Special Relativity

Hendrik van Hees

August 16, 2020

Contents

Chapter 1

Kinematics

1.1 Introduction

While the didactics of quantum theory is difficult, because its interpretation is difficult and partially still controversial and one thus often has to use historical heuristic arguments and then teaches historically interesting but overcome concepts, in the case of Special Relativity Theory (SRT) no such excuse can be found. Here, the didactic sin is just due to the inertia of textbook authors to invent a better way of teaching the subject instead of inventing a new way to introduce it and thus go back to the very early original and textbook literature. This leads to an overemphasis of the "paradoxical" aspects of the theory. Instead of concentrating on the naturalness of the theory, emerging from simple empirical facts about electromagnetic phenomena as in Einstein's original paper [Ein05], which is a masterpiece not only in theoretical physics but also in scientific prose, one investigates the problems of parking a length contracted car in too short garages or the different aging of twins, depending on their travel habits.

Also some other problems, long overcome due to the detailed mathematical analysis of the theory by Minkowski [Min09], are overemphasized or unnecessarily introduced, among them the idea of a "relativistic mass", depending on the velocity of the particle or the Lorentz-transformation properties of thermodynamic quantities like temperature and chemical potential of a gas.

In this article I try to give an introduction into SRT which is as simple as possible and try to avoid an overemphasis of the unintuitive features by treating them with the full machinery of the covariant Minkowski-space formulation. First, the space-time structure is postulated, using the arguments in Einstein's paper but formulating everything in modern four-vector notation a la Minkowski. The usual kinematic effects (relativity of simultaneity, time dilation, and length contraction) will be discussed within this framework and finally also the Lorentz boosts are derived. Minkowski diagrams with the correct geometry, leading to apparently different scales on the temporal and spatial axes of different inertial observers' reference frames, based on the indefinite Minkowski product are introduced and the kinematic effects illustrated with their help.

In the next Sect. the basic classical mechanics of a point particle is treated in a manifestly covariant way, leading to the introduction of proper time and constraints to force laws.

After this, the classical statistics of many-particle systems and the ideal gas is introduced, leading to a clear description of the thermodynamic quantities like temperature, entropy, chemical potential, etc.

Finally also the basics of relativistically covariant electrodynamics is introduced by "translating" the (microscopic) Maxwell equations to the framework, showing that the original Maxwell equations have already been a relativistic theory and thus must have lead inevitably to the discovery of the special-relativistic space-time structure.

1.2 The special-relativistic space-time model

Around 1900 one of the most pressing questions in theoretical physics was the question, how to determine the rest frame of the ether. The ether was thought to be a substance through which electromagnetic waves, as very successfully described by Maxwell's equations, should propagate as sound waves through air. The only problem was that this ether should have very strange properties and that nobody could find a clear empirical evidence for its existence. From a mathematical point of view the problem was that Maxwell's equations are not invariant under Galilei transformations, which describe the change from one inertial reference frame to another, moving with constant relative velocity within Newtonian mechanics. Various transformations, keeping the Maxwell equations of the free field invariant, were known already since the 19th century[1], but only Einstein drew that conclusion that one must reformulate the description of space and time in a comprehensive way, i.e., that the space-time structure of all of physics is different from the Galilei-Newton space-time model successfully used for almost 300 years before.

He started with two postulates:

1. The physical laws look the same in any inertial reference. An inertial reference frame is a frame, where a body stays at rest or moves with constant velocity, if there are no forces acting on that body.
2. The velocity of electromagnetics waves (particularly including light) in a vacuum is independent of the velocity of the light source and that of the observer.

The first postulate is valid in Newtonian physics as well. Since Einstein considered the equivalence of all inertial frames to be valid also for electromagnetic phenomena, which are correctly described by Maxwell's theory, the second postulate must hold true, since any inertial observer will describe the same situation of emission of electromagnetic waves by the same Maxwell equations, and the Maxwell equations contain the speed of light in the vacuum as a parameter[2] [Jac98]. Tacitly, he also made another assumption: Any inertial observer will find that in the physical space the laws of Euclidean geometry is valid, which is true in Newtonian mechanics too. In a way this is also implicit in the assumption of the invariance of Maxwell's equations, which is formulated in terms of vector calculus in three-dimensional Euclidean space.

With these postulates, we can infer the correct space-time description without using the full math of the Maxwell equations. To this end we consider an inertial observer, Alice, sending a short light signal from a point source. According to the above postulates the corresponding spherical wave propagates with a constant speed c around the light source, which we locate at the origin of Alice's Cartesian coordinate system. If she sent her light signal at $t=0$, the wave packet obeys the equation

$$c^2 t^2 - \vec{x}^2 = 0. \tag{1.2.1}$$

Here $\vec{x}$ is any point at the wave packet.

Now according the Einstein's 2nd theorem another observer, Bob, moving with the constant velocity $\vec{v} = c\vec{\beta}$ with respect to Alice will also describe this light wave as a spherical wave from his origin, when

[1] see, `https://en.wikipedia.org/wiki/History_of_Lorentz_transformations`

[2] Here and in the following I use Heaviside-Lorentz units to discuss the Maxwell equations, which are the most natural system of units, particularly in the context of the theory of relativity.

we assume that he meets Alice at $t=0$ at the origin of here Cartesian coordinate system, and the wave packet will also travel with the same speed c as in Alice's frame. This, however, can only be true if not only the spatial coordinates change when transforming from Alice's to Bob's coordinates (and vice versa) but also the time. If we thus denote Bob's time and space coordinates with $\overline{t}$ and $\overline{\vec{x}}$ respectively, he will also describe the wave packet by the same equation (1.2.1), as Alice, but of course using his space-time coordinates:

$$c^2\overline{t}^2 - \overline{\vec{x}}^2 = 0. \tag{1.2.2}$$

Further due to the first postulate, the transformations between Alice's coordinates and Bob's must be linear, because the motion of any force-free body in her inertial reference frame with constant velocity must be also described as such a motion in his inertial reference frame.

Now the most simple way to achieve this, is to use Minkowski's idea of 1908 to describe space and time together as a four-dimensional affine space with four-dimensional vectors. Alice uses her coordinates to map these four-dimensional vectors to $\mathbb{R}^4$: $(x^\mu) = (ct, x^1, x^2, x^3)$. For reasons, which will become clear in a moment we use upper indices $\mu \in \{0,1,2,3\}$ to label the time and space coordinates. The four-vector itself is then given with help of the corresponding basis vectors $\boldsymbol{e}_\mu$ (with a lower index), $\boldsymbol{x} = x^\mu \boldsymbol{e}_\mu$, where we define that one has to sum over any pair of indices appearing twice in an equation from 0 to 3 (Einstein's summation convention). When now Alice defines a bilinear form with her coordinates (x^μ) as

$$\boldsymbol{x} \cdot \boldsymbol{y} = x^0 y^0 - \vec{x} \cdot \vec{y}, \tag{1.2.3}$$

then the Eq. (1.2.1) can be written as

$$\boldsymbol{x} \cdot \boldsymbol{x} =: \boldsymbol{x}^2 = 0. \tag{1.2.4}$$

Obviously for the here introduced "pseudo-Cartesian" basis vectors $\boldsymbol{e}_\mu$ the bilinear form is uniquely defined by

$$\boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu = \eta_{\mu\nu} := \begin{cases} 1 & \text{for} \quad \mu = \nu = 0, \\ -1 & \text{for} \quad \mu = \nu, \quad \mu, \nu \in \{1,2,3\}, \\ 0 & \text{for} \quad \mu \neq \nu, \quad \mu, \nu \in \{0,1,2,3\}. \end{cases} \tag{1.2.5}$$

With this definition we can write (1.2.3), using the Einstein summation convention again,

$$\boldsymbol{x} \cdot \boldsymbol{y} = \eta_{\mu\nu} x^\mu y^\nu. \tag{1.2.6}$$

Then (1.2.2) holds true, if we assume that Bob simply also uses pseudo-Cartesian basis vectors $\overline{\boldsymbol{e}}_\mu$, for which also (1.2.5) holds. Then the linear transformation between Alice's and Bob's coordinates is given by a matrix $({\Lambda^\mu}_\nu) \in \mathbb{R}^4$ such that

$$\boldsymbol{e}_\nu = \overline{\boldsymbol{e}}_\mu {\Lambda^\mu}_\nu \tag{1.2.7}$$

and the requirement that also Bob's basis vectors are pseudo-Cartesian yields

$$\eta_{\rho\sigma} = \boldsymbol{e}_\rho \cdot \boldsymbol{e}_\sigma = {\Lambda^\mu}_\rho \overline{\boldsymbol{e}}_\mu \cdot {\Lambda^\nu}_\sigma \overline{\boldsymbol{e}}_\nu = {\Lambda^\mu}_\rho {\Lambda^\nu}_\sigma \eta_{\mu\nu}. \tag{1.2.8}$$

This is very similar to the definition of an orthogonal matrix in a four-dimensional Euclidean vector space. The only difference is that in the above introduced bilinear form, the Minkowski product, the spatial coordinates occur with a $-$ sign, i.e., the Minkowski product is not positive definite as the Cartesian product. We shall come back to the concrete form of these so-called Lorentz transformations ${\Lambda^\mu}_\nu$, but first it is worth to consider the "geometry" of the now defined affine space-time continuum, the Minkowski space.

To this end we consider the motion of point particles along the 1 axis of Alice's reference frame. These motions we can describe in a t-x^1 or better, measuring time in terms of $x^0 = ct$, in an x^0-x^1 plane. We draw the corresponding axes for Alice's frame in a Cartesian coordinate system, i.e., with the axes perpendicular to each other, but one must keep in mind that this Euclidean angle in this plane has no proper geometrical meaning in Minkowski space. This is so, because instead of the idea of a "distance" between two points in the Euclidean plane, here we must use the Minkowski product to define a kind of "distance" between points. However, this distance is not a proper one, because using the Minkowski product $(\boldsymbol{x}-\boldsymbol{y})\cdot(\boldsymbol{x}-\boldsymbol{y})$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are arbitrary vectors in Alice's x^0-x^1 plane, this "square of the distance" can be positive, negative, and also vanish even, if $\boldsymbol{x} \neq \boldsymbol{y}$. Obviously a light signal as introduced above will always travel along such a **null vector**. For a light signal started at $t = 0$ in the origin, it is described by the vector $\boldsymbol{x} \neq 0$ with $\boldsymbol{x} \cdot \boldsymbol{x} = 0$. For the spacial basis vectors we have $\boldsymbol{e}_j \cdot \boldsymbol{b}_j = -1$ for $j \in \{1,2,3\}$. Thus we call any vector $\boldsymbol{x}$ with $\boldsymbol{x}\cdot\boldsymbol{x} < 0$ a space-like vector. The time axis is described by the basis vector $\boldsymbol{e}_0$, which fulfills $\boldsymbol{e}_0 \cdot \boldsymbol{e}_0 = 1$, and thus any vector $\boldsymbol{x}$ with $\boldsymbol{x} \cdot \boldsymbol{x} > 0$ is called a time-like vector. Now any motion of a point-like object, moving along the $\boldsymbol{e}_1$ direction including the observer (which we idealize as a point-like object as well) is described as a "world line" in the x^0-x^1 plane. Particularly Alice's world line is defined as being at rest at $x^1 = 0$, and thus her world line is described by $\boldsymbol{x}_A = \lambda \boldsymbol{e}_0$ with $\lambda \in \mathbb{R}$, i.e., Alice's world line is just the vertical axis in our Minkowski diagram.

Now, the same arguments can be made with respect to Bob, whose world line is along his time-like basis vector! In Alice's reference frame, he is however moving with constant velocity $v = \beta c$ along Alice's x^1"=axis, i.e., for Alice, his world line is given in terms of her space-time coordinates by

$$\begin{pmatrix} x_B^0 \\ x_B^1 \end{pmatrix} = \begin{pmatrix} ct \\ \beta ct, \end{pmatrix} \tag{1.2.9}$$

where we assume that at $t = 0$ Bob is located at Alice's origin, $x^1 = 0$.

On the other hand, we can write this in a coordinate-independent way as

$$\boldsymbol{x}_B = \lambda \overline{\boldsymbol{e}}_0 = \lambda \boldsymbol{u}_{\mathrm{B}}, \quad \lambda \in \mathbb{R}. \tag{1.2.10}$$

The unit vector

$$(u_{\mathrm{B}}^{\mu}) = \gamma \begin{pmatrix} 1 \\ \beta \end{pmatrix} \quad \text{with} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{1.2.11}$$

is Bob's four-velocity, and in (1.2.11) we have written the components with respect to Alice's basis.

We can also easily construct Bob's spacial basis vector in Alice's 01-plane. It should fulfil the constraints

$$\overline{\boldsymbol{e}}_1 \cdot \overline{\boldsymbol{e}}_0 = 0, \quad \overline{\boldsymbol{e}}_1 \cdot \overline{\boldsymbol{e}}_1 = -1 \tag{1.2.12}$$

In Alice's coordinates it is obviously given by

$$\overline{e}_1^{\mu} = \gamma \begin{pmatrix} \beta \\ 1 \end{pmatrix}. \tag{1.2.13}$$

The unit vectors $\overline{\boldsymbol{e}}_\mu$ particularly define Bob's space-time units in the (ct, x^1) plane, which can also be found by drawing the temporal (upper sign) and spatial (lower sign) unit hyperbolas $(ct)^2-(x_1)^2 = \pm 1$ in the corresponding **Minkowski diagram** (cf. Fig. 1.1)

Now it is already easy to discuss some simple kinematic effects. We begin with two events $\boldsymbol{x}$ and $\boldsymbol{y}$ happening at the same place for Alice at times t_x and t_y. The time difference according to Alice and

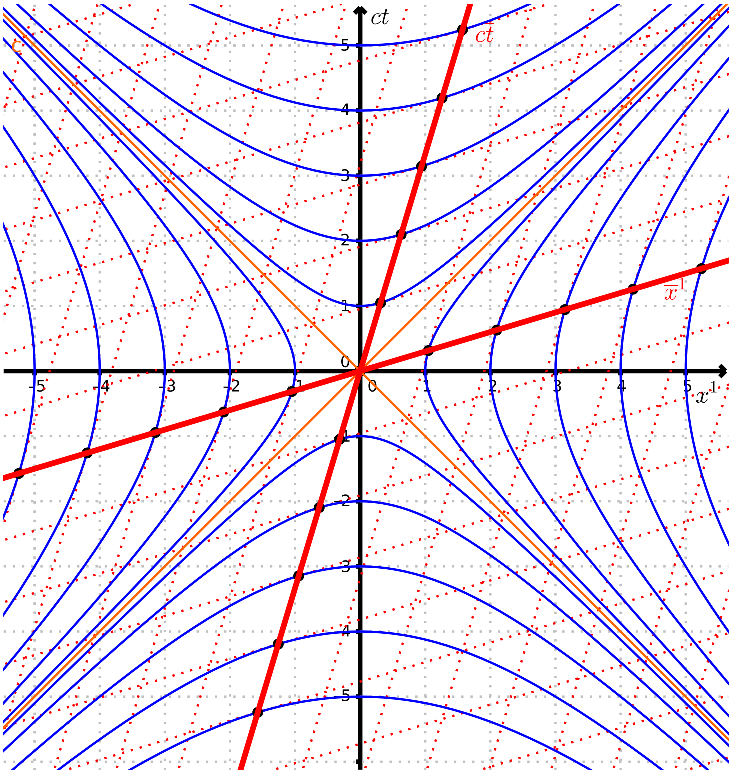


Figure 1.1: *Minkowski diagram, illustrating the construction of Bob's space-time coordinates: The thick red lines are Bob's $c\bar{t}$ and $\bar{x}^1$ axis. The corresponding units are given by the blue hyperbolas $(ct^2)-(x^1)^2 = \pm n^2$ with $n \in \{1,2,\ldots\}$. We also draw the corresponding coordinate grid (light dotted red lines). The orange lines $ct = \pm x^1$ define the light cone in this space-time plane.*

Bob is given by

$$\Delta t_{\mathrm{A}} = \frac{1}{c}\boldsymbol{u}_{\mathrm{A}} \cdot (\boldsymbol{y}-\boldsymbol{x}) = t_y - t_x, \quad \Delta t_{\mathrm{B}} = \frac{1}{c}\boldsymbol{u}_{\mathrm{B}} \cdot (\boldsymbol{y}-\boldsymbol{x}) = \gamma \Delta t_{\mathrm{Alice}}, \tag{1.2.14}$$

i.e., according to Bob's clock the time between the two events is longer by a Lorentz-γ factor. This phenomenon is known as **time dilation**. It is also clear that for Bob the events do not occur at the same position. The distance can be calculated from Bob's spatial coordinates of these events

$$(\bar{y}-\bar{x})^1 = -\bar{\boldsymbol{e}}_1 \cdot (\boldsymbol{y}-\boldsymbol{x}) = -\beta c \gamma \Delta t_{\mathrm{A}} = -\beta c \Delta t_{\mathrm{B}}. \tag{1.2.15}$$

This is indeed trivial, because during the time Δt_{B} between the events Bob travels the distance $\beta c \Delta t_{\mathrm{B}}$ along the Alice's 1-axis.

Next we consider a rigid rod at rest in Alice's reference frame located in her 1-direction. The ends of the rod are described in terms of Alice's coordinates as

$$(x^\mu) = \begin{pmatrix} \lambda_1 \\ x \end{pmatrix}, \quad (y^\mu) = \begin{pmatrix} \lambda_2 \\ y \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}. \tag{1.2.16}$$

She measures the spacial distance by reading off the coordinates at the same time t_{A}, i.e., for $\lambda_2 = \lambda_1$, and she finds

$$(y-x)^\mu = \begin{pmatrix} 0 \\ y-x \end{pmatrix} \Rightarrow L_{\mathrm{A}} = |y^1 - x^1|. \tag{1.2.17}$$

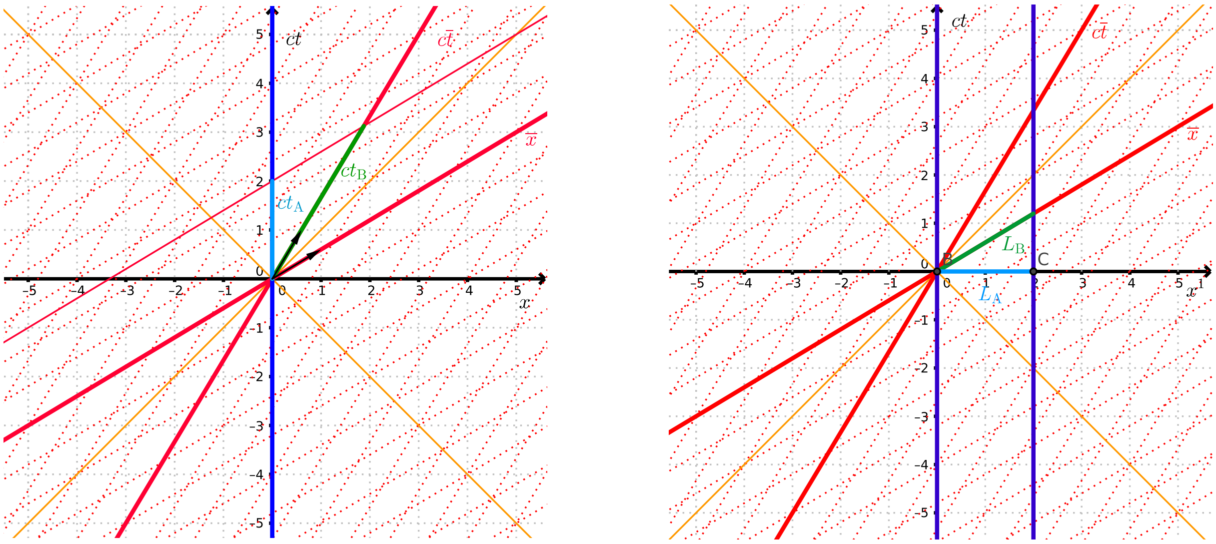


Figure 1.2: ***Left:*** *Minkowski diagram illustrating the time-dilation effect. The time between two events at the same place* $x^1=0$ *in Alice's frame appears longer as measured by Bob, for whom the events are of course also not taking place at the same position.* ***Right:*** *Minkowski diagram illustrating the length-contraction effect: The blue lines are the world lines of the end points of the rod, being at rest with respect to Alice's reference frame. She measures the length to be* $L_A=2$ *a.u. In Bob's frame the rod moves, and its length is measured by Bob by marking his coordinates of the endpoints simultaneously at his time* $\bar{t}=0$*. Comparing the corresponding length* L_B *to the 2a.u. mark, labeled as* $\bar{2}$ *in his reference frame, one realizes that* $L_B<2$ *a.u. One should realize that the unit length in the different reference frames are given by the intersection of the corresponding observers temporal and spatial axes with the corresponding unit hyperbolae.*

Now, Bob will measure a different length, if we define the length of the rod, which of course moves with respect to his frame in the negative 1 direction, to be measured in the analogous way as for the rod at rest: Bob reads off the spatial coordinates of the rod at the same time as measured by his clock! To figure this out in terms of Alice's coordinates (1.2.16), we first evaluate the time interval in terms of the parameters λ_1, λ_2, and L_A (assuming that $y-x=L_A>0$):

$$\begin{aligned}&\Delta t_\mathrm{B}=\boldsymbol{u}_\mathrm{B}\cdot(\boldsymbol{y}-\boldsymbol{x})=\gamma[(\lambda_2-\lambda_1)-\beta L_\mathrm{A}]\stackrel{!}{=}0\\&\Rightarrow\ \lambda_2-\lambda_1=\beta L_\mathrm{A}.\end{aligned}\tag{1.2.18}$$

The condition employed is that Bob measures the length by marking the spatial coordinates at the same time. Now Bob's spatial coordinate of the difference vector with this constraint on $\lambda_2-\lambda_1$ leads to

$$L_\mathrm{Bob}=-\bar{\boldsymbol{e}}_1\cdot(\boldsymbol{y}-\boldsymbol{x})=\gamma L_\mathrm{A}(1-\beta^2)=\frac{L_\mathrm{A}}{\gamma}=\sqrt{1-\beta^2}L_\mathrm{A}.\tag{1.2.19}$$

This phenomenon is known as **length contraction**, because Bob measures the length of the moving rod to be shorter by an inverse Lorentz factor $1/\gamma=\sqrt{1-\beta^2}$.
Usually it is more convenient to define the length of an object as the length in its rest frame, i.e., L_A in our case. For an arbitrarily moving rod this *proper length* can be expressed in covariant terms. If the four-velocity of the rod is $\boldsymbol{u}$, in Alice's coordinates its coordinates are obviously given by $(u^\mu)=(1,0)^\mathrm{T}$, and thus the proper length can be evaluated with help of the Minkowski product between four-vectors,

because the spatial part of the difference vector in the rest frame of the rod (which is in our case Alice's reference frame) is given by

$$(\boldsymbol{y}-\boldsymbol{x})_{\text{S, proper}} = (\boldsymbol{y}-\boldsymbol{x}) - \boldsymbol{u}[\boldsymbol{u}\cdot(\boldsymbol{y}-\boldsymbol{x})] \tag{1.2.20}$$

and thus the proper length of the rod by

$$L_{\text{proper}}^2 = L_{\text{A}}^2 = -(\boldsymbol{y}-\boldsymbol{x})_{\text{S, proper}}\cdot(\boldsymbol{y}-\boldsymbol{x})_{\text{S, proper}} = [\boldsymbol{u}\cdot(\boldsymbol{y}-\boldsymbol{x})]^2 - (\boldsymbol{y}-\boldsymbol{x})\cdot(\boldsymbol{y}-\boldsymbol{x}). \tag{1.2.21}$$

We can also generalize Bob's measurement of the length of a rod that is not oriented along Alice's 1-axis (assuming again the rod being at rest in Alice's frame). In Alice's frame the world lines of the end points of the rod are now given by

$$(x^\mu) = \begin{pmatrix} \lambda_1 \\ \vec{x} \end{pmatrix}, \quad (y^\mu) = \begin{pmatrix} \lambda_2 \\ \vec{y} \end{pmatrix}. \tag{1.2.22}$$

The length of the rod in Alice's frame is given by the proper distance (1.2.20), i.e., by

$$L_A^2 = (\vec{y}-\vec{x})\cdot(\vec{y}-\vec{x}). \tag{1.2.23}$$

For Bob we have to repeat the calculation from above. In terms of Alice's coordinates the time difference between the location of the endpoint of the rod is

$$\underline{u}_{\text{B}}\cdot(\underline{y}-\underline{x}) = \gamma[(\lambda_2-\lambda_1)-\beta(y^1-x^1)] \stackrel{!}{=} 0 \Rightarrow (\lambda_2-\lambda_1) = \beta(y^1-x^1), \tag{1.2.24}$$

and the spatial components as measured by Bob by

$$\begin{aligned} \overline{y}^1-\overline{x}^1 &= -\overline{\boldsymbol{e}}_1\cdot(\boldsymbol{y}-\boldsymbol{x}) = \gamma[(y^1-x^1)-\beta(\lambda_2-\lambda_1)] = \frac{y^1-x^1}{\gamma}, \\ \overline{y}^2-\overline{x}^2 &= -\overline{\boldsymbol{e}}_2\cdot(\boldsymbol{y}-\boldsymbol{x}) = -\boldsymbol{e}_2\cdot(\boldsymbol{y}-\boldsymbol{x}) = y^2-x^2, \\ \overline{y}^3-\overline{x}^3 &= -\overline{\boldsymbol{e}}_3\cdot(\boldsymbol{y}-\boldsymbol{x}) = -\boldsymbol{e}_3\cdot(\boldsymbol{y}-\boldsymbol{x}) = y^3-x^3. \end{aligned} \tag{1.2.25}$$

Here we have made use of the fact that we can choose as the additional two spatial vectors $\overline{\boldsymbol{e}}_2$ and $\overline{\boldsymbol{e}_3}$ for Bob Alice's spatial vectors $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, because they are perpendicular to both $\overline{\boldsymbol{e}}_0$ and $\overline{\boldsymbol{e}}_1$. This choice is in some sense natural, because we just don't rotate the axes perpendicular to the direction of Bob's motion relative to Alice. From (1.2.25) we can evaluate the length of the rod as measured by Bob as

$$L_{\text{B}}^2 = \frac{(y^1-x^1)^2}{\gamma^2} + (y^2-x^2)^2 + (y^3-x^3)^2. \tag{1.2.26}$$

This just means that for Bob the projection of the rod perpendicular to Bob's velocity relative to Alice is the same length as measured by Alice (i.e., in the rest frame of the rod), while the part parallel to Bob's velocity gets length contracted by an inverse Lorentz factor.

