Special Relativity and Maxwell Equations

by

Bernd A. Berg

Department of Physics

Florida State University

Tallahassee, FL 32306, USA.

Chapter 1

A self-contained summary of the theory of special relativity is given, which provides the space-time frame for classical electrodynamics. Historically [2] special relativity emerged out of electromagnetism. Nowadays, is deserves to be emphasized that special relativity restricts severely the possibilities for electromagnetic equations.

1.1 Special Relativity

Let us deal with space and time in vacuum. The conventional time unit is one

$$\textit{second}\ [s]. \tag{1.1}$$

Here, and in the following abbreviations for units are placed in brackets [ ]. For a long time period the second was defined in terms of the rotation of the earth as $\frac{1}{60} \times \frac{1}{60} \times \frac{1}{24}$ of the mean solar day. Nowadays most accurate time measurements rely on atomic clocks. They work by tuning a electric frequency into resonance with some atomic transition. Consequently, the second has been re–defined, such that the frequency of the light between the two hyperfine levels of the ground state of the cesium ^{132}Cs atom is now exactly 9,192,631,770 cycles per second.

Special relativity is founded on two basic postulates:

1. Galilee invariance: The laws of nature are independent of any uniform, translational motion of the reference frame.

This postulate gives rise to a triple infinite set of reference frames moving with constant velocities relative to one another. They are called *inertial frames*. For a freely moving body, i.e. a body which is not acted upon by an external force, inertial systems exist. The differential equations which describe physical laws take the same form in all inertial frames. This *Galilee invariance* was known long before Einstein.

2. The speed c of light in empty space is independent of the motion of its source.

The second Postulate was introduced by Einstein 1905 [2]. It implies that c takes the same constant value in all inertial frames. Transformations between inertial frames are implied which have far reaching physical consequences.

The distance unit

$$1\ meter\ [m] = 100\ centimeters\ [cm] \tag{1.2}$$

was originally defined by two scratches on a bar made of platinum–iridium alloy kept at the International Bureau of Weights and Measures in Sèvres, France. As measurements of the speed of light have become increasingly accurate, it has become most appropriate to exploit Postulate 2 to define the distance unit. The standard meter is now defined [6] as the distance traveled by light in empty space during the time of 1/299,792,458 [s]. This makes the speed of light exactly

$$c = 299,792,458\ [m/s]. \tag{1.3}$$

1.1.1 Natural Units

The units for *second* (1.1 and *meter* (1.2) are not independent, as the speed of light is an universal constant. This allows to define natural units, which are frequently used in nuclear, particle and astro physics. They define

$$c = 1 \tag{1.4}$$

as a dimensionless constant, and

$$1\ [s] = 299,792,458\ [m]$$

holds. The advantage of natural units is that factors of c disappear in calculations. The disadvantage is that, for converting back to conventional units, the appropriate factors have to be recovered by dimensional analysis. For instance, if time is given in seconds $x = t$ in natural units converts to $x = ct$ with x in meters and c given by (1.3).

1.1.2 Definition of distances and synchronization of clocks

The introduced concepts of Galilee invariance and of a constant speed of light are not pure mathematics, but relate to physical measurements. Therefore, I discuss briefly how to employ them to reduce measurements of spatial distances to time measurements. Let us consider an inertial frame K with coordinates $(t, \vec{x})$. We like to place observers at rest at different places $\vec{x}$ in K. The observers are equipped with clocks of identical making, to define the time t at $\vec{x}$. The origin $\vec{x} = 0$, in other notation $(t, \vec{0})$, of K is defined by placing observer O_0 and his clock there. We like to place another observer O_1 at $\vec{x}_1$ to define $(t, \vec{x}_1)$. How can O_1 know to be at $\vec{x}_1$? By using a mirror he can reflect light flashed by observer O_0 at him. Observer O_0 may measure the polar and azimuthal angles (θ, ϕ) at which he emits the light and

$$|\vec{x}_1| = c \triangle t/2\,,$$

where $\triangle t$ is the time light needs to travel to O_1 and back. This determines $\vec{x}_1$ and he can signal this information to O_1. By repeating the measurement, he can make sure that O_1 is not moving with respect to K. For an idealized, force free environment the observers will then never start moving with respect to one another. Next O_1 needs to synchronize his clock. This may be done as follows: Observer O_0 flashes his instant time t_0 at O_1, and O_1 puts his clock to

$$t_1 = t_0 + |\vec{x}_1|/c$$

at the instant receiving the signal. In this way $(t, \vec{x}_1)$ is operationally given. If O_0 flashes again his instant time t_{01} over to O_1, the clock of O_1 will show time $t_{11} = t_{01} + |\vec{x}_1|/c$ at the instant of receiving the signal. In the same way the time t can be defined at any desired point $\vec{x}$ in K.

Now we consider an inertial frame K' with coordinates $(t', \vec{x}')$, moving with constant velocity $\vec{v}$ with respect to K. The origin of K' is defined through a third observer O_0'. What does it mean that O_0' moves with constant velocity $\vec{v}$ with respect to O_0? At times t^e_{01} and t^e_{02} observer O_0 may flash light signals (the superscript e stands for "emit") at O_0', which are reflected and arrive back after time intervals $\triangle t_{01}$ and $\triangle t_{02}$. From principle 2 it follows that the reflected light needs the same time (measured by O_0 in K) to travel from O_0' to O_0, as the light needed to travel from O_0 to O_0'. Hence, O_0 concludes that O_0' received the signal at

$$t_{0i} = t^e_{0i} + \triangle t_{0i}/2,\ \ (i = 1, 2) \tag{1.5}$$

in the O_0 time. This extremely simple equation does look quite complicated for non-relativistic physics, because the speed on the return path would then be distinct from that on the arrival path (consider for instance elastic scattering of a very light particle on a very heavy surface). The constant velocity of light implies that relativistic distance measurements are simpler than such non-relativistic measurements. For observer O_0 the positions $\vec{x}_{01}$ and $\vec{x}_{02}$ are now defined through the angles (θ_{0i}, ϕ_{0i}) and

$$|\vec{x}_{0i}| = \triangle t_{0i}\, c/2,\ \ (i = 1, 2)\,. \tag{1.6}$$

For the assumed force free environment observer O_0 can conclude that O_0' moves with respect to him with uniform velocity

$$\vec{v} = (\vec{x}_{02} - \vec{x}_{01})/(t_{02} - t_{01})\,. \tag{1.7}$$

O_0 may repeat the procedure at later time t_{0i}, $i = 3, 4, \cdots$ to check that O_o' moves indeed with uniform velocity.

Similarly, observer O_0' will find out that O_0 move with velocity $\vec{v}\,' = -\vec{v}$. According to principle 1, observers in K' can now go ahead to define t' for any point $\vec{x}\,'$ in K', and $(t', \vec{x}')$ is operationally defined. Let us further discuss the motion of O_0' as observed by O_0. The equation of motion for the origin of K' is

$$\vec{x}\,(\vec{x}^{\,'} = 0) = \vec{x}_0 + \vec{v}\,t, \tag{1.8}$$

with $\vec{x}_0 = \vec{x}_{01} - \vec{v}\,t_{01}$. The equation for $\vec{x}_0$ expresses the fact that for $t = t_{01}$ observer O_0' is at $\vec{x}_{01}$. Shifting his space convention by a constant vector, observer O_0 can achieve $\vec{x}_0 = 0$, such that equation (1.8) becomes

$$\vec{x}\,(\vec{x}^{\,\prime} = 0) = \vec{v}\,t.$$

Similarly, observer O_0' may choose his space convention such that

$$\vec{x}^{\,\prime}\,(\vec{x} = 0) = -\vec{v}\,t'$$

holds.

