# Special relativity

## 1.1 Introduction

These notes provide a brief introduction to the basic ideas and mathematical formalism of special relativity. A textbook should be referred to for a complete presentation[1].

Classical mechanics is well described by Newtonian relativity which states that Newton's laws of motion are identical in all inertial frames of reference that are connected by a Galilean transformation, $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$, $t' = t$, where $\mathbf{v}$ is the relative velocity of the frames. An inertial frame can be defined by requiring that the momentum of a particle in an inertial frame is constant, unless there is a net force acting on the particle. Any frame moving with constant velocity is an inertial frame. A uniformly accelerating frame can also be a local inertial frame ( a person inside a falling elevator with no contact with objects outside the elevator).

A problem with Newtonian relativity arose with Maxwell's equations. It was found that the equations did not satisfy Galilean invariance. In addition electromagnetic waves appeared to always propagate at the same velocity. It is generally true that wave equations take on a simple form in a frame where the medium which the waves propagate in is at rest, e.g. sound waves or water waves. However, all attempts to detect such a medium for light waves failed. Experiments failed to show the presence of any "lumineferous" ether, and very accurate measurements showed the speed of light to be constant irrespective of motion of the source.

This problem was unresolved before 1905 when Einstein introduced the special theory of relativity. At the time there appeared three possible solutions.
1) Maxwell's equations were wrong.
2) Electromagnetic waves have a preferred reference frame where the ether is at rest.
3) There is a new relativity principle for classical mechanics and electromagnetism. In other words classical mechanics needs correcting.

Einstein chose 3), his postulates were

1) Principle of relativity: The laws of nature are the same in all inertial reference frames.

2) The speed of light $c$ is finite and independent of the motion of the source. The speed of light is the maximum possible speed of propagation.

An inertial reference frame is one in which a moving body that is not acted on by external forces continues moving with constant velocity. Say $K$ is such an inertial frame. Then any frame $K'$ which moves with constant velocity with respect to $K$ is also an inertial reference frame.

We describe the space-time coordinates of a point in $K$ by the 4-vector $x = (x_0, x_1, x_2, x_3)$ where $x_0 = ct$ is the time coordinate multiplied by $c$ and $(x_1, x_2, x_3) = \mathbf{x} = (x, y, z)$ are the spatial coordinates. Suppose frame $K'$ is moving with speed $v$ along the positive $x_1$ direction

[1] There are many good books on relativity or books that include chapters on relativity. A few favorites are listed at the end of these notes[1, 2, 3, 4, 5].

as viewed by an observer in $K$. The transformation between the space-time coordinates in frames $K, K'$ is

$$
\begin{aligned}
x'_0 &= \gamma(x_0 - \beta x_1) && (1.1a)\\
x'_1 &= \gamma(x_1 - \beta x_0) && (1.1b)\\
x'_2 &= x_2 && (1.1c)\\
x'_3 &= x_3. && (1.1d)
\end{aligned}
$$

Here $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$. Note that $0 \leq |\beta| \leq 1$ and $1 \leq \gamma \leq \infty$. The inverse transformation is found by swapping the primed and unprimed variables and changing the sign of $v$ which gives

$$
\begin{aligned}
x_0 &= \gamma(x'_0 + \beta x'_1) && (1.2a)\\
x_1 &= \gamma(x'_1 + \beta x'_0) && (1.2b)\\
x_2 &= x'_2 && (1.2c)\\
x_3 &= x'_3. && (1.2d)
\end{aligned}
$$

When frame $K'$ moves with relative velocity $\mathbf{v}$ in an arbitrary direction as seen in $K$ the transformation is

$$
\begin{aligned}
x'_0 &= \gamma(x_0 - \boldsymbol{\beta}\cdot\mathbf{x}) && (1.3a)\\
\mathbf{x}' &= \mathbf{x} + \frac{\gamma-1}{\beta^2}(\boldsymbol{\beta}\cdot\mathbf{x})\boldsymbol{\beta} - \gamma\boldsymbol{\beta}x_0 && (1.3b)
\end{aligned}
$$

where $\boldsymbol{\beta} = \mathbf{v}/c$.