Finally also the simultaneity of events is a frame-dependent notion. Suppose from Alice's point of view there are two events at different places at the same time, t, i.e.,

$$(x^\mu) = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix}, \quad (y^\mu) = \begin{pmatrix} ct \\ \vec{y} \end{pmatrix}. \tag{1.2.27}$$

Then Bob finds that the events are not simultaneous, but the time difference is given by

$$\overline{t} = \frac{1}{c}\boldsymbol{u}_{\text{B}}\cdot(\boldsymbol{y}-\boldsymbol{x}) = \frac{\gamma\beta}{c}(y^1-x^1). \tag{1.2.28}$$

1.3 The twin “paradox”

The twin paradox is simply an example for the time dilation effect. Suppose Alice stays at rest within her inertial reference frame and consider Bob traveling with a large speed. After some time t (as measured by Alice) they meet again and compare their clocks. Because of the time-dilation effect Bob's clock, measuring Bob's proper time, is behind Alice's, which means that Bob has aged less than Alice during his travel. As an example, assume that Bob starts at $t=0$ traveling on a circle in the x^1x^2-plane with constant angular velocity. In Alice's inertial frame this motion is given by the world line, which we parametrize with Alice's time, t, as a parameter,

$$[x^\mu(t)] = \begin{pmatrix} ct \\ R[\cos(\omega t)-1] \\ R\sin\omega t \\ 0 \end{pmatrix}. \tag{1.3.1}$$

Now it is easy to calculate Bob's proper time as a function of Alice's time,

$$\tau = \frac{1}{c}\int_0^t \mathrm{d}t' \sqrt{\frac{\mathrm{d}x^\mu(t')}{\mathrm{d}t'}\frac{\mathrm{d}x_\mu(t')}{\mathrm{d}t'}} = \int_0^t \mathrm{d}t' \sqrt{1-\frac{\omega^2 R^2}{c^2}} = t\sqrt{1-\frac{\omega^2R^2}{c^2}}. \tag{1.3.2}$$

This means that Bob is aging slower than Alice during his travel by a factor given by the square route which is $1/\gamma = \sqrt{1-v^2/c^2}$ with Bob's speed $v=\omega R$.
Of course, this treatment implies that the time-dilation effect can be calculated by adding the incremental proper-time elements, using a sequence of instantaneous inertial reference frames also for accelerated motion. Indeed, this assumption has been verified by comparing the lifetime of various instable particles running at relativistic speeds in storage rings to their lifetime when they are at rest. Lifetime measurements for unstable particles flying with constant velocity also confirm the relativistic time-dilation factor[3].

1.4 General Lorentz transformations

Now we go back to the question, how to transform Alice's to Bob's coordinates. We start with the **rotation free boosts**, which we have already employed in the previous section. Suppose again that Bob is moving with a constant speed βc along Alice's x^1-axis. Then by definition a rotation-free Lorentz transformation is defined as the linear map from Alice's to Bob's frame, when Bob uses the basis vectors, given in terms of Alice's coordinates

$$(\bar{e}_0^\mu) = \gamma\begin{pmatrix} 1 \\ \beta \\ 0 \\ 0 \end{pmatrix}, \quad (\bar{e}_1^\mu) = \gamma\begin{pmatrix} \beta \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (\bar{e}_2^\mu) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad (\bar{e}_3^\mu) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{1.4.1}$$

[3]For references, see the Wikipedia entry on this subect:
`https://en.wikipedia.org/wiki/Time_dilation_of_moving_particles`

Assuming that the origin of Bob's spatial reference frame coincides at Alice's at Alice's time $t = 0$ and that then also Bob's time $\bar{t} = 0$, we then find for the Bob's coordinates of the event $\boldsymbol{x}$

$$(\bar{x}^\mu) = \begin{pmatrix} \bar{\boldsymbol{e}}_0 \cdot \boldsymbol{x} \\ -\bar{\boldsymbol{e}}_1 \cdot \boldsymbol{x} \\ -\bar{\boldsymbol{e}}_2 \cdot \boldsymbol{x} \\ -\bar{\boldsymbol{e}}_3 \cdot \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x^1) \\ \gamma(x^1 - \beta ct) \\ x^2 \\ x^3 \end{pmatrix}. \tag{1.4.2}$$

This can be written with help of the Lorentz-transformation matrix as

$$\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu \quad \text{with} \quad (\Lambda^\mu{}_\nu) = \hat{\Lambda} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.4.3}$$

It is easy to see that we can generalize this to a motion with a velocity $c\vec{\beta}$ in a a self-explaining (1+3)-dimensional notation as

$$\hat{B}(\vec{\beta}) = \begin{pmatrix} \gamma & -\gamma\vec{\beta}^{\mathrm{T}} \\ -\gamma\vec{\beta} & (\gamma - 1)\hat{\beta}\hat{\beta}^{\mathrm{T}} + \mathbb{1}_3 \end{pmatrix} \tag{1.4.4}$$

Here $\hat{\beta} = \vec{\beta}/|\vec{\beta}|$, and one should note that

$$(\hat{\beta}\hat{\beta}^{\mathrm{T}})^j{}_k = \hat{\beta}^j \hat{\beta}^k. \tag{1.4.5}$$

These so-called rotation-free Lorentz boosts are still not the complete set of Lorentz transformations, which can be defined as the entire matrix group that leaves the Minkowski product invariant, i.e., for which

$$\eta_{\mu\nu} x^\mu x^\nu = \eta_{\mu\nu} \bar{x}^\mu \bar{x}^\nu \quad \text{with} \quad \bar{x}^\rho = \Lambda^\rho{}_\mu x^\mu. \tag{1.4.6}$$

Since this must hold true for any $(x^\mu) \in \mathbb{R}^4$, the sufficient and necessary condition is

$$\eta_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu = \eta_{\mu\nu} \tag{1.4.7}$$

or in matrix notation

$$\hat{\Lambda}^{\mathrm{T}} \hat{\eta} \hat{\Lambda} = \hat{\eta}. \tag{1.4.8}$$

This implies that all these matrices build a group, the Lorentz group, because if one has two Lorentz-transformation matrices, fulfilling (1.4.8) also its product fulfills this equation. Indeed, because of $\hat{\eta}^2 = \mathbb{1}$ we have

$$(\hat{\Lambda}_1 \hat{\Lambda}_2)^{\mathrm{T}} \hat{\eta} (\hat{\Lambda}_1 \hat{\Lambda}_2) \hat{\eta} = \hat{\Lambda}_2^{\mathrm{T}} \hat{\Lambda}_1^{\mathrm{T}} \hat{\eta} \hat{\Lambda}_1 \hat{\Lambda}_2 = \hat{\Lambda}_2^{\mathrm{T}} \hat{\eta} \hat{\Lambda}_2 = \hat{\eta}. \tag{1.4.9}$$

Now obviously, despite the general boosts (1.4.4) also all the spatially orthogonal matrices (rotations and space reflections) that are of the form

$$\hat{R} = \begin{pmatrix} 1 & \vec{0}^{\mathrm{T}} \\ \vec{0} & \hat{O} \end{pmatrix}, \tag{1.4.10}$$

where $\vec{0}^{\mathrm{T}} = (0,0,0)$ and $\hat{O} \in \mathrm{O}(3)$ is an orthogonal $\mathbb{R}^{3\times 3}$ matrix and also all combinations of such matrices.

One can show that any Lorentz matrix can be decomposed in two orthogonal spatial transformations and a boost in 1"=direction in the form

$$\hat{\Lambda} = \hat{R}_1 \hat{B}(\beta \vec{e}_1) \hat{R}_2. \tag{1.4.11}$$

At the first glance one might think that thus the Lorentz transformations are parametrized by 7 parameters, because any rotation matrix is parametrized with an angle and an $\mathbb{R}^3$-unit vector, giving the direction of the rotation according to the right-hand rule and there is an additional parameter β, the boost velocity, but that is not true.

To see this, we give a geometric proof of (1.4.11). To that end we think again about two inertial observers, Alice and Bob, using their Minkowski-space basis vectors. Compared to the derivation of the rotation-free boosts above we just have to generalize to the case that both use arbitrary spatial Cartesian basis vectors for their three space-like basis vectors. The only invariant quantity in the game is Bob's four-velocity vector $\boldsymbol{u}_{\mathrm{B}}$ with respect to Alice. Now we can construct the corresponding Lorentz matrix of the form (1.4.11) by transforming her basis vectors successively in Bob's: First Alice uses a rotation matrix $\hat{R}_1^{-1}$ such that Bob's three-velocity vector $\vec{\beta}_{\mathrm{B}} = \vec{u}_{\mathrm{B}}/u_{\mathrm{B}}^0$ points along her new 1-Axis, leading to the new basis system, where Alice is still at rest

$$\boldsymbol{e}'_\mu = (R_2^{-1})^\rho{}_\mu \boldsymbol{e}_\rho. \tag{1.4.12}$$

Obviously one only needs only *two* angles to parametrize this rotation, namely the polar and azimuthal angle of Alice's old 1-axis with respect to the direction $\hat{\beta}$ of Bob's three-velocity.

Then she performs a boost $\hat{B}^{-1}(\beta \vec{e}_1')$ to transform into Bob's rest frame, leading to another Lorentz basis

$$\boldsymbol{e}''_\mu = \left(B^{-1}(\beta \vec{e}_1')\right)^\rho{}_\mu \boldsymbol{e}'_\rho. \tag{1.4.13}$$

In the corresponding reference frame Bob is at rest, but the spatial vectors of this basis can be in an arbitrary orientation relative to the spatial axes he has chosen. Thus he needs to perform another rotation, parametrized by the three Euler angles between the spatial $\boldsymbol{e}''_\mu$ vectors (perhaps followed by a space reflection, if Bob chooses another orientation than Alice for his spatial Cartesian basis system) to get to his originally chosen basis vectors

$$\overline{\boldsymbol{e}}_\mu = (R_1^{-1})^\rho{}_\mu \boldsymbol{e}''_\rho. \tag{1.4.14}$$

Carefully putting the three steps (1.4.12-1.4.13) together, one finally gets

$$\overline{\boldsymbol{e}}_\mu = (\hat{R}_2^{-1} \hat{B}(\beta \vec{e}_1) \hat{R}_1^{-1})^\nu{}_\mu \boldsymbol{e}_\nu \Rightarrow \boldsymbol{e}_\mu = (\hat{R}_1 \hat{B}(\beta \vec{e}_1) \hat{R}_2)^\nu{}_\mu \overline{\boldsymbol{e}} = \Lambda^\nu{}_\mu \overline{\boldsymbol{e}}. \tag{1.4.15}$$

We see that we need only two parameters for the rotation to adjust Alice's spatial basis vectors such that Bob's velocity points in Alice's new 1-direction, then the boost speed β and a full (proper or improper) rotation, characterized by three parameters like the Euler angles, to get Bob's spatial basis vectors in the ones he has arbitrarily chosen. Thus the most general Lorentz transformation is characterized by 6 parameters.

Another way to characterize a boost is with help of the so-called **rapidity**. The idea is based on the point of view that the boost is a kind of "hyperbolic rotation" in the 01-plane (for a boost in Alice's 1 direction), directly reflecting the demand that Alice's and Bob's reference frames are "orthonormalized" in the sense of the Minkowski product. This is reflected by the property of the Lorentz matrix that its columns as well as its rows can be read as the components of the corresponding four vectors being orthonormalized in the sense of the Minkowski product. For the boost (1.4.3) that means we can parametrize the matrix with help of the rapidity η via

$$\Lambda^0{}_0 = \gamma = \cosh\eta, \quad \Lambda^0{}_1 = \Lambda^1{}_0 = -\beta\gamma = -\sinh\eta, \quad \Lambda^1{}_1 = \cosh\eta. \tag{1.4.16}$$

For the three-velocity $v = \beta c$ of Bob with respect to Alice's frame of reference this parametrization leads to

$$\beta = \frac{\Lambda^1{}_0}{\Lambda^0{}_0} = \frac{\sinh\eta}{\cosh\eta} = \tanh\eta, \tag{1.4.17}$$

and the Lorentz matrix reads

$$(B_1{}^{\mu}{}_{\nu}) = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.4.18}$$

It is easy to show that two Lorentz boosts *in the same direction* leads again to a Lorentz boost in this direction. Multiplying two Lorentz-boost matrices of the form (1.4.17) indeed leads to

$$\hat{B}_1(\eta_1)\hat{B}_2(\eta_2) = \hat{B}_1(\eta_1+\eta_2), \tag{1.4.19}$$

where we have made use of the identities

$$\sinh(\eta_1+\eta_2) = \cosh\eta_1 \sinh\eta_2 + \sinh\eta_1 \cosh\eta_2, \quad \cosh(\eta_1+\eta_2) = \cosh\eta_1 \cosh\eta_2 + \sinh\eta_1 \sinh\eta_2. \tag{1.4.20}$$

The boosts *in a fixed direction* thus build an Abelian subgroup of the Lorentz group group, because obviously not only is the composition of two such boosts another boost but according to (1.4.19) one also has

$$\hat{B}_1(\eta_1)\hat{B}_1(\eta_2) = \hat{B}_1(\eta_2)\hat{B}_1(\eta_1) = \hat{B}_1(\eta_1+\eta_2). \tag{1.4.21}$$

We note without derivation that this does *not* hold true for the composition of boosts with boost velocities in *different directions*! Such boosts are neither commutative nor do they lead to another pure boost in whatever direction, but there is always an additional rotation, the so-called **Wigner rotation** involved.

1.5 Addition of velocities

Here we investigate the question of the velocity of a point particle in different frames of reference. Let us consider a point particle moving with three-velocity $\overline{\vec{w}}$ with respect to Bob's frame of reference and ask the question, which velocity Alice will measure. Again we assume that Bob moves with constant velocity $\vec{v} = \beta c \vec{e}_1$ along Alice's 1 axis and the origin of space and time coordinates coincide in the two reference frames. The most easy way to answer this question obviously is to translate everything into manifestly covariant four-vectors and apply the Lorentz boost. To characterize the particle's velocity we again use the idea of its four-velocity, given by its components with respect to Bob's reference frame

$$(\overline{W}^{\mu}) = \gamma_w \begin{pmatrix} 1 \\ \overline{\vec{\beta}}_w \end{pmatrix}, \quad \text{with} \quad \overline{\vec{\beta}}_w = \frac{1}{c}\overline{\vec{w}}, \quad \gamma_w = \frac{1}{\sqrt{1-\beta_w^2}}. \tag{1.5.1}$$

The components of this four-vector with respect to Alice's frame is found by applying (1.4.6) with $\hat{\Lambda} = \hat{B}_1(v)$

$$(W^\mu) = \hat{\Lambda}^{-1}(\overline{W}^\mu) = \hat{B}_1(-v)\overline{W}^\mu = \gamma_w \begin{pmatrix} \gamma_v(1+\beta_v \overline{\beta}_w^1) \\ \gamma_v(\beta_v + \overline{\beta}_w^1) \\ \overline{\beta}_w^2 \\ \overline{\beta}_w^3 \end{pmatrix}. \tag{1.5.2}$$

Thus for the three-velocity with respect to Alice's frame of reference we find

$$\vec{w} = c\frac{\vec{W}}{W^0} = \frac{c}{1+\beta_v \overline{\beta}_w^1} \begin{pmatrix} \beta_v + \overline{\beta}_w^1 \\ \overline{\beta}_w^2/\gamma_v \\ \overline{\beta}_w^3/\gamma_v \end{pmatrix}. \tag{1.5.3}$$

As we can see that is a pretty complicated law of "addition of two velocities". This is due to the fact that for the component in direction of the Bob's velocity we have the effect of all three basic kinematic effects, i.e., time dilation, length contraction and relativity of simultaneity, while for the perpendicular component there is no length contraction.
We can write (1.5.3) with help of the three-vectors in the form

$$\vec{\beta}_w = \frac{1}{c}\vec{w} = \frac{1}{1+\vec{\beta}_v \cdot \overrightarrow{\overline{\beta}}_w} \left[\vec{\beta}_v + \frac{1}{\gamma_v}\overrightarrow{\overline{\beta}}_w + \left(1-\frac{1}{\gamma_v}\right)\vec{\beta}_v \frac{\vec{\beta}_v \cdot \overrightarrow{\overline{\beta}}_w}{\beta_v^2}\right], \tag{1.5.4}$$

which can be rewritten after some algebra as

$$\vec{\beta}_w = \frac{1}{1+\vec{\beta}_v \cdot \overrightarrow{\overline{\beta}}_w}\left[\vec{\beta}_v + \overrightarrow{\overline{\beta}}_w + \frac{\gamma_v}{1+\gamma_v}\vec{\beta}_v \times (\vec{\beta}_v \times \overrightarrow{\overline{\beta}}_w)\right]. \tag{1.5.5}$$

For the magnitude we find, again after some algebra,

$$|\vec{\beta}_w| = \frac{1}{1+\vec{\beta}_v \cdot \overrightarrow{\overline{\beta}}_w}\sqrt{\left(\vec{\beta}_v + \overrightarrow{\overline{\beta}}_w\right)^2 - \left(\vec{\beta}_v \times \overrightarrow{\overline{\beta}}_w\right)^2}. \tag{1.5.6}$$

1.6 Relative velocity

Closely related to the addition theorem of velocities, detailed in the previous section, is the definition of **relative velocity** between two point particles. This is needed later to give a Poincaré-invariant definition of cross sections for two-particle collisions and a covariant formulation of the Boltzmann transport equation in the relativistic kinetic theory of gases.
The relative velocity of a particle 2 relative to a particle 1 is defined as particle 2's velocity in the rest-frame of particle 1[4]. To obtain the relative velocity we use the same idea as in the previous section. Let

[4] Of course, we have to assume that particle 1 has a mass $m_1 > 0$, because a massless particle has no rest frame.

the four-velocities of particle 1 and particle 2 with respect to an arbitrary inertial reference frame be given as

$$u_1^\mu = \frac{1}{\sqrt{1-\vec{\beta}_1^2}}\begin{pmatrix} 1 \\ \vec{\beta}_1 \end{pmatrix}, \quad u_2^\mu = \frac{1}{\sqrt{1-\vec{\beta}_2^2}}\begin{pmatrix} 1 \\ \vec{\beta}_2 \end{pmatrix}. \tag{1.6.1}$$

Then the rotation-free Lorentz bost to the rest frame of particle 1 is given by

$$({\Lambda^\mu}_\nu) = \hat{B}(\vec{\beta}_1) = \begin{pmatrix} \gamma_1 & -\gamma_1 \vec{\beta}_1^{\mathrm{T}} \\ -\gamma_1 \vec{\beta}_1 & \mathbb{1}_3 + (\gamma_1 - 1)\hat{\beta}_1 \hat{\beta}_1^{\mathrm{T}} \end{pmatrix}. \tag{1.6.2}$$

Indeed, it is easily seen by direct calculation that

$$(\tilde{U}^\mu) = ({\Lambda^\mu}_\nu U^\nu) = (1,0,0,0)^{\mathrm{T}}. \tag{1.6.3}$$

This means we can use the result of the previous section by setting $\overline{W}^\mu = u_2^\mu$ or $\vec{\overline{\beta}}_w = \vec{\beta}_2$ and $\vec{\beta}_v = -\vec{\beta}_1$. Then we get from (1.5.5)

$$\vec{\beta}_{\mathrm{rel}} = \frac{1}{1-\vec{\beta}_1 \cdot \vec{\beta}_2}\left[\vec{\beta}_2 - \vec{\beta}_1 + \frac{\gamma_1}{1+\gamma_1}\vec{\beta}_1 \times (\vec{\beta}_1 \times \vec{\beta}_2)\right] \tag{1.6.4}$$

and from (1.5.6)

$$|\vec{\beta}_{\mathrm{rel}}| = \frac{1}{1-\vec{\beta}_1 \cdot \vec{\beta}_2}\sqrt{\left(\vec{\beta}_2 - \vec{\beta}_1\right)^2 - \left(\vec{\beta}_1 \times \vec{\beta}_2\right)^2}. \tag{1.6.5}$$

1.7 The Lorentz group as a Lie group

In this section we investigate the question, how to build finite Lorentz transformations successively from "infinitesimal ones". For that purpose we consider a general Lorentz transformation, which deviates from the unity matrix only by a small matrix

$${\Lambda^\mu}_\nu = {\delta^\mu}_\nu + \delta {\omega^\mu}_\nu. \tag{1.7.1}$$

Now we apply the characterization of the Lorentz transformation as a "pseudo-orthonormal transformation", which keeps the Minkowski product invariant (1.4.7) up to first order in the $\delta\hat{\omega}$:

$$\eta_{\mu\nu} {\Lambda^\mu}_\rho {\Lambda^\nu}_\sigma = \eta_{\rho\sigma} + (\delta\omega_{\rho\sigma} + \delta\omega_{\sigma\rho}) + \mathscr{O}(\delta^2) \stackrel{!}{=} \eta_{\rho\sigma}, \tag{1.7.2}$$

where we have used the usual rule for lowering indices,

$$\delta\omega_{\rho\sigma} = \eta_{\rho\mu} \delta {\omega^\mu}_\sigma. \tag{1.7.3}$$

So from (1.7.2) up to first order we must have

$$\delta\omega_{\rho\sigma} = -\delta\omega_{\sigma\rho}. \tag{1.7.4}$$

So an infinitesimal Lorentz transformation is characterized by an anti-symmetric matrix $\delta\omega_{\rho\sigma}$ with $(4 \cdot 4 - 4)/2 = 6$ independent real parameters.

For the original infinitesimal matrices with an upper first and a lower second index we have

$$\delta\omega^{\mu}{}_{\nu} = \eta^{\mu\rho}\delta\omega_{\rho\nu} = -\eta^{\mu\rho}\delta\omega_{\nu\rho}. \tag{1.7.5}$$

Now we split these matrices into temporal and spatial components. Obviously we have

$$\delta\omega^{0}{}_{j} = \eta^{0\rho}\delta\omega_{\rho j} = \delta\omega_{0j}, \quad \delta\omega^{j}{}_{0} = \eta^{j\rho}\delta\omega_{\rho 0} = \eta^{jk}\delta\omega_{k0} = -\omega_{j0} = +\omega_{0j}. \tag{1.7.6}$$

This means that $\delta\omega^{0}{}_{j} = \delta\omega^{j}{}_{0}$. Thus we can build the mixed temporal-spatial components of the matrix as linear combinations of the three *symmetric matrices*

$$\left[(\hat{K}^1)^{\mu}{}_{\nu}\right] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \left[(\hat{K}^2)^{\mu}{}_{\nu}\right] = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \left[(\hat{K}^3)^{\mu}{}_{\nu}\right] = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{1.7.7}$$

Putting these matrices formaly in a three-vector with matrix-valued components $\hat{\vec{K}}$ we see that the spatio-temporal part of the infinitesimal Lorentz-transformation matrix can be written as

$$\hat{B}(\delta\vec{\beta}) = \mathbb{1}_4 - \delta\vec{\beta}\cdot\hat{\vec{K}}. \tag{1.7.8}$$

It is intuitively clear that this corresponds to an infinitesimal boost with the boost-velocity vector $\delta\vec{\beta}$. To verify this we need to expand (1.4.4) up to first order around $\vec{\beta} = 0$:

$$\hat{B}(\delta\vec{\beta}) = \mathbb{1}_4 + \delta\vec{\beta}\cdot\left(\frac{\partial}{\partial\vec{\beta}}\hat{B}(\vec{\beta})\right)_{\vec{\beta}=0} + \mathcal{O}(\delta^2). \tag{1.7.9}$$

It is, however simpler, to directly look at the full matrix itself instead of calculating the cumbersome matrix-valued gradient in (1.7.10). Since

$$\gamma = \frac{1}{\sqrt{1-\delta\vec{\beta}^2}} = 1 + \mathcal{O}(\delta^2) \tag{1.7.10}$$

we indeed simply have

$$B(\delta\vec{\beta}) = \begin{pmatrix} 1 & -\delta\vec{\beta}^{\mathrm{T}} \\ -\delta\vec{\beta} & \mathbb{1}_3 \end{pmatrix} + \mathcal{O}(\delta^2) \equiv \mathbb{1}_4 - \delta\vec{\beta}\cdot\hat{\vec{K}} + \mathcal{O}(\delta^2). \tag{1.7.11}$$

For the purely spatial components of the infinitesimal matrix $\delta\hat{\omega}$ we find

$$\delta\omega^{j}{}_{k} = \eta^{j\rho}\delta\omega_{\rho k} = -\delta\omega_{jk}, \quad \Rightarrow \delta\omega^{j}{}_{k} = -\delta\omega_{jk} = \delta\omega_{kj} = -\delta\omega^{k}{}_{j}, \tag{1.7.12}$$

i.e., the spatial part is *antisymmetric* and of course refers to rotations. It can be built from the three antisymmetric matrices

$$\left[(J^1)^{\mu}{}_{\nu}\right] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \left[(J^2)^{\mu}{}_{\nu}\right] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \left[(J^3)^{\mu}{}_{\nu}\right] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.7.13}$$

To see that one creates general rotations from that we investigate, how to obtain the finite transformations from the infinite ones. It is suggestive to think that a finite rotation of a three-vector around an axis given by a unit vector $\vec{n}$ by a rotation angle φ can be built by successive infinitesimal rotations $-\delta\varphi\vec{n}\cdot\hat{\vec{J}}$. This we can write in terms of a differential equation for

$$x'(\varphi) = \hat{R}(\varphi\vec{n})x. \tag{1.7.14}$$

Assuming that our idea is correkt this differential equation should read

$$\frac{\mathrm{d}}{\mathrm{d}\varphi}x' = -\vec{n}\cdot\hat{\vec{J}}x. \tag{1.7.15}$$

Given the initial-value condition $\vec{x}\,'(\varphi = 0) = \vec{x}$ the formal solution reads

$$x'(\varphi) = \exp(-\varphi\vec{n}\cdot\hat{\vec{J}})x. \tag{1.7.16}$$

The matrix-exponential function is defined by the corresponding series

$$\exp\hat{M} = \sum_{k=0}^{\infty}\frac{1}{k!}\hat{M}^k, \quad \hat{M}^0 = \mathbb{1}. \tag{1.7.17}$$

To get this series for (1.7.16) we need the matrix $(\vec{n}\cdot\hat{\vec{J}})^k$. We note that all the matrix multipilications deal only with the purely spatial 3×3-submatrices of the 4×4 Minkowski matrices (1.7.13). For these submatrices we can write

$$(\vec{n}\cdot\hat{\vec{J}})^j{}_k = n^i\epsilon^{ijk}, \tag{1.7.18}$$

where ϵ^{ijk} is the usual three-dimensional Levi-Civita symbol, $\epsilon^{123} = 1$ and totally antisymmetric under permutation of its indices. Then we have

$$[(\vec{n}\cdot\hat{\vec{J}})^2]^j{}_k = n^i\epsilon^{ijk'}n^{i'}\epsilon^{i'k'k} = -\delta^{jk} + n^i n^j. \tag{1.7.19}$$

So we have

$$(\vec{n}\cdot\hat{\vec{J}})^{2k} = (-1)^k\begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_3 - \vec{n}\vec{n}^{\mathrm{T}}\end{pmatrix} = (-1)^k\hat{P}_\perp(\vec{n}), \quad k\in\{1,2,3,\ldots\} \tag{1.7.20}$$

and

$$(\vec{n}\cdot\hat{\vec{J}})^k = (-1)^{k+1}\vec{n}\cdot\hat{\vec{J}}, \quad k\in\{0,1,2,\ldots\}. \tag{1.7.21}$$

Plugging this into (1.7.16) and writing out the series (1.7.17) for the matrix exponential for this case, one finally obtains

$$\hat{D}(\vec{\varphi}) = \exp(-\varphi\vec{n}\cdot\hat{\vec{J}}) = \hat{P}_{||}(\vec{n}) + \cos\varphi\hat{P}_\perp(\vec{n}) - \sin\varphi(\vec{n}\cdot\hat{\vec{J}}). \tag{1.7.22}$$

Applying this to a four-vector,

$$\hat{D}(\vec{\varphi})x = \begin{pmatrix} x^0 \\ \vec{n}(\vec{n}\cdot\vec{x}) + \cos\varphi[\vec{x} - \vec{n}(\vec{n}\cdot\vec{x})] - \sin\varphi(\vec{n}\times\vec{x})\end{pmatrix}, \tag{1.7.23}$$

which indeed describes the change of coordinates when one rotates the spatial basis around the axis $\vec{n}$ with an rotation angle φ.