1.1.3 Lorentz invariance and Minkowski space

Having defined time and space operationally, let us focus on a more abstract discussion. We consider the two inertial frames with uniform relative motion $\vec{v}$: K with coordinates $(t, \vec{x})$ and K' with coordinates $(t', \vec{x}^{\,\prime})$. We demand that at time $t = t' = 0$ their two origins coincide. This can be achieved, as we have just seen. Now, imagine a spherical shell of radiation originating at time $t = 0$ from $\vec{x} = \vec{x}^{\,\prime} = 0$. The propagation of the wavefront is described by

$$c^2t^2 - x^2 - y^2 - z^2 = 0 \quad \text{in } K, \tag{1.9}$$

and by

$$c^2t'^2 - x'^2 - y'^2 - z'^2 = 0 \quad \text{in } K'. \tag{1.10}$$

We define 4-vectors ($\alpha = 0, 1, 2, 3$) by

$$(x^\alpha) = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix} \quad \text{and} \quad (x_\alpha) = (ct,\ -\vec{x})\ . \tag{1.11}$$

By reasons explained in the next section the components x^α are called *contravariant* and the components x_α *covariant*. In matrix notation the contravariant 4-vector (x^α) is represented by a column and the covariant 4-vector (x_α) as a row.

The *Einstein summation convention* is defined by

$$x_\alpha x^\alpha = \sum_{\alpha=0}^{3} x_\alpha x^\alpha, \tag{1.12}$$

and will be employed throughout this script. Equations (1.9) and (1.10) read then $x_\alpha x^\alpha = x'_\alpha x'^\alpha = 0$. *(Homogeneous) Lorentz transformations* are defined as the group of transformations which leave the distance

$$s^2 = x_\alpha x^\alpha \quad \text{invariant.} \tag{1.13}$$

If the initial condition $t' = 0$ and $\vec{x}^{\,\prime}(\vec{x} = 0) = \vec{x}(\vec{x}^{\,\prime}) = 0$ for $t = 0$ is replaced by and arbitrary one, the equation $(x_\alpha - y_\alpha)(x^\alpha - y^\alpha) = (x'_\alpha - y'_\alpha)(x'^\alpha - y'^\alpha)$ still holds. *Inhomogeneous Lorentz* or *Poincaré* transformations are defined as the group of transformations which leave

$$s^2 = (x_\alpha - y_\alpha)(x^\alpha - y^\alpha) \quad \text{invariant.} \tag{1.14}$$

In contrast to the Lorentz transformations the Poincaré transformations include invariance under *translations*

$$x^\alpha \to x^\alpha + a^\alpha \quad \text{and} \quad y^\alpha \to y^\alpha + a^\alpha$$

where a^α is a constant vector. A fruitful concept is that of a 4-dimensional space-time, called *Minkowski space*. Equation (1.14) gives the invariant metric of this space. Compared to the norm of 4-dimensional Euclidean space, the crucial difference is the relative minus sign between time and space components. The *light cone* of a 4-vector x_0^α is defined as the set of vectors x^α which satisfy

$$(x - x_0)^2 = (x_\alpha - x_{0\alpha})\,(x^\alpha - x_0^\alpha) = 0.$$

The light cone separates events which are *timelike* and *spacelike* with respect to x_0^α, namely

$$(x - x_0)^2 > 0 \quad \text{for timelike}$$

and

$$(x - x_0)^2 < 0 \quad \text{for spacelike.}$$

We shall see later, compare equation (1.43), that the time ordering of spacelike points is distinct in different inertial frames, whereas it is the same for timelike points. For the choice $x_0^\alpha = 0$ this Minkowski space situation is depicted in figure 1.1. On the abscissa we have the projection of the three dimensional Euclidean space on $r = |\vec{x}|$. The regions *future* and *past* of this figure are the timelike points of $x_0 = 0$, whereas *elsewhere* are the spacelike points.

To understand special relativity in some depth, we have to explore Lorentz and Poincaré transformations in some details. Before we come to this, it is convenient to introduce some relevant calculus.

1.1.4 Vector and tensor notation

Let us define a general transformation $x \to x'$ through

$$x'^\alpha = x'^\alpha(x) = x'^\alpha\left(x^0, x^1, x^2, x^3\right), \ \alpha = 0, 1, 2, 3. \tag{1.15}$$

This means, x'^α is a function of four variables and, when it is needed, this function is assumed to be sufficiently often differentiable with respect to each of its arguments. In the following we consider the transformation properties of various quantities (scalars, vectors and tensors) under $x \to x'$.

A *scalar* is a single quantity whose value is not changed under the transformation (1.15).

A *4-vector* A^α, $(\alpha = 0, 1, 2, 3)$ is said *contravariant* if its components transform according to

$$A'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta. \tag{1.16}$$

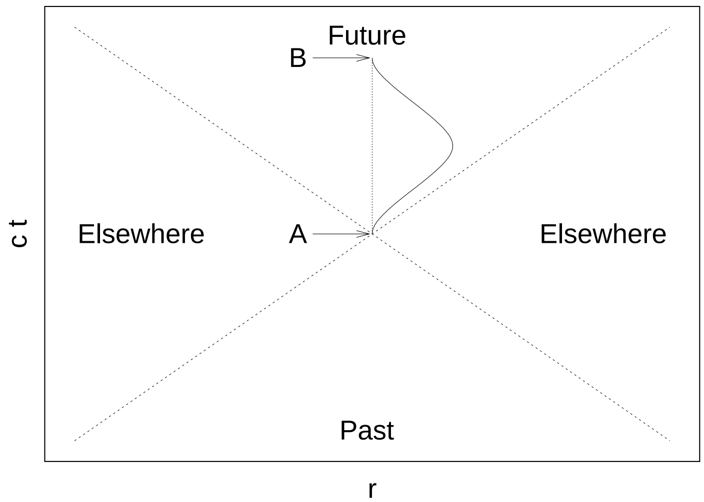

Figure 1.1: Minkowski space: Seen from the spacetime point A at the origin, the spacetime points in the forward light cone are in the future, those in the backward light cone are in the past and the spacelike points are "elsewhere", because their time-ordering depends on the inertial frame chosen. Paths of two clocks which separate at the origin (the straight line one stays at rest) and merge again at a future space-time point B are also indicated. For the paths shown the clock moved along the curved path will, at B, show an elapsed time of about 70% of the elapsed time shown by the other clock (*i.e.* the one which stays at rest).

An example is $A^\alpha = dx^\alpha$, where (1.16) reduces to the well–known rule for the differential of a function of several variable ($f^\alpha(x) = x'^\alpha(x)$):

$$dx'^\alpha = \frac{\partial x'^{\,\alpha}}{\partial x^\beta} dx^\beta .$$

Remark: In this general framework the vector x^α is not always contravariant. When

$$a^\alpha{}_\beta = \frac{\partial x'^{\,\alpha}}{\partial x^{\,\beta}}$$

defines a linear transformation with constant (*i.e.* space-time independent) coefficients $a^\alpha{}_\beta$, the vector x^α itself is contravariant. In the present framework we are only interested in that case (the other leads into general relativity).

A 4-vector is said *covariant* when it transforms like

$$B'_\alpha = \frac{\partial x^\beta}{\partial x'^\alpha} B_\beta . \tag{1.17}$$

An example is

$$B_\alpha = \partial_\alpha = \frac{\partial}{\partial x^\alpha}, \tag{1.18}$$

because of

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

The *inner* or *scalar product* of two vectors is defined as the product of the components of a covariant and a contravariant vector:

$$B \cdot A = B_\alpha A^\alpha. \tag{1.19}$$

With this definition the scalar product is an invariant under the transformation (1.15). This follows from (1.16) and (1.17):

$$B' \cdot A' = \frac{\partial x^\beta}{\partial x'^\alpha}\frac{\partial x'^\alpha}{\partial x^\gamma} B_\beta A^\gamma = \frac{\partial x^\beta}{\partial x^\gamma} B_\beta A^\gamma = \delta^\beta{}_\gamma B_\beta A^\gamma = B \cdot A.$$

Here the *Kronecker delta* is defined by:

$$\delta^\alpha{}_\beta = \delta_\alpha{}^\beta = \begin{cases} 1 & \text{for} \quad \alpha = \beta, \\ 0 & \text{for} \quad \alpha \neq \beta. \end{cases} \tag{1.20}$$

Vectors are *rank* one tensors. *Tensors* of general rank k are quantities with k indices, like for instance

$$T^{\alpha_1 \alpha_2 \dots}{}_{\dots \alpha_i \dots \alpha_k} \, .$$

The convention is that the upper indices transform contravariant and the lower transform covariant. For instance, a contravariant tensor of rank two $F^{\alpha\beta}$ consists of 16 quantities that transform according to

$$F'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma}\frac{\partial x'^\beta}{\partial x^\delta} F^{\gamma\delta}.$$

A covariant tensor of rank two $G_{\alpha\beta}$ transforms as

$$G'_{\alpha\beta} = \frac{\partial x^\gamma}{\partial x'^\alpha}\frac{\partial x^\delta}{\partial x'^\beta} G_{\gamma\delta}.$$

The *inner product* or *contraction* with respect to any pair of indices, either on the same tensor or between different tensors, is defined in analogy with (1.19). One index has to be contravariant and the other covariant.