Any 4-vector transforms in the same way as the above position-time 4-vectors. Thus for relative motion along $x_1$ a 4-vector $A$ transforms as

$$
\begin{aligned}
A'_0 &= \gamma(A_0 - \beta A_1) && (1.4a)\\
A'_1 &= \gamma(A_1 - \beta A_0) && (1.4b)\\
A'_2 &= A_2 && (1.4c)\\
A'_3 &= A_3. && (1.4d)
\end{aligned}
$$

The squared length of a 4-vector is defined as $A_0^2 - |\mathbf{A}|^2$. This quantity is invariant under Lorentz transformations, so $A_0^2 - |\mathbf{A}|^2 = A_0'^2 - |\mathbf{A}'|^2$. Furthermore for any two 4-vectors $A, B$ the scalar product defined as $A_0B_0 - \mathbf{A}\cdot\mathbf{B}$ is a Lorentz invariant, i.e.

$$A_0B_0 - \mathbf{A}\cdot\mathbf{B} = A'_0B'_0 - \mathbf{A}'\cdot\mathbf{B}'$$

for any two inertial reference frames.

The proper length $l_0$ of an object is the length measured in a frame where the object is at rest. The length measured in a frame where the object is moving with speed $v$ is $l = l_0/\gamma < l_0$. This is known as length contraction.

The proper time $\tau$ is the time measured in a frame where the measuring system is at rest. The differential time increment is $d\tau$. The differential time measured in a frame $K$ where the measuring system moves is given by $dt = \gamma(t)d\tau > d\tau$. Hence moving clocks run

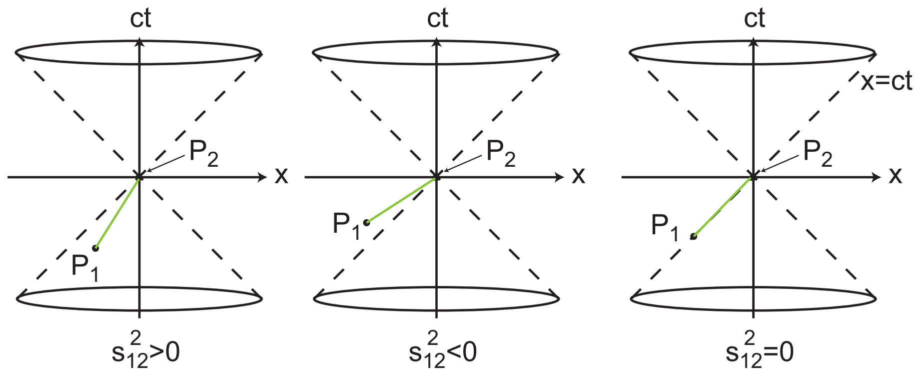


Figure 1.1: Spacetime diagrams for (left to right) timelike separation, spacelike separation, and lightlike separation.

more slowly. This is known as time dilation. The elapsed time in $K$ is related to the elapsed proper time by

$$t_2 - t_1 = \int_{\tau_1}^{\tau_2} \frac{d\tau}{\sqrt{1-\beta^2(\tau)}} = \int_{\tau_1}^{\tau_2} d\tau\gamma(\tau) > \tau_2 - \tau_1.$$

### 1.1.1 Space time

The Lorentz invariant separation between any two points $P_1(t_1, \mathbf{x}_1), P_2(t_2, \mathbf{x}_2)$ is

$$s_{12}^2 = c^2(t_1 - t_2)^2 - |\mathbf{x}_1 - \mathbf{x}_2|^2.$$

When $s_{12}^2 > 0$ points are said to be timelike separated. There exists a frame $K'$ such that $\mathbf{x}_1' = \mathbf{x}_2'$ so $s_{12}^2 = s_{12}'^2 = c^2(t_1 - t_2)^2 > 0$. Timelike separated points can have a causal influence on each other.