Now we consider the Lorentz boosts in a direction $\vec{n}$ in the analogous way, using the generators (1.7.7). First we note that

$$\vec{n}\cdot\hat{\vec{K}} = \begin{pmatrix} 0 & \vec{n}^{\mathrm{T}} \\ \vec{n} & 0_3. \end{pmatrix}, \quad (\vec{n}\cdot\hat{\vec{K}})^2 = \hat{P}_0 + \hat{P}_{||}(\vec{n}) \tag{1.7.24}$$

with

$$\hat{P}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0_3 \end{pmatrix}. \tag{1.7.25}$$

This implies

$$(\vec{n}\cdot\hat{\vec{K}})^k = \begin{cases} \vec{n}\cdot\hat{\vec{K}} & \text{for} \quad k \in \{1,3,5,\ldots\}, \\ \hat{P}_0 + \hat{P}_{||}(\vec{n}) & \text{for} \quad k \in \{2,4,6,\ldots.\}. \end{cases} \tag{1.7.26}$$

Then the exponential series for a boost reads

$$\exp(-\eta\vec{n}\cdot\hat{\vec{K}}) = \cosh\eta\hat{P}_0 - \sinh\eta\vec{n}\cdot\hat{\vec{K}} + (\cosh\eta - 1)\hat{P}_{||}(\vec{n}). \tag{1.7.27}$$

Setting $\gamma = \cosh\eta$, $\beta = \sinh\eta$, we find indeed the matrix for a rotation-free boost

$$\hat{B}(\vec{\beta}) = \exp[-\eta(|\vec{\beta}|)\vec{n}\cdot\hat{\vec{K}}] \quad \text{with} \quad \eta(\beta) = \operatorname{artanh}\beta. \tag{1.7.28}$$

It is clear that the rapidity is the natural parameter for the Lorentz boost, when built from successive infinitesimal Lorentz boosts in the same direction, because the corresponding infinitesimal rapidities add but not the infinitesimal speeds. However, for an infinitesimal boost we have

$$\eta(\delta\beta)\vec{n} = \operatorname{artanh}\delta\beta\vec{n} = \delta\beta + \mathcal{O}(\delta\beta^3)\vec{n} = \delta\vec{\beta} + \mathcal{O}(\delta\beta^3). \tag{1.7.29}$$

So up to $\mathcal{O}(\delta\beta^3)$ we can identify $\eta(|\delta\vec{\beta}|)\vec{n} \simeq \delta\vec{\beta}$.

As we need it in the next section we write down the matrix for an infinitesimal boost followed by an infinitesimal rotation:

$$B(\delta\vec{\beta})R(\delta\vec{\varphi}) = (\mathbb{1}_4 - \delta\vec{\beta}\cdot\hat{\vec{K}})(\mathbb{1}_4 - \delta\vec{\varphi}\cdot\hat{\vec{J}}) = \mathbb{1}_4 - \delta\vec{\beta}\cdot\hat{\vec{K}} - \delta\vec{\varphi}\cdot\hat{\vec{J}} = R(\delta\vec{\varphi})B(\delta\vec{\beta}), \tag{1.7.30}$$

where we have only kept expressions in linear order of the infinitesimal parameters $\delta\vec{\beta}$ and $\delta\vec{\varphi}$. Writing this out leads to

$$B(\delta\vec{\beta})R(\delta\vec{\varphi}) = R(\delta\vec{\varphi})B(\delta\vec{\beta}) = \mathbb{1}_4 + \begin{pmatrix} 0 & -\delta\vec{\beta}^{\mathrm{T}} \\ -\delta\vec{\beta} & \delta\varphi^i\epsilon^{ijk}\vec{e}_j\vec{e}_k^{\mathrm{T}} \end{pmatrix}. \tag{1.7.31}$$

1.8 Fermi-Walker transport and Thomas precession

Let $x^\mu(s)$ denote an arbitrary time-like curve in spacetime, where the x^μ are components of the spacetime four-vector with respect to an arbitrary inertial Minkowski frame. One can think of this spacetime curve as the worldline of an arbitrarily moving massive particle. The parameter s is chosen via the invariant increment

$$\mathrm{d}s = c\mathrm{d}\tau = \sqrt{\mathrm{d}x^\mu \mathrm{d}x_\mu}. \tag{1.8.1}$$

Then the four-velocity is given by

$$u^\mu = \frac{\mathrm{d}}{\mathrm{d}s}x^\mu. \tag{1.8.2}$$

It is a time-like unit tangent vector along the worldline, i.e., by definition one has $u_\mu u^\mu = 1 = \text{const}$. The idea of **Fermi-Walker transport** is to construct a set of local inertial rest frames along the worldline by a **transport prescription** for arbitrary vectors along the world line such that

(a) u^μ is transported as itself along the curve, since it defines at any instant the time-like basis vector $\boldsymbol{e}'_0 = \boldsymbol{u}$ of the comoving instantaneous inertial restframe of the particle.

Any Minkowski-unit vector n^μ with $n^\mu u_\mu = 0$ (which is necessarily space-like since it is Minkowski orthogonal to the time-like vector, i.e., $n_\mu n^\mu = -1$) should be transported along the worldline from s to $s + \mathrm{d}s$ such that it stays

(b) Minkowski-orthogonal to u^μ,

(c) stays a spacelike Minkowski-unit vector, and

(d) it is unrotated with respect to the local inertial restframe at s.

The transport prescription fulfilling the properties (a)-(d) defines the **Fermi-Walker transport**. Then it is easy to construct the set of instantaneous non-rotating Minkowski bases: One just starts at the point $x_0^\mu = x^\mu(s_0)$ with an arbitrary Minkowski basis $\boldsymbol{n}'_{(0)}(s_0) = \boldsymbol{u}(s_0)$ and $\boldsymbol{n}_{(j)}(s_0)'$ ($j \in \{1,2,3\}$) such that $\boldsymbol{n}'_{(\mu)}(s_0) \cdot \boldsymbol{n}'_{(\nu)}(s_0) = \eta_{\mu\nu}$ and then defines $\boldsymbol{n}_\mu(s)$ by Fermi-walker transporting this basis along the curve to obtain a locally rotation free basis along the entire curve.

The differential equation for Fermi-Walker transport of vectors is easily derived making use of the Lie algebra of the Lorentz group discussed in Sect. 1.7. The unit-tangent vector $\boldsymbol{u}$ within an infinitesimal step $\mathrm{d}s$ changes by

$$\mathrm{d}\boldsymbol{u} = \mathrm{d}s\,\boldsymbol{a}, \quad \boldsymbol{a} = \frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{u}. \tag{1.8.3}$$

Obviously it obeys

$$\boldsymbol{u} \cdot \boldsymbol{a} = 0, \tag{1.8.4}$$

because $\boldsymbol{u} \cdot \boldsymbol{u} = 1 = \text{const}$, and thus is spacelike and Minkowski orthogonal to $\boldsymbol{u}$. Now the vector $\boldsymbol{n}(s)$ should be transported along the infinitesimal path such that it is not rotated within the local inertial restframe at the instance s and still stays perpendicular to $\boldsymbol{u}$. Obviously this is achieved if we demand that $\boldsymbol{n}$ is transported along the worldline from s to $\mathrm{d}s$ such that it undergoes an infinitesimal rotation-free Lorentz boost in the $\boldsymbol{u}$-$\boldsymbol{a}$ plane, i.e., according to (1.7.7) and (1.7.8) such that

$$\mathrm{d}n^\mu = \mathrm{d}s(a^\mu u_\nu - u^\mu a_\nu)n^\nu \tag{1.8.5}$$

or

$$\frac{\mathrm{d}\boldsymbol{n}}{\mathrm{d}s} = \boldsymbol{a}(\boldsymbol{u} \cdot \boldsymbol{n}) - \boldsymbol{u}(\boldsymbol{a} \cdot \boldsymbol{n}). \tag{1.8.6}$$

For $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ it is immediately clear that $\mathrm{d}\boldsymbol{n} = -\mathrm{d}s\,\boldsymbol{u}(\boldsymbol{a} \cdot \boldsymbol{n})$ is along the original time direction of the local rest frame at instant s, i.e., it doesn't change its spatial components wrt. this frame, and thus particularly does not undergo any rotation relative to this frame. Applying (1.8.7) to $\boldsymbol{u}$ itself one gets $\mathrm{d}\boldsymbol{u}/\mathrm{d}s = \boldsymbol{a}$ as it must be since that is the definition of $\boldsymbol{a}$. So with Fermi-Walker transporting $\boldsymbol{u}(s_0)$ leads to $\boldsymbol{u}(s)$ when integrating the equation (1.8.6) for $\boldsymbol{n} = \boldsymbol{u}$. Further for any $\boldsymbol{n}$ Fermi-Walker transport does not change the Minkowski length of this vector, because multiplying (1.8.6) with $\boldsymbol{n}$ leads to

$$\boldsymbol{n} \cdot \frac{\mathrm{d}\boldsymbol{n}}{\mathrm{d}s} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{n} \cdot \boldsymbol{n}) = 0. \tag{1.8.7}$$

Last but not least also the scalar product $\boldsymbol{u} \cdot \boldsymbol{n}$ stays constant for any Fermi-Walker transported vector $\boldsymbol{n}$:

$$\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{n} \cdot \boldsymbol{u}) = \frac{\mathrm{d}\boldsymbol{n}}{\mathrm{d}s} \cdot \boldsymbol{u} + \boldsymbol{n} \cdot \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}s} = [\boldsymbol{a}(\boldsymbol{u} \cdot \boldsymbol{n}) - \boldsymbol{u}(\boldsymbol{a} \cdot \boldsymbol{n})] \cdot \boldsymbol{u} + \boldsymbol{n} \cdot \boldsymbol{a}. \tag{1.8.8}$$

Because of (1.8.4) and $\boldsymbol{u} \cdot \boldsymbol{u} = 1$ this finally indeed gives

$$\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{n} \cdot \boldsymbol{u}) = 0. \tag{1.8.9}$$

Thus, the Fermi-Walker transport defined for an arbitrary set of vectors $\boldsymbol{n}(s)$ along the worldline via the differential equation (1.8.6) obeys the properties (a)-(d).

An important application of Fermi-Walker transport is the question, how a classical **spin** of a particle behaves, if this particle is moving along an arbitrary time-like curve and if no torque is applied to it. It is clear that then it should not rotate with respect to the rotation free instantaneous rest frames defined by Fermi-Walker transport as described above. This implies that the spin itself is Ferm-Walker transported. Since by definition spin is Minkowski-orthogonal to $\boldsymbol{u}$ its equation of motion is given by (1.8.6),

$$\frac{\mathrm{d}\boldsymbol{\Sigma}}{\mathrm{d}\tau} = -\boldsymbol{u}\left(\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\tau} \cdot \boldsymbol{\Sigma}\right), \tag{1.8.10}$$

where we used $s = c\tau$ to write the equation in terms of proper time rather than Minkowski length of the curve.

Now we want to show that the spin undergoes rotation with respect to an instantaneous inertial rest frame that is defined as rotation free with respect to an arbitrary fixed inertial frame. Let us denote components of the spin-four-vector with respect to this fixed inertial frame as $\Sigma = (\Sigma^\mu)$ and that with respect to this rest frame as $\Sigma^* = (0, \vec{\Sigma}^*)$, where the latter holds true because of the relativistic definition of spin in (any!) local rest frame of the particle. Thus we have

$$\boldsymbol{u} \cdot \boldsymbol{\Sigma} = u \cdot \Sigma = 0. \tag{1.8.11}$$

Now we want to evaluate $\mathrm{d}\vec{\Sigma}^*/\mathrm{d}\tau$ to see that indeed this three-vector undergoes a rotation. Because of the constraint (1.8.11) it is sufficient to consider the spatial part of (1.8.10), because we can write (1.8.11) as

$$\Sigma^0 = \frac{\vec{u}}{u^0} \cdot \vec{\Sigma} = \vec{\beta} \cdot \vec{\Sigma}, \tag{1.8.12}$$

where we have used that $u^0 = \gamma$ and $\vec{u} = \gamma\vec{\beta}$, where $\vec{\beta}$ is the three-velocity of the particle in the fixed intertial frame. Now we want write the spatial part of (1.8.10) in terms of $\vec{\Sigma}$ and $\vec{\beta}$. To simplify notation, we write a dot for the derivative with respect to proper time, τ:

$$\dot{\vec{\Sigma}} = -\vec{u}(\dot{\gamma}\Sigma^0 - \dot{\vec{u}} \cdot \vec{\Sigma}) = -\vec{u}(\dot{\gamma}\vec{\beta} \cdot \vec{\Sigma} - \dot{\vec{u}} \cdot \vec{\Sigma}). \tag{1.8.13}$$

We note that

$$\dot{\gamma} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{\sqrt{1-\vec{\beta}^2}} = \gamma^3\vec{\beta} \cdot \dot{\vec{\beta}}, \quad \dot{\vec{u}} = \dot{\gamma}\vec{\beta} + \gamma\dot{\vec{\beta}}. \tag{1.8.14}$$

Plugging this into (1.8.13) after some algebra one finds

$$\dot{\vec{\Sigma}} = \gamma^2(\dot{\vec{\beta}} \cdot \vec{\Sigma})\vec{\beta}. \tag{1.8.15}$$

Now we use (1.4.4) to evaluate $\vec{\Sigma}^*$, i.e., the spin components in the instantaneous rest frame which is unrotated with respect to the fixed inertial frame. With (1.8.12) we find

$$\Sigma^* = \begin{pmatrix} 0 \\ \vec{\Sigma}^* \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\vec{\beta}^{\mathrm{T}} \\ -\gamma\vec{\beta} & (\gamma-1)/\beta^2\vec{\beta}\vec{\beta}^{\mathrm{T}} + \mathbb{1} \end{pmatrix} \begin{pmatrix} \vec{\beta}\cdot\vec{\Sigma} \\ \vec{\Sigma} \end{pmatrix} = \begin{pmatrix} 0 \\ \vec{\Sigma} - \gamma\vec{\beta}(\vec{\beta}\cdot\vec{\Sigma})/(\gamma+1) \end{pmatrix}. \tag{1.8.16}$$

For the final expression we have used

$$\frac{\gamma-1}{\beta^2} - \gamma = \frac{\gamma(1-\beta^2)-1}{\beta^2} = \frac{1/\gamma-1}{1-1/\gamma^2} = -\frac{\gamma(1-\gamma)}{1-\gamma^2} = -\frac{\gamma}{\gamma+1}. \tag{1.8.17}$$

Thus we have with (1.8.15)

$$\begin{aligned} \dot{\vec{\Sigma}}^* &= \frac{\gamma^2\vec{\beta}}{\gamma+1}(\dot{\vec{\beta}}\cdot\vec{\Sigma}) - (\vec{\beta}\cdot\vec{\Sigma})\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\gamma\vec{\beta}}{\gamma+1}\right) \\ &= \frac{\vec{\beta}\gamma^2}{\gamma+1}\left[(\dot{\vec{\beta}}\cdot\vec{\Sigma}) - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\dot{\vec{\beta}})(\vec{\beta}\cdot\vec{\Sigma})\right] - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\vec{\Sigma})\dot{\vec{\beta}}, \end{aligned} \tag{1.8.18}$$

where we have used

$$\dot{\gamma} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{\sqrt{1-\vec{\beta}^2}} = \gamma^3(\vec{\beta}\cdot\dot{\vec{\beta}}). \tag{1.8.19}$$

The next step is to express the scalar products involving $\vec{\Sigma}$ as expressions with $\vec{\Sigma}^*$. From (1.8.16) we find

$$\vec{\beta}\cdot\vec{\Sigma}^* = (\vec{\beta}\cdot\vec{\Sigma})\left(1-\frac{\gamma\beta^2}{\gamma+1}\right) = (\vec{\beta}\cdot\vec{\Sigma})\frac{1+1/\gamma}{\gamma+1} = \frac{\vec{\beta}\cdot\vec{\Sigma}}{\gamma} \Rightarrow \vec{\beta}\cdot\vec{\Sigma} = \gamma(\vec{\beta}\cdot\vec{\Sigma}^*). \tag{1.8.20}$$

Further we have again with (1.8.16)

$$\begin{aligned} \dot{\vec{\beta}}\cdot\vec{\Sigma}^* &= \dot{\vec{\beta}}\cdot\vec{\Sigma} - (\vec{\beta}\cdot\dot{\vec{\beta}})(\vec{\beta}\cdot\vec{\Sigma})\frac{\gamma}{\gamma+1} \\ &= \dot{\vec{\beta}}\cdot\vec{\Sigma} - (\vec{\beta}\cdot\dot{\vec{\beta}})(\vec{\beta}\cdot\vec{\Sigma}^*)\frac{\gamma^2}{\gamma+1} \end{aligned} \tag{1.8.21}$$

and thus

$$\dot{\vec{\beta}}\cdot\vec{\Sigma} = \dot{\vec{\beta}}\cdot\vec{\Sigma}^* + \frac{\gamma^2}{\gamma+1}(\vec{\beta}\cdot\dot{\vec{\beta}})(\vec{\beta}\cdot\vec{\Sigma}^*). \tag{1.8.22}$$

Now we apply (1.8.20) and (1.8.22) in the bracket of the fist term in (1.8.18), and find after some algebra

$$(\dot{\vec{\beta}}\cdot\vec{\Sigma}) - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\dot{\vec{\beta}})(\vec{\beta}\cdot\vec{\Sigma}) = \dot{\vec{\beta}}\cdot\vec{\Sigma}^*. \tag{1.8.23}$$

Plugging this into (1.8.18) and apply again (1.8.20) also in the second term, we finally get

$$\dot{\vec{\Sigma}}^* = \frac{\gamma^2}{\gamma+1}[(\dot{\vec{\beta}}\cdot\vec{\Sigma}^*)\vec{\beta} - (\vec{\beta}\cdot\vec{\Sigma}^*)\dot{\vec{\beta}}] = \frac{\gamma^2}{\gamma+1}(\dot{\vec{\beta}}\times\vec{\beta})\times\vec{\Sigma}^* = \vec{\Omega}_{\mathrm{T}}\times\vec{\Sigma}^*, \tag{1.8.24}$$

i.e., in the instantaneous restframe that is rotation free to the fixed inertial frame the spin rotates with the **Thomas precession** angular velocity

$$\vec{\Omega}_{\mathrm{T}} = \frac{\gamma^2}{\gamma+1} \dot{\vec{\beta}} \times \vec{\beta} \underset{|\vec{\beta}| \ll 1}{\cong} \frac{1}{2} \dot{\vec{\beta}} \times \vec{\beta}. \tag{1.8.25}$$

The reason for this rotation is that $\vec{\beta}(\tau)$ and $\vec{\beta}(\tau + \mathrm{d}\tau)$ need not be collinear, and thus the Lorentz boosts $\Lambda_{\mathrm{B}}[\vec{\beta}(\tau)]$ can be seen built from infinitesimal Lorentz boosts in different directions, but these lead not to a boost again but to a boost followed by a rotation, the socalled **Wigner rotation**, and this additional rotation figures into (1.8.24) leading to the Thomas precession with angular velocity (1.8.25). This was important to derive the correct gyromagnetic factor of the electron in non-relativistic quantum theory, leading to the correct fine structure (due to spin-orbit interaction) of the spectral lines of hydrogen with help of the **Pauli equation**. Although the Thomas precession was already derived by Silberstein [Sil14], it had to be rediscovered by Thomas to solve the puzzle concerning a factor of 1/2 in the corresponding Hamilton operator of the spin-orbit interaction [Tho26].

Chapter 2

Mechanics

2.1 Particle dynamics

In this Section we leave the purely kinematic considerations and start to formulate dynamical laws. As in Newtonian mechanics the most simple dynamical system is the motion of a single particle in some field. However, in relativistic physics the notion of fields becomes much more important and in some sense a fundamental issue, while in Newtonian mechanics nothing prevents the assumption of "action at a distance", i.e., forces among point particles, depending only on the positions of these particles at a given instant of time. In the theory of relativity this concepts becomes problematic, because what it means is that, if one defines the "action at a distance" in Alice's frame of reference, due to the relativity of simultaneity, this action is not simultaneous in Bob's, moving with respect to Alice's reference frame. In any case the relation between the particles due to the action at a distance is among space-like separated events. It is also not clear, in which sense the interaction between the particles is "causal", because if in one frame a space-like distance may have a positive time component it can have a negative one in another frame.

The way out is to assume that forces are acting on a point particle only through the mediation of fields, i.e., only locally at the space-time position of this point particle. The mutual interaction is then understood by the creation of the field due to one or more particles at the place of the point particle under consideration, and the fields themselves are described by dynamical laws, depending on their various sources.

The paradigmatic example for such a model is **Maxwell electromagnetism**. As it turned out through the work of many physicists at the end of the 19^{th} and the beginning of the 20^{th} century, culminating in Einstein's famous paper of (1905) [Ein05], Maxwell electrodynamics is a relativistic classical field theory of "massless four-vector fields" (see Sect. 3.1).

In this Section we restrict ourselves to an approximation, where we consider the motion of a single particle in a given electromagnetic field. We can assume that the latter is somehow caused by some given somehow far away charge-current distributions, which can be considered as unaffected by the charge under consideration. In addition we assume that we can also neglect the radiation of electromagnetic waves from this "test charge". This leads to the quite difficult problem of **radiation reaction**, which is beyond the scope of this introduction to Special Relativity.

First of all it is convenient to represent the motion of our test particle in a manifestly covariant way. There are two ways to achieve this goal: The first is to start from a description of the particle as a trajectory in Minkowski space $x(\lambda)$, where λ is an arbitrary **scalar world-line** parameter. The second is to use Hamilton's principle and making the action functional a scalar, which then leads to Lorentz covariant equations, which are however, not manifestly covariant. The third way is to combine both ideas, leading to manifestly covariant equations of motion.

The first ansatz, however leads to the problem that instead of the usual three spatial components $\vec{x}$ with respect to some inertial reference frame we introduce four (apparently independent) degrees of freedom. On the other hand instead of an arbitrary world-line parameter one can use the **particle's proper time**[1]. In terms of the coordinate time of an arbitrary inertial reference frame it is defined by

$$\mathrm{d}s = c\mathrm{d}\tau = \mathrm{d}t\sqrt{\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}}, \tag{2.1.1}$$

which is obviously a Lorentz scalar. From its definition it follows that derivatives of $\boldsymbol{x}(\tau)$, with respect to τ always gives four-vectors, but the first derivative is constrained, because

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = c\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s} = c\boldsymbol{u}, \tag{2.1.2}$$

where $\boldsymbol{u}$ is the normalized four-velocity introduced above. But now we indeed have the constraint

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau}\cdot\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = c^2, \tag{2.1.3}$$

as follows from the definition of the proper time (2.1.1). This shows that of the four components of the four-velocity only three are independent and thus we can easily formulate equations of motion in the spirit of Newton's 2[nd] law in a covariant way. To this end we define the **four-momentum** of the particle as

$$\boldsymbol{p} = mc\boldsymbol{u}, \tag{2.1.4}$$

where m is the **invariant mass** of the particle. Indeed with respect to an arbitrary reference frame its components read

$$(p^\mu) = mc\gamma\begin{pmatrix} 1 \\ \vec{\beta} \end{pmatrix}, \tag{2.1.5}$$

where $c\vec{\beta}$ is the velocity of the particle in the given reference frame and $\gamma = (1-\vec{\beta}^2)^{-1/2}$ the Lorentz factor. Now we consider the space-like part for a reference frame, where the particle's speed is very small compared to the speed of light, i.e., $\beta = |\vec{\beta}| \ll 1$ then we have for the spatial components

$$\vec{p} = m\vec{v}\gamma = m\vec{v}\left(1+\frac{\beta^2}{2}+\mathcal{O}(\beta^4)\right). \tag{2.1.6}$$

So up to corrections of order $\mathcal{O}(\beta^2)$ the definition of the momentum merges with the usual Newtonian one in the limit $\beta \ll 1$ as it should be.

Now we can also make sense of the temporal component of (2.1.5):

$$p^0 = mc\gamma = mc\left(1+\frac{\beta^2}{2}+\mathcal{O}(\beta^4)\right) \tag{2.1.7}$$

[1]Here we restrict ourselves to massive particles. Massless particles are special, of limited importance, and thus will not be discussed in this manuscripts.

or defining

$$E = cp^0 = mc^2\gamma = mc^2 + \frac{m}{2}v^2 + mc^2\mathcal{O}(\beta^4), \tag{2.1.8}$$

we see that the spatial component is related to the kinetic energy of the particle in Newtonian physics. The only difference is that, up to higher corrections in powers of β, it is convenient in relativistic physics to include a constant rest energy

$$E_0 = mc^2 \tag{2.1.9}$$

into the total energy of the particle. So the generalization of the kinetic energy to relativistically moving particles is

$$E_{\text{kin}} = E - E_0 = mc^2(\gamma - 1). \tag{2.1.10}$$

Since $\gamma \to \infty$ for $\beta \to 1$ we see that we need an infinite amount of energy to accelerate a particle from rest to the speed of light, and this is impossible in practice. This is consistent with the notion that a physically sensible trajectory must be described by a strictly time-like world line, which is tacitly assumed by using the definition (2.1.1) for proper time, because if the world line would not be time-like the square root would become imaginary, which does not make sense for the definition of a time variable.

Now we can formulate Newton's 2$^{\text{nd}}$ Law for the motion of a particle in a four-dimensionally covariant way as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}p^\mu = mc\frac{\mathrm{d}}{\mathrm{d}\tau}u^\mu = m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}x^\mu = K^\mu, \tag{2.1.11}$$

where K^μ describes the force in terms of a four-vector, the **Minkowski four-vector**.