A tensor $S^{\dots\alpha\dots\beta\dots}$ is the symmetric in α and β when

$$S^{\dots\alpha\dots\beta\dots} = S^{\dots\beta\dots\alpha\dots}.$$

A tensor $A^{\dots\alpha\dots\beta\dots}$ is the antisymmetric in α and β when

$$A^{\dots\alpha\dots\beta\dots} = -A^{\dots\beta\dots\alpha\dots}.$$

Let $S^{\alpha\beta}$ be a symmetric and $A^{\alpha\beta}$ be an antisymmetric tensor. It holds

$$S^{\dots\alpha\dots\beta\dots} A_{\dots\alpha\dots\beta\dots} = 0. \tag{1.21}$$

Proof:

$$S^{\dots\alpha\dots\beta\dots}A_{\dots\alpha\dots\beta\dots} = -S^{\dots\beta\dots\alpha\dots}A_{\dots\beta\dots\alpha\dots} = -S^{\dots\alpha\dots\beta\dots}A_{\dots\alpha\dots\beta\dots},$$

and consequently zero. The first step exploits symmetry and antisymmetry, and the second step renames the summation indices. Every tensor can be written as a sum of its symmetric and antisymmetric parts

$$T^{\dots\alpha\dots\beta\dots} = T_S^{\dots\alpha\dots\beta\dots} + T_A^{\dots\alpha\dots\beta\dots} \tag{1.22}$$

by simply defining

$$T_S^{\dots\alpha\dots\beta\dots} = \frac{1}{2}\left(T^{\dots\alpha\dots\beta\dots} + T^{\dots\beta\dots\alpha\dots}\right) \quad \text{and} \quad T_A^{\dots\alpha\dots\beta\dots} = \frac{1}{2}\left(T^{\dots\alpha\dots\beta\dots} - T^{\dots\beta\dots\alpha\dots}\right) . \tag{1.23}$$

So far the results and definitions are general. We now specialize to Poincaré transformations. The specific geometry of the space–time of special relativity is defined by the invariant distance s^2, see eqn.(1.14). In differential form, the infinitesimal interval ds that defines the norm of our space is

$$(ds)^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2. \tag{1.24}$$

Here we have used superscripts on the coordinates in accordance to our insight that dx^α is a contravariant vector. Introducing a *metric tensor* $g_{\alpha\beta}$ we re–write equation (1.24) as

$$(ds)^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta. \tag{1.25}$$

Comparing (1.24) and (1.25) we see that for special relativity $g_{\alpha\beta}$ is diagonal:

$$g_{00} = 1,\ g_{11} = g_{22} = g_{33} = -1 \quad \text{and} \quad g_{\alpha\beta} = 0 \text{ for } \alpha \neq \beta. \tag{1.26}$$

Comparing (1.25) with the invariant scalar product (1.19), we conclude that

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

The covariant metric tensor lowers the indices, i.e. transforms a contravariant into a covariant vector. Correspondingly the contravariant metric tensor $g^{\alpha\beta}$ is defined to raise indices:

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

The last two equations and the symmetry of $g_{\alpha\beta}$ imply

$$g_{\alpha\gamma}\, g^{\gamma\beta} = \delta_\alpha^{\ \beta}$$

for the contraction of contravariant and covariant metric tensors. This is solved by $g^{\alpha\beta}$ being the normalized co–factor of $g_{\alpha\beta}$. For the diagonal matrix (1.26) the result is simply

$$g^{\alpha\beta} = g_{\alpha\beta}. \tag{1.27}$$

Consequently the equations

$$A^\alpha = \begin{pmatrix} A^0 \\ \vec{A} \end{pmatrix}, \quad A_\alpha = (A^0, -\vec{A})$$

and, compare (1.18),

$$(\partial_\alpha) = \left(\frac{\partial}{c\partial t}, \nabla \right), \quad (\partial^\alpha) = \begin{pmatrix} \frac{\partial}{c\partial t} \\ \\ -\nabla \end{pmatrix}. \tag{1.28}$$

hold. It follows that the 4-divergence of a 4-vector

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

and the d'Alembert (4-dimensional Laplace) operator

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

are invariants. Sometimes the notation $\triangle = \nabla^2$ is used for the (3-dimensional) Laplace operator.

1.1.5 Lorentz transformations

Let us now construct the Lorentz group. We seek a group of linear transformations

$$x'^\alpha = a^\alpha{}_\beta \, x^\beta, \quad (\Rightarrow \frac{\partial x'^\alpha}{\partial x^\beta} = a^\alpha{}_\beta) \tag{1.29}$$

such that the scalar product stays invariant:

$$x'_\alpha \, x'^\alpha = a_\alpha{}^\beta x_\beta \, a^\alpha{}_\gamma x^\gamma = x_\alpha \, x^\alpha = \delta^\beta{}_\gamma x_\beta x^\gamma .$$

As the $x_\beta x^\gamma$ are independent, this yields

$$a_\alpha{}^\beta \, a^\alpha{}_\gamma = \delta^\beta{}_\gamma \;\Leftrightarrow\; a_{\alpha\beta} \, a^\alpha{}_\gamma = g_{\beta\gamma} \;\Leftrightarrow\; a^\delta{}_\beta \, g_{\delta\alpha} \, a^\alpha{}_\gamma = g_{\beta\gamma} \,.$$

In matrix notation

$$\tilde{A} g A = g, \tag{1.30}$$

where $g = (g_{\beta\alpha})$ is given by (1.26),

$$A = (a^\beta{}_\alpha) = \begin{pmatrix} a^0{}_0 & a^0{}_1 & a^0{}_2 & a^0{}_3 \\ a^1{}_0 & a^1{}_1 & a^1{}_2 & a^1{}_3 \\ a^2{}_0 & a^2{}_1 & a^2{}_2 & a^2{}_3 \\ a^3{}_0 & a^3{}_1 & a^3{}_2 & a^3{}_3 \end{pmatrix}, \tag{1.31}$$

and $\tilde{A} = (\tilde{a}_\beta{}^\alpha)$ with $\tilde{a}_\beta{}^\alpha = a^\alpha{}_\beta$ is the transpose of the matrix $A = (a^\beta{}_\alpha)$, explicitly

$$\tilde{A} = (\tilde{a}_\beta{}^\alpha) = \begin{pmatrix} \tilde{a}_0{}^0 & \tilde{a}_0{}^1 & \tilde{a}_0{}^2 & \tilde{a}_0{}^3 \\ \tilde{a}_1{}^0 & \tilde{a}_1{}^1 & \tilde{a}_1{}^2 & \tilde{a}_1{}^3 \\ \tilde{a}_2{}^0 & \tilde{a}_2{}^1 & \tilde{a}_2{}^2 & \tilde{a}_2{}^3 \\ \tilde{a}_3{}^0 & \tilde{a}_3{}^1 & \tilde{a}_3{}^2 & \tilde{a}_3{}^3 \end{pmatrix} = \begin{pmatrix} a^0{}_0 & a^1{}_0 & a^2{}_0 & a^3{}_0 \\ a^0{}_1 & a^1{}_1 & a^2{}_1 & a^3{}_1 \\ a^0{}_2 & a^1{}_2 & a^2{}_2 & a^3{}_2 \\ a^0{}_3 & a^1{}_3 & a^2{}_3 & a^3{}_3 \end{pmatrix} . \tag{1.32}$$

Certain properties of the transformation matrix A can be deduced from (1.30). Taking the determinant of both sides gives us $\det(\tilde{A}gA) = \det(g)\det(A)^2 = \det(g)$. Since $\det(g) = -1$, we obtain

$$\det(A) = \pm 1. \tag{1.33}$$

One distinguishes two classes of transformations. *Proper* Lorentz transformations are continuously connected with the identity transformation $A = \mathbf{1}$. All other Lorentz transformations are *improper*. Proper transformations have necessarily $\det(A) = 1$. For improper Lorentz transformations it is sufficient, but not necessary, to have $\det(A) = -1$. For instance $A = -\mathbf{1}$ (space and time inversion) is an improper Lorentz transformation with $\det(A) = +1$.