When $s_{12}^2 < 0$ points are said to be spacelike separated. There exists a frame $K'$ such that $t_1' = t_2'$ so $s_{12}^2 = s_{12}'^2 = -|\mathbf{x}_1 - \mathbf{x}_2|^2 < 0$. Spacelike separated points do not have any causal influence on each other.

When $s_{12}^2 = 0$ points are said to be light like separated. They can only be connected by light rays.

### 1.1.2 Addition of velocity and acceleration

Consider inertial frames $K$ and $K'$ where $K'$ is moving with velocity $\mathbf{v}$ with respect to $K$. A point $P$ is moving with velocity vector $\mathbf{u}'$ in $K'$. The velocity components in each frame are $u_i = c dx_i/dx_0$ and $u_i' = c dx_i'/dx_0'$.

Using the Lorentz transformations for the differentials $dx_i, dx_i'$ the velocity seen in $K$ is found to be

$$\begin{aligned} u_\parallel &= \frac{u_\parallel' + v}{1 + \frac{\mathbf{v}\cdot\mathbf{u}'}{c^2}} \\ \mathbf{u}_\perp &= \frac{\mathbf{u}_\perp'}{\gamma_v\left(1 + \frac{\mathbf{v}\cdot\mathbf{u}'}{c^2}\right)}. \end{aligned}$$

Here $u_\parallel, \mathbf{u}_\perp$ refer to the velocity components parallel and perpendicular to $\mathbf{v}$.

It can also be shown that the acceleration seen in the two frames is related by

$$\begin{aligned}
\mathbf{a}_\parallel &= \frac{1}{\left[\gamma_v\left(1+\frac{\mathbf{v}\cdot\mathbf{u}'}{c^2}\right)\right]^3}\mathbf{a}'_\parallel \\
\mathbf{a}_\perp &= \frac{1}{\gamma_v^2\left(1+\frac{\mathbf{v}\cdot\mathbf{u}'}{c^2}\right)^3}\left[\mathbf{a}'_\perp + \frac{\mathbf{v}}{c^2}\times(\mathbf{a}'\times\mathbf{u}')\right].
\end{aligned}$$

### 1.1.3 4-velocity and 4-momentum

The velocity addition law shows that the usual velocity does not transform as a 4-vector. We define the velocity 4-vector $U = (U_0, \mathbf{U})$ which transforms as a Lorentz 4-vector by

$$U_0 = \frac{dx_0}{d\tau} = \frac{dx_0}{dt}\frac{dt}{d\tau} = \gamma_u c \tag{1.5a}$$

$$\mathbf{U} = \frac{d\mathbf{x}}{d\tau} = \frac{d\mathbf{x}}{dt}\frac{dt}{d\tau} = \gamma_u \mathbf{u} \tag{1.5b}$$

where $\gamma_u = 1/\sqrt{1-u^2/c^2}$.

The components of the 4-velocity $U$ are proportional to the total energy and momentum. The mass of a particle $m$ is a Lorentz invariant quantity, i.e. it is a Lorentz scalar. The 4-momentum is defined as

$$p = (p_0, \mathbf{p}) = mU = (mU_0, m\mathbf{U}) = (m\gamma_u c, m\gamma_u \mathbf{u}).$$

The total energy of a particle is (writing $\gamma$ instead of $\gamma_u$ for brevity)

$$E = \gamma mc^2 = \frac{mc^2}{\sqrt{1-u^2/c^2}}$$

so the 4-momentum can be written as

$$p = (E/c, \mathbf{p}).$$

Using $\gamma m = \mathbf{p}/\mathbf{u} = E/c^2$ we can also write $\mathbf{u} = c^2\mathbf{p}/E$.