It is, however not totally arbitrary, because of the constraint (2.1.3). Taking the derivative of this constraint with respect to τ we find

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\tau} \cdot \boldsymbol{u} = 0, \tag{2.1.12}$$

i.e., multiplying (2.1.11) with u^μ leads to

$$\boldsymbol{K} \cdot \boldsymbol{u} = \eta_{\mu\nu}K^\mu u^\nu = 0. \tag{2.1.13}$$

Now we can make sense of the time component of the equation of motion (2.1.11), which is automatically fulfilled if the spatial components of this equation is solved and the Minkowski force fulfills the necessary constraint (2.1.13), because we have

$$\eta_{\mu\nu}K^\mu u^\nu = u^0K^0 - \vec{u}\cdot\vec{k} = 0 \Rightarrow u^0K^0 = \vec{u}\cdot\vec{k}. \tag{2.1.14}$$

Thus we can rewrite the 0-component of (2.1.11) as

$$mc\frac{\mathrm{d}u^0}{\mathrm{d}\tau} = K^0 = \frac{\vec{u}}{u^0}\cdot\vec{K} = \vec{\beta}\cdot\vec{K}. \tag{2.1.15}$$

Because of $\mathrm{d}\tau = \sqrt{1-\vec{\beta}^2}\mathrm{d}t = \mathrm{d}t/u^0$ we have

$$mc\frac{\mathrm{d}u^0}{\mathrm{d}t} = \vec{\beta}\cdot\frac{\vec{K}}{u^0} \tag{2.1.16}$$

Multiplying this equation with c and setting $\vec{K} = u^0 \vec{F}$ we see that according to the definition of the energy (2.1.9) and kinetic energy (2.1.10)

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}E_{\mathrm{kin}}}{\mathrm{d}t} = \vec{v} \cdot \vec{F}. \tag{2.1.17}$$

Since from the spatial part of (2.1.11) we have the non-covariant form of the equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = \frac{1}{u_0}\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{p} = \frac{1}{u_0}\vec{K} = \vec{F}, \tag{2.1.18}$$

which shows that $\vec{F}$ is a force corresponding to the Newtonian notion of force as the time derivative of the particle's three-momentum, and thus (2.1.17) is just the usual **work-energy theorem**. This shows that the redundant temporal part of the covariant equation of motion (2.1.11) is just the work-energy theorem in covariant form.

We end our discussion of elementary one-particle dynamics with a note on **"massless particles"**. One should, however, warn the readers that this is a quite academic example, because there is nothing in nature that corresponds literally to a massless classical particle. Sometimes it is stated that this model describes **photons**, but this is a very naive idea about a photon, which is well-defined only within relativistic quantum field theory, and a careful analysis of this true meaning of photons shows that there is no really meaningful classical limit in terms of a massless classical particle. Nevertheless it is an interesting mathematical problem to state possible equations of motion for massless particles. The definition of a massless classical particle is available through the limit $m \rightarrow 0$, which is not as trivial as it seems. It is clear that this limit makes only sense if we define it in terms of energy and momentum, i.e., via

$$p_\mu p^\nu = m^2 c^2 = 0 \Rightarrow p^0 = \frac{E}{c} = |\vec{p}|. \tag{2.1.19}$$

This implies that the three-velocity of a massless particle has the constant magnitude c, because

$$\vec{v} = c\frac{\vec{p}}{p^0} \stackrel{(2.1.19)}{\Rightarrow} |\vec{v}| = c. \tag{2.1.20}$$

This shows that for such a particle a four-velocity as it is defined above for massive particles does not make any sense, because along the trajectory of the particle we have

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}t^2\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t} = \mathrm{d}t^2(c^2 - \vec{v}\cdot\vec{v}) = 0 \Rightarrow \mathrm{d}\tau = 0, \tag{2.1.21}$$

i.e., there is no natural definition of a "proper time". This is also clear from the motivation for the introduction of "proper time" as the time measured by an observer comoving with the particle, i.e., an observer for whom the particle is always at rest, but we have just shown that in any inertial reference frame the particle is moving with the speed of light c, i.e., there cannot be an instantaneous inertial rest frame for a massless particle.

So to formulate the equations of motion in a manifestly covariant way, we must introduce an arbitrary scalar parameter λ and state the equation of motion as

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}\lambda} = K^\mu. \tag{2.1.22}$$

Since $p_\mu p^\mu = 0$, the constraint on possible Minkowski forces now reads

$$p_\mu K^\mu = 0. \tag{2.1.23}$$

So again, we only have to solve the equations of motion for the three spatial equations (2.1.22). Since in general K^μ depends on both p^μ and x^μ we just need to close these equations of motion with the relation of the momentum to the spacetime coordinates. Thus it is more convenient to use the time of the reference frame as a parameter, $\lambda = t$ and write

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \vec{K} \tag{2.1.24}$$

and close the equations of motion simply via (2.1.20)

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \vec{v} = c\frac{\vec{p}}{|\vec{p}|}. \tag{2.1.25}$$

2.2 Motion of a particle in an electromagnetic field

Now it is pretty simple to guess equations of state in a manifestly covariant way. One important example is the equation of motion of a charged point particle in an electromagnetic field. As will be shown in Sect. 3.1 the electromagnetic field can be described by a antisymmetric tensor field $F^{\mu\nu}(x^\rho) = -F^{\nu\mu}(x^\rho)$, here written in terms of its components with respect to an arbitrary inertial frame of reference ("Alice"). The only property relevant for the discussion in this Section is its behavior under Lorentz transformations. The here written contravariant (specified by superscript indices) tensor components simply transform for each index like contravariant vector components, i.e., Bob will describe the same field by the components

$$\overline{F}^{\mu\nu}(\overline{x}^\rho) = {\Lambda^\mu}_\alpha {\Lambda^\nu}_\beta F^{\alpha\beta}(x^\gamma) \quad \text{with} \quad x^\gamma = {(\Lambda^{-1})^\gamma}_\rho \overline{x}^\rho. \tag{2.2.1}$$

Now it is very easy to postulate the equations of motion of a particle in this field. With the electric charge q, describing the coupling strength of the particle to the electromagnetic field it reads

$$m\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = K^\mu = qF^{\mu\nu}(x^\rho)u_\nu \quad \text{with} \quad u^\mu = \frac{1}{c}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}, \quad u_\nu = \eta_{\mu\nu}u^\mu. \tag{2.2.2}$$

This form shows immediately that it is a covariant equation of motion and because $F^{\mu\nu} = -F^{\nu\mu}$ it fulfills the constraint (2.1.13).

To bring it in a more familiar form we write the antisymmetric tensor components in terms of the usual three-dimensional electric and magnetic field components $\vec{E}$ and $\vec{B}$:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{2.2.3}$$

That this is the right choice of the names for the components becomes clear when writing the Minkowski force on the right-hand side of Eq. (2.2.2) explicitly

$$(K^\mu) = qF^{\mu\nu}u_\nu = q\begin{pmatrix} \vec{E}\cdot\vec{u} \\ u_0\vec{E} + \vec{u}\times\vec{B} \end{pmatrix}. \tag{2.2.4}$$

This becomes even more familiar, when we rewrite the derivatives of the equation of motion in terms of the usual coordinate time, $t = x^0/c$. To this end we write the equations of motion in two first-order equations by introducing the four-momentum (2.1.4). Then the equations of motion read

$$\left(\frac{\mathrm{d}}{\mathrm{d}\tau}p^\mu\right) = (K^\mu) = q\begin{pmatrix} \vec{E}\cdot\vec{u} \\ u_0\vec{E}+\vec{u}\times\vec{B}\end{pmatrix}, \quad u^\mu = \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}\tau}x^\mu. \tag{2.2.5}$$

Now we have by definition $u^0 = u_0 = \gamma = (1-\beta^2)^{-1/2}$ with $\vec{\beta} = \vec{v}/c$, $\vec{v} = \vec{u}/u^0$. Thus the spatial part of the first equation of motion reads

$$\frac{1}{u_0}\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{p} = \frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = q\vec{E} + \frac{\vec{v}}{c}\times\vec{B}, \tag{2.2.6}$$

which precisely reads like the equation of motion of a particle in an electromagnetic field $(\vec{E},\vec{B})$ in non-relativistic mechanics. However, one must note that it is not the same equation, because of the additional γ factor in the definition of $\vec{p}$. For the relation to the coordinates we need the second equation in (2.2.5):

$$\vec{u} = \frac{1}{mc}\vec{p} = \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{x}. \tag{2.2.7}$$

Dividing this by $u^0 = p^0/(mc)$ we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{x} = c\frac{\vec{p}}{p^0}. \tag{2.2.8}$$

Now we note that the temporal component in (2.2.5) is automatically fulfilled for any solution of the spatial equation in terms of (2.2.6). To see this we multiply (2.2.6) with $\vec{v}$, leading to

$$\vec{v}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = q\vec{v}\cdot\vec{E}, \tag{2.2.9}$$

but this is precisely the temporal component of the first equation in (2.2.5) divided by u^0. So the "on-shell" condition $p_\mu p^\mu = m^2c^2 = \text{const}$ is compatible with the equations of motion by construction, and thus after solving (2.2.6) we can evaluate $p^0 = \sqrt{m^2c^2+\vec{p}^2}$ and then solve (2.2.8). As usual we just need appropriate initial conditions for $\vec{p}$ and $\vec{x}$ to make the solutions consistent. Of course, in the general case the solution is not that easily obtained, because the fields $\vec{E}$ and $\vec{B}$ depend on t and $\vec{x}$. So it's a full system of equations of motion as in Newtonian mechanics.

2.2.1 Massive particle in a homogeneous electric field

The most simple example is of course the case of a homogeneous electric field $\vec{E} = \text{const}$ and $\vec{B} = 0$. Assume $\vec{E} = (E,0,0)^{\mathrm{T}}$ and suppose we have the initial condition $\vec{p}(0) = 0$ and $\vec{x}(0) = 0$. Then we can very easily solve (2.2.6)

$$\frac{\mathrm{d}}{\mathrm{d}t}p^1 = \frac{q}{m}E \Rightarrow p^1(t) = qEt. \tag{2.2.10}$$

The other components of the momentum are conserved and thus we have for our initial conditions $p^2 = p^3 = 0$. Now we have from the on-shell condition

$$p^0(t) = \sqrt{m^2c^2+(qEt)^2} \tag{2.2.11}$$

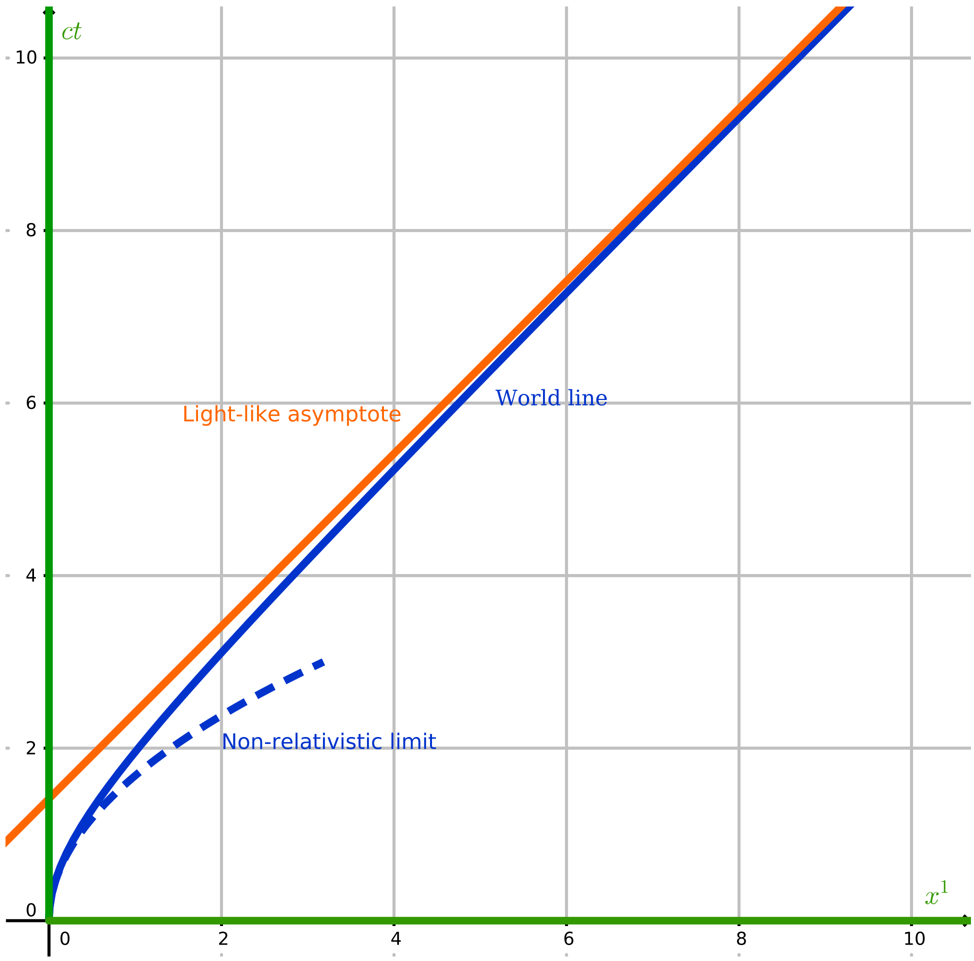


Figure 2.1: *Minkowski diagram for the "hyperbolic motion" of a charged particle due to a homogeneous electric field. The corresponding world line has a light-like asymptote, which indicates that even a light signal sent from x^1 at a time $t >$ cannot reach an observer being accelerated with the charged particle. In this sense this light-like asymptote is a* ***horizon****, the so-called* ***Rindler horizon****.*

and thus

$$v^1(t) = \frac{\mathrm{d}x^1}{\mathrm{d}t} = \frac{c p^1}{p^0} = \frac{qcEt}{\sqrt{m^2c^2 + q^2E^2t^2}}. \tag{2.2.12}$$

This is easily integrated to

$$x^1(t) = \frac{mc^2}{qE}\left(\sqrt{1 + \frac{q^2E^2t^2}{m^2c^2}} - 1\right). \tag{2.2.13}$$

Here we have written the solution in a form, which lets us discuss easily the non-relativistic limit, which is valid at early times, where $|v^1| \ll c$ which obviously is the case at early times, i.e., for $|qEt| \ll mc$ or $t \ll mc/(qE)$. Then we can expand the square root in (2.2.13), leading to

$$x^1(t) = \frac{mc^2}{qE}\left(\frac{q^2E^2t^2}{2m^2c^2} + \mathcal{O}\left[\left(\frac{qEt}{mc}\right)^4\right]\right) = \frac{qEt^2}{2m} + \cdots \tag{2.2.14}$$

In leading order for the early-time evolution we have the expected parabola of Newtonian mechanics in the space-time diagram. For large times $t \gg mc/(qE)$ the world line comes closer and closer to a light-like world line. To estimate this behavior more accurately we rearrange (2.2.13) to

$$x^1(t) = \frac{mc^2}{qE}\left(\frac{qEt}{mc}\sqrt{1+\frac{m^2c^2}{q^2E^2t^2}}-1\right) \underset{t\to\infty}{\cong} c\left(t-\frac{mc}{qE}\right). \tag{2.2.15}$$

This also implies that any light signal of the kind discussed in Sect. 1.2 cannot reach the particle, if sent from the origin later than $t_{\mathrm{H}} = \frac{mc}{qE}$ (see the Minkowski diagram 2.1).

2.2.2 Massive particle in a homogeneous magnetic field

Now we turn to the case of the motion in a homogeneous magnetic field, i.e., $\vec{E} = 0$, $\vec{B} = B\vec{e}_3 = \text{const.}$ Here it is more convenient to start from the manifestly covariant equations of motion (2.2.5). We need to consider only the spatial part of the equation of motion since the temporal equation is then fulfilled automatically as detailed above. We write the equation of motion in the form

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{u} = \frac{q}{mc}\vec{u}\times\vec{B} = \frac{q}{mc}\begin{pmatrix} Bu^2 \\ -Bu^1 \\ 0 \end{pmatrix}. \tag{2.2.16}$$

The 3"=component is immediately integrated

$$u^3(\tau) = u_0^3 = \text{const.} \tag{2.2.17}$$

The motion in the (12)-plane can be found by introducing the auxilliary complex variable

$$\xi = u^1 + \mathrm{i}u^2 \Rightarrow \frac{\mathrm{d}\xi}{\mathrm{d}\tau} = \frac{qB}{mc}(u^2 - \mathrm{i}u^1) = -\mathrm{i}\omega\xi, \tag{2.2.18}$$

where we have defined the **cyclotron frequency** $\omega = qB/(mc)$. This equation of motion is easily integrated by

$$\xi(\tau) = \xi_0 \exp(-\mathrm{i}\omega\tau). \tag{2.2.19}$$

Separating back into real and imaginary part gives

$$\begin{aligned} u^1(\tau) &= \mathrm{Re}\,\xi(\tau) = u_0^1\cos(\omega\tau) + u_0^2\sin(\omega\tau), \\ u^2(\tau) &= \mathrm{Im}\,\xi(\tau) = -u_0^1\sin(\omega\tau) + u_0^2\cos(\omega\tau). \end{aligned} \tag{2.2.20}$$

The temporal component of the four-velocity is given by the constraint $u_\mu u^\mu = 1$, i.e.,

$$u^0(\tau) = \sqrt{1+\vec{u}^2(\tau)} = \sqrt{1+|\xi(\tau)|^2+(u_0^3)^2} = \sqrt{1+\vec{u}_0^2} = \text{const.} \tag{2.2.21}$$

This shows that indeed also the temporal equation of motion (2.2.5) is fulfilled, as it must be.

Further integration of (2.2.20) and (2.2.18) leads to

$$[x^\mu(\tau)] = c\begin{pmatrix} \sqrt{1+\vec{u}_0^2}\,\tau \\ -u_0^1/\omega\sin(\omega\tau) + u_0^2/\omega[\cos(\omega\tau)-1] + x_0^1 \\ u_0^1/\omega[\cos(\omega\tau)-1] + u_0^2/\omega\sin(\omega\tau) + x_0^2 \\ u_0^3\tau + x_0^3 \end{pmatrix}. \tag{2.2.22}$$

To interpret the trajectory in an easier way we can also integrate (2.2.19) to get the projection of the motion in the (12)-plane

$$\zeta(\tau) = x^1(\tau) + \mathrm{i}x^2(\tau) = \frac{\mathrm{i}\xi_0 c}{\omega}[\exp(-\mathrm{i}\omega\tau) - 1] + \zeta_0, \tag{2.2.23}$$

i.e., this projection describes a circle of the radius

$$R = \frac{c|\xi_0|}{|\omega|} = \frac{c\sqrt{(u_0^1)^2 + (u_0^2)^2}}{|\omega|}. \tag{2.2.24}$$

It is also easy to express the solution (2.2.20) in terms of the reference-frame time t, since

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = u^0 \;\Rightarrow\; t = \tau\sqrt{1 + \vec{u}_0^2} = \text{const.} \tag{2.2.25}$$

or in terms of the three-velocity $\vec{\beta} = \vec{v}/c = \vec{u}/u^0$,

$$u^0 = \frac{1}{\sqrt{1 - \vec{\beta}^2}} = \gamma = \frac{1}{\sqrt{1 - \vec{\beta}_0^2}} \;\Rightarrow\; \tau = t\sqrt{1 - \vec{\beta}_0^2}. \tag{2.2.26}$$

With $\vec{v} = c\vec{u}/u^0$ this leads via substitution in (2.2.17) and (2.2.20)

$$\vec{v}(t) = \begin{pmatrix} v_0^1 \cos(\omega_{\text{lab}} t) + v_0^2 \sin(\omega_{\text{lab}} t) \\ -v_0^1 \sin(\omega_{\text{lab}} t) + v_0^2 \cos(\omega_{\text{lab}} t) \\ v_0^3 \end{pmatrix} \quad \text{with} \quad \omega_{\text{lab}} = \omega\sqrt{1 - \vec{\beta}_0^2}, \tag{2.2.27}$$

and substitution of (2.2.27) in (2.2.22)

$$[x^\mu(t)] = \begin{pmatrix} ct \\ -v_0^1/\omega_{\text{lab}} \sin(\omega_{\text{lab}} t) + v_0^2/\omega_{\text{lab}}[\cos(\omega_{\text{lab}} t) - 1] + x_0^1 \\ v_0^1/\omega_{\text{lab}}[\cos(\omega_{\text{lab}} t) - 1] + v_0^2/\omega_{\text{lab}} \sin(\omega_{\text{lab}} t) + x_0^2 \\ v_0^3 t + x_0^3 \end{pmatrix}. \tag{2.2.28}$$

The radius of the circle in the projection of the trajectory in lab-frame parameters is given by

$$R = \frac{\sqrt{(v_0^1)^2 + (v_0^2)^2}}{|\omega_{\text{lab}}|} = \frac{mc\sqrt{(v_0^1)^2 + (v_0^2)^2}}{qB\sqrt{1 - (\beta_0^1)^2 + (\beta_0^2)^2}}. \tag{2.2.29}$$

2.3 Bell's space-ship paradox

Now we can also discuss another famous "paradox", known as **Bell's spaceship paradox**[2] [Bel87]. There Bell tells the story that once a discussion started at CERN canteen about the following question: Suppose there are three space-ships A, B, C initially at rest in A's inertial frame of reference with B and C at equal distance from A. Then A sends a light signal, reaching B and C simultaneously at A's

[2] I thank the Physics Forums community for a very lively and fruitful discussion on this subject in the thread https://www.physicsforums.com/threads/bell-spaceship-paradox-quantitatively.828670/
Particularly helpful were the comments by PeterDonis, DaleSpam, and bcrowell.

coordinate time $t_A = 0$, which starts a mechanism in both space ships to accelerate in precisely the same way along A's 1-axis. Then, from the point of view of A, they move always staying at the same distance from each other, because they both move exactly with the same velocity at each time t_A. Now suppose B and C where loosely connected by some "rigid rope" at the beginning. Now the question is whether the rope will simply stay between B and C as it is or whether it will break. The short answer given in the book, referring to length contraction is, in my opinion, not very convincing, and thus we shall look at the problem in a concrete analytical example from an elementary as possible point of view.

To answer this question, we have to look at the situation from the point of view of the instantaneous restframe of B or C looking not only at the spaceships but also at the attached "rigid rod". First we have to define what we understand under a rigid body. Here we refer to **Born's notion of a rigid body**, according to which a rigid body is such that around any of its points there always exists an instantaneous inertial reference frame, where any infinitesimal volume element does not change its shape. We note right away that this is, even in this local form, a very restrictive definition, and such a body cannot really exist according to relativity: Soon after Born has given his definition of a rigid body in special relativity [Bor09] Herglotz and Noether [Her10, Noe10] have proven that such a body's motion is completely determined by the motion of one of its points along an arbitrarily given timelike worldline. Then the motion is purely translative for all of its points or the body rotates with constant angular velocity around a point at rest or uniform motion with respect to an inertial reference frame[3]. Thus such a "Born-rigid" body has only three degrees of freedom rather than six, as expected from the non-relativistic theory of the spinning top. In fact, Max von Laue has given a simple argument that a real many-body system in relativistic continuum mechanics must always have an infinite number of degrees of freedom: Due to the finite limiting speed (the speed of light in vacuo) to determine the initial state of motion of such a continuous medium one has to give the positions and momenta of each of its points. Particularly accelerated rigid motion of several point masses (or a continuum of such points) is only possible, if one carefully preadjusts the acceleration of all points, and as we shall see at the most elementary example in this section there are severe restrictions on the possibility of such a preadjustment.

To this end we use the example of constant proper acceleration for both spaceships in Bell's paradox. This is the motion, we have already discussed above as the motion of point mass in an electrostatic homogeneous field. We solve the equation of motion again using the manifestly covariant formalism, which means we use only four vectors and the proper times of B and C. This simplifies the further discussion, particularly the Lorentz transformations from A's observational frame to the instantaneous rest frames of either B and C.

We define the constant proper acceleration as α, which is in the above given example of the motion of a charged particle in an external homogeneous electric field $\alpha = eE/m$. Using the relation $p^\mu = mcu^\mu$ between momentum and four-velocity That means we can write (2.2.5) as

$$\frac{\mathrm{d}u^0}{\mathrm{d}\tau} = \frac{\alpha}{c} u^1, \quad \frac{\mathrm{d}u^1}{\mathrm{d}\tau} = \frac{\alpha}{c} u^0. \tag{2.3.1}$$

For the setting of the Bell space-ship paradox u^μ are the components of the four-velocity of both space ships B and C, and we assume that the proper times of both space ships are $\tau = 0$ when A's coordinate time $t = 0$, and before the acceleration starts, the space ships are at rest wrt. A, i.e., $u^1(0) = 0$, which implies that $u^0(0) = 1$. To solve (2.4.1) we take the τ-derivative of the 2$^{\text{nd}}$ equation (2.4.1) and use the first one

$$\frac{\mathrm{d}^2 u^1}{\mathrm{d}\tau^2} = \frac{\alpha}{c}\frac{\mathrm{d}u^0}{\mathrm{d}\tau} = \frac{\alpha^2}{c^2} u^1. \tag{2.3.2}$$

[3] For a derivation, using continuum-mechanical methods, see Sect. 2.5.