Next the number of parameters, needed to specify completely a transformation in the group, follows from (1.30). Since A and g are 4×4 matrices, we have 16 equations for $4^2 = 16$ elements of A. But they are not all independent because of symmetry under transposition. The off-diagonal equations are identical in pairs. Therefore, we have $4+6 = 10$ linearly independent equations for the 16 elements of A. This means that there are *six free parameters*. In other words, the Lorentz group is a six–parameter group.

In the 19th century *Lie* invented the subsequent procedure to handle these parameters. Let us now consider only proper Lorentz transformations. To construct A explicitly, Lie makes the ansatz

$$A = e^L = \sum_{n=0}^{\infty} \frac{L^n}{n!},$$

where L is a 4×4 matrix. The determinant of A is

$$\det(A) = \det(e^L) = e^{\mathrm{Tr}(L)}. \tag{1.34}$$

It may be noted that $\det(A) = +1$ implies that L is traceless. Equation (1.30) can be written

$$g\tilde{A}g = A^{-1}. \tag{1.35}$$

From the definition of L, $\tilde{L}$ and the fact that $g^2 = \mathbf{1}$ we have (note $(g\tilde{L}g)^n = g\tilde{L}^n g$ and $\mathbf{1} = \left(\sum_{n=0}^{\infty} L^n/n!\right)\left(\sum_{n=0}^{\infty}(-L)^n/n!\right)$)

$$\tilde{A} = e^{\tilde{L}}, \quad g\tilde{A}g = e^{g\tilde{L}g} \quad \text{and} \quad A^{-1} = e^{-L} \ .$$

Therefore, (1.35) is equivalent to

$$g\tilde{L}g = -L \quad \text{or} \quad \widetilde{(gL)} = -gL.$$

The matrix gL is thus antisymmetric and it is left as an exercise to show that the general form of L is:

$$L = \begin{pmatrix} 0 & l^0{}_1 & l^0{}_2 & l^0{}_3 \\ l^0{}_1 & 0 & l^1{}_2 & l^1{}_3 \\ l^0{}_2 & -l^1{}_2 & 0 & l^2{}_3 \\ l^0{}_3 & -l^1{}_3 & -l^2{}_3 & 0 \end{pmatrix}. \tag{1.36}$$

It is customary to expand L in terms of six *infinitesimal generators*:

$$L = -\sum_{i=1}^{3}(\omega_i S_i + \zeta_i K_i) \quad \text{and} \quad A = e^{-\sum_{i=1}^{3}(\omega_i S_i + \zeta_i K_i)}. \tag{1.37}$$

The matrices are defined by

$$S_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.38}$$

and

$$K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{1.39}$$

They satisfy the following *Lie algebra* commutation relations:

$$[S_i, S_j] = \sum_{k=1}^{3} \epsilon^{ijk} S_k, \quad [S_i, K_j] = \sum_{k=1}^{3} \epsilon^{ijk} K_k, \quad [K_i, K_j] = -\sum_{k=1}^{3} \epsilon^{ijk} S_k,$$

where the commutator of two matrices is defined by $[A, B] = AB - BA$. Further ϵ_{ijk} is the completely antisymmetric Levi–Cevita tensor. Its definition in n–dimensions is

$$\epsilon^{i_1 i_2 \dots i_n} = \begin{cases} +1 \ \text{ for } \ (i_1, i_2, ..., i_n) \ \text{ being an even permutation of } \ (1, 2, ..., n), \\ -1 \ \text{ for } \ (i_1, i_2, ..., i_n) \ \text{ being an odd permutation of } \ (1, 2, ..., n), \\ 0 \ \text{ otherwise.} \end{cases} \tag{1.40}$$

To get the physical interpretation of equation (1.37) for A, it is suitable to work out simple examples. First, let $\vec{\zeta} = \omega_1 = \omega_2 = 0$ and $\omega_3 = \omega$. Then (this is left as exercise)

$$A = e^{-\omega S_3} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\omega & \sin\omega & 0 \\ 0 & -\sin\omega & \cos\omega & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.41}$$

which describes a rotation by the *angle* ω (in the clockwise sense) around the $\hat{e}_3$ axis. Next, let $\vec{\omega} = \zeta_2 = \zeta_3 = 0$ and $\zeta_1 = \zeta$. Then

$$A = e^{-\zeta K_1} = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.42}$$

is obtained, where ζ is known as the *boost parameter* or *rapidity*. The structure is reminiscent to a rotation, but with hyperbolic functions instead of circular, basically because of the relative negative sign between the space and time terms in eqn.(1.24). “Rotations” in the $x^0 - x^i$ planes are boosts and governed by an hyperbolic gemometry, whereas rotations in the $x^i - x^j$ $(i \neq j)$ planes are governed by the ordinary Euclidean geometry.

Finally, note that the parameters ω_i, ζ_i, $(i = 1, 2, 3)$ turn out to be real, as equation (1.29) implies that the elements of A have to be real. In the next subsection relativistic kinematics is discussed in more details.

1.1.6 Basic relativistic kinematics

The matrix (1.42) gives the Lorentz boost transformation

$$x'^0 = x^0 \cosh(\zeta) - x^1 \sinh(\zeta), \tag{1.43}$$

$$x'^1 = -x^0 \sinh(\zeta) + x^1 \cosh(\zeta), \tag{1.44}$$

$$x'^i = x^i, \; (i = 2, 3). \tag{1.45}$$

An interesting feature of equation (1.43) is that for spacelike points, say $x^1 > x^0 > 0$, a value ζ_0 for the rapidity exists, such that

$$\mathrm{sign}(x'^0) = -\mathrm{sign}(x^0)$$

for $\zeta > \zeta_0$, *i.e.* the time-ordering becomes reversed, whereas for timelike points such a reversal of the time-ordering is impossible. In figure 1.1 this is emphasized by calling the spacelike (with respect to $x_0 = 0$) region *elsewhere* in contrast to *future* and *past*.

Seen from K', the origin of K ($\Rightarrow x^1 = 0$) moves with uniform velocity $-v$. Using $x'^0 = ct'$, we obtain (x'^1 and x'^0 corresponding to the origin of K)

$$-\frac{v}{c} = \frac{x'^1}{x'^0} = \frac{-x^0 \sinh(\zeta)}{x^0 \cosh(\zeta)}.$$

The following definition of β and γ are frequently used conventions in the literature.

$$\beta = \frac{v}{c} = \tanh(\zeta) \;\text{ and }\; \gamma = \frac{1}{\sqrt{1-\beta^2}} = \cosh(\zeta) \;\;\Rightarrow\;\; \gamma\beta = \sinh(\zeta). \tag{1.46}$$

Hence, the Lorentz transformations (1.43), (1.44) follow in their familiar form

$$x'^0 = \gamma\,(\;\; x^0 - \beta x^1), \tag{1.47}$$

$$x'^1 = \gamma\,(-\beta x^0 + x^1). \tag{1.48}$$

To find the transformation law of an arbitrary vector $\vec{A}$ in case of a general relative velocity $\vec{v}$, it is convenient to decompose $\vec{A}$ into components parallel and perpendicular to $\vec{\beta} = \vec{v}/c$:

$$\vec{A} = A^{\parallel}\hat{\beta} + \vec{A}^{\perp} \;\text{ with }\; A^{\parallel} = \hat{\beta}\vec{A}.$$

