1 Special relativity

1.1 Minkowski space

Inertial frames and the principle of relativity Newton's *Lex Prima* (or the Galilean law of inertia) states: *Each force-less mass point stays at rest or moves on a straight line at constant speed.* In a Cartesian *inertial* coordinate system, Newton's lex prima becomes

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \frac{\mathrm{d}^2y}{\mathrm{d}t^2} = \frac{\mathrm{d}^2z}{\mathrm{d}t^2} = 0\,. \tag{1.1}$$

Most often, we call such a coordinate system just an *inertial frame*. Newton's first law is not just a trivial consequence of its second one, but may be seen as practical definition of those reference frames for which his following laws are valid.

Since the symmetries of Euclidean space, translations $\boldsymbol{a}$ and rotations R, leave Eq. (1.1) invariant, all frames connected by $\boldsymbol{x}' = R\boldsymbol{x} + \boldsymbol{a}$ to an inertial frame are inertial frames too. Additionally, proper Galilean transformations $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{v}t$ connect inertial frames moving with relative speed $\boldsymbol{v}$. The principle of relativity states that the physics in all inertial frames is the same.

If we consider two frames with relative velocity along the x direction, then the most general linear transformation between the two frames is

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} At + Bx \\ Dt + Ex \\ y \\ z \end{pmatrix} = \begin{pmatrix} At + Bx \\ A(x - vt) \\ y \\ z \end{pmatrix}. \tag{1.2}$$

Newtonian physics assumes the existence of an absolute time, $t = t'$, and thus $A = 1$ and $B = 0$. This leads to the "common sense" addition law of velocities. Time differences Δt and space differences

$$\Delta\boldsymbol{x}_{12}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \tag{1.3}$$

are separately invariant.

Lorentz transformations In special relativity, we replace the Galilean transformations as symmetry group of space and time by Lorentz transformations Λ. The latter are all those coordinate transformations $x^\mu \to \tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu$ that keep the squared distance

$$\Delta s^2 \equiv (ct_1 - ct_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2 \tag{1.4}$$

between two space-time events $x_1^\mu = (ct_1, \boldsymbol{x}_1)$ and $x_2^\mu = (ct_2, \boldsymbol{x}_2)$ invariant. The distance of two infinitesimally close space-time events is called the line-element $\mathrm{d}s$ of the space-time. In Minkowski space, it is given by

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \tag{1.5}$$

using a Cartesian inertial frame.

If we replace t by $-\mathrm{i}t$ in Δs^2, the difference between two space-time events becomes (minus) the normal Euclidean distance. Similarly, the identity $\cos^2\alpha + \sin^2\alpha = 1$ for an imaginary angle $\eta = \mathrm{i}\alpha$ becomes $\cosh^2\eta - \sinh^2\eta = 1$. Thus a close correspondence exists between rotations R_{ij} in Euclidean space which leave $\Delta\boldsymbol{x}^2$ invariant and Lorentz transformations $\Lambda^\mu{}_\nu$ which leave Δs^2 invariant. We try therefore as a guess for a boost along the x direction

$$\tilde{ct} = ct\cosh\eta + x\sinh\eta\,, \tag{1.6}$$

$$\tilde{x} = ct\sinh\eta + x\cosh\eta\,, \tag{1.7}$$

with $\tilde{y} = y$ and $\tilde{z} = z$. Direct calculation shows that Δs^2 is invariant as desired. Consider now in the system $\tilde{K}$ the origin of the system K. Then $x = 0$ and

$$\tilde{x} = ct\sinh\eta \quad and \quad \tilde{ct} = ct\cosh\eta\,. \tag{1.8}$$

Dividing the two equations gives $\tilde{x}/\tilde{ct} = \tanh\eta$. Since $\beta = \tilde{x}/\tilde{ct}$ is the relative velocity of the two systems measured in units of c, the imaginary "rotation angle η" equals the rapidity

$$\eta = \operatorname{arctanh}\beta\,. \tag{1.9}$$

Note that the rapidity η is a more natural variable than v or β to characterise a Lorentz boost, because η is additive: Boosting a particle with rapidity $\boldsymbol{\eta}_1$ by $\boldsymbol{\eta}$ leads to the rapidity $\boldsymbol{\eta}_2 = \boldsymbol{\eta}_1 + \boldsymbol{\eta}$. Using the following identities,

$$\cosh\eta = \frac{1}{\sqrt{1-\tanh^2\eta}} = \frac{1}{\sqrt{1-\beta^2}} \equiv \gamma \tag{1.10}$$

$$\sinh\eta = \frac{\tanh\eta}{\sqrt{1-\tanh^2\eta}} = \frac{\beta}{\sqrt{1-\beta^2}} = \gamma\beta \tag{1.11}$$

in (1.6) gives the standard form of the Lorentz transformations,

$$\tilde{x} = \frac{x+vt}{\sqrt{1-\beta^2}} = \gamma(x+\beta ct) \tag{1.12}$$

$$c\tilde{t} = \frac{ct+vx/c}{\sqrt{1-\beta^2}} = \gamma(ct+\beta x)\,. \tag{1.13}$$

