{xx}(v).$$

Therefore

$$a_{xt} = -v\alpha(v).$$

Let's say

$$a_{tt}(v) = \gamma(v), \quad a_{tx}(v) = -\gamma(v)\zeta(v), \quad a_{yy}(v) = a_{zz}(v) = \phi(v).$$

We see

$$\begin{aligned} \gamma(v) &= \gamma(-v), & -\gamma(v)\zeta(v) &= \gamma(-v)\zeta(-v), \\ \zeta(v) &= -\zeta(-v), & \phi(v) &= \phi(-v), \end{aligned}$$

and

$$t' = \gamma(v) (t - \zeta(v) x).$$

Now we can write the transformation in two 2-dimensional expressions.

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} \alpha(v) & -v\alpha(v) \\ -\gamma(v)\zeta(v) & \gamma(v) \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}, \quad \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} \phi(v) & 0 \\ 0 & \phi(v) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} \quad (\text{B.8})$$

Let's assume the reciprocity of the transformation. That means the transformation from S' to S moving at velocity $-v$ relative to S' will be

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \alpha(-v) & v\alpha(-v) \\ -\gamma(-v)\zeta(-v) & \gamma(-v) \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}, \quad \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \phi(-v) & 0 \\ 0 & \phi(-v) \end{pmatrix} \begin{pmatrix} y' \\ z' \end{pmatrix}$$

Recalling that $\alpha(v) = \alpha(-v)$, $\gamma(v) = \gamma(-v)$, $\zeta(v) = -\zeta(-v)$, $\phi(v) = \phi(-v)$, we find

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \alpha(v) & v\alpha(v) \\ \gamma(v)\zeta(v) & \gamma(v) \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}, \quad \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \phi(v) & 0 \\ 0 & \phi(v) \end{pmatrix} \begin{pmatrix} y' \\ z' \end{pmatrix}.$$

Plugging these into Eqs. (B.8), we find (omitting (v) for shorthand)

$$\begin{pmatrix} \alpha & -v\alpha \\ -\gamma\zeta & \gamma \end{pmatrix} \begin{pmatrix} \alpha & v\alpha \\ \gamma\zeta & \gamma \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \phi & 0 \\ 0 & \phi \end{pmatrix} \begin{pmatrix} \phi & 0 \\ 0 & \phi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore

$$\alpha^2 - v\alpha\gamma\zeta = 1, \quad v\alpha^2 - v\alpha\gamma = 0 \quad \rightarrow \quad \alpha = \gamma, \quad \gamma^2(1 - v\zeta) = 1, \quad (\text{B.9})$$

$$\phi^2 = 1 \quad \rightarrow \quad \phi = \pm 1. \quad (\text{B.10})$$

The transformation should be the identity when $v = 0$, (Eq. (B.5)), we should choose $\phi = 1$. To summarize what we got so far,

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \gamma(v) \begin{pmatrix} 1 & -v \\ -\zeta(v) & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}, \quad \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}.$$

Now we require that two consecutive transformations yield the same class of transformation. That is

$$\gamma(u) \begin{pmatrix} 1 & -u \\ -\zeta(u) & 1 \end{pmatrix} \gamma(v) \begin{pmatrix} 1 & -v \\ -\zeta(v) & 1 \end{pmatrix} = \gamma(w) \begin{pmatrix} 1 & -w \\ -\zeta(w) & 1 \end{pmatrix}.$$

Proceeding with the calculation yields

$$\gamma(u) \gamma(v) \begin{pmatrix} 1 + u \zeta(v) & -u - v \\ -\zeta(u) - \zeta(v) & 1 + v \zeta(u) \end{pmatrix} = \gamma(w) \begin{pmatrix} 1 & -w \\ -\zeta(w) & 1 \end{pmatrix}. \quad (\text{B.11})$$

From the diagonal elements,

$$\begin{aligned} \gamma(u) \gamma(v) (1 + u \zeta(v)) &= \gamma(w), & \gamma(u) \gamma(v) (1 + v \zeta(u)) &= \gamma(w), \\ \rightarrow 1 + u \zeta(v) &= 1 + v \zeta(u) = \gamma(w) / \gamma(u) \gamma(v), \\ \rightarrow u \zeta(v) &= v \zeta(u) \rightarrow \zeta(v)/v = \zeta(u)/u, \end{aligned}$$

we see that $\zeta(v)/v$ is constant because $\zeta(v)/v = \zeta(u)/u$ for any v and u . Let's say $\zeta(v)/v = K$. From Eq. (B.9),

$$\gamma(v)^2 (1 - v \zeta(v)) = 1 \quad \rightarrow \quad \gamma(v) = \frac{1}{\sqrt{1 - v^2 K}}$$