The invariant length of the 4-momentum is

$$\begin{aligned}
p_0^2 - \mathbf{p}\cdot\mathbf{p} &= E^2/c^2 - \mathbf{p}\cdot\mathbf{p} \\
&= m^2\gamma^2c^2 - m^2\gamma^2\mathbf{u}\cdot\mathbf{u} \\
&= m^2\gamma^2c^2(1-|\mathbf{u}|^2/c^2) \\
&= m^2c^2.
\end{aligned}$$

Thus

$$E^2 = m^2c^4 + c^2|\mathbf{p}|^2$$

or

$$E = \sqrt{(mc^2)^2 + (c\mathbf{p})^2}.$$

Here $\mathbf{p} = m\gamma\mathbf{u}$ is the spatial part of the 4-momentum.

## 1.2 Mathematical elaboration

The theory of special relativity can be presented in a more elegant and powerful form using tensor analysis. The length squared of a space-time coordinate is

$$s^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2.$$

The kinematics of special relativity concerns the study of the transformations that leave $s^2$ invariant. These transformations are referred to as the homogeneous Lorentz group consisting of rotations in 3-dimensional space plus the Lorentz transformations considered above. In general we are interested in tensor quantities which are defined by their transformation properties when $x \to x'$ due to the application of an element of the Lorentz group.

A Lorentz scalar is a tensor of rank 0 whose value is not changed by a Lorentz transformation. Lorentz scalars include mass $m$, the norm squared $s^2$, and the elementary charge $q$.

Tensors of rank 1 are vectors. We distinguish between covariant and contravariant vectors. A contravariant vector is written with superscripts as

$$A^\alpha = (A^0, A^1, A^2, A^3) = (A^0, A^i), \quad i = 1, 2, 3.$$

In general greek dummy indices will take the values $0, 1, 2, 3$ while latin dummy indices will run over $1, 2, 3$. Using $\mathbf{A}$ without a superscript or subscript to denote the spatial components we can also write $A^\alpha = (A^0, \mathbf{A})$.

A Lorentz transformation affects the change $A^\alpha \to A'^\alpha$ and $x^\alpha \to x'^\alpha = x'^\alpha(x^0, x^1, x^2, x^3)$. The components of a contravariant vector transform as

$$\begin{aligned} A'^\alpha &= \sum_\beta \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \\ &= \frac{\partial x'^\alpha}{\partial x^0} A^0 + \frac{\partial x'^\alpha}{\partial x^1} A^1 + \frac{\partial x'^\alpha}{\partial x^2} A^2 + \frac{\partial x'^\alpha}{\partial x^3} A^3. \end{aligned} \tag{1.6}$$

A covariant vector which is written with subscripts $A_\alpha = (A_0, A_1, A_2, A_3)$ (we will see below that $A_\alpha = (A^0, -\mathbf{A})$) transforms as

$$\begin{aligned} A'_\alpha &= \sum_\beta \frac{\partial x^\beta}{\partial x'^\alpha} A_\beta \\ &= \frac{\partial x^0}{\partial x'^\alpha} A_0 + \frac{\partial x^1}{\partial x'^\alpha} A_1 + \frac{\partial x^2}{\partial x'^\alpha} A_2 + \frac{\partial x^3}{\partial x'^\alpha} A_3. \end{aligned} \tag{1.7}$$

The contravariant components of a vector correspond to the projections of the vector onto the coordinate axes. The covariant components correspond to projections onto a reciprocal coordinate system. It would perhaps have been more sensible if the names covariant and contravariant were interchanged.

A tensor of rank 2 can be contravariant, covariant, or mixed. A contravariant tensor transforms according to

$$F'^{\alpha\beta} = \sum_{\gamma,\delta} \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x'^\beta}{\partial x^\delta} F^{\gamma\delta}, \tag{1.8}$$

a covariant tensor transforms as

$$F'_{\alpha\beta} = \sum_{\gamma,\delta} \frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\delta}{\partial x'^\beta} F_{\gamma\delta}, \tag{1.9}$$

and a mixed tensor transforms as

$$F'^\alpha{}_\beta = \sum_{\gamma,\delta} \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x^\delta}{\partial x'^\beta} F^\gamma{}_\delta. \tag{1.10}$$