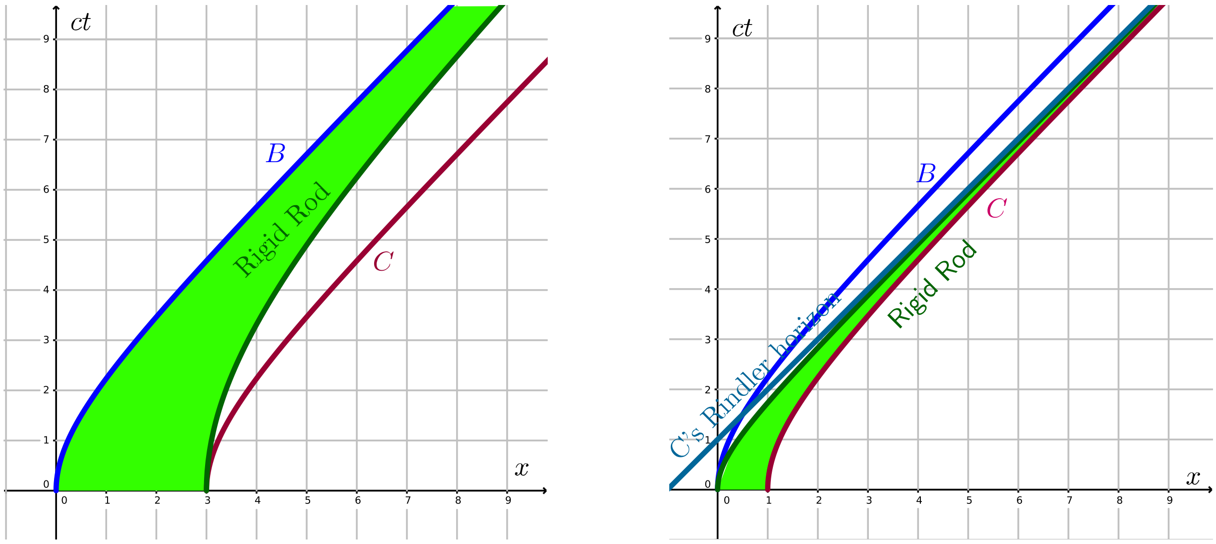


Figure 2.2: *Minkowski diagram illustrating Bell's space-ship paradox.* ***Left:*** *Comparing the world tube of a rigid rod of proper length L_A attached to the front of spaceship B to that of spaceship C, one sees that the front end of the rod always stays behind spaceship C, i.e., a real elastic string connecting the spaceships must stretch and eventually break.* ***Right:*** *Comparing the world tube of a rigid rod of proper length L_A attached to the rear of spaceship C to that of spaceship B one sees that the rear end of the rod drifts further and further away from spaceship B. At given proper acceleration α of spaceship C the length of the rod is restricted such that it must not reach or exceed spaceship C's Rindler horizon, i.e., one has the constraint $\alpha L_A < c^2$.*

The general solution reads

$$u^1(\tau) = C_1 \cosh\left(\frac{\alpha\tau}{c}\right) + C_2 \sinh\left(\frac{\alpha\tau}{c}\right). \tag{2.3.3}$$

Because of the initial condition $u^1(0) = 0$ we have $C_1 = 0$, which implies using the 2nd equation in (2.4.1)

$$u^0 = \frac{c}{\alpha}\frac{\mathrm{d}u^1}{\mathrm{d}\tau} = C_2 \cosh\left(\frac{\alpha\tau}{c}\right) \; u^0(0) = 1 \Rightarrow C_2 = 1. \tag{2.3.4}$$

Thus our solution reads

$$u^0(\tau) = \cosh\left(\frac{\alpha\tau}{c}\right), \quad u^1(\tau) = \sinh\left(\frac{\alpha\tau}{c}\right). \tag{2.3.5}$$

We get the components of the position four-vectors by integrating the corresponding initial-value problem

$$\frac{\mathrm{d}x_B^\mu}{\mathrm{d}\tau} = c u^\mu, \quad \frac{\mathrm{d}x_C^\mu}{\mathrm{d}\tau} = c u^\mu, \quad [x_B^\mu(0)] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad [x_C^\mu(0)] = \begin{pmatrix} 0 \\ L_A \end{pmatrix}, \tag{2.3.6}$$

where we have only written the relevant components $x^0_{B/C}$ and $x^1_{B/C}$. Using (2.4.5) gives by simple integration

$$\begin{aligned} x_B^0 &= c t_B = \frac{c^2}{\alpha}\sinh\left(\frac{\alpha\tau}{c}\right), \quad x_B^1 = \frac{c^2}{\alpha}\left[\cosh\left(\frac{\alpha\tau}{c}\right) - 1\right], \\ x_C^0 &= c t_C = \frac{c^2}{\alpha}\sinh\left(\frac{\alpha\tau}{c}\right), \quad x_C^1 = \frac{c^2}{\alpha}\left[\cosh\left(\frac{\alpha\tau}{c}\right) - 1\right] + L_A. \end{aligned} \tag{2.3.7}$$

Note that all these equations are only valid for $\tau > 0$. For $\tau < 0$ B and C are at rest wrt. A by definition, i.e.,

$$(x_B^\mu) = \begin{pmatrix} c\tau \\ 0 \end{pmatrix}, \quad (x_C^\mu) = \begin{pmatrix} c\tau \\ L_A \end{pmatrix} \quad \text{for} \quad \tau < 0. \tag{2.3.8}$$

Now we consider first the case that a Born-rigid rod of proper length L_A is fixed with one of its end at spaceship B. To determine its other endpoint's world line, we simply have to transform into the instantaneous restframe of B. Because the four-velocity of B is (2.3.5), obviously the corresponding Lorentz boost in the $(x^0 x^1)$ plane is given by

$$\hat{\Lambda}(\tau_0) = \begin{pmatrix} u^0(\tau_0) & -u^1(\tau_0) \\ -u^1(\tau_0) & u^0(\tau_0) \end{pmatrix}. \tag{2.3.9}$$

Indeed doing the matrix multiplication explicitly, one immediately finds

$$[\tilde{u}^\mu(\tau_0)] = \hat{\Lambda}(\tau_0)[u^\nu(\tau_0)] = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{2.3.10}$$

Then the world line of the other end of the rod of proper length L_A is given as pointing towards the positive $\tilde{1}$ direction in this instantneous restframe, i.e.,

$$[\tilde{x}_{\text{rod}}^\mu(\tau_0) - \tilde{x}_B^\mu(\tau_0)] = \begin{pmatrix} 0 \\ L_A \end{pmatrix}. \tag{2.3.11}$$

Thus in A's frame we have for the endpoint of this rigid rod

$$[x_{\text{rod}}^\mu(\tau_0)] = [x_B^\mu(\tau_0)] + \hat{\Lambda}^{-1}(\tau_0)\begin{pmatrix} 0 \\ L_A \end{pmatrix} = \begin{pmatrix} (c^2/\alpha + L_A)\sinh(\alpha\tau_0/c) \\ (c^2/\alpha + L_A)\cosh(\alpha\tau_0/c) - c^2/\alpha \end{pmatrix}. \tag{2.3.12}$$

We shall show now that this is again a motion with a constant proper acceleration. Using a dot as derivative with respect to the world-line parameter τ (the proper time of spaceships B and C) we first get for the endpoint's velocity with respect to A's frame

$$\beta_{\text{rod}}(\tau_0) = \frac{\dot{x}_{\text{rod}}^1(\tau_0)}{\dot{x}_{\text{rod}}^1(\tau_0)} = \tanh\left(\frac{\alpha\tau_0}{c}\right) \tag{2.3.13}$$

and thus

$$u_{\text{rod}}^\mu(\tau_0) = u_B^\mu(\tau_0). \tag{2.3.14}$$

Finally the proper acceleration is given by

$$\alpha_{\text{rod}}(\tau_0) = c\frac{\mathrm{d}}{\mathrm{d}t}u_{\text{rod}}^1(\tau_0) = c^2\frac{\dot{u}_{\text{rod}}^1(\tau_0)}{\dot{x}_{\text{rod}}^0(\tau_0)} = \frac{c^2\alpha}{c^2 + \alpha L_A}, \tag{2.3.15}$$

which is always less than the proper acceleration of the spaceships B and C. Particularly this shows that the endpoint of the rod never touches spaceship C after B and C started accelerating. This is also clearly seen in the left space-time diagram of Fig. 2.2.

The very same analysis can now be performed by assuming that the rigid rod of proper length L_A is fixed with one end at spaceship C. For its other end we find the world line of the rod's endpoint

$$[x_{\text{rod}}^\mu(\tau_0)] = \left[x_C^\mu(\tau_0) - \hat{\Lambda}^{-1}(\tau_0)\begin{pmatrix} 0 \\ L_A \end{pmatrix}\right] = \begin{pmatrix} (c^2/\alpha - L_A)\sinh(\alpha\tau_0/c) \\ (c^2/\alpha - L_A)[\cosh(\alpha\tau_0/c) - 1] \end{pmatrix}. \tag{2.3.16}$$

In the very same way as above deriving (2.3.15) we find for the proper acceleration of this endpoint

$$\alpha_{\text{rod}}(\tau_0) = \frac{c^2\alpha}{c^2 - \alpha L_A}. \tag{2.3.17}$$

This result shows a remarkable restriction to the possibility of rigid motion: As long as $\alpha L_A < c^2$ there is no problem with the above result, and the proper acceleration of the rigid rod of proper length L_A attached at the rear of spaceship C is always less than the proper acceleration of the rocket. Consequently spaceship B stays more and more behind the rod's end. For $\alpha L_A \rightarrow c^2 - 0^+$ the rod's end's properacceleration $\alpha_{\text{rod}} \rightarrow \infty$, and this is impossible. So one cannot attach a rigid rod of proper length L_A to the rear of spaceship C with $\alpha L_A \geq c^2$, which restricts with given proper length L_A the allowed acceleration of the rod. The constraint $\alpha L_A < c^2$ for the acceleration of the rigid rod means that the endpoint of the rod must always stay on the same side of spaceship C's Rindler horizon, which is given by the asumptotic form of (2.3.7) for $\tau \rightarrow \infty$. Then we have $\cosh(\alpha\tau/c) \cong \sinh(\alpha\tau/c) \cong \exp(\alpha\tau/c)/2$ and thus the equation for the Rindler horizon

$$x^1_{\text{Rindler}}(t) = ct - \frac{c^2 - \alpha L_A}{\alpha}. \tag{2.3.18}$$

For $t = 0$ this becomes positive, when $\alpha L_A > c^2$, and then one cannot accelerate a rigid rod initially connecting B and C, because then it would immediately break when the acceleration starts at $t = 0$.

This finally resolves the paradox: Using a non-rigid rod of initial proper length L_A to connect spaceships B and C it must break, because it would have to stretch to keep B and C connected.

A coordinate-independent characterization of the stretch of the rope in terms of the timelike congruences, see [Cro15].

2.4 The action principle

The formulation of classical mechanics in terms of the "**Hamilton Principle of Least Action**" is of high importance for the development of modern physics since it allows for a detailed analysis of the equations of motion in terms of symmetries, using Noether's theorem. This explains, why it is also an important heuristic tool to find concrete equations of motion. Also above, we have simply stated the electromagnetic force. In principle, there is not way to somehow rigorously "derive" the fundamental equations of motion, but has to make an "educated" guess and then compare the resulting predictions of the model with experiments. The Hamilton Principle is a very clever way to make such educated guesses.

2.4.1 The frame-dependent (1+3)-formalism

The idea underlying the Hamilton Principle is to formulate the equations of motion as a **variational principle**. The equations of motion are then the **stationary points** of the **action functional**. In Newtonian mechanics the action functional for a single particle is defined on the set of trajectories in space via a **Lagrange function**

$$A[q] = \int_{t_1}^{t_2} \mathrm{d}t L(q, \dot{q}, t), \tag{2.4.1}$$

where $q \in \mathbb{R}^3$ are an arbitrary set of **generalized coordinates** (not necessarily cartesian once) and the $\dot{q} \in \mathbb{R}^3$ the corresponding **generalized velocities**. Hamilton's principle then states that the equations of motion are the trajectories $q(t)$ that make the action stationary, assuming **fixed boundary conditions** $q(t_1)=q_1$ and $q(t_2)=q_2$.

To derive the general equations of motion we assume that q is the solution of these equations and take small variations δq around this solution. Because the time, t, is not varied, i.e., by assumption we set $\delta t = 0$, in the Hamilton principle we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta q = \delta \dot{q}. \tag{2.4.2}$$

We note that in the following taking *partial* derivatives of functions of q, $\dot{q}$, and t means that we treat the variables t, $\dot{q}$, and t as independent variables, while a total time derivatives takes q and $\dot{q}$ as functions of time. With this usual notation the variation of the action (2.4.1) up to first order in δq reads

$$\delta A = \int_{t_1}^{t_2} \mathrm{d}t \left[\delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}}\right]. \tag{2.4.3}$$

Now using (2.4.2) we get by integration by parts

$$\int_{t_1}^{t_2} \mathrm{d}t\, \delta \dot{q} \frac{\partial L}{\partial \dot{q}} = \left[\delta q \frac{\partial L}{\partial \dot{q}}\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \mathrm{d}t\, \delta q \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) = -\int_{t_1}^{t_2} \mathrm{d}t\, \delta q \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right). \tag{2.4.4}$$

In the final step we have used the assumption that the δq vanish at the boundaries of the integral. Using this in (2.4.3) gives

$$\delta A = \int_{t_1}^{t_2} \mathrm{d}t\, \delta q \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\right]. \tag{2.4.5}$$

Now the action principle demands that $\delta A = 0$ for all δq, and this can be true only, if the square bracket under the integral itself vanishes along the trajectory of the particle. This leads to the equation of motion in terms of the **Euler-Lagrange equations** of the variational principle:

$$\frac{\delta A}{\delta q} = \frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0. \tag{2.4.6}$$

Now it is clear that we like to derive relativistically covariant equations of motion. As we have already seen above it is not always necessary or convenient to express the equations of motion in manifestly covariant form, i.e., to express the trajectories as world lines in terms of the position four-vector and the derivative with respect to a scalar "world parameter" or the proper time τ of the particle. Of course, it is reassuring if one can bring the equations of motion in such a covariant form. As we also have seen above we only have to pay the prize that we need a constraint, because not all four components of a four-vector are independent quantities but only three. As we have seen above, this constraint can be formulated by (2.1.12) or, expressed in terms of the four-force by (2.1.13).

One of the advantages of the Hamilton principle is that it is simpler to implement Lorentz invariance of the equations of motion, be it in manifest or indirect form by simply demanding that the action is a Lorentz scalar. We can find action functionals by the considering how to build such Lorentz scalars out of the building position four-vector and its derivative with any world parameter we find convenient. The most simple choice is to describe everything in terms of the components of the position four-vector with respect to a fixed inertial frame of reference and use the time of this frame as the world parameter. This, of course leads to equations of motion in a non-covariant way.

We start with the **kinetic term**, i.e., with the action of the "free particle". Its Euler-Lagrange equation must of course lead to uniform motion as the solution, and this implies that the free Lagrangian should be a function of $\dot{\vec{x}}$ only. To find a Lorentz-invariant action we can then use the components position-four vector, x^μ, of the particle. Their time derivatives are of course no four-vectors anymore since time takes part in the Lorentz transformation to another inertial reference frame, but we only need an expression $\mathrm{d}t f(\dot{\vec{x}})$ that is invariant. Obviously the only covariant expression $\propto \mathrm{d}t$ we can build from the time derivatives of $\vec{x}$ is

$$\mathrm{d}t\sqrt{\eta_{\mu\nu}\dot{x}^\mu\dot{y}^\mu} = \mathrm{d}t\, c\sqrt{1-\vec{\beta}^2}, \quad \vec{\beta} = \frac{1}{c}\dot{\vec{x}}. \tag{2.4.7}$$

Since the Lagrangian should have the dimension of an energy, obviously the right choice for the free particle is

$$L_0 = -mc^2\sqrt{1-\vec{\beta}^2}. \tag{2.4.8}$$

The choice of the sign has been taken such that in the non-relativistic limit $|\vec{\beta}| \ll 1$ we get (up to a constant additive factor) the non-relativistic expression. Indeed

$$L_0 = -mc^2 + \frac{m}{2}\vec{v}^2 + \mathcal{O}(\beta^4), \tag{2.4.9}$$

i.e., up to the constant rest energy $E_0 = mc^2$ we obtain the non-relativistic kinetic energy for the free Lagrangian in the non-relativistic limit.

The Euler-Lagrange equations of motion for the free particle, using (2.4.8) read

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_0}{\partial \dot{\vec{x}}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\dot{\vec{x}}}{\sqrt{1-\vec{\beta}^2}}\right) \stackrel{!}{=} 0. \tag{2.4.10}$$

This gives

$$\vec{u} = \frac{\dot{\vec{x}}}{\sqrt{1-\vec{\beta}^2}} = \text{const}. \tag{2.4.11}$$

Since then also

$$\vec{u}^2 = \frac{c^2\vec{\beta}^2}{1-\vec{\beta}^2} = \text{const} \Rightarrow \vec{\beta}^2 = \text{const} \tag{2.4.12}$$

we find indeed the correct solution $\vec{v} = \dot{\vec{x}} = \text{const}$ for the free-particle motion.

Now it is very simple to find expressions for the motion of the particle in external fields by adding a corresponding interaction Lagrangian. The most straight-forward one is for the motion in a four-vector field A^μ. An expression for the interaction part of the Lagrangian, leading to an invariant contribution to the action, is

$$L_{\text{int}} = -\frac{q}{c}\dot{x}^\mu A_\mu. \tag{2.4.13}$$

With the full Lagrangian

$$L = L_0 + L_{\text{int}} = -mc^2\sqrt{1-\dot{\vec{x}}^2/c^2} - \frac{q}{c}\dot{x}^\mu A_\mu = -mc^2\sqrt{1-\dot{\vec{x}}^2/c^2} - qA^0 + \frac{q}{c}\dot{\vec{x}}\cdot\vec{A}. \tag{2.4.14}$$

we find the equations of motion via

$$\frac{\partial L}{\partial \dot{\vec{x}}} = \frac{m\dot{\vec{x}}}{1-\dot{\vec{x}}^2/c^2} + \frac{q}{c}\vec{A}, \tag{2.4.15}$$

where $\vec{A} = (A^i)^4$. Further we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\vec{x}}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\dot{\vec{x}}}{\sqrt{1-\dot{\vec{x}}^2/c^2}}\right) + \frac{q}{c}\partial_t \vec{A} + \frac{q}{c}(\dot{\vec{x}}\cdot\vec{\nabla})\vec{A}. \tag{2.4.16}$$

Further we need

$$\frac{\partial L}{\partial x^k} = -q\frac{\partial A^0}{\partial x^k} + \frac{q}{c}\sum_{j=1}^{3}\dot{x}^j\frac{\partial A^j}{\partial x^k}. \tag{2.4.17}$$

Then the Euler-Lagrange equations (2.4.6) can be brought to the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\dot{x}^k}{\sqrt{1-\dot{\vec{x}}^2/c^2}}\right) = -q\left(\frac{\partial A^0}{\partial x^k} + \frac{1}{c}\frac{\partial A^k}{\partial t}\right) + \frac{q}{c}\sum_{j=1}^{3}\dot{x}^j\left(\frac{\partial A^j}{\partial x^k} - \frac{\partial A^k}{\partial x^j}\right). \tag{2.4.18}$$

Now we can rewrite the 2^{nd} term on the right-hand side by using the definition of the curl of a vector field

$$(\vec{\nabla}\times\vec{A})^l =: B^l = \sum_{j,k=1}^{3}\epsilon^{lmn}\frac{\partial A^n}{\partial x^m}. \tag{2.4.19}$$

By using the relation for the 3D-Levi-Civita symbol

$$\sum_{l=1}^{3}\epsilon^{lmn}\epsilon^{lkj} = \delta^{mk}\delta^{nj} - \delta^{mj}\delta^{nk} \tag{2.4.20}$$

we find

$$\sum_{l=1}^{3}\epsilon^{lkj}B^l = \frac{\partial A^j}{\partial x^k} - \frac{\partial A^k}{\partial x^j}, \tag{2.4.21}$$

and thus

$$\sum_{j=1}^{3}\dot{x}^j\left(\frac{\partial A^j}{\partial x^k} - \frac{\partial A^k}{\partial x^j}\right) = \sum_{j,l=1}^{3}\epsilon^{lkj}\dot{x}^j B^l = \sum_{j,l=1}^{3}\epsilon^{kjl}\dot{x}^j B^l = (\dot{\vec{x}}\times\vec{B})^k. \tag{2.4.22}$$

Also setting

$$E^k = -\frac{\partial A^0}{\partial x^k} - \frac{1}{c}\frac{\partial A^k}{\partial t} \quad \text{or} \quad \vec{E} = -\vec{\nabla}A^0 - \frac{1}{c}\partial_t\vec{A}, \tag{2.4.23}$$

finally the equations of motion (2.4.18) read

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\dot{\vec{x}}}{\sqrt{1-\dot{\vec{x}}^2/c^2}}\right) = q\vec{E} + \frac{q}{c}\vec{v}\times\vec{B}. \tag{2.4.24}$$

This is precisely the equation of motion for a particle in an electromagnetic field $(\vec{E},\vec{B})$ given already in Eq. (2.2.6) due to the definition of relativistic momentum (2.1.5).

[4]Here and in the following, when we work in the non-covariant (1+3)-formalism within a fixed inertial reference frame, a symbol like $\vec{A}$ is always derived from the spatial components of a four-vector quantity with upper indices (contravariant components). Note, however, that in this formalism there are three-vectors (transforming as vectors under rotations) that need not be the spatial components of true four-vectors. An example is the usual three-velocity $\vec{v} = \dot{\vec{x}}$.

2.4.2 Manifestly covariant formulation

It is now also immediately clear, how to write the action principle in a manifestly covariant form. To this end we just have to introduce an arbitrary **scalar "world time"** λ. By construction our action is invariant under redefinition of this parameter. So the action for the motion in an electromagnetic field reads

$$A[x^\mu] = \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda L = \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \left[-mc\sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu} - \frac{q}{c}\eta_{\mu\nu}\dot{x}^\mu A^\nu(x) \right]. \tag{2.4.25}$$

Here, the dot is our short-hand notation for the derivative with respect to λ. The Lagrange function is thus given by

$$L = -mc\sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu} - \frac{q}{c}\eta_{\mu\nu}\dot{x}^\mu A^\nu(x). \tag{2.4.26}$$

The equations of motion follow from the Hamilton principle, now assuming that the endpoints of the world line are fixed, i.e., $\delta x^\mu(\lambda_1) = \delta x^\mu(\lambda_2) = 0$. This gives the **manifestly covariant Euler-Lagrange equations**

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = \frac{\partial L}{\partial x^\mu}. \tag{2.4.27}$$

For the Lagrangian (2.4.26) we are lead to the equations of motion

$$mc\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\dot{x}_\mu}{\sqrt{\dot{x}^\alpha \dot{x}_\alpha}} = \frac{q}{c}F_{\alpha\mu}\dot{x}^\mu \quad \text{with} \quad F_{\alpha\mu} = \partial_\alpha A_\mu - \partial_\mu A_\nu. \tag{2.4.28}$$

Now, by construction, the Lagrangian is a homogeneous function of 1st order of $\dot{x}^\mu$, i.e.,

$$\dot{x}^\mu \frac{\partial L}{\partial \dot{x}^\mu} = L, \tag{2.4.29}$$

which can easily be shown by direct inspection of the Lagrangian (2.4.26). Taking the derivative of this identity with respect to λ leads to

$$\ddot{x}^\mu \frac{\partial L}{\partial \dot{x}^\mu} + \dot{x}^\mu \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = \dot{x}^\mu \frac{\partial L}{\partial x^\mu} + \ddot{x}^\mu \frac{\partial L}{\partial \dot{x}^\mu} \Rightarrow \dot{x}^\mu \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = \dot{x}^\mu \frac{\partial L}{\partial x^\mu}. \tag{2.4.30}$$

This holds *for any world line* of the particle, i.e., not only for the solution of the equations of motion (2.4.27). This implies that only three of the four space-time variables x^μ are independent as it should be.

This enables us to choose the proper time τ for λ, leading to the constraint

$$\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = \dot{x}_\mu \dot{x}^\mu = c^2, \tag{2.4.31}$$

because then $\dot{x}^\mu = cu^\mu$. Indeed for (2.4.28) this leads to the equation of motion (2.2.2) for a particle in the electromagnetic field,

$$m\ddot{x}_\alpha = \frac{q}{c}F_{\alpha\mu}\dot{x}^\mu. \tag{2.4.32}$$

It is now simple to also derive other covariant equations of motion, e.g., the motion in a **scalar field**. For the interaction Lagrangian we have two choices to be a homogeneous function of order 1 of $\dot{x}^\mu$:

$$L_1 = -g\sqrt{\dot{x}_\mu \dot{x}^\nu}\Phi(x), \quad L_2 = -g'\dot{x}^\mu \partial_\mu \Phi = \frac{\mathrm{d}}{\mathrm{d}\lambda}\Phi, \tag{2.4.33}$$

because we want the Lagrangian to be a Lorentz scalar itself and the action parametrization invariant. The second choice, however does not lead to any consequence in the equation of motion, because it is a toal time derivative of a function of the coordinates only, and this part of the Lagrangian leads to a contribution of the action

$$A_2[x^\mu] = -g'\{\Phi[x(\lambda_2)] - \Phi[x(\lambda_1)]\} \Rightarrow \delta A_2 = -g'\left[\delta x^\mu \partial_\mu \Phi(x)\right]_{\lambda_1}^{\lambda_2} \equiv 0. \tag{2.4.34}$$

Thus the Lagrangian for the motion of a particle in a scalar field reads

$$L = -\sqrt{\dot{x}_\mu \dot{x}^\mu}[mc + g\Phi(x)]. \tag{2.4.35}$$

The variational principle leads, after setting $\lambda = \tau$, to the equations of motion

$$(m + g\Phi)\ddot{x}^\mu = \frac{g}{c}(c^2\eta^{\mu\nu} - \dot{x}^\mu \dot{x}^\nu)\partial_\nu \Phi. \tag{2.4.36}$$

Obviously by contraction with $\dot{x}_\mu = c u_\mu$ one again finds that the constraint $u_\mu u^\mu = 1 = \text{const}$ is compatible with this equation of motion since

$$\dot{x}_\mu(c^2\eta^{\mu\nu} - \dot{x}^\mu \dot{x}^\nu) = 0, \tag{2.4.37}$$

because $\dot{x}^\mu \dot{x}_\mu = c^2$ by definition.

2.4.3 Alternative Lagrange formalism

There is an alternative treatment of the manifestly covariant action principle, which has some advantages compared to the "square-root formulation" (i.e., the free-particle part $\propto \sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\mu}$), which simplifies the analysis greatly. For that purpose one gives up the parametrization invariance and chooses right from the beginning the proper time as scalar world-line parameter. The naive application of this idea seems to lead to a problem, since the free Lagrangian of the "square-root formulation" seems to become a constant. For an arbitrary scalar world-time parameter λ the Lagrangian for the motion of a particle in an electromagnetic field is (2.4.2),

$$L = -mc\sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu} - \frac{q}{c}\eta_{\mu\nu}\dot{x}^\mu A^\nu(x). \tag{2.4.38}$$

where the dot means the derivative with respect to the arbitrary world-line parameter λ in the previous section. For a massive particle, whose trajectories are timelike, we can of course choose the proper time, τ, defined by

$$c\dot{\tau} = \sqrt{\dot{x}^\mu \dot{x}_\mu}, \tag{2.4.39}$$

but obviously it cannot be used as an independent parameter for the worldline of the particle, because it is itself a functional of this world line. By definition it leads to the contraint

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x_\mu}{\mathrm{d}\tau} = \frac{1}{\dot{\tau}^2}\dot{x}^\mu \dot{x}_\mu = c^2. \tag{2.4.40}$$

Thus in order to be allowed to use τ as an independent integration parameter in the action principle we have to introduce this condition as a constraint. With a corresponding Lagrange parameter $\mu(\tau)$, we thus consider the Lagrangian.

$$\tilde{L} = L - \frac{\mu}{2}(\dot{x}^\mu \dot{x}_\mu - c^2) = -mc^2 - \frac{q}{c}\dot{x}^\mu A_\mu - \frac{\mu}{2}(\dot{x}^\mu \dot{x}_\mu - c^2), \tag{2.4.41}$$

where now the dot stands for the derivative with respect to τ.

To see the meaning of the Lagrange parameter we derive the equations of motion. Thanks to our Lagrange parameter we are now allowed to vary the x^μ and μ as independent functions of τ, which leads to the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial \tilde{L}}{\partial \dot{x}^\mu} - \frac{\partial \tilde{L}}{\partial x^\mu} = 0, \quad \frac{\partial \tilde{L}}{\partial \mu} = 0. \tag{2.4.42}$$

For (2.4.41) after some algebra this results in

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(\mu \dot{x}^\mu) = \mu \ddot{x}^\mu + \dot{\mu}\dot{x}^\mu = \frac{q}{c} F^{\mu\nu}\dot{x}_\nu, \tag{2.4.43}$$

$$\dot{x}_\mu \dot{x}^\mu = c^2. \tag{2.4.44}$$

Note that, by construction, the variation with respect to the Lagrange parameter leads to the constraint equation. Now contracting the first equation with $\dot{x}_\mu$ we get

$$\mu \ddot{x}^\mu \dot{x}_\mu + \dot{\mu} c^2 = 0, \tag{2.4.45}$$

where we have used the constraint equation (2.4.44). Taking the derivative of (2.4.44) leads to $\ddot{x}^\mu \dot{x}_\mu = 0$, and thus (2.4.45) implies $\dot{\mu} = 0$, i.e., $\mu = \text{const}$.