Therefore, the transformation, Eq. (B.8), becomes

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \frac{1}{\sqrt{1 - v^2 K}} \begin{pmatrix} 1 & -v \\ -vK & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}, \quad \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} \quad (\text{B.12})$$

From Eq. (B.11),

$$\frac{1 + u v K}{\sqrt{(1 - u^2 K)(1 - v^2 K)}} \begin{pmatrix} 1 & -\frac{u + v}{1 + u v K} \\ -\frac{(u + v)K}{1 + u v K} & 1 \end{pmatrix} = \frac{1}{\sqrt{1 - w^2 K}} \begin{pmatrix} 1 & -w \\ -wK & 1 \end{pmatrix}.$$

Now the inverse of the forefront factor of above is

$$\begin{aligned} \frac{\sqrt{(1 - u^2 K)(1 - v^2 K)}}{1 + u v K} &= \sqrt{\frac{1 + u^2 v^2 K^2 - (u^2 + v^2)K}{1 + u^2 v^2 K^2 + 2 u v K}}, \\ &= \sqrt{\frac{1 + u^2 v^2 K^2 + 2 u v K - (u^2 + v^2 + 2 u v)K}{1 + u^2 v^2 K^2 + 2 u v K}}, \\ &= \sqrt{1 - \frac{(u + v)^2 K}{(1 + u v K)^2}}. \end{aligned}$$

Therefore

$$w = \frac{u + v}{1 + uvK}. \quad (\text{B.13})$$

In the case where $K = 0$, above becomes Newtonian velocity addition law.

$$w = u + v$$

In the case where $K > 0$, Eq. (B.13) becomes the relativistic velocity addition law with $K = 1/c^2$.

$$w = \frac{u + v}{1 + uv/c^2}$$

Finally

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \frac{1}{\sqrt{1 - (v/c)^2}} \begin{pmatrix} 1 & -v \\ -v/c^2 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}, \quad \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

or using ct instead of t makes the matrix symmetric and the both element of the vector have the same dimension.

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \frac{1}{\sqrt{1 - (v/c)^2}} \begin{pmatrix} 1 & -v/c \\ -v/c & 1 \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix}, \quad \begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}.$$

B.5 Mermin's derivation of velocity addition law

Mermin [10] demonstrated the way to observe the violation of classical velocity addition law, and derived the relativistic one without using the constancy of the speed of light. In this case, it does not have to be light; any object fast enough can demonstrate the violation of classical velocity addition law. Mermin's derivation also includes an arbitrary parameter K , which is also identified as $1/c^2$.

B.6 Vector notation

Let's consider an inertial frame K with coordinates $(ct, \mathbf{r})$, and another inertial frame K' with coordinates $(ct', \mathbf{r}')$. Suppose that K' is moving at velocity $\mathbf{V}$ relative to K and the origin of K' coincides with the origin of K at $t = 0$ and $t' = 0$. One trivial condition is that a point $\mathbf{r} = \mathbf{V}t$ is transformed to $\mathbf{r}' = 0$. From the homogeneity of space and time, we see that the transformation must be linear. From the isotropy of space, we see that vector components parallel to $\mathbf{V}$ are scaled differently than vector components normal to $\mathbf{V}$. Therefore the transformation should be written with real parameters $\eta, \alpha, \gamma, \zeta$,

$$\mathbf{r}' = \eta \left(1 - \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot)}{\beta^2} \right) (\mathbf{r} - \boldsymbol{\beta} ct) + \alpha \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot)}{\beta^2} (\mathbf{r} - \boldsymbol{\beta} ct), \quad ct' = \gamma \left(ct - \zeta \frac{(\boldsymbol{\beta} \cdot)}{\beta} \mathbf{r} \right),$$

where we used $\boldsymbol{\beta} = \mathbf{V}/c$ for shorthand. Note that

$$\frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2}$$

is a projection operator which projects a vector onto the direction of $\mathbf{V}$.