Then the Lorentz transformation law is simply

$$A'^{\,0} = A^0 \cosh(\zeta) - A^{\|} \sinh(\zeta) = \gamma(\ \ A^0 - \beta A^{\|}), \tag{1.49}$$

$$A'^{\,\|} = -A^0 \sinh(\zeta) + A^{\|} \cosh(\zeta) = \gamma(-\beta A^0 + A^{\|}), \tag{1.50}$$

$$\vec{A}'^{\perp} = \vec{A}^{\perp}. \tag{1.51}$$

Here I have reserved the notation $A_{\|}$ and $\vec{A}_{\perp}$ for use in connection with covariant vectors: $A_{\|} = -A^{\|}$ and $\vec{A}_{\perp} = -\vec{A}^{\perp}$. We proceed deriving the addition theorem of velocities. Assume a particle moves with respect to K' with velocity $\vec{u}\,'$:

$$x'^{\,i} = c^{-1} u'^{\,i} x'^{\,0}.$$

Equations (1.47), (1.48) imply

$$\gamma\,(x^1 - \beta x^0) = c^{-1} u'^{\,1} \gamma\,(x^0 - \beta x^1).$$

Sorting with respect to x^1 and x^0 gives

$$\gamma\,\left(1 + \frac{u'^{\,1} v}{c^2}\right)\, x^1 = c^{-1} \gamma\,(u'^1 + v)\, x^0\,.$$

Using the definition of the velocity in K, $\vec{x} = c^{-1} \vec{u}\, x^0$, gives

$$u^1 = c\,\frac{x^1}{x^0} = \frac{u'^1 + v}{1 + \frac{u'^1 v}{c^2}}\,. \tag{1.52}$$

Along similar lines, we obtain for the two other components

$$u^i = \frac{u'^i}{\gamma\left(1 + \frac{u'^1 v}{c^2}\right)}, \quad (i = 2, 3). \tag{1.53}$$

To derive these equations, $\vec{v}$ was chosen along to the x^1-axis. For general $\vec{v}$ one only has to decompose $\vec{u}$ into its components parallel and perpendicular to the $\vec{v}$

$$\vec{u} = u^{\|}\hat{v} + \vec{u}^{\perp}\,,$$

where $\hat{v}$ is the unit vector in $\vec{v}$ direction, and obtains

$$u^{\|} = \frac{u'^{\,\|} + v}{1 + \frac{u'^{\,\|} v}{c^2}} \quad \text{and} \quad \vec{u}^{\perp} = \frac{\vec{u}'^{\,\perp}}{\gamma\left(1 + \frac{u'^{\,\|} v}{c^2}\right)}. \tag{1.54}$$

From this addition theorem of velocities it is obvious that the velocity itself is not part of of a 4-vector. The relativistic generalization is given in subsection (1.1.9). It is left as exercise to derive the addition theorem for the rapidity used in equation (1.46). The results of this problem is central for the interpretation of the rapidities as angles of a hyperbolic geometry.

1.1.7 Proper time

We now explore in more details Minkowski's earlier mentioned concept of a 4-dimensional space-time. The Minkowski space allows to depict *world lines* of particles. A useful concept for a particle (or observer) traveling along its world line is its *proper time* or *eigenzeit*. Assume the particle moves with velocity $\vec{v}(t)$, then $d\vec{x} = \vec{\beta} dx^0$ holds, and the infinitesimal invariant along its world line is

$$(ds)^2 = (dx^0)^2 - (d\vec{x})^2 = (c\,dt)^2\,(1-\beta^2)\ . \tag{1.55}$$

Each instantaneous rest frame of the particle is an inertial frame. The increment of time $d\tau$ in such an instantaneous rest frame is thus a Lorentz invariant quantity which takes the form

$$d\tau = dt\,\sqrt{1-\beta^2} = dt\,\gamma^{-1} = dt/\cosh\zeta\ , \tag{1.56}$$

where τ is called proper time. It is the time which a co-moving clock shows. As $\gamma(\tau) \geq 1$ it follows

$$t_2 - t_1 = \int_{\tau_1}^{\tau_2} \gamma(\tau) d\tau \geq \tau_2 - \tau_1. \tag{1.57}$$

This phenomenon is known as *time dilatation*. A moving clock runs more slowly than a stationary clock. Note that equation (1.57) applies to general paths of a clock, including those with acceleration. Relevant is that the time coordinates t_2 and t_1 refer to an inertial system. Two experimental examples are: (i) Time of flight of unstable particles in high energy scattering experiments, where these particles move at velocities close to the speed of light. (ii) Explicit verification through travel with atomic clocks on air planes [4, 9].

1.1.8 Plane waves and the relativistic Doppler effect

Let us choose coordinates with respect to an inertial frame K. In complex notation a plane wave is defined by the equation

$$W(x) = W(x^0, \vec{x}) = W_0\ \exp[\,i\,(k^0\,x^0 - \vec{k}\,\vec{x})\,]\,, \tag{1.58}$$

where $W_0 = U_0 + i\,V_0$ is a complex *amplitude*. The vector $\vec{k}$ is called *wave vector*. It becomes a 4-vector (k^α) by identifying

$$k^0 = \omega/c \tag{1.59}$$

as its zero-component, where ω is the *angular frequency* of the wave. Waves of the form (1.58) may either propagate in a medium (water, air, shock waves, etc.) or in vacuum (light waves, particle waves in quantum mechanics). We are interested in the latter case, as the other defines a preferred inertial frame, namely the one where the medium is at rest. The *phase* of the wave is defined by

$$\Phi(x) = \Phi(x^0, \vec{x}) = k^0\,x^0 - \vec{k}\,\vec{x} = \omega\,t - \vec{k}\,\vec{x}\,. \tag{1.60}$$

When (k^α) is a 4-vector, it follows that the phase is a scalar, *invariant* under Lorentz transformations

$$\Phi'(x') = k'_\alpha \, x'^{\,\alpha} = k_\alpha \, x^\alpha = \Phi(x)\,. \tag{1.61}$$

That this is correct can be seen as follows: For an observer at a fixed position $\vec{x}$ (note the term $\vec{k}\,\vec{x}$ is then constant) the wave performs a periodic motion with *period*

$$T = \frac{2\pi}{\omega} = \frac{1}{\nu}\,, \tag{1.62}$$

where ν is the *frequency*. In particular, the phase (and hence the wave) takes identical values on the two-dimensional hyperplanes perpendicular to $\vec{k}$. Namely, let $\hat{k}$ be the unit vector in $\vec{k}$ direction, by decomposing $\vec{x}$ into components parallel and perpendicular to $\vec{k}$, $\vec{x} = x^{\|}\,\hat{k} + \vec{x}^{\,\perp}$, the phase becomes

$$\Phi = \omega\, t - k\, x^{\|}\,, \tag{1.63}$$

where $k = |\vec{k}|$ is the length of the vector $\vec{k}$. Phases which differ by multiples of 2π give the same values for the wave W. For example, when we take $V_0 = 0$, the real part of the wave becomes

$$W_x = U_0\, \cos(\omega\, t - k\, x^{\|})$$

and $\Phi = 0,\ n\,2\pi,\ n = \pm1, \pm2, ...$ describes the wave crests. From (1.63) it follows that the crests pass by our observer with speed $\vec{u} = u\,\hat{k}$, where

$$u = \frac{\omega}{k} \quad \text{as for } \Phi = 0 \text{ we have} \quad x^{\|} = \frac{\omega}{k}\, t\,. \tag{1.64}$$

Let our observer count the number of wave crests passing by. How has then the wave (1.58) to be described in another inertial frame K'? An observer in K' who counts the number of wave crests, passing through the same space-time point at which our first observer already counts, must get at same number. After all, the coordinates are just labels and the physics is the same in all systems. When in frame K the wave takes its maximum at the space-time point (x^α) it must also be at its maximum in K' at the same space-time point in appropriately transformed coordinates (x'^α). More generally, this is true for every value of the phase, which hence has to be a scalar.