### 1.2.1 Matrix form of Lorentz transformations

Using tensor notation we can write the Lorentz transformations as a matrix. We define a column vector $x^\alpha = (x^0, x^1, x^2, x^3)^T$. It can then be shown that an arbitray Lorentz transformation can be written as a matrix $\mathbf{\Lambda}$ such that $x' = \mathbf{\Lambda} x$ or $x'^\alpha = \sum_\beta \Lambda^\alpha_\beta x^\beta$. The transformation corresponding to a rotation in the clockwise sense about the $x^3$ axis by an angle $\theta$ is

$$\mathbf{\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The tranformation due to a relative velocity $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)$ is (a derivation can be found in Jackson)

$$\mathbf{\Lambda} = \begin{pmatrix} \gamma & -\gamma\beta_1 & -\gamma\beta_2 & -\gamma\beta_3 \\ -\gamma\beta_1 & 1 + \frac{(\gamma-1)\beta_1^2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} \\ -\gamma\beta_2 & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & 1 + \frac{(\gamma-1)\beta_2^2}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} \\ -\gamma\beta_3 & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} & 1 + \frac{(\gamma-1)\beta_3^2}{\beta^2} \end{pmatrix}.$$

The same transformation applies to any 4-vector, not just the coordinates $x$. Thus we can write

$$\begin{aligned} A'^\alpha &= \sum_\beta \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \\ &= \sum_\beta \Lambda^\alpha_\beta A^\beta. \end{aligned} \tag{1.11}$$

Comparing these two forms of the transformation we see that

$$\frac{\partial x'^\alpha}{\partial x^\beta} = \Lambda^\alpha_\beta.$$

The transformation inverse to (1.11) is

$$A^\beta = \sum_\alpha \Lambda^\beta_\alpha A'^\alpha. \tag{1.12}$$

Using this we can show that the transformations associated with $\Lambda_\alpha^\beta$ and $\Lambda_\beta^\alpha$ are inverses of each other. We have

$$\begin{aligned} A'^\alpha &= \sum_\beta \Lambda_\beta^\alpha A^\beta \\ &= \sum_\beta \Lambda_\beta^\alpha \sum_\gamma \Lambda_\gamma^\beta A'^\gamma \\ &= \sum_\gamma \left( \sum_\beta \Lambda_\beta^\alpha \Lambda_\gamma^\beta \right) A'^\gamma \end{aligned} \tag{1.13}$$

Equation (1.13) is true provided

$$\sum_\beta \Lambda_\beta^\alpha \Lambda_\gamma^\beta = \delta_\gamma^\alpha.$$

### 1.2.2 Transformation of derivatives

We need to also consider the transformation rules for derivatives. Using the chain rule the derivative with respect to a contravariant component can be written as

$$\frac{\partial}{\partial x'^\alpha} = \sum_\beta \frac{\partial}{\partial x^\beta} \frac{\partial x^\beta}{\partial x'^\alpha} = \sum_\beta \frac{\partial x^\beta}{\partial x'^\alpha} \frac{\partial}{\partial x^\beta}. \tag{1.14}$$

Comparing with Eq. (1.7) we see that differentiation with respect to a contravariant component of the coordinate vector transforms like a covariant vector. It is also the case that differentiation with respect to a covariant component of the coordinate vector transforms like a contravariant vector. A differential element also transforms contravariantly since

$$dx'^\alpha = \sum_\beta \frac{\partial x'^\alpha}{\partial x^\beta} dx^\beta. \tag{1.15}$$

Comparing with Eq. (1.6) we see that the transformation rule for a differential element is that of a contravariant vector.

The transformation properties of derivatives are made more apparent by using the short-hand notation

$$\begin{aligned} \partial_\alpha \equiv \frac{\partial}{\partial x^\alpha} &= \left( \frac{\partial}{\partial x^0}, \nabla \right) \\ \partial^\alpha \equiv \frac{\partial}{\partial x_\alpha} &= \left( \frac{\partial}{\partial x^0}, -\nabla \right) \end{aligned}$$

which shows that $\frac{\partial}{\partial x_\alpha} = \partial^\alpha$ is like a contravariant vector and $\frac{\partial}{\partial x^\alpha} = \partial_\alpha$ is like a covariant vector.