$$\frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2} \mathbf{r} = \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2}$$

Proceeding with the calculation yields,

$$\mathbf{r}' = \eta \left(1 - \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2} \right) \mathbf{r} + \alpha \left(\frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2} - \boldsymbol{\beta} ct \right), \quad ct' = \gamma \left(ct - \zeta \frac{(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta} \right),$$

or

$$\mathbf{r}' = \eta \mathbf{r}_\perp + \alpha (\mathbf{r}_\parallel - \boldsymbol{\beta} ct), \quad ct' = \gamma \left(ct - \zeta \frac{(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta} \right),$$

where

$$\mathbf{r}_\perp = \left(1 - \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2} \right) \mathbf{r}, \quad \mathbf{r}_\parallel = \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2}.$$

Now we invoke the condition that interval is invariant, $(ct')^2 - (r')^2 = (ct)^2 - r^2$.

$$\begin{aligned} (ct')^2 &= \gamma^2 \left((ct)^2 + \zeta^2 \frac{(\boldsymbol{\beta} \cdot \mathbf{r})^2}{\beta^2} - 2 ct \zeta \frac{(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta} \right), \\ &= \gamma^2 (ct)^2 + \gamma^2 \zeta^2 \frac{(\boldsymbol{\beta} \cdot \mathbf{r})^2}{\beta^2} - 2 ct \gamma^2 \zeta \frac{(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta}, \\ (r')^2 &= \eta^2 r_\perp^2 + \alpha^2 (\mathbf{r}_\parallel - \boldsymbol{\beta} ct)^2, \\ &= \eta^2 \left(\mathbf{r} - \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2} \right)^2 + \alpha^2 r_\parallel^2 + \alpha^2 \beta^2 (ct)^2 - 2 ct \alpha^2 (\boldsymbol{\beta} \cdot \mathbf{r}_\parallel), \\ &= \alpha^2 \beta^2 (ct)^2 + \eta^2 r^2 + (\alpha^2 - \eta^2) \frac{(\boldsymbol{\beta} \cdot \mathbf{r})^2}{\beta^2} - 2 ct \alpha^2 \beta \frac{(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta} \end{aligned}$$

To make $(ct')^2 - (r')^2$ equal to $(ct)^2 - r^2$, we want

$$\gamma^2 - \alpha^2 \beta^2 = 1, \quad \eta^2 = 1, \quad \gamma^2 \zeta^2 - \alpha^2 + \eta^2 = 0, \quad \gamma^2 \zeta - \alpha^2 \beta = 0. \quad (\text{B.14})$$

From the last equation, we find

$$\alpha^2 = \frac{\gamma^2 \zeta}{\beta}.$$

Inserting this and $\eta^2 = 1$ into the first and the third equations of Eqs. (B.14), we find

$$\zeta \beta = 1 - \frac{1}{\gamma^2}, \quad \gamma^2 \zeta \beta (1 - \zeta \beta) = \beta^2.$$

Inserting the first equation to the second, we find

$$\gamma^2 = \frac{1}{1 - \beta^2}, \quad \zeta = \beta, \quad \alpha^2 = \gamma^2, \quad \eta^2 = 1.$$

Since the transformation has to be the identity when $\mathbf{V} = 0$, we choose positive values.

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}}, \quad \alpha = \gamma, \quad \eta = 1.$$

Finally the transformation is

$$\mathbf{r}' = \mathbf{r}_\perp + \gamma (\mathbf{r}_\parallel - \beta ct), \quad ct' = \gamma (ct - (\boldsymbol{\beta} \cdot \mathbf{r})),$$

where

$$\mathbf{r}_\perp = (\mathbf{r} - \mathbf{r}_\parallel), \quad \mathbf{r}_\parallel = \frac{\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{r})}{\beta^2}, \quad \boldsymbol{\beta} = \mathbf{V}/c, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}.$$

Appendix C Details of (1.50)

We start from this.

$$\mathbf{f} = \left(\gamma \mathbf{f}'_\perp + \gamma^2 \mathbf{f}'_\parallel \right) \left(1 - \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) \right) + \gamma^2 \boldsymbol{\beta} \left(\left(\frac{\mathbf{v}_\perp/\gamma}{c} + \frac{\mathbf{v}_\parallel}{c} - \boldsymbol{\beta} \right) \cdot \mathbf{f}' \right)$$