We note that this holds for any equations of motion derived from the parameter-independent action principle discussed in the previous section, because the modified Lagrangian derived from it reads

$$\tilde{L} = -mc^2 - \frac{\mu}{2}(\dot{x}^\mu \dot{x}_\mu - c^2) + L_{\text{int}}, \tag{2.4.46}$$

where L_{int} is a homogeneous function of degree one with respect to $\dot{x}^\mu$, i.e.,

$$\dot{x}^\mu \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} = L_{\text{int}}. \tag{2.4.47}$$

Since $\tilde{L}$ is not explictly dependent on τ the "Hamiltonian"

$$\tilde{H} = p_\mu \dot{x}^\mu - \tilde{L} \quad \text{with} \quad p_\mu = \frac{\partial \tilde{L}}{\partial \dot{x}^\mu} \tag{2.4.48}$$

is conserved along the trajectory of the particle solving the Euler-Lagrange equations (2.4.43) and (2.4.44), because

$$\frac{\mathrm{d}\tilde{H}}{\mathrm{d}\tau} = \dot{p}_\mu \dot{x}^\mu + p_\mu \ddot{x}^\mu - \ddot{x}^\mu \frac{\partial \tilde{L}}{\partial \dot{x}^\mu} - \dot{x}^\mu \frac{\partial \tilde{L}}{\partial x^\mu} - \dot{\mu}\frac{\partial \tilde{L}}{\partial \mu} = \dot{p}_\mu \dot{x}^\mu - \dot{x}^\mu \frac{\partial \tilde{L}}{\partial x^\mu} - \dot{\mu}\frac{\partial \tilde{L}}{\partial \mu}. \tag{2.4.49}$$

Applying (2.4.43) and (2.4.44) one indeed finds

$$\frac{\mathrm{d}\tilde{H}}{\mathrm{d}\tau} = 0. \tag{2.4.50}$$

Because of (2.4.47) we have

$$\tilde{H} = -\frac{\mu}{2}\dot{x}^\mu \dot{x}_\mu \stackrel{(2.4.44)}{=} -\frac{\mu}{2}c^2 = \text{const} \Rightarrow \mu = \text{const}. \tag{2.4.51}$$

Comparing to the equations of motion in the parametrization-invariant formalism of Sect. 2.4.3 leads again to

$$\mu = m. \tag{2.4.52}$$

Finally we note that we can just use

$$\tilde{L}' = \frac{m}{2}\dot{x}_\mu \dot{x}^\mu - L_{\text{int}} \tag{2.4.53}$$

together with the constraint equation

$$\dot{x}_\mu \dot{x}^\mu = c^2 \tag{2.4.54}$$

to derive equations of motion equivalent to the parametrization-invariant formalism by varying the x^μ as independent degrees of freedom. The consistency with the constraint equation is automatically guaranteed since with

$$p_\mu := \frac{\partial \tilde{L}'}{\partial \dot{x}^\mu} - \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} = m\dot{x}_\mu - \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu}. \tag{2.4.55}$$

where we have introduced the additional sign on the left-hand side for convenience. Now the „Hamiltonian" reads

$$\tilde{H}' = p_\mu \dot{x}^\mu - \tilde{L}' = \frac{m}{2}\dot{x}^\mu \dot{x}_\mu, \tag{2.4.56}$$

where we have used that since L_{int} is a homogeneous function of degree 1 wrt. the $\dot{x}^\mu$ we find

$$\dot{x}^\mu \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} = L_{\text{int}}. \tag{2.4.57}$$

Since $\tilde{L}'$ does not explicitly depend on τ, $\tilde{H} = \text{const}$ along the trajectory of the particles defined by the Euler-Lagrange equations of motion. This implies that indeed (2.4.54) is consistent with the equations of motion. It just defines the absolute normalization of τ to make it the usual proper time of the particle.

2.5 Born rigid bodies

As we shall see later, a theory for the motion of a closed system of interacting point-particles is problematic, and a resolution of these problems is rather to use a description in terms of **continuum mechanics**. Continuum mechanics has to be understood as an effective theory, coarse-graining over macroscopic small but microscopic large regions of space. Then the medium is described as a continuum, which is parametrized by three **material coordinates** ξ^j ($j \in \{1,2,3\}$) describing, e.g., the locations of macroscopically small spatial cells at some fixed coordinate time $t = 0$ within an inertial reference frame or some spatial standard configuration of the body. Then the world line of the fluid cell with material coordinates can be parametrized by

$$x^\mu = x^\mu(\xi^j, s), \tag{2.5.1}$$

where $s = c\tau$ with the τ denoting the proper time of this fluid cell. This defines a congruence of time-like worldlines describing the motion of the continuous medium. This description is analogous to the parametrization of a fluid flow with **Lagrange coordinates** in non-relativistic hydrodynamics. The often more convenient description in terms of the **Euler coordinates** is achieved by considering the inverse of the function (2.5.1), i.e.,

$$\xi^j = \xi^j(x^\mu), \tag{2.5.2}$$

which tells the observer which fluid cell identified by ξ^j is at position $\vec{x}$ at time t. This is a very general description of any continuum-mechanical system. Whether we describe a fluid or a solid depends on the dynamics, which can be formulated elegantly in terms of the Hamilton formalism. In this section we restrict ourselves to **Born rigid bodies** and follow the lucid derivation in [DeW11]. For better understanding we treat the corresponding non-relativistic case first.

2.5.1 Non-relativistic rigid bodies

The non-relativistic case simplifies significantly due to the Newtonian space-time model with its absolute time. We can simply use time to parametrize the trajectories of the matter cells. It is also a priori clear that a rigid body is described by fixing one of its points $\vec{x}_0(t) = \vec{x}(\vec{\xi}_0, t)$ and a body-fixed Cartesian basis. The body has 6 degrees of freedom: 3 tranlatory degrees of freedom described by $\vec{x}_0(t)$ and the rotation of the body-fixed basis against a space-fixed Cartesian basis (usually parametrized with help of **Euler angles**. To prepare for the more complicated relativistic case, we derive this result, using the continuum-mechanical description. The vector connecting two infinitesimally close points separated by the shift $\delta\vec{\xi}$ of the material coordinates is dimply given by

$$\delta x_j = \delta\xi_k \frac{\partial x_j}{\partial \xi_k}, \tag{2.5.3}$$

where we use the usual Euclidean Ricci calculus with only lower indices. The **condition for rigidity** now is that

$$\delta\vec{x}^2 = \frac{\partial x_j}{\partial \xi_k}\frac{\partial x_j}{\partial \xi_l}\delta\xi_k\delta\xi_l = \text{const.} \tag{2.5.4}$$

Defining $\vec{v}(\vec{\xi}, t) = \partial_t \vec{x}(\vec{\xi}, t)$, this equation reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\vec{x}^2 = \left(\frac{\partial v_j}{\partial \xi_k}\frac{\partial x_j}{\partial \xi_l} + \frac{\partial v_j}{\partial \xi_l}\frac{\partial x_j}{\partial \xi_k}\right)\delta\xi_k\delta\xi_l \stackrel{!}{=} 0. \tag{2.5.5}$$

Since the $\delta\xi_k$ are arbitrary variations, the bracket must identically vanish. Now we express this equation in terms of Eulerian coordinates, i.e., using

$$\frac{\partial v_j}{\partial \xi_k} = \frac{\partial v_j}{\partial x_{k'}}\frac{\partial x_{k'}}{\partial \xi_k}, \tag{2.5.6}$$

we get

$$\frac{\partial v_j}{\partial x_{k'}}\left(\frac{\partial x_{k'}}{\partial \xi_k}\frac{\partial x_j}{\partial \xi^l} + \frac{\partial x_j}{\partial \xi_k}\frac{\partial x_{k'}}{\partial \xi^l}\right) = 0. \tag{2.5.7}$$

Now we exchange the summation variables j and k' for the 2nd term on the left-hand side, which yields

$$\left(\frac{\partial v_j}{\partial x_{k'}} + \frac{\partial v_{k'}}{\partial x_j}\right)\frac{\partial x_j}{\partial \xi_l}\frac{\partial x_{k'}}{\partial \xi_k} = 0. \tag{2.5.8}$$

Since the Jacobian matrix $J_{jl} = \partial x_j / \partial \xi_l$ of the transformation from Lagrange to Euler coordinates is by assumption invertible, this is equivalent to the equation

$$\sigma_{jk} = \frac{\partial v_j}{\partial x_k} + \frac{\partial v_k}{\partial x_j} = 0. \tag{2.5.9}$$

Writing $\partial_k = \partial/\partial x_k$ we get

$$\partial_l \sigma_{jk} = 0. \tag{2.5.10}$$

It is easy to show that from this

$$\partial_l \sigma_{jk} + \partial_j \sigma_{kl} - \partial_k \sigma_{lj} = 0 \Rightarrow \partial_j \partial_l v_k = 0. \tag{2.5.11}$$

Now we use an arbitrary fixed material point with material coordinates $\vec{\xi}_0$ within the body, $\vec{x}_0(t) = \vec{x}(\vec{\xi}_0, t)$ and define $\vec{v}_0(t) = \partial_t \vec{x}_0(t)$. Then the general solution of (2.5.12) can be written as This can now be easily integrated twice, leading to

$$v_k = \omega_{jk}(t)[x_j - x_{0j}(t)] + v_{0k}(t). \tag{2.5.12}$$

With (2.5.9) we further find the constraint

$$\omega_{jk}(t) = -\omega_{kj}(t) \Rightarrow \omega_{jk}(t) = \epsilon_{ijk}\omega_i(t), \tag{2.5.13}$$

where we have used that we can write any antisymmetric matrix as the Hodge dual of a vector. Then (2.5.12) reads

$$\vec{v}(t,\vec{x}) = \vec{v}_0(t) + \partial_t \vec{y}(\vec{\xi}, t) = \vec{\omega}(t) \times \vec{y}(\vec{\xi}, t) + \vec{v}_0(t), \tag{2.5.14}$$

where $\vec{y}(\vec{\xi}, t) = \vec{x}(\vec{\xi}, t) - \vec{x}_0(t)$. Then we find

$$\vec{y}(\vec{\xi}, t) = \hat{D}(t)\vec{y}(\vec{\xi}, 0), \tag{2.5.15}$$

where $\hat{D}(t) \in \mathrm{SO}(3)$. This proves the above heuristic argument about the three translational and three rotational degrees of freedom that describe the motion of a non-relativistic rigid body.

2.5.2 Relativistic case

2.6 Thermodynamics

As we shall see later, a theory for the motion of a closed system of interacting point-particles is problematic, and a resolution of these problems is rather to use a description in terms of **continuum mechanics**. The most simple example is the

Chapter 3

Classical Electromagnetism

3.1 Heuristic foundations

We start with the usual form of "microscopic electromagnetism", defined by the **Maxwell equations** for the electric and magnetic field components $\vec{E}(t,\vec{x})$ and $\vec{B}(t,\vec{x})$. Throughout this manuscript we use rationalized Gauss units, also known as **Heaviside-Lorentz units**. This is the usual choice in relativistic high-energy physics. In this system of units the Maxwell equations read

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{3.1.1}$$

$$\vec{\nabla} \times \vec{E} + \frac{1}{c}\partial_t \vec{B} = 0, \tag{3.1.2}$$

$$\vec{\nabla} \cdot \vec{E} = \rho, \tag{3.1.3}$$

$$\vec{\nabla} \times \vec{B} - \frac{1}{c}\partial_t \vec{E} = \frac{1}{c}\vec{j}. \tag{3.1.4}$$

Eqs. (3.1.1) and (3.1.2) are the homogeneous Maxwell equations. They provide merely constraints on the field components rather than being dynamical equations for the field, as we shall see in the following analysis in terms of the relativistic formulation we are aiming at. Nevertheless the homogeneous equations have profound physical content. As any fundamental laws of nature they cannot be derived but summarize empirical evidence. **Gauss's Law for the magnetic field** (3.1.1) summarizes the empirical lack of evidence for the existence of **magnetic monopoles**, i.e., there is no generic magnetic charge seen in nature. **Faraday's Law of induction** (3.1.2) means that the solenoidal part of the electric field is due to the changes of the magnetic field with time.

The inhomogeneous Maxwell equations, **Gauss's Law for the electric field** (3.1.3) and the **Maxwell-Ampere Law** (3.1.4), describe the dynamical laws determining the electromagnetic field due to the sources, i.e., the **electric charge distribution**, $\rho(t,\vec{x})$, which is the amount of electric charge per unit volume at position $\vec{x}$, and the **electric current density**, $\vec{j}(t,\vec{x})$. For an arbitrary surface element with surface-normal vector $\mathrm{d}^2\vec{f}$ around position $\vec{x}$, $\mathrm{d}I = \mathrm{d}^2\vec{f} \cdot \vec{j}(t,\vec{x})$ is the amount of charge per unit time flowing through the surface element with the meaning of its sign given by the (arbitrary!) orientation of the surface-normal vector.

The physical meaning of the electromagnetic field is provided by providing the forces on electric charges. We have already discussed this in the context of point-particle mechanics in Sect. 2.4.2, where we derived in (2.4.24) the equation of motion of a point particle with electric charge q in an **external electromag-**

netic field $(\vec{E},\vec{B})$ as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\dot{\vec{x}}}{\sqrt{1-\dot{\vec{x}}^2/c^2}}\right)=q\vec{E}+\frac{q}{c}\vec{v}\times\vec{B} \tag{3.1.5}$$

or, in manifestly covariant form (2.4.32)

$$m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}x_\alpha=\frac{q}{c}F_{\alpha\mu}\frac{\mathrm{d}}{\mathrm{d}\tau}x^\mu. \tag{3.1.6}$$

Here, τ is the proper time of the partice along his world line. The spatial part is also readily derived by using $\mathrm{d}\tau=\mathrm{d}t\sqrt{1-\dot{\vec{x}}^2/c^2}$ from (3.1.5) by deviding this equation by this square-root factor. Using $u^\mu=1/c\mathrm{d}x^\mu/\mathrm{d}\tau$ then gives

$$m\mathrm{d}_\tau^2\vec{x}=q(u^0\vec{E}+\vec{u}\times\vec{B}). \tag{3.1.7}$$

In any way, (3.1.5) provides the **Lorentz force** for the motion of a particle in an electromagnetic field.

However, one should be aware that this is only an approximate equation of motion since it does not take into account the interaction of the charged particle with its *own electric field*. As we shall see this **radiation-reaction force** provides an unsolvable problem in relativistic dynamics, which finally leads to the conclusion that the notion of a strict **point particle** is at least problematic in relativistic theories.

Thus in the following, we shall first discuss the usually treated problem of finding the electromagnetic field for a given (continuous!) charge-current distribution from Maxwell's equations (3.1.1-3.1.4) in a relativistic covariant way. Only later we shall also address the problem of a completely closed self-consistent dynamical system of continuous charge-particle distributions and the electromagnetic field. We shall deal with the problem of a "point particle" as one example in describing it as a small Born-rigid body of finite extent and then discussing possible approximations for the carefully taken limits of the finite extension to a "point".

Thus, in the following we consider the charge-current distributions in terms of a **continuous fluid**, described by the charge-current distribution, which is given in terms of the fluid-velocity field $\vec{v}(t,\vec{x})$ as

$$\vec{j}(t,\vec{x})=\rho(t,\vec{x})\vec{v}(t,\vec{x}). \tag{3.1.8}$$

Then the **force density** is given by

$$\vec{f}(t,\vec{x})=\rho(t,\vec{x})\vec{E}(t,\vec{x})+\frac{1}{c}\vec{j}\times\vec{B}, \tag{3.1.9}$$

i.e., the force $\mathrm{d}\vec{F}$ on a charged-fluid cell around $\vec{x}$ due to the presence of the electromagnetic field is given by $\mathrm{d}\vec{F}=\mathrm{d}^3x\vec{f}(t,\vec{x})$.

In the next section we aim at a **manifestly covariant description** of these heuristic ideas.

3.2 Manifestly covariant formulation of electrodynamics

We start with the discussion of the charge and current distribution. It is remarkable that already from the Maxwell equations (3.1.1-3.1.4) we can conclude the **conservation of electric charge**, without the necessity to discuss the details of the dynamics of the charged fluid. Indeed taking the divergence of (3.1.4) and using (3.1.3) we find

$$\frac{1}{c}\vec{\nabla}\cdot\vec{j}=-\frac{1}{c}\vec{\nabla}\cdot\partial_t\vec{E}=-\frac{1}{c}\partial_t\vec{\nabla}\cdot\vec{E}=-\frac{1}{c}\partial_t\rho\ \Rightarrow\ \partial_t\rho+\vec{\nabla}\cdot\vec{j}=0. \tag{3.2.1}$$

The latter equation is the **continuity equation**. Its physical meaning becomes clear when integrating (3.2.1) over an arbitrary volume at rest. Then we find, using Gauss's theorem

$$\int_V \mathrm{d}^3 x \partial_t \rho(t,\vec{x}) = \frac{\mathrm{d}}{\mathrm{d}t} \int_V \mathrm{d}^3 x \rho(t,\vec{x}) = \frac{\mathrm{d}}{\mathrm{d}t} Q_V(t) = -\int_V \mathrm{d}^3 x \vec{\nabla} \cdot \vec{j}(t,\vec{x}) = -\int_{\partial V} \mathrm{d}^2 \vec{f} \cdot \vec{j}(t,\vec{x}). \quad (3.2.2)$$

Here ∂V is the (closed) boundary surface of V with the convention to orient the surface-normal vectors $\mathrm{d}^2 \vec{f}$ to point *out of the volume*. Now $Q_V(t)$ is the total charge of the fluid contained in the fixed volume V at time t. According to the equation its change $\dot{Q}_V$ is entirely given by the total current flowing through the boundary surface.

Now (3.2.1) can be brought in a form that we shall prove to be **covariant under Lorentz transformations**. To that end we realize that $(x^\mu) = (ct, \vec{x})$ is a four-vector and $\partial_t = c\partial_0$, where we define the four-dimensional Nabla operator as

$$\partial_\mu = \frac{\partial}{\partial x^\mu}. \quad (3.2.3)$$

We will prove in a moment that putting the index as a lower index when deriving with respect to x^μ provides the correct transformation property of derivatives of scalar, vector, and more general tensor **fields**. In terms of this 4-dimensional Nabla operator we can write (3.2.1) as

$$\partial_\mu j^\mu(t,\vec{x}) = 0, \quad (3.2.4)$$

where we define the **current four-vector field**

$$j^\mu(x) = \begin{pmatrix} c\rho(t,\vec{x}) \\ \vec{j}(t,\vec{x}) \end{pmatrix}. \quad (3.2.5)$$

To confirm that (3.2.3) really defines an operator that acts formally like co-variant vector components we start with the definition of a tensor field by its components as a function of the space-time four-vector components, $T^{\mu\nu\rho\cdots}(x)$.

Now, as detailed in Sect. 1.4, the contravariant components of the space-time four-vector transform under a Lorentz transformation as

$$\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu, \quad \eta_{\mu\rho} \Lambda^\mu{}_\nu \Lambda^\rho{}_\sigma = \eta_{\nu\sigma} \Rightarrow \eta_{\nu\sigma} (\Lambda^{-1})^\sigma{}_\beta = \eta_{\mu\beta} \Lambda^\mu{}_\nu \Rightarrow (\Lambda^{-1})^\alpha{}_\beta = \eta_{\mu\beta} \eta^{\nu\alpha} \Lambda^\mu{}_\nu. \quad (3.2.6)$$

From this the transformation of the co-variant components read

$$\bar{x}_\rho = \eta_{\mu\rho} \Lambda^\mu{}_\nu x^\nu = \eta_{\mu\rho} \eta^{\nu\sigma} \Lambda^\mu{}_\nu x_\sigma = (\Lambda^{-1})^\sigma{}_\rho x_\sigma. \quad (3.2.7)$$

Now by definition the transformation rule for the tensor-field components reads

$$\bar{T}^{\mu\nu\rho\cdots}(\bar{x}) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \cdots T^{\alpha\beta\gamma\cdots}(x). \quad (3.2.8)$$

From this we immediately find the transformation rule for the partial derivatives of these components to be

$$\bar{\partial}_\sigma \bar{T}^{\mu\nu\rho\cdots}(\bar{x}) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \cdots \frac{\partial x^\delta}{\partial \bar{x}^\sigma} \partial_\delta T^{\alpha\beta\gamma\cdots}(x) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \cdots (\Lambda^{-1})^\delta{}_\sigma \partial_\delta T^{\alpha\beta\gamma\cdots}(x), \quad (3.2.9)$$

which indeed means that the derivative ∂_ν transforms like the components of a new tensor field with one more covariant index.

This implies that the left-hand side of (3.2.5), the **four-divergence** of the vector field (j^μ), is a scalar field, and thus the continuity equation a Lorentz-invariant property, as it should be for a relativistic physical law.

To also express the components of the electromagnetic fields in a manifestly covariant way, next we aim at writing them in terms of appropriate four-tensor components. To that end it is easier to first exploit the homogeneous Maxwell equations (3.1.1) and (3.1.2) to introduce the **electromagnetic potentials**. From (3.1.1) we know that $\vec{B}$ has a **vector potential** $\vec{A}$ such that

$$\vec{B}(t,\vec{x}) = \vec{\nabla} \times \vec{A}(t,\vec{x}). \tag{3.2.10}$$

Using this in (3.1.2) we find

$$\vec{\nabla} \times \left(\vec{E} + \frac{1}{c}\partial_t \vec{A}\right) = 0, \tag{3.2.11}$$

from which we can conclude that the expression in the brackets is the gradient of a **scalar potential** Φ,

$$\vec{E} + \frac{1}{c}\partial_t \vec{A} = -\vec{\nabla}\Phi \Rightarrow \vec{E}(t,\vec{x}) = -\vec{\nabla}\Phi(t,\vec{x}) - \frac{1}{c}\partial_t \vec{A} = -\vec{\nabla}\Phi(t,\vec{x}) - \partial_0 \vec{A}. \tag{3.2.12}$$

This indicates that we can combine the potentials to a **four-vector potential**,

$$[A^\mu(x)] = \begin{pmatrix} \Phi(x) \\ \vec{A}(x) \end{pmatrix}, \tag{3.2.13}$$

and that then the electric and magnetic field components are found as the tensor components of a kind of four-dimensional curl, i.e., the antisymmetrized derivatives of A^μ, which leads to the **Faraday tensor**

$$F_{\alpha\beta}(x) = \partial_\alpha A_\beta(x) - \partial_\beta A_\alpha(x). \tag{3.2.14}$$

To find the relation of the tensor components $F_{\alpha\beta}$ with the usual electric and magnetic components $\vec{E}$ and $\vec{B}$ of the non-covariant formalism, we have to be careful with the signs. We als note that $\vec{E}$ and $\vec{B}$ are *not* the spatial components of a four-vector. In the following we write always lower indices for these vector components, i.e., $\vec{E} = (E_1, E_2, E_3)$ and $\vec{B} = (B_1, B_2, B_3)$. This is the case for $\vec{A}$, according to (3.2.13). Since $F_{\alpha\beta} = -F_{\beta\alpha}$ we just need to calculate the six components F_{a0} and F_{ab} with $a, b \in \{1,2,3\}$:

$$\begin{aligned} F_{0a} &= \partial_0 A_b - \partial_b A_0 = -\partial_0 A^b - \partial_b A^0 = -\frac{1}{c}\partial_t A^b - \partial_b \Phi = E_b, \\ F_{ab} &= \partial_a A_b - \partial_b A_a = -\partial_a A^b + \partial_b A^c = \epsilon_{bac}(\vec{\nabla} \times \vec{A})_c = -\epsilon_{abc} B_c. \end{aligned} \tag{3.2.15}$$

Writing the the Faraday-tensor components in terms of a matrix we get

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}. \tag{3.2.16}$$

Raising both indices we get

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix} \tag{3.2.17}$$

Now the Maxwell equations are first-order differential equations of the field components. Since we expect that the Maxwell equations are in fact a relativistic field theory, it should be possible to write these equations in terms of manifestly covariant differential manipulations. The most simple possibility is to contract the Faraday tensor with the four-dimensional Nabla operator, i.e.,

$$\partial_\mu F^{\mu\nu} = \begin{pmatrix} \vec{\nabla} \cdot \vec{E} \\ -\partial_0 \vec{E} + \vec{\nabla} \times \vec{B} \end{pmatrix} = \begin{pmatrix} \rho \\ \vec{j}/c \end{pmatrix} = \frac{1}{c}(j^\nu). \tag{3.2.18}$$

In the second step we have used the inhomogeneous Maxwell equations (3.1.3-3.1.4) and thus brought them already in a manifestly covariant form.

To also find the covariant form of the homogeneous Maxwell equations we note that the **Levi"=Civita" symbol**, $\epsilon^{\alpha\beta\gamma\delta}$, which is defined to be totally antisymmetric under exchange of any pair of its indices and $\epsilon^{0123} = 1$, are invariant under proper (and thus particularly under proper isochronous) Lorentz transformations, because

$$\bar{\epsilon}^{\mu\nu\rho\sigma} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \Lambda^\nu{}_\delta \epsilon^{\alpha\beta\gamma\delta} = \epsilon^{\mu\nu\rho\sigma} \det(\hat{\Lambda}) = \epsilon^{\mu\nu\rho\sigma}. \tag{3.2.19}$$

In the same way we find for the covariant components

$$\epsilon_{\mu\nu\rho\sigma} = \epsilon^{\mu\nu\rho\sigma} \det(\hat{\eta}) = -\epsilon^{\mu\nu\rho\sigma}. \tag{3.2.20}$$

Finally we note that

$$\epsilon_{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\rho\sigma} = -\delta^{\rho\mu\nu\sigma}_{\alpha\beta\gamma\delta}, \tag{3.2.21}$$

where

$$\delta^{\mu\nu\rho\sigma}_{\alpha\beta\gamma\delta} = \det \begin{pmatrix} \delta^\mu_\alpha & \delta^\mu_\beta & \delta^\mu_\gamma & \delta^\mu_\delta \\ \delta^\nu_\alpha & \delta^\nu_\beta & \delta^\nu_\gamma & \delta^\nu_\delta \\ \delta^\rho_\alpha & \delta^\rho_\beta & \delta^\rho_\gamma & \delta^\rho_\delta \\ \delta^\sigma_\alpha & \delta^\sigma_\beta & \delta^\sigma_\gamma & \delta^\sigma_\delta \end{pmatrix}. \tag{3.2.22}$$

This means that $\delta^{\mu\nu\rho\sigma}_{\alpha\beta\gamma\delta}$ is antisymmstric under exchange of any lower or any upper index pair, and $\delta^{0123}_{0123} = 1$. It is also clear that this symbol defines invariant tensor components under any linear transformation, particularly also under arbitrary Lorentz transformations.