As (k^α) is a 4-vector the transformation law for angular frequency and wave vector is just a special case of equations (1.49), (1.50) and (1.51)

$$k'^{\,0} = k^0 \cosh(\zeta) - k^{\|} \sinh(\zeta) = \gamma(k^0 - \beta k^{\|})\,, \tag{1.65}$$

$$k'^{\,\|} = -k^0 \sinh(\zeta) + k^{\|} \cosh(\zeta) = \gamma(k^{\|} - \beta k^0)\,, \tag{1.66}$$

$$\vec{k}'^{\,\perp} = \vec{k}^{\perp}\,, \tag{1.67}$$

where the notation $k^{\|}$ and $k^{\perp}$ is with respect to the relative velocity of the two frames, $\vec{v}$. These transformation equations for the frequency and the wave vector describe the relativistic Doppler effect. To illustrate their meaning, let us specialize to the case of a light source, which recedes along its wave vector from the observer, i.e. $\vec{v} \parallel \vec{k}$. The equation for the wave speed (1.64) implies

$$c = \frac{\omega}{k} \;\Rightarrow\; k = |\vec{k}| = \frac{\omega}{c} = k^0$$

and (1.65) becomes

$$k'^{\,0} = \gamma\,(k^0 - \beta\,k) = \gamma\,(1-\beta)\,k^0 = k^0\,\sqrt{\frac{1-\beta}{1+\beta}}$$

or

$$\omega' = \frac{\nu'}{2\pi} = \omega\,\sqrt{\frac{1-\beta}{1+\beta}} = \frac{\nu}{2\pi}\,\sqrt{\frac{1-\beta}{1+\beta}}\,.$$

Now, $c = \nu\lambda = \nu'\lambda'$, where λ is the wavelength in K and λ' the wavelength in K'. Consequently, we have

$$\lambda' = \lambda\,\sqrt{\frac{1+\beta}{1-\beta}}\,.$$

The wave length of light from the receding source is larger as it is for a source at rest. This is an example of the red-shift, which is of major importance, for instance when analyzing spectral lines in astrophysics. Using the method of section 1, a single light signal suffices now to obtain position and speed of a distant perfectly reflecting mirror.

1.1.9 Relativistic dynamics

On a basic level this section deals with the relativistic generalization of energy, momentum and their conservation laws. So far we have introduced two units, meter to measure distances and seconds to measure time. Both are related through a fundamental constant, the speed of light, so that there is really only one independent unit up to now. In the definition of the momentum a new, independent dimensional quantity enters, the mass of a particle. This unit is defined through the gravitational law, which is out of the scope of this article. Ideally, one would like to defines mass of a body just as multiples of the mass of an elementary particle, say an electron or proton. However, this has remained too inaccurate. The mass unit has so far resisted modernization and the mass unit

$$1\ kilogram\ [kg] \ = \ 1000\ gram\ [g]$$

is still defined through a one kilogram standard object a cylinder of platinum alloy which is kept at the International Bureau of Weights and Measures at Sévres, France.

Let us consider a point-like particle in its rest-frame and denote its mass there by m_0. In any other frame the rest-mass of the particle is still m_0, which in this way is defined as a scalar. It may be noted that most books in particle and nuclear physics simply use m to denote the rest-mass, whereas many books on special relativity employ the ugly notation to use m for a mass which is proportional to the energy, *i.e.* the zero component of the energy-momentum vector introduced below. To avoid confusion, we use m_0 for the rest mass.

In the non-relativistic limit the momentum is defined by $\vec{p} = m_0\vec{u}$. We give up this relation now and define $\vec{p}$ as part of a relativistic 4-vector (p^α). Consider a particle at rest in frame K. Then the non-relativistic limit is correct, and $\vec{p} = 0$. Assume now that frame K' is moving with a small velocity $\vec{v}$ with respect to K. Then $\vec{p}\,' = -m_0\vec{v}$ has to hold approximately. On the other hand, the transformation laws (1.49),(1.50),(1.51) for vectors (note $\vec{p}\,\|\,\vec{\beta} = \vec{v}/c$) imply

$$\vec{p}^{\,'} = \gamma\,(\vec{p} - \vec{\beta}\,p^0)\,,$$

where $\vec{p} = 0$. In the nonrelativstic limit $\gamma\,\beta \to \beta$ and consistency requires $p^0 = m_0\,c$ in the rest frame, so that we get $\vec{p}^{\,'} = -m_0\,\gamma\vec{v}$.

Consequently, for a particle moving with velocity $\vec{u}$ in frame K

$$\vec{p} = m_0\,\gamma\,\vec{u} \tag{1.68}$$

is the correct relation between relativistic momentum and velocity. From the invariance of the scalar product, $p_\alpha p^\alpha = (p^0)^2 - \vec{p}^{\,2} = p'_\alpha p'^\alpha = m_0^2 c^2$,

$$p^{\,0} = +\sqrt{m_0^2 c^2 + \vec{p}^{\,\prime 2}} \tag{1.69}$$

follows, which is of course consistent with calculating p^0 via the Lorentz transformation law (1.49). It should be noted that $c\,p^0$ has the dimension of an energy, i.e. the relativistic energy of a particle is

$$E = c\,p^0 = +\sqrt{m_0^2 c^4 + c^2\vec{p}^{\,2}} = m_0 c^2 + \frac{\vec{p}^{\,2}}{2m_0} + \,...\;, \tag{1.70}$$

where the second term is just the non-relativistic kinetic energy $T = \vec{p}^{\;2}/(2m_0)$. The first term shows that (rest) mass and energy can be transformed into one another [3]. In processes where the mass is conserved we just do not notice it. At this point it may be noted that special relativity books like [7] tend to use the notation $m = c\,p^0$ as definition of the mass, whereas in particle and nuclear physics the mass of a particle is its rest mass. Together with (1.70) the first definition yields the equation $E = m\,c^2$. Here the second definition is favored, because it defines the mass of a particle as a invariant scalar and the essence of Einstein's equation is captured correctly by

$$E_0 = m_0\,c^2\,,$$

where E_0 is the energy of a massive body (or particle) in its rest frame. The particle and nuclear physics literature does not use a subscript $_0$ and m denotes the rest mass.

Non-relativistic momentum conservation $\vec{p}_1 + \vec{p}_2 = \vec{q}_1 + \vec{q}_2$, where $\vec{p}_i$, $(i = 1,2)$ are the momenta of two incoming, and $\vec{q}_i$, $(i = 1,2)$ are the momenta of two outgoing particles, becomes relativistic *energy–momentum conservation*:

$$p_1^\alpha + p_2^\alpha = q_1^\alpha + q_2^\alpha. \tag{1.71}$$

Useful formulas in relativistic dynamics are

$$\gamma = \frac{p_0}{m_0\, c} = \frac{E}{m_0\, c^2} \quad \text{and} \quad \beta = \frac{|\vec{p}|}{p^0}. \tag{1.72}$$

Proof: $(p^0)^2 = c^2[(1-\beta^2)\, m_0^2 + \beta^2 m_0^2]/(1-\beta^2) = c^2\gamma^2 m_0^2$ and $\vec{p}^2 = \beta^2 c^2 \gamma^2 m_0^2 = \beta^2 (p^0)^2$. Further, the contravariant generalization of the velocity vector is given by

$$U^\alpha = \frac{dx^\alpha}{d\tau} = \gamma u^\alpha \quad \text{with} \quad u^0 = c\,, \tag{1.73}$$

compare the definition of the infinitesimal proper time (1.56). The relativistic generalization of the force is then the 4-vector

$$f^\alpha = \frac{d\,p^\alpha}{d\,\tau} = m_0\, \frac{dU^\alpha}{d\tau}\,, \tag{1.74}$$

where the last equality can only be used for particle with non-zero rest mass.

1.2 Maxwell Equations

As before all considerations are in vacuum, as for fields in a medium a preferred reference system exists. Maxwell's equations in their standard form in vacuum are

$$\nabla \vec{E} = 4\pi\rho, \quad \nabla \times \vec{B} - \frac{1}{c}\frac{\partial \vec{E}}{\partial t} = \frac{4\pi}{c}\vec{J}, \tag{1.75}$$

and

$$\nabla \vec{B} = 0, \quad \nabla \times \vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0. \tag{1.76}$$

Here ∇ is the Nabla operator. Note that $\nabla\vec{E} = \nabla \cdot \vec{E}$, $\vec{a}\,\vec{b} = \vec{a}\cdot\vec{b}$ etc. throughout the script. $\vec{E}$ is the *electric field* and $\vec{B}$ the *magnetic field* in vacuum. When matter gets involved one introduces the *applied electric field* $\vec{D}$ and the *applied magnetic field* $\vec{H}$. Here we follow the convention of, for instance, Tipler [8] and use the notation magnetic field for the measured field $\vec{B}$, in precisely the same way as it is done for the electric field $\vec{E}$, It should be noted that this is at odds with the notation in the book by Jackson [5], where (historically correct, but quite confusingly) $\vec{H}$ is called magnetic field and $\vec{B}$ magnetic flux or magnetic induction.