The first term,

$$\begin{aligned} \left(\gamma \mathbf{f}'_\perp + \gamma^2 \mathbf{f}'_\parallel \right) \left(1 - \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) \right) &= \gamma \mathbf{f}'_\perp + \gamma^2 \mathbf{f}'_\parallel - \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) \gamma \mathbf{f}'_\perp - \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) \gamma^2 \mathbf{f}'_\parallel, \\ &= \gamma \mathbf{f}'_\perp + \gamma^2 \mathbf{f}'_\parallel - \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) \gamma \mathbf{f}' - \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) (\gamma^2 - \gamma) \mathbf{f}'_\parallel \end{aligned}$$

The second term,

$$\begin{aligned} \gamma^2 \boldsymbol{\beta} \left(\left(\frac{\mathbf{v}_\perp/\gamma}{c} + \frac{\mathbf{v}_\parallel}{c} - \boldsymbol{\beta} \right) \cdot \mathbf{f}' \right) &= \boldsymbol{\beta} \left(\left(\gamma \frac{\mathbf{v}_\perp}{c} + \gamma^2 \frac{\mathbf{v}_\parallel}{c} - \gamma^2 \boldsymbol{\beta} \right) \cdot \mathbf{f}' \right), \\ &= \boldsymbol{\beta} \left(\left(\gamma \frac{\mathbf{v}}{c} + (\gamma^2 - \gamma) \frac{\mathbf{v}_\parallel}{c} - \gamma^2 \boldsymbol{\beta} \right) \cdot \mathbf{f}' \right), \\ &= \boldsymbol{\beta} \left(\left(\frac{\mathbf{v}}{c} \cdot \gamma \mathbf{f}' \right) + \frac{(\boldsymbol{\beta} \cdot \mathbf{v}/c)}{\beta^2} (\gamma^2 - \gamma) (\boldsymbol{\beta} \cdot \mathbf{f}') - \gamma^2 (\boldsymbol{\beta} \cdot \mathbf{f}') \right), \\ &= \boldsymbol{\beta} \left(\frac{\mathbf{v}}{c} \cdot \gamma \mathbf{f}' \right) + (\gamma^2 - \gamma) \left(\boldsymbol{\beta} \cdot \frac{\mathbf{v}}{c} \right) \mathbf{f}'_\parallel - \gamma^2 \beta^2 \mathbf{f}'_\parallel. \end{aligned}$$

Adding them together, noting $\gamma^2(1 - \beta^2) = 1$,

$$\mathbf{f} = \mathbf{f}'_{\parallel} + \gamma \mathbf{f}'_{\perp} + \left(\frac{\mathbf{v}}{c} \cdot \gamma \mathbf{f}' \right) \boldsymbol{\beta} - \left(\frac{\mathbf{v}}{c} \cdot \boldsymbol{\beta} \right) \gamma \mathbf{f}'.$$

Using $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}$, we find

$$\mathbf{f} = \mathbf{f}'_{\parallel} + \gamma \mathbf{f}'_{\perp} + \frac{\mathbf{v}}{c} \times (\boldsymbol{\beta} \times \gamma \mathbf{f}'). \quad (1.50)$$

Appendix D Vectors and tensors

We follow the notation used in the Feynman's Lectures on Physics [1].

D.1 Vectors

We use following notation for ordinary vectors. We use $i, j, k, \dots$ for index x, y, z .

$$\mathbf{a} = a_i = (a_x \quad a_y \quad a_z)$$

Below is inner product of ordinary vectors. Summation rule⁵ applies.

$$\mathbf{a} \cdot \mathbf{b} = a_i b_i = a_x b_x + a_y b_y + a_z b_z$$

We use following notation for four-vectors. We use $\mu, \nu, \rho, \dots$ for index t, x, y, z .

$$\mathbf{a} = (a_t, \mathbf{a}) = a_{\mu} = (a_t \quad a_x \quad a_y \quad a_z)$$

Below is inner product of four-vectors. Summation rule applies.