The inverse transformation is obtained by replacing $v \to -v$ and exchanging quantities with and without tilde.

We can interpret the line-element $\mathrm{d}s^2$ as a scalar product, if we introduce the metric tensor $\eta_{\mu\nu}$ with elements

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.14}$$

and a scalar product of two vectors with coordinates a^μ and b^μ as

$$a\cdot b \equiv \eta_{\mu\nu}a^\mu b^\nu = a_\mu b^\mu = a^\mu b_\mu\,. \tag{1.15}$$

Here we used also Einstein's summation convention, cf. the box for details. In the last part of (1.15), we "lowered an index:" $a_\mu = \eta_{\mu\nu} a^\mu$ or $b_\mu = \eta_{\mu\nu} b^\mu$. Next we introduce the opposite operation of rasing an index by $a^\mu = \eta^{\mu\nu} a_\mu$. Since raising and lowering are inverse operations, we have $\eta_{\mu\nu}\eta^{\nu\sigma} = \delta_\mu^\sigma$. Thus the elements of $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$ form inverse matrices, which agree with (1.14) for a Cartesian intertial coordinate frame in Minkowski space.

Einstein's summation convention:

1. Two equal indices, of which one has to be an upper and one an lower index, imply summation. We use Greek letters for indices from zero to three, $\mu = 0, 1, 2, 3$, and Latin letters for indices from one to three, $i = 1, 2, 3$. Thus

$$a_\mu b^\mu \equiv \sum_{\mu=0}^{3} a_\mu b^\mu = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 = a^0 b^0 - \boldsymbol{a} \cdot \boldsymbol{b} = a^0 b^0 - a^i b^i \,.$$

2. Summation are dummy indices and can be freely exchanged; the free indices of the LHS and RHS of an equation have to agree. Hence

$$8 = a_\mu^\mu = c_{\mu\nu} d^{\mu\nu} = c_{\mu\sigma} d^{\mu\sigma}$$

is okay, while $a_\mu = b^\mu$ or $a^\mu = b^{\mu\nu}$ compares apples to oranges.

In Minkowski space, we call a four-vector any four-tupel V^μ that transforms as $\tilde{V}^\mu = \Lambda^\mu{}_\nu V^\nu$. By convention, we associate three-vectors with the spatial part of vectors with upper indices, e.g. we set $x^\mu = \{ct, x, y, z\}$ or $A^\mu = \{\phi, \boldsymbol{A}\}$. Lowering then the index by contraction with the metric tensor results in a minus sign of the spatial components of a four-vector, $x_\mu = \eta_{\mu\nu} x^\mu = \{ct, -x, -y, -z\}$ or $A_\mu = \{\phi, -\boldsymbol{A}\}$. Summing over a pair of Lorentz indices, always one index occurs in an upper and one in a lower position. Additionally to four-vectors, we will meet tensors $T^{\mu_1 \cdots \mu_n}$ of rank n which transform as $\tilde{T}^{\mu_1 \cdots \mu_n} = \Lambda^{\mu_1}{}_{\nu_1} \cdots \Lambda^{\mu_n}{}_{\nu_n} T^{\nu_1 \cdots \nu_n}$. Every tensor index can be raised and lowered, using the metric tensors $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$.

Special tensors are the Kronecker delta, $\delta_\mu^\nu = \eta_\mu^\nu$ with $\delta_\mu^\nu = 1$ for $\mu = \nu$ and 0 otherwise, and the Levi–Civita tensor $\varepsilon_{\mu\nu\rho\sigma}$. The latter tensor is completely antisymmetric and has in four dimensions the elements $+1$ for an even permutation of 0123, -1 for odd permutations and zero otherwise. In three dimensions, we define the Levi–Civita tensor by $\varepsilon^{123} = \varepsilon_{123} = 1$.