### 1.2.3 Manipulating tensor indices

The transformations given in Eqs. (1.6-1.14) are tedious to write out component by component. A great economy of expression is achieved using the Einstein summation convention that any pair of dummy indices is summed over when one of the indices is raised and the other is lowered. Thus

$$
\begin{aligned}
A'^{\alpha} &= \partial_\beta x'^{\alpha} A^{\beta} && (1.16a)\\
A'_{\alpha} &= \partial'_\alpha x^{\beta} A_{\beta} && (1.16b)\\
F'^{\alpha\beta} &= \partial_\gamma x'^{\alpha} \partial_\delta x'^{\beta} F^{\gamma\delta}, && (1.16c)\\
F'_{\alpha\beta} &= \partial'_\alpha x^{\gamma} \partial'_\beta x^{\delta} F_{\gamma\delta}, && (1.16d)\\
F'^{\alpha}{}_{\beta} &= \partial_\gamma x'^{\alpha} \partial'_\beta x^{\delta} F^{\gamma}{}_{\delta}. && (1.16e)
\end{aligned}
$$

The notation $\partial'$ means take the derivative with respect to a primed coordinate. The scalar product of two vectors is

$$A \cdot B \equiv A_\alpha B^\alpha = A^\alpha B_\alpha.$$

The scalar product is invariant since

$$A' \cdot B' = \partial'_\alpha x^\beta A_\beta \, \partial_\gamma x'^\alpha B^\gamma = \partial_\gamma x^\beta A_\beta B^\gamma = \delta^\beta{}_\gamma A_\beta B^\gamma = A_\gamma B^\gamma = A \cdot B.$$

In differential form the norm squared is

$$(ds)^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2.$$

This norm or metric is a special case of the general length element

$$(ds)^2 = g_{\alpha\beta} dx^\alpha dx^\beta.$$

The geometry is defined by the metric tensor $g_{\alpha\beta} = g_{\beta\alpha}$. In special relativity we use

$$g_{00} = 1, g_{11} = g_{22} = g_{33} = -1 \qquad (1.17)$$

with all off-diagonal elements zero. The contravariant metric tensor is $g^{\alpha\beta} = g_{\alpha\beta}$ and $g_{\alpha\gamma} g^{\gamma\beta} = \delta^\beta_\alpha = 0$ for $\alpha \neq \beta$ and $\delta^\alpha_\alpha = 1$ for $\alpha = 0, 1, 2, 3$.

We can transform from contravariant to covariant vectors by contraction with the metric tensor. Thus

$$x_\alpha = g_{\alpha\beta} x^\beta$$

and

$$x^\alpha = g^{\alpha\beta} x_\beta.$$

An analogous procedure can be used to change the index on a higher rank tensor.

In any tensor equation the two sides must contain identical and identically placed (with regards to up or down) free indices (those indices which are not dummy indices). Free indices can be shifted up or down, if it is done simultaneously in all terms of the equation. It is "illegal" to equate covariant and contravariant components of different tensors. Doing so, which might happen accidentally in a particular coordinate system, would not remain a valid equality upon transformation to a different reference frame.

With the above definition of the metric tensor of special relativity we see that the covariant and contravariant components of a vector are related by $A_0 = A^0$ and $A_i = -A^i$ for $i = 1, 2, 3$. The scalar product is

$$A \cdot B = A_\alpha B^\alpha = g_{\alpha\beta} A^\beta B^\alpha = A^0 B^0 - \mathbf{A} \cdot \mathbf{B}$$

which agrees with our earlier definition.