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= (a_t, \mathbf{a}) \cdot (b_t, \mathbf{b}), \\ &= a_{\mu} b_{\mu}, \\ &= a_t b_t - \mathbf{a} \cdot \mathbf{b}, \\ &= a_t b_t - a_x b_x - a_y b_y - a_z b_z \end{aligned}$$

D.2 Tensors

Direct product of two vectors is called second-rank tensor.

$$T_{\mu\nu} = a_{\mu} b_{\nu}$$

A vector is a first-rank tensor. There are third and higher rank tensors, but we will not use them. Here we talk about second-rank tensors. In fact, we only need a special

⁵When a term has a repeated index, sum them over the index. Like shown in the text.

type of tensor: anti-symmetric tensor. That is covered in the next section. Readers in a hurry can skip this section and go to D.3 directly.

It appears a linear combination of those direct products is also a tensor.

$$T_{\mu\nu} = \sum_n a_{n,\mu} b_{n,\nu}$$

A tensor whose 16 elements are given can be represented as follows.

$$T_{\mu\nu} = \sum_{lm} T_{lm} e_{l,\mu} e_{m,\nu}$$

where l and m are one of t, x, y, z and $e_{l,\mu}$ is the base unit vector:

$$e_{t,\mu} = (1 \ 0 \ 0 \ 0), \quad e_{x,\mu} = (0 \ 1 \ 0 \ 0), \quad \dots$$

We take usual sum for l and m . Four-vector summation rule only applies to $\mu, \nu, \dots$ Therefore, inner products are calculated as follows.

$$\begin{aligned} u_\mu e_{t,\mu} &= u_t e_{t,t} - u_x e_{t,x} - u_y e_{t,y} - u_z e_{t,z} = u_t, \\ u_\mu e_{x,\mu} &= u_t e_{x,t} - u_x e_{x,x} - u_y e_{x,y} - u_z e_{x,z} = -u_x, \\ u_\mu e_{y,\mu} &= u_t e_{y,t} - u_x e_{y,x} - u_y e_{y,y} - u_z e_{y,z} = -u_y, \\ u_\mu e_{z,\mu} &= u_t e_{z,t} - u_x e_{z,x} - u_y e_{z,y} - u_z e_{z,z} = -u_z. \end{aligned}$$

Giving four-vector u_μ to $T_{\mu\nu}$,

$$\begin{aligned} u_\nu T_{\mu\nu} &= \sum_{lm} u_\nu e_{l,\mu} e_{m,\nu} T_{lm} \\ &= \sum_l e_{l,\mu} (u_\nu e_{t,\nu} T_{lt} + u_\nu e_{x,\nu} T_{lx} + u_\nu e_{y,\nu} T_{ly} + u_\nu e_{z,\nu} T_{lz}), \\ &= \sum_l e_{l,\mu} (u_t T_{lt} - u_x T_{lx} - u_y T_{ly} - u_z T_{lz}), \\ &= u_t T_{\mu t} - u_x T_{\mu x} - u_y T_{\mu y} - u_z T_{\mu z}. \end{aligned}$$

Similarly,

$$\begin{aligned} T_{\mu\nu} u_\nu &= T_{\mu t} u_t - T_{\mu x} u_x - T_{\mu y} u_y - T_{\mu z} u_z, \\ u_\nu T_{\nu\mu} &= u_t T_{t\mu} - u_x T_{x\mu} - u_y T_{y\mu} - u_z T_{z\mu}. \end{aligned}$$

Scalar product:

$$\begin{aligned}
T_{\mu\nu}U_{\mu\nu} &= \sum_{lm} T_{lm} e_{l,\mu} e_{m,\nu} \sum_{l'm'} U_{l'm'} e_{l',\mu} e_{m',\nu}, \\
&= \sum_{lm} \sum_{l'm'} T_{lm} U_{l'm'} e_{l,\mu} e_{l',\mu} e_{m,\nu} e_{m',\nu}, \\
&= \sum_{lm} \sum_{l'} T_{lm} U_{l'm} e_{l,\mu} e_{l',\mu} e_{m,\nu} (e_{t,\nu} + e_{x,\nu} + e_{y,\nu} + e_{z,\nu}), \\
&= \sum_{lm} \sum_{l'} T_{lm} U_{l'm} e_{l,\mu} e_{l',\mu} (\delta_{tm} - \delta_{xm} - \delta_{ym} - \delta_{zm}), \\
&= \sum_{lm} T_{lm} U_{lm} (\delta_{tl} - \delta_{xl} - \delta_{yl} - \delta_{zl}) (\delta_{tm} - \delta_{xm} - \delta_{ym} - \delta_{zm}). \tag{D.1}
\end{aligned}$$