Next consider differential operators. Forming the differential of a function f defined on Minkowski space x^μ,

$$\mathrm{d}f = \frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y + \frac{\partial f}{\partial z}\,\mathrm{d}z = \frac{\partial f}{\partial x^\mu}\,\mathrm{d}x^\mu \,, \tag{1.16}$$

we see that an upper index in the denominator counts as lower index, and vice versa. We define the four-dimensional nabla operator as

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) .$$

Note the "missing" minus sign in the spatial components, which is consistent with $\partial_\mu = \frac{\partial}{\partial x^\mu}$ and the rule for the differential in Eq. (1.16). The d'Alembert or wave operator is

$$\Box \equiv \eta_{\mu\nu} \partial^\mu \partial^\nu = \partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta \,. \tag{1.17}$$

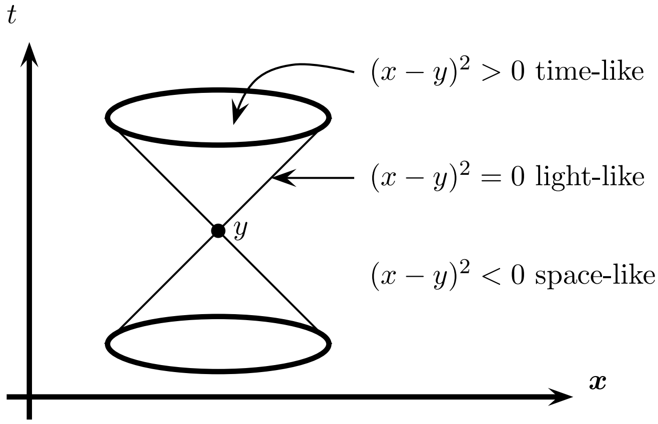

Figure 1.1: Light-cone at the point y generated by light-like vectors. Contained in the light-cone are the time-like vectors, outside the space-like ones.

This operator is a scalar, i.e. all the Lorentz indices are contracted, and thus invariant under Lorentz transformations.

Since the metric $\eta_{\mu\nu}$ is indefinite, the norm of a vector a^μ can be

$$a_\mu a^\mu > 0, \quad \text{time-like,} \tag{1.18}$$
$$a_\mu a^\mu = 0, \quad \text{light-like or null-vector,} \tag{1.19}$$
$$a_\mu a^\mu < 0, \quad \text{space-like.} \tag{1.20}$$

The cone of all light-like vectors starting from a point P is called *light-cone*, cf. Fig. 1.1. The time-like region inside the light-cone consists of two parts, past and future. Only events inside the past light-cone can influence the physics at point P, while P can influence only its future light-cone.

The line describing the position of an observer is called *world-line*. The *proper-time* τ is the time displayed by a clock moving with the observer. In the rest system of the observer, $\mathrm{d}\tau = (\mathrm{d}t, \mathbf{0}) = \mathrm{d}s/c$, and proper- and coordinate-time agree, $\tau = t$. In general, we have to integrate the line-element,

$$\tau_{12} = c\int_1^2 \mathrm{d}s = \int_1^2 [\eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu]^{1/2} = \int_1^2 \mathrm{d}t[1-(v/c)^2]^{1/2} < t_2 - t_1\,, \tag{1.21}$$

to obtain the proper-time. The last part of this equation, where we introduced the three-velocity $v^i = \mathrm{d}x^i/\mathrm{d}t$ of the clock, shows explicitly the relativistic effect of time dilation, as well as the connection between coordinate time t and the proper-time τ of a moving clock, $\mathrm{d}\tau = (1-(v/c)^2)^{1/2}\mathrm{d}t \equiv \mathrm{d}t/\gamma$.

1.2 Mechanics

From now on, we set $c = \hbar = 1$.

Four-velocity and four-momentum What is the relativistic generalization of the three-velocity $\boldsymbol{v} = \mathrm{d}\boldsymbol{x}/\mathrm{d}t$? The nominator $\mathrm{d}\boldsymbol{x}$ has already the right behaviour to become part of a four-vector, if the denominator would be invariant. We use therefore instead of $\mathrm{d}t$ the invariant proper time $\mathrm{d}\tau$ and write

$$u^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\,. \tag{1.22}$$

The four-velocity is thus the tangent vector to the world-line $x^\alpha(\tau)$ parametrised by the proper-time τ of a particle. Written explicitly, we have

$$u^0 = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{1}{\sqrt{1-v^2}} = \gamma \tag{1.23}$$

and

$$u^i = \frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{v^i}{\sqrt{1-v^2}} = \gamma v^i\,. \tag{1.24}$$

Hence the four-velocity is $u^\alpha = (\gamma, \gamma\boldsymbol{v})$ and its norm is

$$u\cdot u = u^0u^0 - \boldsymbol{uu} = \gamma^2 - \gamma^2 v^2 = \gamma^2(1-v^2) = 1\,. \tag{1.25}$$

The fact that its norm is constant proves that u^α is a four-vector.