The 4-divergence of a 4-vector is

$$\partial^\alpha A_\alpha = \partial_\alpha A^\alpha = \frac{\partial A^0}{\partial x^0} + \nabla \cdot \mathbf{A}$$

which is a Lorentz invariant (scalar). The 4-dimensional Laplacian is

$$\Box \equiv \partial_\alpha \partial^\alpha = \frac{\partial^2}{\partial x^{0^2}} - \nabla^2$$

which is the wave equation operator in vacuum.

## 1.3 Covariant form of the Maxwell equations

Using tensor notation we can write the Maxwell equations in manifestly covariant form. We start from the Lorentz force law

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right)$$

for a particle of charge $q$. The momentum is the space part of the 4-momentum $p^\alpha = (p_0, \mathbf{p}) = m(U_0, \mathbf{U})$. To make the force independent of coordinate system we differentiate with respect to the proper time $\tau$ and use $d\tau = dt/\gamma$ to get

$$\frac{d\mathbf{p}}{d\tau} = \frac{q}{c}\left(U_0 \mathbf{E} + \mathbf{U} \times \mathbf{B}\right). \tag{1.18}$$

The corresponding time part of $dp/d\tau$ is found from the time rate of change of the particle energy $E$ which is given by

$$\frac{dE}{dt} = \mathbf{v} \cdot q\mathbf{E}.$$

Thus

$$\frac{dp_0}{d\tau} = \frac{1}{c}\frac{dE}{d\tau} = \frac{1}{c}\mathbf{U} \cdot q\mathbf{E} = \frac{q}{c}\mathbf{U} \cdot \mathbf{E}. \tag{1.19}$$

In order to make the force and energy change equations Lorentz covariant the right hand sides of (1.18,1.19) must be components of a 4-vector. Experiment indicates that charge is a Lorentz invariant. Thus to satisfy (1.19) the right hand side must be of the form $F_{0\beta} U^\beta$.

In order to find the field strength tensor $F^{\alpha\beta}$ we use the Maxwell equations. The continuity equation is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

The charge and current form a 4-vector $J^\alpha = (c\rho, \mathbf{J})$ so the continuity equation can be written compactly as

$$\partial_\alpha J^\alpha = 0.$$

The electric and magnetic fields can be written in terms of the vector and scalar potentials $\mathbf{A}$ and $\phi$. In the Lorenz gauge the potentials satisfy

$$\begin{aligned} \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} &= \frac{4\pi}{c}\mathbf{J} \\ \frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} - \nabla^2 \Phi &= 4\pi\rho \end{aligned}$$

with the Lorenz condition

$$\frac{1}{c}\frac{\partial \Phi}{\partial t} + \nabla \cdot A = 0.$$

The right hand sides are the components of a 4-vector so to ensure Lorenz covariance we require that the potentials form a 4-vector potential $A^\alpha = (\Phi, \mathbf{A})$. The wave equations and the Lorenz condition can then be written as

$$\begin{aligned} \Box A^\alpha &= \frac{4\pi}{c} J^\alpha \\ \partial_\alpha A^\alpha &= 0. \end{aligned}$$

The potentials determine the fields through

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi \tag{1.20a}$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{1.20b}$$

These equations can be written in tensor notation as

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha$$

where the antisymmetric field strength tensor is

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}.$$

The covariant form of the tensor is

$$F_{\alpha\beta} = g_{\alpha\gamma} g_{\delta\beta} F^{\gamma\delta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

which corresponds to the transformation $\mathbf{E} \rightarrow -\mathbf{E}$. It is also useful to define the dual field strength tensor

$$\mathcal{F}^{\alpha\beta} = \frac{1}{2}\epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix}$$

which corresponds to the replacements $\mathbf{E} \to \mathbf{B}$ and $\mathbf{B} \to -\mathbf{E}$ in $F^{\alpha\beta}$. Here $\epsilon^{\alpha\beta\gamma\delta}$ is the 4-dimensional generalization of $\epsilon_{ijk}$.