D.3 Anti-symmetric tensor

An anti-symmetric tensor is a second-rank tensor whose component flips its sign when its index order is flipped, as follows.

$$A_{\mu\nu} = -A_{\nu\mu}$$

It has six independent components.

$$A_{\mu\nu} = \begin{pmatrix} 0 & A_{tx} & A_{ty} & A_{tz} \\ A_{xt} & 0 & A_{xy} & A_{xz} \\ A_{yt} & A_{yx} & 0 & A_{yz} \\ A_{zt} & A_{zx} & A_{zy} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -A_{xt} & -A_{yt} & -A_{zt} \\ A_{xt} & 0 & -A_{yx} & A_{xz} \\ A_{yt} & A_{yx} & 0 & -A_{zy} \\ A_{zt} & -A_{xz} & A_{zy} & 0 \end{pmatrix}$$

Let us work on the case when it is written with two four-vectors as below.

$$A_{\mu\nu} = a_\mu b_\nu - a_\nu b_\mu$$

Because any anti-symmetric tensor can be written with a linear combination of these, extending below discussion to general case is straight forward.

D.3.1 Two ordinary vectors

$\mathbf{p} = (A_{xt} \ A_{yt} \ A_{zt})$ forms an ordinary vector because it is linear combination of two ordinary vectors (or linear combination of those).

$$\begin{aligned}
\mathbf{p} &= (A_{xt} \ A_{yt} \ A_{zt}) \\
&= (a_x b_t - a_t b_x \quad a_y b_t - a_t b_y \quad a_z b_t - a_t b_z), \\
&= (a_x \quad a_y \quad a_z) b_t - a_t (b_x \quad b_y \quad b_z), \\
&= \mathbf{a} b_t - a_t \mathbf{b}
\end{aligned}$$

$\mathbf{q} = (A_{zy} \ A_{xz} \ A_{yx})$ forms an ordinary vector too, because it is vector product of two ordinary vectors (or linear combination of those).

$$\begin{aligned}\mathbf{q} &= (A_{zy} \ A_{xz} \ A_{yx}), \\ &= ((a_z b_y - a_y b_z) \ (a_x b_z - a_z b_x) \ (a_y b_x - a_x b_y)), \\ &= -\mathbf{a} \times \mathbf{b}\end{aligned}$$

We use $(\mathbf{p}, \mathbf{q})$ to represent an anti-symmetric tensor $\mathbf{A} = A_{\mu\nu}$.

$$(\mathbf{p}, \mathbf{q}) = \mathbf{A} = A_{\mu\nu} = \begin{pmatrix} 0 & -p_x & -p_y & -p_z \\ p_x & 0 & -q_z & q_y \\ p_y & q_z & 0 & -q_x \\ p_z & -q_y & q_x & 0 \end{pmatrix} \quad (\text{D.2})$$

When both $\mathbf{a}$ and $\mathbf{b}$ are polar, $\mathbf{p}$ is polar and $\mathbf{q}$ is axial.

Note that, in our notation, when it is a pair of a scalar and an ordinary vector such as $(a_t, \mathbf{a})$, it is a four-vector and that when it is a pair of two ordinary vectors such as $(\mathbf{p}, \mathbf{q})$, it is an anti-symmetric tensor.

D.3.2 Linear map

Giving a four-vector $\mathbf{w}$ to an anti-symmetric tensor $\mathbf{A}$ yields a four-vector. We write this as follows.