How to guess physical tensors: We may guess just by considering the different number of degrees, which physical 3-dim. quantities should be combined into 4-dim. tensors:

1. A 4-vector a^μ has $4 = 3+1$ components, i.e. combines a 3-vector $\boldsymbol{a}$ and a 3-scalar s, that are related by a physical law. For instance current conservation

   $$\partial_t\rho + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$$

   suggests to combine $(\rho, \boldsymbol{j}) = j^\mu$ and $\partial_\mu = (\partial_t, \boldsymbol{\nabla})$ into four-vectors (consistent with our definition of the nabla operator). Similarly, we combine the scalar potential ϕ and the vector potential $\boldsymbol{A}$ into a four-vector $A^\mu = (\phi, \boldsymbol{A})$.
2. An antisymmetric tensor $F_{\mu\nu}$ has $3+2+1 = 6$ components, i.e. combines two 3-vectors (more precisely a pure vector like $\boldsymbol{E}$ and an axial vector like $\boldsymbol{B}$).
3. A general tensor $T_{\mu\nu}$ has $16 = 9+2\cdot3+1$ components, i.e. combines from a 3-dim. point of view a 3-dim tensor, 2 vectors and a scalar. A symmetric $T_{\mu\nu}$ has $10 = 4+3+2+1$ components, and the two 3-dim. vectors agree.

Energy and momentum After having constructed the four-velocity, the simplest guess for the four-momentum is

$$p^\alpha = mu^\alpha = (\gamma m, \gamma m\boldsymbol{v})\,. \tag{1.26}$$

For small velocities, $v \ll 1$, we obtain

$$\boldsymbol{p} = \left(1 + \frac{v^2}{2} - \ldots\right) m\boldsymbol{v} \tag{1.27}$$

$$p^0 = m + \frac{mv^2}{2} - \ldots = m + E_{\mathrm{kin,nr}} + \ldots \tag{1.28}$$

Thus we can interpret the components as $p^\alpha = (E, \boldsymbol{p})$. The norm follows with (1.25) immediately as

$$p\cdot p = m^2\,. \tag{1.29}$$

Solving for the energy, we obtain

$$E = \pm\sqrt{m^2 + \boldsymbol{p}^2} \tag{1.30}$$

including the famous $E = mc^2$ as special case for a particle at rest. Note that (1.30) predicts the existence of solutions with negative energy—undermining the stability of the universe. According Feynman, we should view these negative energy solutions as positive energy solutions moving backward in time, $\exp(-\mathrm{i}(-\sqrt{m^2 + \boldsymbol{p}^2})t) = \exp[-\mathrm{i}(+\sqrt{m^2 + \boldsymbol{p}^2})(-t)]$.

Four-forces We postulate now that in relativistic mechanics Newton's law becomes

$$f^\alpha = \frac{\mathrm{d}p^\alpha}{\mathrm{d}\tau} \tag{1.31}$$

where we introduced the four-force f^α. Since both u^α and p^α consist of only three independent component, we expect that there exists also a constraint on the four-force f^α. We form the scalar product

$$u \cdot f = u \cdot \frac{\mathrm{d}(mu)}{\mathrm{d}\tau} = u \cdot u \frac{\mathrm{d}m}{\mathrm{d}\tau} + mu \cdot \frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{\mathrm{d}m}{\mathrm{d}\tau}\,. \tag{1.32}$$

In the last step we used twice that $u \cdot u = 1$. Since all electrons ever observed have the same mass, no force should exist which changes m. As a consequence, we have to ask that all physical acceptable force-laws satisfy $u \cdot f = 0$; such forces are called *pure forces*.

1.3 Relativistic kinematics

Translation symmetry of Minkowski space implies 4-momentum conservations: Thus for a reaction with n particles in the initial and m particles in the final state, the four-momentum satisfies

$$\sum_{i=1}^{n} p_i^\mu = \sum_{f=1}^{m} p_f^\mu\,, \qquad \mu = \{0,1,2,3\}\,.$$

The description of reaction is simplified if Lorentz invariant quantities are used. For a $1 + 2 \to 3 + 4$ scattering process, it is useful to introduce Mandelstam variables s, t, and u defined by

$$s = (p_1 + p_2)^2 = (p_2 + p_4)^2\,, \tag{1.33}$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2\,, \tag{1.34}$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2\,. \tag{1.35}$$

Since $s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2$, the scattering amplitude $\mathcal{A}$ describing such a process depends only on two variables, e.g $\mathcal{A}(s,t)$ or $\mathcal{A}(s,\vartheta)$.