The Maxwell equations in cgs units are the two equations with sources

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 4\pi\rho \\ \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} &= \frac{4\pi}{c}\mathbf{J} \end{aligned}$$

and the homogeneous equations

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} &= 0. \end{aligned}$$

In manifestly covariant form these can be written as

$$\begin{aligned} \partial_\alpha F^{\alpha\beta} &= \frac{4\pi}{c} J^\beta \\ \partial_\alpha \mathcal{F}^{\alpha\beta} &= 0. \end{aligned}$$

The last equation can also be written as

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0.$$

Finally the Lorentz force and rate of change of energy can be written as

$$\frac{dp^\alpha}{d\tau} = m\frac{dU^\alpha}{d\tau} = \frac{q}{c} F^{\alpha\beta} U_\beta.$$

This manifestly covariant formulation of electromagnetism is very useful for calculating how the electromagnetic fields transform between Lorentz frames as well as finding the fields of a charge moving at relativistic speeds.

## 1.4 Transformation of the Electromagnetic field

We can follow either of two routes to transform the electromagnetic field between Lorentz frames. Either we transform the 4-potential $A^\alpha = (\Phi, \mathbf{A})$ and then use Eqs. (1.20) to find the fields, or we can transform the field strength tensor directly using Eqs. (1.16).

For transformation from $K$ to $K'$ which moves with velocity $\mathbf{v} = c\boldsymbol{\beta}$ with respect to $K$ the transformations are

$$\begin{aligned} \mathbf{E}' &= \gamma(\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B}) - \frac{\gamma^2}{\gamma+1}\boldsymbol{\beta}(\boldsymbol{\beta}\cdot\mathbf{E}) \\ \mathbf{B}' &= \gamma(\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E}) - \frac{\gamma^2}{\gamma+1}\boldsymbol{\beta}(\boldsymbol{\beta}\cdot\mathbf{B}). \end{aligned} \tag{1.21a}$$

The electric and magnetic fields are interrelated, and are not really separate entities. For the particular case of relative motion along the $x_1$ axis the field transformations are

$$\begin{aligned} E_1' &= E_1 & B_1' &= B_1 \\ E_2' &= \gamma(E_2 - \beta B_3) & B_2' &= \gamma(B_2 + \beta E_3) \\ E_3' &= \gamma(E_3 + \beta B_2) & B_3' &= \gamma(B_3 - \beta E_2). \end{aligned}$$

Suppose that in frame $K$ all charges are at rest so there an electric field $\mathbf{E}$ but no magnetic field ($\mathbf{B} = 0$). Then

$$\mathbf{E} = \gamma(\mathbf{E}' - \boldsymbol{\beta} \times \mathbf{B}') - \frac{\gamma^2}{\gamma + 1}\boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{E}')$$

so

$$\begin{aligned} \mathbf{B}' &= -\gamma\boldsymbol{\beta} \times \mathbf{E} \\ &= -\gamma\boldsymbol{\beta} \times \gamma(\mathbf{E}' - \boldsymbol{\beta} \times \mathbf{B}') \\ &= -\gamma^2\boldsymbol{\beta} \times \mathbf{E}' + \gamma^2\boldsymbol{\beta} \times \boldsymbol{\beta} \times \mathbf{B}' \\ &= -\boldsymbol{\beta} \times \mathbf{E}' \end{aligned} \tag{1.22}$$

which provides a simple relation between $\mathbf{B}'$ and $\mathbf{E}'$.

# Bibliography


[1] W. Pauli, *Theory of Relativity*, (Pergamon, Oxford, 1967).

[2] J. D. Jackson, *Classical Electrodynamics, 3rd edition*, (Academic Press, 1999).

[3] L. D. Landau and E. M. Lifschitz, *Classical theory of fields, 4th edition*, (Pergamon Press, Oxford, 1975).

[4] E. Schrödinger, *Space-time structure*, (Cambridge Unviersity Press, Cambridge, 1950).

[5] C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, Vol. 1 $2^{\text{nd}}$ Ed., (McGraw-Hill, New York , 1973).