$$\begin{aligned}\mathbf{wA} &= (w_t, \mathbf{w}) (\mathbf{p}, \mathbf{q}) = w_\nu A_{\mu\nu} \\ &= w_t A_{\mu t} - w_x A_{\mu x} - w_y A_{\mu y} - w_z A_{\mu z}.\end{aligned}$$

t, x, y, z components are, respectively,

$$\begin{aligned}w_\nu A_{t\nu} &= w_t A_{tt} - w_x A_{tx} - w_y A_{ty} - w_z A_{tz}, \\ &= w_x p_x + w_y p_y + w_z p_z = \mathbf{w} \cdot \mathbf{p}, \\ w_\nu A_{x\nu} &= w_t A_{xt} - w_x A_{xx} - w_y A_{xy} - w_z A_{xz}, \\ &= w_t p_x + w_y q_z - w_z q_y, \\ w_\nu A_{y\nu} &= w_t A_{yt} - w_x A_{yx} - w_y A_{yy} - w_z A_{yz}, \\ &= w_t p_y + w_z q_x - w_x q_z, \\ w_\nu A_{z\nu} &= w_t A_{zt} - w_x A_{zx} - w_y A_{zy} - w_z A_{zz}, \\ &= w_t p_z + w_x q_y - w_y q_x.\end{aligned}$$

Therefore

$$\mathbf{wA} = (w_t, \mathbf{w}) (\mathbf{p}, \mathbf{q}) = w_\nu A_{\mu\nu} = (\mathbf{w} \cdot \mathbf{p}, \ w_t \mathbf{p} + \mathbf{w} \times \mathbf{q}). \quad (\text{D.3})$$

D.3.3 Lorentz transformation

Let us use a Cartesian coordinate whose x-axis coincides with $\boldsymbol{\beta}$, i.e., $\boldsymbol{\beta} = (\beta_x \ 0 \ 0)$. The Lorentz transformation of $\mathbf{a}$ and $\mathbf{b}$ is, respectively,

$$a'_x = \gamma(a_x - \beta_x a_t), \quad a'_y = a_y, \quad a'_z = a_z, \quad a'_t = \gamma(a_t - \beta_x a_x),$$

and

$$b'_x = \gamma(b_x - \beta_x b_t), \quad b'_y = b_y, \quad b'_z = b_z, \quad b'_t = \gamma(b_t - \beta_x b_x).$$

For $\mathbf{p} = (A_{xt} \ A_{yt} \ A_{zt})$,

$$\begin{aligned} p'_x &= A'_{xt} = a'_x b'_t - a'_t b'_x, \\ &= \gamma^2 (a_x - \beta_x a_t) (b_t - \beta_x b_x) - \gamma^2 (a_t - \beta_x a_x) (b_x - \beta_x b_t), \\ &= \gamma^2 (1 - \beta_x^2) a_x b_t - \gamma^2 (1 - \beta_x^2) a_t b_x = a_x b_t - a_t b_x \\ &= A_{xt} = p_x, \\ p'_y &= A'_{yt} = a'_y b'_t - a'_t b'_y, \\ &= \gamma a_y (b_t - \beta_x b_x) - \gamma (a_t - \beta_x a_x) b_y, \\ &= \gamma (a_y b_t - a_t b_y - \beta_x a_y b_x + \beta_x a_x b_y) \\ &= \gamma (A_{yt} - \beta_x A_{yx}) = \gamma (p_y - \beta_x q_z), \\ p'_z &= A'_{zt} = \gamma (A_{zt} + \beta_x A_{xz}) = \gamma (p_z + \beta_x q_y). \end{aligned}$$

Similarly, for $\mathbf{q} = (A_{zy} \ A_{xz} \ A_{yx})$,

$$\begin{aligned} q'_x &= A'_{zy} = A_{zy} = q_x, \\ q'_y &= A'_{xz} = \gamma (A_{xz} + \beta_x A_{zt}) = \gamma (q_y + \beta_x p_z), \\ q'_z &= A'_{yx} = \gamma (A_{yx} - \beta_x A_{yt}) = \gamma (q_z - \beta_x p_y). \end{aligned}$$

To convert these to the vector notation, we need to recover missing terms. For example, when z component has term $\beta_x q_y$, we need to add $-\beta_y q_x$ to make it a vector. Then,

$$\mathbf{p}'_{\parallel} = \mathbf{p}_{\parallel}, \quad \mathbf{p}'_{\perp} = \gamma (\mathbf{p} + \boldsymbol{\beta} \times \mathbf{q})_{\perp}, \quad (\text{D.4})$$

$$\mathbf{q}'_{\parallel} = \mathbf{q}_{\parallel}, \quad \mathbf{q}'_{\perp} = \gamma (\mathbf{q} - \boldsymbol{\beta} \times \mathbf{p})_{\perp}. \quad (\text{D.5})$$