In the center-of-mass system (cms), the net 3-momentum of the colliding particles balance, $p_1 = (E_1^*, \boldsymbol{p}^*)$ and $p_2 = (E_2^*, -\boldsymbol{p}^*)$ and thus

$$s = (p_1 + p_2)^2 = (E_1^* + E_2^*)^2 - (\boldsymbol{p}^* - \boldsymbol{p}^*)^2 = (E_1^* + E_2^*)^2. \tag{1.36}$$

Hence s is the squared cms energy. Often one is interested to know the threshold energy $\sqrt{s_{\min}}$, i.e. the minimal possible energy that a certain reaction can happen. Consider e.g. the process $pp \to ppp\bar{p}$, where one of the initial protons at rest is hit by the second proton with energy E_p. Then

$$s_{\min} = 2m_p^2 + 2E_p m_p = 16m_p^2 \tag{1.37}$$

and thus $E_p \geq 7m_p$.

The second important quantity characterizing a scattering process is the four-momentum transfer t. As an example, consider electron-proton scattering $e^- + p \to e^- + p$. Then

$$t = (p_e - p_e')^2 = 2m_e^2 - 2E_e E_e'(1 - \beta_e \beta_e' \cos\vartheta)\,. \tag{1.38}$$

For high energies, $\beta_e, \beta_e' \to 1$, and

$$t \approx -2E_e E_e'(1 - \cos\vartheta) = -4E_e E_e' \sin^2\vartheta/2\,. \tag{1.39}$$

We will see later that four-momentum conservation holds at each vertex of a Feynman diagram. Thus the variable t corresponds to the squared momentum q of the exchanged virtual photon, $t = q^2 = (p_e - p_e')^2 < 0$. The virtual particle does not fulfil the relativistic energy-momentum relation. The energy-time uncertainty relation $\Delta E \Delta t \gtrsim 1$ allows such a violation, if the virtual particle is exchanged only during a short enough time. Note also that the angular dependence of Rutherford scattering, $\mathrm{d}\sigma/\mathrm{d}\Omega \propto 1/\sin^4\vartheta/2$ is obtained if $\mathrm{d}\sigma/\mathrm{d}\Omega \propto 1/t^2$. Thus the virtual photon seems to be described by an expression $\propto 1/q^2$.

1.4 Electrodynamics

Field-strength tensor We want to convert the three-dimensional Lorentz force,

$$\boldsymbol{F}_L = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \tag{1.40}$$

into a valid relativistic expression. The LHS should become the four-vector $f_\alpha = \mathrm{d}p_\alpha/\mathrm{d}\tau$. The RHS has to become a yet unknown tensor $F^{\mu\nu}$ of rank two contracted with the four-velocity,

$$f^\alpha = qF^{\alpha\beta}u_\beta\,. \tag{1.41}$$

The requirement of a pure force means

$$u_\alpha f^\alpha = qF^{\alpha\beta}u_\alpha u_\beta = 0\,, \tag{1.42}$$

and thus the tensor must be antisymmetric, $F^{\alpha\beta} = -F^{\beta\alpha}$.

The force law becomes simplest in a frame with $u^\alpha = (1, 0, 0, 0)$. Then $f^i = qF^{i0}$ which should equal $\boldsymbol{F} = q\boldsymbol{E}$. Thus $F^{k0} = E^k = -F^{0k}$. The antisymmetry of the field-strength tensor implies $f^0 = qF^{00} = 0$. In a moving frame, we split $F^{\alpha\beta}u_\beta$ first into $F^{\alpha\beta}u_\beta = F^{\alpha 0}u^0 - F^{\alpha j}u^j$ and consider then e.g. the z component,

$$F_z = \gamma\frac{\mathrm{d}p_z}{\mathrm{d}t} = q\gamma(E_z - F^{31}v^1 - F^{32}v^2) \stackrel{!}{=} q(E_z + B_y v_x - B_x v_y). \tag{1.43}$$

Thus $F^{31} = -B_y$ and $F^{32} = B_x$. Hence the components of the field-strength tensor are (the first index denotes raws, the second columns)