With help of (3.2.21) and the properties of the determinant (3.2.22) it is easy to evaluate all possible contractions of two Levi"=Civita tensors,

$$\epsilon_{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\rho\delta} = -\det \begin{pmatrix} \delta^\mu_\alpha & \delta^\mu_\beta & \delta^\mu_\gamma \\ \delta^\nu_\alpha & \delta^\nu_\beta & \delta^\nu_\gamma \\ \delta^\rho_\alpha & \delta^\rho_\beta & \delta^\rho_\gamma \end{pmatrix} = -\delta^{\mu\nu\rho}_{\alpha\beta\gamma}, \tag{3.2.23}$$

$$\epsilon_{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\gamma\delta} = -2\det \begin{pmatrix} \delta^\mu_\alpha & \delta^\mu_\beta \\ \delta^\nu_\alpha & \delta^\nu_\beta \end{pmatrix} = -2\delta^{\mu\nu}_{\alpha\beta}, \tag{3.2.24}$$

$$\epsilon_{\alpha\beta\gamma\delta}\epsilon^{\mu\beta\gamma\delta} = -6\delta^\mu_\alpha. \tag{3.2.25}$$

The Levi-ivita tensor allows us to define for each completely antisymmetric tensor of ranks 2, 3, or 4 another antisymmetric tensor by contraction with the Levi-Civita tensor, e.g., for the Faraday tensor

$$^\dagger F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}. \tag{3.2.26}$$

The tensor defined by the components ${}^{\dagger}F^{\mu\nu}$ are called the **Hodge dual** of the Faraday tensor. Using (3.2.23-3.2.25) we find that the operation of forming the Hodge dual of an antisymmetric tensor is a one-to-one mapping in the space of the antisymmetric tensors. The inverse is given, up to a sign, by the Hodge-dual operator itself. E.g., for the Faraday tensor we get

$$\begin{aligned}
{}^{\dagger}({}^{\dagger}F)_{\alpha\beta} &= \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \\
&= +\frac{1}{4}\epsilon_{\alpha\beta\mu\nu}\epsilon^{\rho\sigma\mu\nu}F_{\rho\sigma} \\
&= -\frac{1}{2}(\delta^{\rho}_{\alpha}\delta^{\sigma}_{\beta} - \delta^{\sigma}_{\alpha}\delta^{\rho}_{\beta})F_{\rho\sigma} \\
&= -\delta^{\rho}_{\alpha}\delta^{\sigma}_{\beta}F_{\rho\sigma} = -F_{\alpha\beta}.
\end{aligned} \tag{3.2.27}$$

To see the relation of ${}^{\dagger}F^{\alpha\beta}$ to $\vec{E}$ and $\vec{B}$ we first note that $\epsilon^{0\mu\nu\rho}$ is only non-zero for $\mu, \nu, \rho \in \{1, 2, 3\}$. This means that

$${}^{\dagger}F^{0\nu} = \frac{1}{2}\epsilon^{0\nu\alpha\beta}F_{\alpha\beta} = -B_{\nu}, \quad \nu \in \{1, 2, 3\} \tag{3.2.28}$$

In the following we define that latin indices always run over the set $\{1, 2, 3\}$ and greek ones over $\{0, 1, 2, 3\}$. For $m, n \in \{1, 2, 3\}$ we then have

$${}^{\dagger}F^{mn} = \frac{1}{2}\epsilon^{mn\alpha\beta}F_{\alpha\beta} = \epsilon^{mn0b}F_{0b} = \epsilon^{mnb}E_{b}. \tag{3.2.29}$$

In matrix notation this reads

$${}^{\dagger}F^{\mu\nu} = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{pmatrix}. \tag{3.2.30}$$

We can express this also in a different way: We can identify the 6 components of $F^{\mu\nu}$ with the ordered pair of three-vectors $(\vec{E}, \vec{B})$. Then ${}^{\dagger}F^{\mu\nu}$ is identified with $(\vec{B}, -\vec{E})$.
Thus we get for the contraction with the four-Nabla operator from (3.2.18)

$$\partial_{\mu}{}^{\dagger}F^{\mu\nu} = \begin{pmatrix} \vec{\nabla} \cdot \vec{B} \\ -\partial_0\vec{B} - \vec{\nabla} \times \vec{E} \end{pmatrix} = 0, \tag{3.2.31}$$

where in the last step we have used the homogeneous Maxwell equations (3.1.1-3.1.2). So the Maxwell equations are equivalent to the manifestly covariant relativistic field equations for the Faraday tensor and its dual (3.2.18) and (3.2.31):

$$\partial_{\mu}F^{\mu\nu} = \frac{1}{c}j^{\nu}, \quad \partial_{\mu}{}^{\dagger}F_{\mu\nu} = 0. \tag{3.2.32}$$

We just note that we can establish a four-dimensional definition of an electric-field four-vector and a magnetic-field four-vector. It depends on the four-velocity $u = (u^{\mu})$ of an inertial observer who measures these fields,

$$E^{\mu} = F^{\mu\nu}u_{\nu} = \begin{pmatrix} \vec{u} \cdot \vec{E} \\ u^0\vec{E} + \vec{u} \times \vec{B} \end{pmatrix}, \quad B^{\mu} = {}^{\dagger}F^{\mu\nu}u_{\nu} = \begin{pmatrix} \vec{u} \cdot \vec{B} \\ u^0\vec{B} - \vec{u} \times \vec{E} \end{pmatrix}. \tag{3.2.33}$$

Indeed, for the observer at rest in the reference frame at hand, we have $u = (1,0,0,0)^{\mathrm{T}}$ and for this observer we have $(E^\mu) = (0, \vec{E})$ and $(B^\mu) = (0, \vec{B})$.
(3.2.33) gives the components of the four-vectors of the electric and magnetic field with respect to the reference frame, where this observer moves with four-velocity u. In his own rest frame he finds the field components by applyting the appropriate boost matrix

$$\hat{\Lambda} = \begin{pmatrix} u^0 & -\vec{u}^{\mathrm{T}} \\ -\vec{u} & \mathbb{1}_3 + (u^0 - 1)\vec{u}\vec{u}^{\mathrm{T}}/\vec{u}^2 \end{pmatrix} \tag{3.2.34}$$

to the vectors given in (3.2.33),

$$(\bar{E}^\mu) = (\Lambda^\mu{}_\nu E^\nu) = \begin{pmatrix} 0 \\ \bar{\vec{E}} \end{pmatrix}, \quad (\bar{B}^\mu) = (\Lambda^\mu{}_\nu B^\nu) = \begin{pmatrix} 0 \\ \bar{\vec{B}} \end{pmatrix}, \tag{3.2.35}$$

where after some simple algebra, using $u^0 = \gamma = 1/\sqrt{1-\beta^2}$ and $\underline{u} = \gamma\vec{\beta}$ with $\vec{\beta} = \vec{v}/c$,

$$\bar{\vec{E}} = \vec{E}_{||} + \gamma(\vec{E}_\perp + \vec{\beta} \times \vec{B}) = \gamma(\vec{E} + \vec{\beta} \times \vec{B}) - \frac{\gamma^2}{1+\gamma}\vec{\beta}(\vec{\beta} \cdot \vec{E}), \tag{3.2.36}$$

$$\bar{\vec{B}} = \vec{B}_{||} + \gamma(\vec{B}_\perp - \vec{\beta} \times \vec{E}) = \gamma(\vec{B} - \vec{\beta} \times \vec{E}) - \frac{\gamma^2}{1+\gamma}\vec{\beta}(\vec{\beta} \cdot \vec{B}). \tag{3.2.37}$$

Here we have used the projection of the 3D-field components to a part longitudinal and perpendicular with respect to $\vec{\beta}$,

$$\vec{V}_{||} = \frac{\vec{\beta} \cdot \vec{V}}{\beta^2}\vec{\beta}, \quad \vec{V}_\perp = \vec{V} - \vec{V}_{||}, \tag{3.2.38}$$

and

$$\vec{E}_{||} = \gamma\vec{E} + \frac{1-\gamma}{\vec{\beta}^2}\vec{\beta}(\vec{\beta} \cdot \vec{E}) = \gamma\vec{E} + \frac{1-\gamma}{1-1/\gamma^2} = \gamma\vec{E} - \frac{\gamma^2}{1+\gamma}\vec{\beta}(\vec{\beta} \cdot \vec{E}), \tag{3.2.39}$$

and the analogous equation for $\vec{B}_{||}$.
From (3.2.33) with some algebra one can derive that

$$F_{\mu\nu} = E_\mu u_\nu - E_\nu u_\mu - \epsilon_{\mu\nu\rho\sigma} B^\rho u^\sigma. \tag{3.2.40}$$

Now the homogeneous Maxwell equations (3.2.18) read

$$\partial_\mu F^{\mu\nu} = \partial_\mu E^\mu u^\nu - u^\mu \partial_\mu E^\nu - \epsilon^{\mu\nu\rho\sigma}(\partial_\mu B_\rho u_\sigma) = \frac{1}{c} j^\nu. \tag{3.2.41}$$

The dual tensor is

$${}^\dagger F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} = \epsilon^{\mu\nu\rho\sigma} E_\rho u_\sigma + B^\mu u^\nu - B^\nu u^\mu, \tag{3.2.42}$$

and the inhomogeneous Maxwell equations (3.2.31) become

$$\partial_\mu {}^\dagger F^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\partial_\mu E_\rho u_\sigma + \partial_\mu B^\mu u^\nu - u^\mu \partial_\mu B^\nu. \tag{3.2.43}$$

With the tensors $F^{\mu\nu}$ and ${}^\dagger F^{\mu\nu}$ we can build two independen **invariants** of the fields,

$$F_{\mu\nu} F^{\mu\nu} = -{}^\dagger F_{\mu\nu} {}^\dagger F^{\mu\nu} = 2(\vec{E}^2 - \vec{B}^2), \quad F_{\mu\nu} {}^\dagger F^{\mu\nu} = 4\vec{E} \cdot \vec{B}. \tag{3.2.44}$$

This leads to the idea to build a complex three-vector, the **Riemann-Silberstein vector**,

$$\vec{F} = \vec{E} + \mathrm{i}\vec{B}. \tag{3.2.45}$$

Keeping the usual scalar product from Euclidean $\mathbb{R}^3$, this leads to

$$\vec{F}^2 = \vec{E}^2 - \vec{B}^2 - 2\mathrm{i}\vec{E}\cdot\vec{B}. \tag{3.2.46}$$

This shows that the invariants of the electromagnetic field (3.2.44) under Lorentz transformation occur as real and imaginary parts of $\vec{F}^2$. This indicates that the transformation of the field components under proper orthochronous Lorentz transformations (3.2.36-3.2.37) can be expressed as a complexified rotation matrices $\hat{O} \in \mathrm{SO}(3,\mathbb{C})$, i.e., matrices $\hat{O} \in \mathbb{C}^{3\times 3}$ with $\det \hat{O} = 1$ and $\hat{O}\hat{O}^\dagger = \hat{O}^\dagger\hat{O} = \mathbb{1}$. Of course, the rotations are given by the subgroup $SO(3)$, i.e., the usual rotation matrix. Indeed both the electric and magnetic field components transform as usual three-vector fields under rotations, and thus this holds also for $\vec{F}$. Now from (3.2.36-3.2.37) we find

$$\vec{\tilde{F}} = \hat{\beta}(\hat{\beta}\cdot\vec{F}) - \mathrm{i}\beta\gamma\hat{\beta}\times\vec{F} + \gamma[\hat{\beta}\times(\vec{F}\times\vec{\beta})] = \hat{D}_{\hat{\beta}}(\mathrm{i}\eta)\vec{F}, \tag{3.2.47}$$

where $\hat{\beta} = \vec{\beta}/\beta$ and $\eta = \operatorname{artanh}\beta$. For real arguments $\hat{D}_{\vec{n}}(\alpha)$ denotes transformation of the three-vector components under usual rotations of the spatial basis vectors with rotation axis $\vec{n}$ and rotation angle α, cf. the spatial part of the four-vector in (1.7.23). Thus, the Riemann-Silberstein vector indeed transforms under proper orthochronous Lorentz transformations according to the complexified group $\mathrm{SO}(3,\mathbb{C})$. The usual rotations are represented by the subgroup $\mathrm{SO}(3,\mathbb{C})$ and rotation-free boosts by those with imaginary rotation angles $\mathrm{i}\eta$, where η is the **rapidity** of the boost, introduced in (1.7.28).

To derive the Maxwell equations In terms of the Riemann-Silberstein vector we can simply use the three-dimensional form (3.1.1-3.1.4). Taking the divergence reads

$$\vec{\nabla}\cdot\vec{F} = \rho \tag{3.2.48}$$

and the curl

$$\vec{\nabla}\times\vec{F} = \vec{\nabla}\times(\vec{E}+\mathrm{i}\vec{B}) = -\frac{1}{c}\partial_t\vec{B} + \frac{\mathrm{i}}{c}\partial_t\vec{E} + \frac{\mathrm{i}}{c}\vec{j} = \frac{\mathrm{i}}{c}(\partial_t\vec{F} + \vec{j}) \tag{3.2.49}$$

or

$$\vec{\nabla}\times\vec{F} - \frac{\mathrm{i}}{c}\partial_t\vec{F} = \frac{\mathrm{i}}{c}\vec{j}. \tag{3.2.50}$$

3.2.1 The Doppler effect for light

As an application of the above derived transformation laws of the electromagnetic field components we consider, how the frequency and intensity of a plane wave changes when its source moves with constant velocity $\vec{v} = \beta c\vec{e}_1$ in the reference frame Σ. Then the source is at rest in the reference frame Σ'.

We assume that the wave vector $\vec{k} = k(\cos\alpha, \sin\alpha, 0)$. Then also in the frame Σ' we have $\vec{k}' = k'(\cos\alpha', \sin\alpha', 0)$. The general plane-wave solutions of the Maxwell equations can be written as

$$\vec{E}(t,\vec{x}) = (E_{01}\vec{\epsilon}_1 + E_{02}\vec{\epsilon}_2)\cos(k\cdot x), \tag{3.2.51}$$

where

$$\vec{\epsilon}_1 = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.2.52}$$

The same equations

3.3 Gauge invariance and retarded potentials

As we have already seen in the previous Sect. the homogeneous Maxwell equations guarantee the existence of the four-vector potential A_μ with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{3.3.1}$$

and for any scalar field χ

$$A'_\mu = A_\mu + \partial_\mu \chi \tag{3.3.2}$$

we get the same Faraday tensor (3.3.1) since $\partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi = 0$. This is called **gauge invariance** of the electromagnetic potential.

The Plugging (3.3.1) into the inhomogeneous Maxwell equation (3.2.18) we get the equation of motion for the four-potential,

$$\partial^\mu F_{\mu\nu} = \partial^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) = \Box A_\nu - \partial_\nu \partial_\mu A^\mu = \frac{1}{c} j^\nu. \tag{3.3.3}$$

Here we have introduced the **D'Alembert operator**,

$$\Box = \partial_\mu \partial^\mu = \frac{1}{c^2} \partial_t^2 - \vec{\nabla}^2 = \frac{1}{c^2} \partial_t^2 - \Delta. \tag{3.3.4}$$

Now due to the gauge invariance the electromagnetic potential is determined only up to the gradient of a scalar field. This means that we can apply one additional constraint on A_μ, which of course has to be consistent with the equation (3.3.3).

Obviously the equations (3.3.3) become separated for the four components of A^μ if one chooses the **Lorenz-gauge condition**,

$$\partial_\mu A^\mu = 0, \tag{3.3.5}$$

which has also the advantage of being a covariant equation, i.e., if one evaluates the four-potential in one inertial frame of reference subject to fulfill (3.3.5) this condition is also fulfilled in any other inertial frame.

Thus with the Lorenz-gauge condition (3.3.5) the equation of motion for the vector potential (3.3.3) simplifies to

$$\Box A^\nu = \frac{1}{c} j^\nu. \tag{3.3.6}$$

This equation is consistent with the gauge condition, because

$$0 = \Box \partial_\nu A^\nu = \frac{1}{c} \partial_\nu j^\nu, \tag{3.3.7}$$

and this is fulfilled according to the conservation of electric charge, which leads to the continuity equation (3.2.4). This condition of charge conservation has to be fufilled thus not only as an integrability condition for the Maxwell equations but also to fulfill gauge invariance.

Now we investigate, how to solve (3.3.3) with the sources, $j^\mu(t, \vec{x})$ considered to be given. Then this is a set of wave equations, which are thanks to the choice of the Lorenz gauge uncoupled. It's a (hyperbolic) linear partial differential equation of second order in time and thus its general solution can be written as the sum of a particular solution of the inhomogeneous equation and the general solution of the homogeneous equation. To make the solution unique one needs initial conditions, i.e., the specification of $A^\mu(t=0,\vec{x}) = A_0^\mu(\vec{x})$ and $\partial_t A^\mu(t=0,\vec{x}) = \partial_t A^\mu(t=0,\vec{x})$ or, equivalently in terms of the physical fields, $F^{\mu\nu}(t=0,\vec{x}) = F_0^{\mu\nu}(\vec{x})$.

We start to look for a solution of the inhomogeneous wave equation of a scalar field. This is then applicable to the four separated equations for the components of A^μ. So we try to solve

$$\Box\Phi(x) = J(x), \tag{3.3.8}$$

with Φ and J scalar fields. The usual way to solve such an inhomogeneous linear equation is to look for a **Green's function** of the corresponding differential operator on the left-hand side, i.e.,

$$\Box_x D(x,x') = \delta^{(4)}(x-x'). \tag{3.3.9}$$

Having found such a solution, a particular solution of the inhomogeneous equation (3.3.8) is given by

$$\Phi(x) = \int_{\mathbb{R}^4} \mathrm{d}^4 x' D(x,x') J(x'). \tag{3.3.10}$$

This already provides a hint, how to specify the Green's function needed for our purposes further: The integral over $x'^0 = ct'$ integrates over the entire "history" of the source. Since we observe the field $\Phi(x)$ at a given time, $t = x^0/c$, due to the **causality principle**, should be determined only by the "past" of this history, and the future states of $J(x')$, i.e., at times $t' > t$ should not contribute. One should note that at this point we bring in the notion of a **time direction**, i.e., we distinguish past from future by envoking the argument with the causality princple, defining the **causal time arrow**. Thus, we look for the socalled **retarded Green's function**, subject to the constraint

$$D(x,x') \equiv D(x-x') = \Theta(x^0 - x'^0) d(x-x'). \tag{3.3.11}$$

In this ansatz we have already taken into account that (3.3.9) is invariant under space-time translations, i.e., $D(x,x')$ can only depend on the difference $x - x'$ of its arguments. Thus setting $y = x - x'$ we have to solve

$$\Box D(y) = \delta^{(4)}(y). \tag{3.3.12}$$

To this end we express D as a Fourier transform wrt. the spatial components,

$$D(y) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 k}{(2\pi)^3} \exp(\mathrm{i}\vec{k}\cdot\vec{x}) \tilde{D}(y^0,\vec{k}). \tag{3.3.13}$$

Since

$$\delta^{(4)}(y) = \delta(y^0) \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 k}{(2\pi)^3} \exp(\mathrm{i}\vec{k}\cdot\vec{x}), \tag{3.3.14}$$

(3.3.12) becomes

$$\partial_0^2 \tilde{D}(y^0,\vec{k}) + \vec{k}^2 \tilde{D}(y^0,\vec{k}) = \delta(y^0). \tag{3.3.15}$$

Plugging in (3.3.11) we find $\tilde{D}(y^0,\vec{k}) = \Theta(y^0)\tilde{d}(y^0,\vec{k})$ and then using $\partial_0\Theta(y^0) = \delta(y_0)$ this leads with (3.3.15) to

$$\begin{aligned} &\partial_0 \delta(y^0)\tilde{d}(y^0,\vec{k}) + 2\delta(y^0)\partial_0\tilde{d}(y^0,k) + \Theta(y^0)[\partial_0^2\tilde{d}(y^0,\vec{k}) + k^2\tilde{d}(y^0,\vec{k})] \\ &\quad = \Theta(y^0)[\partial_0^2\tilde{d}(y^0,\vec{k}) + k^2\tilde{d}(y^0,\vec{k})] + \delta(y^0)\partial_0\tilde{d}(y^0,\vec{k}) = \delta(y^0). \end{aligned} \tag{3.3.16}$$

The solution of this equation is obviously satisfied, if $\tilde{d}(0^+,\vec{k})=0$ and $\partial_0\tilde{d}(0^+,\vec{k})=1$ and

$$\partial_0^2\tilde{d}(y^0,\vec{k})+k^2\tilde{d}(y^0,\vec{k})=0 \Rightarrow \tilde{d}(y^0,\vec{k})=\frac{1}{K}\sin(y^0K), \tag{3.3.17}$$

where $K=|\vec{k}|$. Plugging this into (3.3.14) and using spherical coordinates with $\vec{y}=r\vec{e}_3$, which is no restriction of generality because of the invariance of the solution (3.3.17) under arbitrary rotations of $\vec{k}$,

$$\begin{aligned}
d(y) &= \int_0^\infty \mathrm{d}K \int_0^\pi \mathrm{d}\vartheta \int_0^{2\pi} \mathrm{d}\varphi \frac{1}{(2\pi)^3} K\sin\vartheta\sin(Ky^0)\exp(\mathrm{i}Kr\cos\vartheta) \\
&= \frac{1}{(2\pi)^2}\int_0^\infty \mathrm{d}K \int_{-1}^1 \mathrm{d}u K\sin(Ky^0)\exp(\mathrm{i}Kru) \\
&= \frac{1}{(2\pi)^2 r}\int_0^\infty \mathrm{d}K 2\sin(Ky^0)\sin(Kr) \\
&= \frac{1}{4\pi^2 r}\int_{\mathbb{R}} \mathrm{d}K\sin(Ky^0)\sin(Kr) \\
&= \frac{1}{16\pi^2}\int_{\mathbb{R}} [\exp(\mathrm{i}Ky^0)-\exp(-\mathrm{i}Ky^0)][\exp(\mathrm{i}Kr)-\exp(-\mathrm{i}Kr)] \\
&= \frac{1}{4\pi r}[\delta(y^0-r)+\delta(y^0+r)].
\end{aligned} \tag{3.3.18}$$

Finally, the retarded propgator reads

$$D(y)=\frac{1}{4\pi r}\Theta(y^0)\delta(y^0-r)=\frac{1}{4\pi r}\delta(y^0-r). \tag{3.3.19}$$

In the last step we have used that $r>0$ anyway. To show that this is a scalar under orthochronous proper Lorentz transformations, we rewrite the δ distribution as

$$D(y)=\frac{1}{2\pi}\Theta(y^0)\delta[(y^0)^2-r^2]=\frac{1}{2\pi}\Theta(y^0)\delta(y^2), \tag{3.3.20}$$

which is a manifestly covariant expression. For *orthochronous* Lorentz transformations that also applies to the factor $\Theta(y^0)$. Using the four-velocity vector of an observer at rest in our computational frame, $(u^\mu)=(1,0,0,0)$, we can write $y^0=u\cdot y$ and $r^2=(u\cdot u)^2-u^2$ and $\Theta(y^0)=(1+\operatorname{sign}y^0)/2=[1+(u\cdot y)/|(u\cdot y)|]/2$, and then we find the covariant form under *all* $\mathrm{O}(1,3)$ transformations,

$$D(y)=\frac{1}{4\pi}\left(1+\frac{u\cdot y}{|u\cdot y|}\right)\delta(y^2).$$

It is of course simpler to use the form (3.3.19). Using this retarded Green's function of the d'Alembert operator to the solution of (3.3.3) we find the **retarded potentials**,

$$A^\mu(x)=\int_{\mathbb{R}^4}\mathrm{d}^4x'\frac{\delta(x^0-x'^0-|\vec{x}-\vec{x}'|)}{4\pi c|\vec{x}-\vec{x}'|}j^\mu(x')=\int_{\mathbb{R}^3}\mathrm{d}^3x'\frac{j^\mu(x^0-|\vec{x}-\vec{x}'|,\vec{x}')}{4\pi c|\vec{x}-\vec{x}'|}. \tag{3.3.21}$$

Rewriting this in terms of $t=x^0/c$ we get

$$A^\mu(x)=\int_{\mathbb{R}^3}\mathrm{d}^3x'\frac{j^\mu(t-|\vec{x}-\vec{x}'|/c,\vec{x}')}{4\pi c|\vec{x}-\vec{x}'|}. \tag{3.3.22}$$

The latter equations makes the physical interpretation of the **retardation effect** simple: The contribution of the sources at position $\vec{x}'$ to the fields observed at position $\vec{x}$ originates from the state of the sources at the **retarded time**, $\tau_{\text{ret}} = t - |\vec{x} - \vec{x}'|/c$, which is earlier by the running time of light from $\vec{x}'$ to $\vec{x}$. This reflects the **relativistic causality structure**, limiting the speed of signal propagation to the speed of light in a vacuum.

To see that (3.3.22) is a valid solution for potentials, leading to a Faraday tensor that solves all Maxwell equations, we must check that the Lorenz-gauge condition (3.3.5), because only then the solution of (3.3.6) leads to a valid solution of the Maxwell equations. To this end, we write the solution in unintegrated form,

$$A^{\mu}(x) = \int_{\mathbb{R}^4} \mathrm{d}^4 x' D(x-x') \frac{j^{\mu}(x')}{c}, \tag{3.3.23}$$

and then, via an integration by parts assuming that the boundary terms do not play any role, because the current vanishes sufficiently fast at infinity, we indeed find

$$\begin{aligned} \partial_{\mu} A^{\mu}(x) &= \int_{\mathbb{R}^4} \mathrm{d}^4 x' \frac{\partial}{\partial x^{\mu}} D(x-x') \frac{j^{\mu}(x')}{c} \\ &= -\int_{\mathbb{R}^4} \mathrm{d}^4 x' \frac{\partial}{\partial x'^{\mu}} D(x-x') \frac{j^{\mu}(x')}{c} \\ &= \int_{\mathbb{R}^4} \mathrm{d}^4 x' D(x-x') \frac{\partial}{\partial x'^{\mu}} \frac{j^{\mu}(x')}{c} = 0, \end{aligned} \tag{3.3.24}$$

where in the last step we have made use of the continuity equation (3.2.4) again. Thus the retarded solution (3.3.22) indeed fulfills the Lorenz"=gauge condition (3.3.5) and the wave equations (3.3.6) and thus provides a valid electromagnetic potential.