Equations (1.75) are the *inhomogeneous* and equations (1.76) are the *homogeneous* Maxwell equations in vacuum.

The charge density ρ (charge per unit volume) and the current density $\vec{J}$ (charge passing through a unit area per time unit) are obviously given once a charge unit is defined through some measurement prescription. From a theoretical point of view the electrical charge unit is best defined by the magnitude of the charge of a single electron (fundamental charge unit). In more conventional units this reads [1]

$$|q_e| = 4.80320420(19) \times 10^{-10}\,[esu] = 1.602176462(63) \times 10^{-19}\, Coulomb\ [C] \tag{1.77}$$

where the errors are given in parenthesis. Definitions of the electrostatic unit $[esu]$ and the $Coulomb$ $[C]$ through measurement prescriptions are given in elementary physics textbooks like [8].

The choice of constants in the inhomogeneous Maxwell equations defines units for the electric and magnetic field. The given conventions $4\pi\rho$ and $(4\pi/c)\vec{J}$ are customarily used in connection with Gaussian units, where the charge is defined in electrostatic units (esu).

In the next subsections the concepts of fields and currents are discussed in the relativistic context and the electromagnetic field equations follow in the last subsection.

1.2.1 Fields and currents

A tensor field is just a tensor function which depends on the coordinates of Minkowski space:

$$T^{\dots\alpha\dots\beta\dots} = T^{\dots\alpha\dots\beta\dots}(x)\,.$$

It is called *static* when there is no time dependence. For instance $\vec{E}(\vec{x})$ in electrostatics would be a static vector field in three dimensions. We are here, of course, primarily interested in contravariant or covariant fields in four dimensions, like vector fields $A^\alpha(x)$.

Suppose n electric charge units are contained it a small volume v, such that we can talk about the position $\vec{x}$ of this volume. The corresponding electrical charge density at the position of that volume is then just $\rho = n/v$ and the electrical current is defined as the charge that passes per unit time through a surface element of such a volume. We demand now that the electric charge density ρ and the electric current $\vec{J}$ form a 4-vector:

$$(J^\alpha) = \begin{pmatrix} c\rho \\ \vec{J} \end{pmatrix}.$$

Here, the factor c is introduced by dimensional reasons and we have suppressed the space-time dependence, *i.e.* $J^\alpha = J^\alpha(x)$ forms a vector field. It is left as a problem to write down the 4-current for a point particle of elementary charge q_e.

The *continuity equation* takes the simple, covariant form

$$\partial_\alpha J^\alpha = 0\,. \tag{1.78}$$

Finally, the charge of a particle in its rest frame is an invariant:

$$c^2 q_0^2 = J_\alpha J^\alpha\,.$$

1.2.2 The inhomogeneous Maxwell equations

The inhomogeneous Maxwell equations are obtained by writing down the simplest covariant equation which yields a 4-vector as first order derivatives of six fields. From undergraduate E&M we remember the electric and magnetic fields, $\vec{E}$ and $\vec{B}$, as the six central fields of electrodynamics. We now like to describe them in covariant form. A 4-vector is unsuitable as we like to describe six quantities E^x, E^y, E^z and B^x, B^y, B^z. Next, we may try a rank two tensor $F^{\alpha\beta}$. Then we have $4 \times 4 = 16$ quantities at our disposal. These are now too many. But, one may observe that an symmetric tensor stays symmetric under Lorentz transformation and an antisymmetric tensor stays antisymmetric. Hence, instead of looking at the full second rank tensor one has to consider its symmetric and antisymmetric parts separately.

By requesting $F^{\alpha\beta}$ to be antisymmetric,

$$F^{\alpha\beta} = -F^{\beta\alpha}, \tag{1.79}$$

this number is reduced to precisely six. The diagonal elements do now vanish,

$$F^{00} = F^{11} = F^{22} = F^{33} = 0\,.$$

The other elements follow through (1.79) from $F^{\alpha\beta}$ with $\alpha < \beta$. As desired, this gives $(16-4)/2 = 6$ independent elements to start with.

Up to an over–all factor, which is chosen by convention, the only way to obtain a 4-vector through differentiation of $F^{\alpha\beta}$ is

$$\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c} J^\beta\ . \tag{1.80}$$

This is the *inhomogeneous Maxwell equation* in covariant form. Note that it determines the physical dimensions of the electric fields, the factor $c^{-1}4\pi$ on the right-hand side corresponds to Gaussian units. The continuity equation (1.78) is a simple consequence of the inhomogeneous Maxwell equation

$$\frac{4\pi}{c}\partial_\beta J^\beta = \partial_\beta\partial_\alpha F^{\alpha\beta} = 0$$

because the contraction with the symmetric tensor $(\partial_\beta\partial_\alpha)$ with the antisymmetric tensor $F^{\alpha\beta}$ is zero.

Let us choose $\beta = 0, 1, 2, 3$ respectively, comparison of equation (1.80) with the inhomogeneous Maxwell equations in their standard form (1.75) yields

$$\left(F^{\alpha\beta}\right) = \begin{pmatrix} 0 & -E^x & -E^y & -E^z \\ E^x & 0 & -B^z & B^y \\ E^y & B^z & 0 & -B^x \\ E^z & -B^y & B^x & 0 \end{pmatrix}. \tag{1.81}$$

Or, in components

$$F^{i0} = E^i \ \text{ and } \ F^{ij} = -\sum_k \epsilon^{ijk}\, B^k \Leftrightarrow B^k = -\frac{1}{2}\sum_i\sum_j \epsilon^{kij} F^{ij}\ . \tag{1.82}$$

Next, $F_{\alpha\beta} = g_{\alpha\gamma}g_{\beta\delta}F^{\gamma\delta}$ implies:

$$F_{0i} = -F^{0i},\ F_{00} = F^{00} = 0,\ F_{ii} = F^{ii} = 0, \quad \text{and} \quad F_{ij} = F^{ij}.$$

Consequently,

$$(F_{\alpha\beta}) = \begin{pmatrix} 0 & E^x & E^y & E^z \\ -E^x & 0 & -B^z & B^y \\ -E^y & B^z & 0 & -B^x \\ -E^z & -B^y & B^x & 0 \end{pmatrix}. \tag{1.83}$$

1.2.3 Four-potential and homogeneous Maxwell equations

We remember that the electromagnetic fields may be written as derivatives of appropriate potentials. The only covariant option are terms like $\partial^\alpha A^\beta$. To make $F^{\alpha\beta}$ antisymmetric, we have to subtract $\partial^\beta A^\alpha$:

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha. \tag{1.84}$$

It is amazing to note that the homogeneous Maxwell equations follow now for free from (1.84). The dual electromagnetic tensor is defined

$${}^*F^{\alpha\beta} = \frac{1}{2}\epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta}, \tag{1.85}$$

and it holds

$$\partial_\alpha {}^*F^{\alpha\beta} = 0. \tag{1.86}$$

Proof:

$$\partial_\alpha {}^*F^{\alpha\beta} = \frac{1}{2}\left(\epsilon^{\alpha\beta\gamma\delta}\partial_\alpha\partial_\gamma A_\delta - \epsilon^{\alpha\beta\gamma\delta}\partial_\alpha\partial_\delta A_\gamma\right) = 0.$$

This first term is zero due to (1.21), because $\epsilon^{\alpha\beta\gamma\delta}$ is antisymmetric in (α, γ), whereas the derivative $\partial_\alpha\partial_\gamma$ is symmetric in (α, γ). Similarly the other term is zero. The homogeneous Maxwell equation is related to the fact that the right-hand side of equation (1.84) expresses six fields in terms of a single 4-vector. An equivalent way to write it is the equation

$$\partial^\alpha F^{\beta\gamma} + \partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta} = 0\ . \tag{1.87}$$

The proof is left as an exercise to the reader.