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \tag{1.44}$$

and

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} . \tag{1.45}$$

Next we have to determine the connection between $F^{\mu\nu}$ and A^μ. The field-strength tensor contains first derivatives of the potential, i.e. terms like $\partial^\mu A^\nu$, and is antisymmetric. Thus the only possible choice (up to a proportionality constant) is

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu . \tag{1.46}$$

Setting the proportionality constant equal to one is correct, since the first raw of $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\nu$ reads with $A_\mu = (\phi, -A_k)$ and $\partial_\nu = \partial/\partial x^\mu = (\partial/\partial t, \nabla_k)$ as

$$F_{0k} = E_k = \partial_0(-A_k) - \partial_k A_0 = -\frac{\partial A_k}{\partial t} - \nabla_k \phi \tag{1.47}$$

or $\boldsymbol{E} = -\partial_t \boldsymbol{A} - \nabla\phi$. We go in the opposite direction for $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. In components, we have e.g.

$$B_x = B^1 = \partial_2 A^3 - \partial_3 A^2 = \partial_3 A_2 - \partial_2 A_3 = F_{32} \tag{1.48}$$

and similarly for the other components.

Maxwell Now we can rewrite the Maxwell equations as

$$\partial_\alpha F^{\alpha\beta} = j^\beta \tag{1.49}$$

and

$$\partial_\alpha F^{\beta\gamma} + \partial_\beta F^{\gamma\alpha} + \partial_\gamma F^{\alpha\beta} = 0 . \tag{1.50}$$

The first (inhomogeneous) equation describes the time evolution of the electromagnetic field, while the second (homogeneous) equation is a constraint. If we use the potential A^μ instead of the field-strength, it is automatically satisfied: Inserting the definition,

$$\partial_\alpha(\partial^\beta A^\gamma - \partial^\gamma A^\beta) + \partial_\beta(\partial^\gamma A^\alpha - \partial^\alpha A^\gamma) + \partial_\gamma F^{\alpha\beta}(\partial^{A\beta} - \partial^\beta A^\alpha) = 0 , \tag{1.51}$$

and using $\partial^\alpha \partial^\beta = \partial^\beta \partial^\alpha$, we see that all terms cancel: Working with A^μ instead of $F^{\mu\nu}$, we have to solve only the inhomogeneous Maxwell equation.

We obtain a wave equation for the photon, inserting the definition of $\mathcal{A}^\mu$ into the inhomogeneous Maxwell equation,

$$\partial_\mu F^{\mu\nu} = \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \Box A^\nu - \partial_\mu \partial^\nu A^\mu = j^\nu . \tag{1.52}$$

Current conservation and gauge invariance We take the divergence of the inhomogeneous Maxwell equation (1.49),

$$\partial_\nu \partial_\mu F^{\mu\nu} = \partial_\nu j^\nu . \tag{1.53}$$

Since $\partial_\nu \partial_\mu$ is symmetric and $F^{\mu\nu}$ antisymmetric, the summation of the two factors has to be zero,

$$\partial_\nu \partial_\mu F^{\mu\nu} = -\partial_\nu \partial_\mu F^{\nu\mu} = -\partial_\mu \partial_\nu F^{\nu\mu} = -\partial_\nu \partial_\mu F^{\mu\nu} . \tag{1.54}$$

Thus current conservation,

$$\partial_\nu j^\nu = 0\,, \tag{1.55}$$

is built-in into the Maxwell equations. Or in other words: the photon couples to a conserved current, i.e. the electric charge is conserved.

Consider next the transformation of $F^{\mu\nu}$ under a gauge transformation,

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \chi\,. \tag{1.56}$$

Assuming that χ is smooth, it is $\partial_\nu\partial_\mu\chi = \partial_\mu\partial_\nu\chi$ and thus

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = F_{\mu\nu} + \partial_\mu\partial_\nu\chi - \partial_\nu\partial_\mu\chi = F_{\mu\nu}\,. \tag{1.57}$$

Thus the gauge invariance of $F^{\mu\nu}$ is again closely connected to the fact that it is an antisymmetric tensor, formed by derivatives of A^μ.

Note that Maxwell's equations admit *local* not global gauge transformations, $\chi = \chi(x)$. A local symmetry implies a redundancy in our description of physics: All A^μ connected by (1.56) lead to the same field-strength, and thus classically to the same Lorentz force and the same trajectories of test particles.