The inverse transformation can be obtained by inverting the sign of $\boldsymbol{\beta}$.

$$\mathbf{p}_{\parallel} = \mathbf{p}'_{\parallel}, \quad \mathbf{p}_{\perp} = \gamma (\mathbf{p}' - \boldsymbol{\beta} \times \mathbf{q}')_{\perp}, \quad (\text{D.6})$$

$$\mathbf{q}_{\parallel} = \mathbf{q}'_{\parallel}, \quad \mathbf{q}_{\perp} = \gamma (\mathbf{q}' + \boldsymbol{\beta} \times \mathbf{p}')_{\perp}. \quad (\text{D.7})$$

D.3.4 Dual tensor

We see symmetry between $\mathbf{p}$ and $\mathbf{q}$ in their Lorentz transformation ((D.4), (D.5)), i.e., following substitution will give the same Lorentz transformation.

$$\mathbf{p} \rightarrow \mathbf{q}, \quad \mathbf{q} \rightarrow -\mathbf{p}. \quad (\text{D.8})$$

Therefore, $\tilde{\mathbf{A}} = (\mathbf{q}, -\mathbf{p})$ gives the same transformation for $\mathbf{p}$ and $\mathbf{q}$ as $\mathbf{A} = (\mathbf{p}, \mathbf{q})$.

$$\tilde{\mathbf{A}} = (\mathbf{q}, -\mathbf{p}) = \tilde{A}_{\mu\nu} = \begin{pmatrix} 0 & -q_x & -q_y & -q_z \\ q_x & 0 & p_z & -p_y \\ q_y & -p_z & 0 & p_x \\ q_z & p_y & -p_x & 0 \end{pmatrix} \quad (\text{D.9})$$

Dual tensor is valid only under “proper” (not inverting) coordinate transformations, since we are substituting polar vector with axial vector, vice versa.

Note that if you regard $\mathbf{p}$ as the x-axis and $\mathbf{q}$ as the y-axis, (D.8) can be regarded as 90 degree rotation.

D.3.5 Scalar

Using (D.1), we find

$$\frac{1}{2} A_{\mu\nu} A_{\mu\nu} = -p^2 + q^2, \quad \frac{1}{4} A_{\mu\nu} \tilde{A}_{\mu\nu} = -\mathbf{p} \cdot \mathbf{q}. \quad (\text{D.10})$$

When $p^2 = q^2$, it is true for any inertial frame. When $\mathbf{p} \cdot \mathbf{q} = 0$, you can always find a inertial frame where either $\mathbf{p} = 0$ or $\mathbf{q} = 0$.

References

- [1] Feynman, Leighton, and Sands, *The Feynman Lectures on Physics*. Addison Wesley, 2006. [Online]. Available: <https://www.feynmanlectures.caltech.edu/index.html>
- [2] S. Kimura, “Feynman Units,” September 2025. [Online]. Available: <http://shkmr.github.io/memo/une.pdf>
- [3] K. Cahill, *Physical Mathematics*. Cambridge University Press, 2013.
- [4] S. Kimura, “Using Force to Derive Maxwell’s Equations,” August 2026. [Online]. Available: <https://shkmr.github.io/memo/ufdm.pdf>
- [5] R. P. Feynman, *Quantum Electrodynamics*. W. A. Benjamin, 1961. [Online]. Available: <https://archive.org/details/quantumelectrody00feyn>
- [6] J. J. Sakurai, *Modern Quantum Mechanics*. Addison-Wesley, 1994. [Online]. Available: <https://archive.org/details/ModernQuantumMechanicsJ.J.Sakurai>
- [7] J. D. Jackson, *Classical Electrodynamics*. Wiley, 1998.
- [8] S. Datta, “A Revisit to Lorentz Transformation without Light,” *arXiv:2212.03706 [physics.class-ph]*, December 2022. [Online]. Available: <https://arxiv.org/abs/2212.03706>
- [9] J.-M. Lévy-Leblond, “One more derivation of the Lorentz transformation,” *Am. J. Phys*, vol. 44, pp. 271–277, 1976.
- [10] N. D. Mermin, “Relativity without light,” *Am. J. Phys*, vol. 52, p. 119, 1984, (reprinted with correction and postscript in “*Boojums all the way through*”, Cambridge University Press, 1990).