Let us mention that the homogeneous Maxwell equation (1.86) or (1.87), and hence our demand that the field can be written in the form (1.84), excludes magnetic monopoles.

The elements of the dual tensor may be calculated from their definition (1.85). For example,

$${}^*F^{02} = \epsilon^{0213} F_{13} = -F_{13} = -B^y,$$

where the first step exploits the anti-symmetries $\epsilon^{0231} = -\epsilon^{0213}$ and $F_{31} = -F_{13}$. Calculating six components, and exploiting antisymmetry of ${}^*F^{\alpha\beta}$, we arrive at

$$({}^*F^{\alpha\beta}) = \begin{pmatrix} 0 & -B^x & -B^y & -B^z \\ B^x & 0 & E^z & -E^y \\ B^y & -E^z & 0 & E^x \\ B^z & E^y & -E^x & 0 \end{pmatrix}. \tag{1.88}$$

The homogeneous Maxwell equations in their form (1.76) provide a non–trivial consistency check for (1.86), which is of course passed. It may be noted that, in contrast to the inhomogeneous equations, the homogeneous equations determine the relations with the $\vec{E}$ and $\vec{B}$ fields only up to an over-all $\pm$ sign, because there is no current on the right-hand side.

A notable observation is that equation (1.84) does not determine the potential uniquely. Under the transformation

$$A^\alpha \mapsto A'^\alpha = A^\alpha + \partial^\alpha \psi, \tag{1.89}$$

where $\psi = \psi(x)$ is an arbitrary scalar function, the electromagnetic field tensor is invariant: $F'^{\alpha\beta} = F^{\alpha\beta}$, as follows immediately from $\partial^\alpha\partial^\beta\psi - \partial^\beta\partial^\alpha\psi = 0$. The transformations (1.89) are called *gauge transformation*[1]. The choice of a convenient gauge is at the heart of many application.

1.2.4 Lorentz transformation for the electric and magnetic fields

The electric $\vec{E}$ and magnetic $\vec{B}$ fields are *not* components of a Lorentz four-vector, but part of the rank two the electromagnetic field ($F^{\alpha\beta}$) given by (1.81). As for any Lorentz tensor, we immediately know its behavior under Lorentz transformation

$$F'^{\alpha\beta} = a^\alpha_{\ \gamma}\, a^\beta_{\ \delta}\, F^{\gamma\delta} \ . \tag{1.90}$$

Using the explicit form (1.42) of $A = (a^\alpha_{\ \beta})$ for boosts and (1.81), it is left as an exercise for the reader to derive the transformation laws

$$\vec{E}' = \gamma \left(\vec{E} + \vec{\beta} \times \vec{B}\right) - \frac{\gamma^2}{\gamma+1}\, \vec{\beta} \left(\vec{\beta}\, \vec{E}\right) \ , \tag{1.91}$$

and

$$\vec{B}' = \gamma \left(\vec{B} - \vec{\beta} \times \vec{E}\right) - \frac{\gamma^2}{\gamma+1}\, \vec{\beta} \left(\vec{\beta}\, \vec{B}\right) \ . \tag{1.92}$$

1.2.5 Lorentz force

Relativistic dynamics of a point particle (more generally any mass distribution) gets related to the theory of electromagnetic fields, because an electromagnetic field causes a change of the 4-momentum of a charged particle. On a deeper level this phenomenon is related to the conservation of energy and momentum and the fact that an electromagnetic carries energy as well as momentum. Here we are content with finding the Lorentz covariant form, assuming we know already that such the approximate relationship.

We consider a charged point particle in an electromagnetic field $F^{\alpha\beta}$. Here *external* means from sources other than the point particle itself and that the influence of the point

[1] In quantum field theory these are the gauge transformations of 2. kind. Gauge transformations of 1. kind transform fields by a constant phase, whereas for gauge transformation of the 2. kind a space–time dependent function is encountered.

particle on these other sources (possibly causing a change of the field $F^{\alpha\beta}$) is neglected. The infinitesimal change of the 4-momentum of a point point particle is dp^α and assumed to be proportional to (i) its charge q and (ii) the external electromagnetic field $F^{\alpha\beta}$. This means, we have to contract $F^{\alpha\beta}$ with some infinitesimal covariant vector to get dp^α. The simplest choice is dx_β, what means that the amount of 4-momentum change is assumed to be proportional to the space-time length at which the particle experiences the electromagnetic field. Hence, we have determined dp^α up to a proportionality constant, which depends on the choice of units. Gaussian units are defined by choosing c^{-1} for this proportionality constant and we have

$$dp^\alpha = \pm \frac{q}{c} F^{\alpha\beta}\, dx_\beta. \tag{1.93}$$

It is a consequence of energy conservation, in this context known as Lenz's law, that the force between charges of equal sign has to be repulsive. This corresponds to the plus sign and we arrive at

$$dp^\alpha = \frac{q}{c} F^{\alpha\beta}\, dx_\beta. \tag{1.94}$$

Experimental measurements are of course in agreement with this sign. But the remarkable point is that energy conservation and the general structure of the theory already imply that the force between charges of equal sign has to be *repulsive*. Therefore, despite the similarity of the Coulomb's inverse square force law with Newton's law it impossible to build a theory of gravity along the lines of this chapter, *i.e.* to use the 4-momentum p^α as source in the inhomogeneous equation (1.80). The resulting force would necessarily be repulsive. Experiments show also that positive and negative electric charges exist and deeper insight about their origin comes from the relativistic Lagrange formulation, which ultimately has to include Dirac's equation for electrons and leads then to Quantum Electrodynamics.

Taking the derivative with respect to the proper time, we obtain the 4-force acting on a charged particle, called *Lorentz force*,

$$f^\alpha = \frac{dp^\alpha}{d\tau} = \frac{q}{c} F^{\alpha\beta} U_\beta. \tag{1.95}$$

As in equation (1.74) $f^\alpha = m_0\, dU^\alpha/d\tau$ holds for non-zero rest mass and the definition of the contravariant velocity is given by equation (1.73).

Using the representation (1.81) of the electromagnetic field the time component of the relativistic Lorentz force, which describes the change in energy, is

$$f^0 = \frac{dp^0}{d\tau} = -\frac{q}{c}\left(\vec{E}\vec{U}\right) . \tag{1.96}$$

To get the space component of the Lorentz force we use besides (1.81) equation (1.82) which give the equality

$$\frac{q}{c}\sum_{j=1}^{3} F^{ij} U_j = -\frac{q}{c}\sum_{j=1}^{3}\sum_{k=1}^{3} \epsilon^{ijk} B^k U_j$$

The space components combine into the well-known equation

$$\vec{f} = q\,\gamma\,\vec{E} + \frac{q}{c}\,\vec{U} \times \vec{B} \tag{1.97}$$

where our derivation reveals that the relativistic velocity (1.73) of the charge q and not its velocity $\vec{v}$ enters the force equation. This allows, for instance, correct force calculations for fast flying electrons in a magnetic field. The equation (1.97) for $\vec{f}$ may now be used to define a measurement prescription for an electric charge unit.

Bibliography

[1] P.J. Mohr and B.N. Taylor, *CODATA Recommended Values of the Fundamental Physical Constants: 1998*, J. of Physical and Chemical Reference Data, to appear. See the website of the National Institute of Standards and Technology (NIST) at `physics.nist.gov/constants`.

[2] A. Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik 17 (1905) 891–921.

[3] A. Einstein, *Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?*, Annalen der Physik 18 (1906) 639–641.

[4] C. Hefele and R. Keating, Science 177 (1972) 166, 168.

[5] J.D. Jackson, *Classical Electrodynamics*, Second Edition, John Wiley & Sons, 1975.

[6] B.W. Petley, Nature 303 (1983) 373.

[7] W. Rindler, *Introduction to Special Relativity*, Clarendon Press, Oxford 1982.

[8] P.A. Tipler, *Physics for Scientists and Engineers*, Worth Publishers, 1995.

[9] R.F.C. Vessot and M.W. Levine, GRG 10 (1979) 181.