The gauge freedom allows us to choose e.g. the Lorentz gauge $\partial_\mu A^\mu = 0$ in Eq. (1.52), simplifying the wave equation to

$$\Box A^\nu = j^\nu\,. \tag{1.58}$$

1.A Appendix: Comments and exercises on index notation

We are mainly concerned with vectors and tensors of rank two. In this case we can express all equations by matrix operations. For instance, lowering the index of a vector, $A_\mu = \eta_{\mu\nu}A^\mu$, becomes

$$A_\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} = \begin{pmatrix} A^0 \\ -A^1 \\ -A^2 \\ -A^3 \end{pmatrix}.$$

Raising and lowering indices is the inverse, and thus $\eta_{\mu\nu}\eta^{\nu\sigma} = \delta_\mu^\sigma$. In matrix notation,

$$\eta\eta^{-1} = \mathbf{1}.$$

We can view $\eta_{\mu\nu}\eta^{\nu\sigma} = \delta_\mu^\sigma$ as the operation of raising an index of $\eta_{\mu\nu}$ (or lowering an index of $\eta^{\mu\nu}$): in both cases, we see that the Kronecker delta corresponds to the metric tensor with mixed indices, $\delta_\mu^\sigma = \eta_\mu^\sigma$.

The expression for the line-element becomes

$$\begin{aligned} \mathrm{d}s^2 &= \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}x^\mu \eta_{\mu\nu}\mathrm{d}x^\nu = \left(\mathrm{d}x^0, \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3\right) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \mathrm{d}x^0 \\ \mathrm{d}x^1 \\ \mathrm{d}x^2 \\ \mathrm{d}x^3 \end{pmatrix} \\ &= (\mathrm{d}x^0)^2 - (\mathrm{d}x^1)^2 - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2. \end{aligned}$$

For a second-rank tensor, raising one index gives

$$T_{\mu\nu} = \eta_{\mu\rho}T^\rho{}_\nu = T^\rho{}_\nu \eta_{\mu\rho} \neq \eta_{\mu\rho}T_\nu{}^\rho = T_{\nu\mu}$$

Note that the order of tensors does not matter, but the order of indices does. If we move to matrix notation, we have to restore the right order. Raising next the second index,

$$T_{\mu\nu} = \eta_{\mu\rho}\eta_{\nu\sigma}T^{\rho\sigma}$$

we have to re-order it as $T_{\mu\nu} = \eta_{\mu\rho}T^{\rho\sigma}\eta_{\sigma\nu}$ in matrix notation (using that η is symmetric). We apply this to the field-strength tensor: Starting from $F^{\mu\nu}$, we want to construct $F_{\mu\nu}$,

$$\begin{aligned} F_{\mu\nu} &= \eta_{\mu\rho}F^{\rho\sigma}\eta_{\sigma\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}. \end{aligned}$$

Note the general behaviour: The F^{00} element and the 3-tensor F^{ik} are multiplied by 1^2 and $(-1)^2$, respectively and do not change sign. The 3-vector F^{0k} is multiplied by $(-1)(+1)$ and does change sign.

Next we want to construct a Lorentz scalar out of $F^{\mu\nu}$. A Lorentz scalar has no indices, so we contract the two indices, $\eta_{\mu\nu}F^{\mu\nu} = F_\mu{}^\mu$. This is invariant, but zero (and thus not useful) because $F^{\mu\nu}$ is antisymmetric. As next try, we construct a Lorentz scalar S using two F's: Multiplying the two matrices $\tilde{F}_{\mu\nu}$ and $F^{\mu\nu}$, and taking then the trace, gives

$$S = F_{\mu\nu}F^{\mu\nu} = -\mathrm{tr}\{F_{\mu\nu}F^{\nu\rho}\} = -\mathrm{tr}\begin{pmatrix} \boldsymbol{E}\cdot\boldsymbol{E} & & & \\ & E_x^2 - B_z^2 - B_y^2 & & \\ & & E_y^2 - B_z^2 - B_x^2 & \\ & & & E_z^2 - B_y^2 - B_x^2 \end{pmatrix}$$

i.e. $S = -2(\boldsymbol{E}\cdot\boldsymbol{E} - \boldsymbol{B}\cdot\boldsymbol{B})$. Note the minus, since we have to change the order of indices in the second F.

Note also that S has to be a bilinear in $\boldsymbol{E}$ and $\boldsymbol{B}$ and invariant under rotations. Thus the only possible terms entering S are the scalar products $\boldsymbol{E}\cdot\boldsymbol{E}$, $\boldsymbol{B}\cdot\boldsymbol{B}$ and $\boldsymbol{E}\cdot\boldsymbol{B}$. Since $\boldsymbol{B}$ is a polar (or axial) vector, $P\boldsymbol{B} = -\boldsymbol{B}$, the last term is a pseudo-scalar and cannot enter the scalar S.

Now we become more ambitious, looking at a tensor with 4 indices, the Levi–Civita or completely antisymmetric tensor $\varepsilon^{\alpha\beta\gamma\delta}$ in four dimensions, with

$$\varepsilon_{0123} = +1, \tag{1.72}$$

and all even permutations, -1 for odd permutations and zero otherwise. We lower its indices,

$$\varepsilon^{\alpha\beta\gamma\delta} = \varepsilon_{\bar\alpha\bar\beta\bar\gamma\bar\delta}\eta^{\bar\alpha\alpha}\eta^{\bar\beta\beta}\eta^{\bar\gamma\gamma}\eta^{\bar\delta\delta}$$

and consider the 0123 element using that the metric is diagonal,

$$\varepsilon^{0123} = +1\eta^{00}\eta^{11}\eta^{22}\eta^{33} = -1. \tag{1.73}$$

Thus in 4 dimensions, $\varepsilon^{\alpha\beta\gamma\delta}$ and $\varepsilon_{\alpha\beta\gamma\delta}$ have opposite signs.

We can use the Levi-Civita tensor to define the dual field-strength tensor

$$\tilde{F}^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}F_{\gamma\delta}.$$

How to find the elements of this? Using simply the definitions,

$$\tilde{F}_{01} = \frac{1}{2}\Big(\underbrace{\varepsilon_{0123}}_{1}\underbrace{F^{23}}_{-B_x} + \varepsilon_{0132}F^{32}\Big) = -B_x$$

$$\tilde{F}_{12} = \frac{1}{2}\Big(\varepsilon_{1203}F^{03} + \varepsilon_{1230}F^{30}\Big) = -E_z$$

etc., gives

$$\tilde{F}_{\mu\nu} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & -E_z & E_y \\ B_y & E_z & 0 & -E_x \\ B_z & -E_y & E_x & 0 \end{pmatrix} \quad \text{and} \quad \tilde{F}^{\mu\nu} = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z & E_y \\ -B_y & E_z & 0 & -E_x \\ -B_z & -E_y & E_x & 0 \end{pmatrix}.$$

The dual field-strength tensor is useful, because the homogeneous Maxwell equation becomes simply

$$\partial_\alpha F^{\beta\gamma} + \partial_\beta F^{\gamma\alpha} + \partial_\gamma F^{\alpha\beta} = \partial_\alpha \tilde{F}^{\alpha\beta} = 0\,. \tag{1.74}$$

Inserting the potential, we obtain zero,

$$\partial_\alpha \tilde{F}^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}\partial_\alpha F_{\gamma\delta} = \varepsilon^{\alpha\beta\gamma\delta}\partial_\alpha\partial_\gamma A_\delta = 0\,, \tag{1.75}$$

because we contract a symmetric tensor ($\partial_\alpha\partial_\gamma$) with an anti-symmetric one ($\varepsilon^{\alpha\beta\gamma\delta}$).

Having $F^{\mu\nu}$ and $\tilde{F}_{\mu\nu}$, we can form another (pseudo-) scalar, $A = \tilde{F}_{\mu\nu}F^{\mu\nu}$. Multiplying the two matrices $\tilde{F}_{\mu\nu}$ and $F^{\mu\nu}$, and taking then the trace, gives

$$\tilde{F}_{\mu\nu}F^{\mu\nu} = -\mathrm{t}r\{\tilde{F}_{\mu\nu}F^{\nu\rho}\} = \mathrm{t}r\begin{pmatrix} \boldsymbol{B}\cdot\boldsymbol{E} & & & \\ & \boldsymbol{B}\cdot\boldsymbol{E} & & \\ & & \boldsymbol{B}\cdot\boldsymbol{E} & \\ & & & \boldsymbol{B}\cdot\boldsymbol{E} \end{pmatrix}$$

i.e. $\tilde{F}_{\mu\nu}F^{\mu\nu} = 4\boldsymbol{E}\cdot\boldsymbol{B}$. We know that $\boldsymbol{E}\cdot\boldsymbol{B}$ is a pseudo-scalar. This tells us that including the Levi-Civita tensor converts a tensor into a pseudo-tensor, which changes sign under a parity transformation, $P\boldsymbol{x} = -\boldsymbol{x}$. (This analogous to $B_i = \varepsilon_{ijk}\partial_j A_k$, which converts two pure vectors into an axial one.)