3.4 The retarded fields

With (3.3.22) we have a valid solution of the Maxwell equations for the four-potential, and to find the corresponding field components $F_{\mu\nu}$ or, equivalently, $\vec{E}$ and $\vec{B}$ we could just take the derivatives, $F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu}$. It is, however, simpler to use the non-covariant notation,

$$\vec{E}(t,\vec{x}) = -\vec{\nabla} A^0(t,\vec{x}) - \frac{1}{c} \partial_t \vec{A}(t,\vec{x}), \quad \vec{B}(t,\vec{x}) = \vec{\nabla} \times \vec{A}(t,\vec{x}). \tag{3.4.1}$$

To evaluate the derivatives, it is generally more convenient to work with the four-dimensional integral (3.3.24), taking the derivatives and only afterwards perform the x'^0 (or t' integral) to get rid of the δ distribution:

$$A^{\mu}(t,\vec{x}) = \int_{\mathbb{R}} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \frac{\delta(t-t'-|\vec{x}-\vec{x}'|/c)}{4\pi c |\vec{x}-\vec{x}'|} j^{\mu}(t',\vec{x}'). \tag{3.4.2}$$

Now we evaluate the derivatives we need in (3.4.1). Introducing the abbreviation $\vec{R} = \vec{r} - \vec{r}'$ and $R = |\vec{r} - \vec{r}'| = |\vec{R}|$, we find

$$\begin{aligned}\partial_t \vec{A}(t,\vec{x}) &= \int_{\mathbb{R}} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \frac{\partial_t \delta(t-t'-R/c)}{4\pi c R} \vec{j}(t',\vec{x}') \\ &= -\int_{\mathbb{R}} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \frac{\partial_t' \delta(t-t'-R/c)}{4\pi c R} \vec{j}(t',\vec{x}') \\ &= \int_{\mathbb{R}} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \frac{\delta(t-t'-R/c)}{4\pi c R} \partial_{t'} \vec{j}(t',\vec{x}') \\ &= \int_{\mathbb{R}^3} \mathrm{d}^3 x' \frac{1}{4\pi c R} \left[\partial_t' \vec{j}(t',\vec{x}')\right]_{t'=t-R/c}.\end{aligned} \tag{3.4.3}$$

Further we have

$$\begin{aligned}\vec{\nabla} A^0 &= \int_{\mathbb{R}} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \rho(t',\vec{x}') \vec{\nabla} \frac{\delta(t-t'-R/c)}{4\pi R} \\ &= \int_{\mathbb{R}^3} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \rho(t',\vec{x}') \left[-\partial_t \delta(t-t'-R/c) \frac{\vec{R}}{4\pi c R^2} - \delta(t-t'-R/c) \frac{\vec{R}}{4\pi R^3}\right] \\ &= \int_{\mathbb{R}^3} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \rho(t',\vec{x}') \left[\partial_t' \delta(t-t'-R/c) \frac{\vec{R}}{4\pi c R^2} - \delta(t-t'-R/c) \frac{\vec{R}}{4\pi R^3}\right] \\ &= -\int_{\mathbb{R}^3} \mathrm{d}^3 x' \left[\frac{\partial_t' \rho(t',\vec{x}') \vec{R}}{4\pi c R^2} + \frac{\rho(t',\vec{x}') \vec{R}}{4\pi R^3}\right]_{t'=t-R/c}\end{aligned} \tag{3.4.4}$$

and

$$\begin{aligned}\vec{\nabla} \times \vec{A}(t,\vec{x}) &= \int_{\mathbb{R}} \mathrm{d}t' \int_{\mathbb{R}^3} \mathrm{d}^3 x' \left[\partial_t' \delta(t-t'-R/c) \frac{\vec{R}}{4\pi c^2 R^2} - \delta(t-t'-R/c) \frac{\vec{R}}{4\pi c R^3}\right] \vec{j}(t',\vec{x}') \\ &= + \int_{\mathbb{R}^3} \mathrm{d}^3 x' \left[\frac{\partial_t' \vec{j}(t',\vec{x}') \times \vec{R}}{4\pi c^2 R^2} + \frac{\vec{j}(t',\vec{x}') \times \vec{R}}{4\pi c R^3}\right]_{t'=t-R/c}.\end{aligned} \tag{3.4.5}$$

Plugging (3.4.3-3.4.5) into (3.4.1) we find the **Jefimenko equations**,

$$\vec{E}(t,\vec{x}) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \mathrm{d}^3 x' \left[\frac{\rho(t',\vec{x}')\vec{R}}{R^3} + \frac{\partial_t' \rho(t',\vec{x}') \vec{R}}{c R^2} - \frac{\partial_t' \vec{j}(t',\vec{x}')}{c^2 R}\right]_{t'=t-R/c}, \tag{3.4.6}$$

$$\vec{B}(t,\vec{x}) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \mathrm{d}^3 x' \left[\frac{\vec{j}(t',\vec{x}') \times \vec{R}}{c R^3} + \frac{\partial_t' \vec{j}(t',\vec{x}') \times \vec{R}}{c^2 R^2}\right]_{t'=t-R/c}. \tag{3.4.7}$$

3.5 The Liénard-Wiechert potential

Now we analyze the special case of a point particle. As we shall see there is no principle problem in answering the question, which electromagnetic field is created by an arbitrarily moving massive particle. So let $y^\mu(\tau)$ describe the world line of the particle as a function of its proper time τ. This world line is necessarily time-like everywhere, i.e., the four-velocity $u^\mu = \dot{y}^\mu/c$ everywhere fulfills the constraint $u_\mu u^\mu = 1$[1].

First of all we need a description of the charge and the current density. For a charged fluid in continuum description with the fluid-flow field $\vec{v}(t,\vec{x})$ we have $\vec{j}(t,\vec{x}) = \rho(t,\vec{x})\vec{v}(t,\vec{x})$, where ρ is the current density as measured in the computational frame. This can be written in manifestly covariant form as follows: The charge-current four-vector $j^\mu = (c\rho,\vec{j}) = (c\rho,\rho c\vec{\beta})$ is a time-like vector field, and thus $j_\mu j^\mu = \rho^2 c^2(1-\vec{\beta}^2) > 0$ is a scalar. Thus defining $\rho^* = \operatorname{sign}\rho\sqrt{j_\mu j^\mu}$, we have

$$(j^\mu) = c\rho^*\gamma \begin{pmatrix} 1 \\ \vec{\beta} \end{pmatrix} = c\rho^* u^\mu. \tag{3.5.1}$$

From this we see that ρ^* is the charge density as seen in the instantaneous rest frame of the fluid cell. A more convenient covariant definition of this **proper charge density** is

$$\rho^* = u_\mu j^\mu. \tag{3.5.2}$$

Now for a point charge q we have

$$j^\mu(x) = q\frac{\mathrm{d}}{\mathrm{d}t} y^\mu(t)\delta^{(3)}[\vec{x}-\vec{y}(t)] = qc\int_{\mathbb{R}} \mathrm{d}t' \frac{\mathrm{d}}{\mathrm{d}t'} y^\mu(t')\delta^{(4)}[x-y(t')]. \tag{3.5.3}$$

We can rewrite this in manifestly covariant form as

$$j^\mu(x) = qc\int_{\mathbb{R}} \mathrm{d}\tau \frac{\mathrm{d}}{\mathrm{d}\tau} y^\mu(\tau)\delta^{(4)}[x-y(\tau)] = qc^2\int_{\mathbb{R}} \mathrm{d}\tau\, u^\mu(\tau)\delta^{(4)}[x-y(\tau)]. \tag{3.5.4}$$

Now we check the continuity equation. To that end we note that

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\delta^{(4)}[x-y(\tau)] = -\frac{\mathrm{d}}{\mathrm{d}\tau} y^\mu(\tau)\partial_\mu\delta^{(4)}[x-y(\tau)] = -c\,u^\mu(\tau)\partial_\mu\delta^{(4)}[x-y(\tau)]. \tag{3.5.5}$$

From this we immediately find through partial integration,

$$\partial_\mu j^\mu(x) = qc^2\int_{\mathbb{R}} \mathrm{d}\tau\, u^\mu(\tau)\partial_\mu\delta^{(4)}[x-y(\tau)] = -qc\int_{\mathbb{R}} \mathrm{d}\tau \frac{\mathrm{d}}{\mathrm{d}\tau}\delta^{(4)}[x-y(\tau)] = 0. \tag{3.5.6}$$

One can of course, also verify the validity of the continuity equation, using the (1+3)-form in (3.5.3), according to which

$$\begin{aligned}
&\vec{\nabla}\cdot\vec{j} = q\frac{\mathrm{d}}{\mathrm{d}t}\vec{y}(t)\cdot\vec{\nabla}\delta^{(3)}[\vec{x}-\vec{y}(t)], \\
&\partial_0 j^0 = q\frac{\mathrm{d}}{\mathrm{d}t}\delta^{(3)}[\vec{x}-\vec{y}(t)] = -q\frac{\mathrm{d}}{\mathrm{d}t}\vec{y}(t)\cdot\vec{\nabla}\delta^{(3)}[\vec{x}-\vec{y}(t)] \\
\Rightarrow\ &\partial_\mu j^\mu = 0.
\end{aligned} \tag{3.5.7}$$

[1] As we shall see, some problems occur, if $|\vec{\beta}| = |\vec{u}|/u^0 = |\mathrm{d}\vec{y}/\mathrm{d}t|/c$ approaches asymptotically 1 for $\tau \to \pm\infty$, as for, e.g., hyperbolic motion. In this cases one has to carefully consider regularized problems and the corresponding limits towards the problem with the $|\vec{\beta}| \to 1$-asymptotic behavior.

Now the four-potential is easily evaluated using (3.3.23) and the explicit expressions (3.3.19) or (3.3.20) for the retarded Green's function of the D'Alembert operator. Using the covariant form for the four-current density (3.5.4) we integrate out the δ distribution, using $\delta[f(\tau)] = \sum_{\tau_0} \delta(\tau-\tau_0)/|\dot{r}(\tau_0)|$ to get

$$\begin{aligned}
A^\mu(x) &= \frac{qc}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\tau\, u^\mu(\tau)\theta[x_0 - y_0(\tau)]\delta\{[x-y(\tau)]^2\} && (3.5.8)\\
&= \frac{q}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}\tau \frac{u^\mu(\tau)}{u(\tau)\cdot[x-y(\tau)]}\delta(\tau-\tau_{\text{ret}}) && (3.5.9)\\
&= \frac{q}{4\pi}\frac{u^\mu(\tau_{\text{ret}})}{u(\tau_{\text{ret}})\cdot[x-y(\tau_{\text{ret}})]}, && (3.5.10)
\end{aligned}$$

where the "retarded proper time" of the particle, τ_{ret}, is the solution of the equation

$$x^0 - y^0(\tau_{\text{ret}}) = |\vec{x} - \vec{y}(\tau_{\text{ret}})|. \tag{3.5.11}$$

Of course, the four-vector $R(\tau_{\text{ret}}, x) = x - y(\tau_{\text{ret}})$ is **light-like**, i.e., $R\cdot R = R_\mu R^\mu = 0$. One should note that in (3.5.8) thus $\tau_{\text{ret}} = \tau_{\text{ret}}(x)$, i.e., when taking space-time derivatives of A^μ to obtain the Faraday tensor, using this fully integrated form (3.5.9), one has to take this x-dependence of τ_{ret} into account. As we shall demonstrate below, it's again more advantageous to use the unintegrated expression (3.5.8).

This means that we can determine the retarded proper time in a Minkowski diagram by darwing the particle's world line $y(\tau)$ and then construct the light cone around the space-time point, where the fields are observed, x, by constructing the light-cone originating at x and determine the event, where the world-line of the particle crosses the past light-cone. We note that there can be only one unique intersection point of the particle's worldline with this past light-cone since

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(R\cdot R) = 2R\cdot \mathrm{d}_\tau R = -2cR\cdot u(\tau) \tag{3.5.12}$$

Since $u(\tau)$ is time like and $R(\tau_{\text{ret}})$ is light-like with $R^0(\tau_{\text{ret}}) > 0$ the function $R(\tau,x)|_{x=\text{const}}$ is strictly monotonously decreasing in the neighborhood of $\tau = \tau_{\text{ret}}$, and thus the intersection point of the particle's world line is unique. One should note that this argument breaks down, if the particle's four-velocity becomes (asymptotically) light-like, as in the case of hyperbolic motion.

To get the electromagnetic field we have to evaluate $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Instead of using the final result (3.5.10) it is more convenient to use the form (3.5.8). We start to note that for the time-derivative, there is in principle a contribution due to the derivative of the Heaviside unit-step function, $\partial_0 \Theta[x^0 - y^0(\tau)] = \delta[x^0 - y^0(\tau)]$. From the δ-distribution that implies that this only gives a divergent contribution for the field at the corresponding (retarded) position of the particle. Since the fields are undefined at the world line of the particle we exclude this contribution.

Thus we only need to consider

$$\begin{aligned}
\partial_\mu \delta\{[x-y(\tau)]^2\} &= 2[x_\mu - y_\mu(\tau)]\delta'\{[x-y(\tau)]^2\}\\
&= 2[x_\mu - y_\mu(\tau)]\left(\frac{\mathrm{d}}{\mathrm{d}\tau}[x-y(\tau)]^2\right)^{-1}\frac{\mathrm{d}}{\mathrm{d}\tau}\delta\{[x-y(\tau)]^2\}\\
&= -\frac{x_\mu - y_\mu(\tau)}{cu\cdot[x-y(\tau)]}\frac{\mathrm{d}}{\mathrm{d}\tau}\delta\{[x-y(\tau)]^2\}.
\end{aligned} \tag{3.5.13}$$

Inserting this into the corresponding derivative of (3.5.8), and integration by parts, ignoring the singular distribution through the derivative of the Heaviside unit-step function as argued above, we finally get

$$\begin{aligned} \partial_\mu A_\nu(x) &= \frac{q}{2\pi}\int_{\mathbb{R}} \mathrm{d}\tau \frac{\mathrm{d}}{\mathrm{d}\tau}\left\{\frac{[x_\mu - y_\mu(\tau)]\dot{u}_\nu(\tau)}{u\cdot[x-y(\tau)]}\right\}\Theta[x^0 - y^0(\tau)]\delta\{[x-y(\tau)]^2\} \\ &= \left[\frac{q}{4\pi c u(\tau)\cdot[x-y(\tau)]}\frac{\mathrm{d}}{\mathrm{d}\tau}\left\{\frac{[x_\mu - y_\mu(\tau)]\dot{u}_\nu(\tau)}{cu\cdot[x-y(\tau)]}\right\}\right]_{\tau=\tau_{\text{ret}}}, \end{aligned} \tag{3.5.14}$$

where in the final step we performed the same evaluation of the integral over τ which lead from (3.5.8) to (3.5.10). Antisymmetrization of this expression wrt. the indices μ and ν finally yields

$$F_{\mu\nu}(x) = \left[\frac{q}{4\pi c u(\tau)\cdot[x-y(\tau)]}\frac{\mathrm{d}}{\mathrm{d}\tau}\left\{\frac{[x_\mu - y_\mu(\tau)]\dot{u}_\nu(\tau) - [x_\nu - y_\nu(\tau)]\dot{u}_\mu(\tau)}{cu\cdot[x-y(\tau)]}\right\}\right]_{\tau=\tau_{\text{ret}}}. \tag{3.5.15}$$

3.6 Field of a uniformly moving point charge

As the most simple example for an electromagnetic field we consider the charge of a **uniformly moving point charge**. Its trajectory in terms of its proper time is given by

$$y(\tau) = cu\tau, \quad u = \text{const.} \tag{3.6.1}$$

For given x the retarded time is determined by the equation

$$[x - y(\tau)]^2 = 0, \quad x^0 > y^0(\tau). \tag{3.6.2}$$

Inserting (3.6.1), evaluating the square and simplifying a bit, we find the quadratic equation for τ

$$\tau^2 - \frac{2u\cdot x}{c}\tau + \frac{x^2}{c^2} = 0 \tag{3.6.3}$$

with the two solutions

$$\tau_\pm = \frac{u\cdot x}{c} \pm \frac{1}{c}\sqrt{(u\cdot x)^2 - x^2}. \tag{3.6.4}$$

It is clear that only the lower sign can lead to the retarded proper time,

$$\tau_{\text{ret}} = \tau_- = \frac{u\cdot x}{c} - \frac{1}{c}\sqrt{(u\cdot x)^2 - x^2}. \tag{3.6.5}$$

The denominator in (3.5.10) then is

$$u\cdot[x - y(\tau_{\text{ret}})] = u\cdot x - [u\cdot x - \sqrt{(u\cdot x) - x^2}] = \sqrt{(u\cdot x) - x^2}. \tag{3.6.6}$$

Thus the four-potential is

$$A^\mu(x) = \frac{qu^\mu}{4\pi\sqrt{(u\cdot x)^2 - x^2}}. \tag{3.6.7}$$

That this is the correct result can be easily veryfied by noting that for a particle at rest, we have $(u^\mu) = (1,0,0,0)^{\mathrm{T}}$ and $(u\cdot x)^2 - x^2 = \vec{x}^2 = r^2$. In the **rest frame** of the particle we thus find

$$A^0(x) = \frac{q}{4\pi r}, \quad \vec{A}(x) = 0. \tag{3.6.8}$$

This is the well-known electrostatic Coulomb potential, which is indeed the solution for a point charge at rest located in the origin of the spatial coordinate system. Since (3.6.7) is the only covariant generalization of (3.6.8) which can be built with the only four-vectors u and x present in this problem, it must be the solution for a uniformly moving point charge. One can also find it by Applying a boost-transformation matrix to the rest-frame solution (3.6.8).

The fields are most simply determined by taking the derivatives. In (1+3)-notation we get after some algebra

$$\vec{E} = -\frac{1}{c}\partial_t \vec{A} - \vec{\nabla} A^0 = \frac{q}{4\pi} \frac{\gamma(\vec{x} - \vec{v}t)}{[(u \cdot x)^2 - x^2]^{3/2}}, \quad \vec{B} = \vec{\nabla} \times \vec{A} = \frac{q}{4\pi} \frac{\vec{u} \times \vec{x}}{[(u \cdot x)^2 - x^2]^{3/2}} = \vec{\beta} \times \vec{E}. \tag{3.6.9}$$

In the final step we have used $\vec{u} = \gamma \vec{\beta} = \gamma \vec{v}/c$.

To put this in a more familar form, note that

$$(u \cdot x)^2 - x^2 = \gamma^2 (ct - \vec{\beta} \cdot \vec{x})^2 - c^2 t^2 + \vec{x}^2 = \gamma^2 (\vec{x}_{\parallel} - \vec{v}t)^2 + \vec{x}_{\perp}^2, \tag{3.6.10}$$

where

$$\vec{x}_{\parallel} = \frac{\vec{\beta} \cdot \vec{x}}{\beta^2} \vec{\beta}, \quad \vec{x}_{\perp} = \vec{x} - \vec{x}_{\parallel} \tag{3.6.11}$$

are the projections of the position vector $\vec{x}$ to the direction of the particle's velocity $\vec{v}$ and perpendicular to it, leading to

$$\vec{E} = \frac{q}{4\pi} \frac{\gamma(\vec{x} - \vec{v}t)}{[\gamma^2 (\vec{x}_{\parallel} - \vec{v}t)^2 + \vec{x}_{\perp}^2]^{3/2}}. \tag{3.6.12}$$

Note that the retarded solution finally leads to the somewhat surprising result that the electric field is in radial direction relative to the *instantaneous* position $\vec{y} = \vec{v}t$ of the charge, i.e., $\vec{E} \propto \vec{x} - \vec{v}t$.

3.7 Lagrange formalism for fields

In this section we shall analyze the electromagnetic field, using **Hamilton's least-action principle** in its formulation by **Lagrange**, which has the advantage to be manifestly covariant. After a general introduction to the formulation of field equations in terms of the action principle, we start with the free electromagnetic field, i.e., with situations in absence of charges and currents. The main goal of the analysis is the use of **Noether's theorem** to the special-relativistic space-time symmetries of Minkowski space to define the corresponding conservation laws for energy, momentum, angular momentum and center of momentum motion.

3.7.1 The action principle for fields and Noether's Theorem

We start with the general formulation of equations of motion in terms of the action principle for some fields Φ_k. Those need not be fields to appear in the relativistic context, but we shall discuss the need of relativistically sensible models, i.e., field dynamics which is consistent with the relativistic space-time model of Minkowski space, at the end of this Section. Nevertheless we shall use the relativistic notation for space-time variables.

Since fields are quantities describing continuous phenomena, i.e., quantities being defined as functions of time and location, $\Phi_i(t, \vec{x}) \equiv \Phi_i(x)$, we expect the **Lagrangian** to be given by a density. From the case of point particles, as discussed in Sect. 2.4, we expect it to be of the form

$$L = \int_{\mathbb{R}^3} \mathrm{d}^3 x \mathscr{L}(\Phi_i, \partial_\mu \Phi_i; t, \vec{x}). \tag{3.7.1}$$

Here $\mathscr{L}$ is called the **Lagrange density**. Then, as in for point-particle systems we define the **action** as

$$A[\Phi_i] = \int_{t_1}^{t_2} \mathrm{d}t L[\Phi_i, t] = \frac{1}{c} \int_{V^{(4)}} \mathrm{d}^4 x \mathscr{L}(\Phi_i, \partial_\mu \Phi_i; t, \vec{x}). \tag{3.7.2}$$

The equations of motion are given by the least-action principle or rather by those field configurations which make the **action functional stationary**. The variation is over all fields $\delta\Phi_i(x)$ with the fields fixed at the initial and final times t_1 and t_2, i.e., $\delta\Phi_i(x) = 0$ for $x^0 = ct_1$ and $x^0 = ct_2$. The space-time variables are *not* varied as in the least-action principle for point mechanics. From a formal point of view the spatial variables $\vec{x}$ are just continuous "labels" for infinitely many field degrees of freedom and thus not varied in the variations of the Hamilton principle, and the time is not varied by definition either.
The variation is then easily calulated as

$$\begin{aligned} c\delta A[\Phi_i] &= \int_{V^{(4)}} \mathrm{d}^4 x \left[\delta\Phi_i \frac{\partial \mathscr{L}}{\partial \Phi_i} + \delta(\partial_\mu \Phi_i) \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_i)} \right] \\ &= \int_{V^{(4)}} \mathrm{d}^4 x \delta\Phi_i \left[\frac{\partial \mathscr{L}}{\partial \Phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_i)} \right]. \end{aligned} \tag{3.7.3}$$

Here the Einstein summation relation applies to both the indices of space-time componentents μ and the field indices i. In the second step we have integrated by parts. The boundary contributions vanish due to the vanishing of the field variations at the boundaries at $t \in (t_1, t_2)$. The spatial boundary is assumed anyway at infinity, and we also assume that boundary terms in the action integral vanish at infinity sufficiently quickly.
Now the Hamilton principle states that $\delta A = 0$ for all field variations $\delta\Phi_i$, and this implies the **field equations of motion**

$$\frac{\partial \mathscr{L}}{\partial \Phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_i)} = 0. \tag{3.7.4}$$

These are the **Euler-Lagrange equations** of the variational principle for fields.
We note that the variation of the action is invariant under the transformation

$$\mathscr{L}' = \mathscr{L} + \partial_\mu \Omega^\mu(\Phi_i, x) \tag{3.7.5}$$

for an arbitrary four-vector function Ω^μ of the fields and the space-time coordinates x. To see this we note that

$$\Delta\mathscr{L} = \partial_\mu \Omega^\mu = \partial_\mu \Phi_i \frac{\partial \Omega^\mu}{\partial \Phi_i} + [\partial_\mu \Omega^\mu(\Phi_i, x)]_{\text{expl}}. \tag{3.7.6}$$

The last term refers to the space-time derivatives from the explicit x dependence of Ω^μ. From this we get

$$\frac{\partial \Delta\mathscr{L}}{\partial(\partial_\mu \Phi_i)} = \frac{\partial \Omega^\mu}{\partial \Phi_i} \Rightarrow \partial_\mu \frac{\partial \Delta\mathscr{L}}{\partial(\partial_\mu \Phi_i)} = \partial_\mu \frac{\partial \Omega^\mu}{\partial \Phi_i}. \tag{3.7.7}$$

On the other hand

$$\frac{\partial \Delta \mathscr{L}}{\partial \Phi_i} = \frac{\partial}{\partial \Phi_i} \partial_\mu \Omega^\mu = \partial_\mu \frac{\partial \Omega^\mu}{\partial \Phi_i}. \tag{3.7.8}$$

Eqs. (3.7.7) and (3.7.8) thus imply that

$$\frac{\partial \Delta \mathscr{L}}{\partial \Phi_i} - \partial_\mu \frac{\partial \Delta \mathscr{L}}{\partial(\partial_\mu \Phi_i)} \equiv 0, \tag{3.7.9}$$

i.e., the addition of $\Delta\mathscr{L}$ does not contribute to the variation of the action, and thus the Lagrangians $\mathscr{L}$ and $\mathscr{L}'$ are leading to the same field equations of motion.

One of the most important advantages to use the action principle for describing the dynamics is **Noether's Theorem**, which connects symmetries of the (variation of the) action with **conservation laws**, i.e., for each one-parameter Lie group which keeps the variation of the action invariant, there is a **conserved quantity**. As we shall see, in the case of field theories, derived from a local Lagrange density the conservation law holds in **local form**.

To derive Noether's theorem we consider **infinitesimal transformations** of the fields and the space-time coordinates. The latter is particularly important since we want to apply the formalism to the proper homogeneous Poincaré group. Let the group depend on parameters α^k. Then we can write an infinitesimal transformation of this kind in the form

$$x^\mu \to x'^\mu = x^\mu - \delta\alpha^k \xi_k^\mu(x), \quad \Phi_i \to \Phi'_i(x') = \Phi_i(x) + \delta\alpha^k \Xi_{ik}(\Phi_i, x). \tag{3.7.10}$$

To evaluate the infinitesimal change of the action under these transformations,

$$\delta A = \int_{\mathbb{R}^4} \mathrm{d}^4 x' \mathscr{L}(\Phi'_i, \partial'_\mu \Phi'_i, x') - \int_{\mathbb{R}^4} \mathrm{d}^4 x \mathscr{L}(\Phi_i, \partial_\mu \Phi_i, x), \tag{3.7.11}$$

we need

$$\delta(\partial_\mu \Phi_i) = \partial'_\mu \Phi'_i - \partial_\mu \Phi_i = \frac{\partial x^\nu}{\partial x'^\mu} \partial_\nu \left[\Phi_i + \delta\alpha^k \Xi_{ik}\right] - \partial_\mu \Phi_i. \tag{3.7.12}$$

Now we need the inverse matrix of

$$J^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} = \delta^\mu_\nu - \delta\alpha^k \partial_\nu \xi^\mu_k. \tag{3.7.13}$$

It is easy to see that up to linear order in the $\delta\alpha^k$

$$\frac{\partial x^\nu}{\partial x'^\mu} = \delta^\nu_\mu + \delta\alpha^k \partial_\mu \xi^\nu_k + \mathscr{O}(\delta\alpha^2). \tag{3.7.14}$$

To prove this we calculate, using (3.7.13)

$$J^\mu{}_\rho \frac{\partial x^\rho}{\partial x'^\sigma} = \left(\delta^\mu_\rho - \delta\alpha^k \partial_\rho \xi^\mu_k\right)\left(\delta^\rho_\sigma + \delta\alpha^k \partial_\sigma \xi^\rho_k\right) = \delta^\mu_\sigma + \mathscr{O}(\delta\alpha^2). \tag{3.7.15}$$

Thus (3.7.14) is verified. Plugging this into (3.7.12) we find after some alegbra

$$\delta(\partial_\mu \Phi_i) = \delta\alpha^k \left[\partial_\mu \Xi_{ik} + \partial_\mu \xi^\nu_k \partial_\nu \Phi_i\right]. \tag{3.7.16}$$

Then we need

$$\mathrm{d}^4 x' = \mathrm{d}^4 x \det\left(\frac{\partial x'}{\partial x}\right) = \mathrm{d}^4 x [1 - \delta\alpha^k \partial_\mu \xi^\mu_k + \mathscr{O}(\delta\alpha^2)]. \tag{3.7.17}$$

The latter follows from taking the determinant of (3.7.12) and only keeping terms up to linear order in the $\delta\alpha^k$ which occur in the product of the diagonal elements of the matrix. All other terms in the determinant obviously are at least of order $\mathscr{O}(\delta\alpha^2)$.
Plugging all this into (3.7.11) we find after some algebra

$$\delta A = \delta\alpha^k \int_{\mathbb{R}^4} \mathrm{d}^4 x \left[\frac{\partial \mathscr{L}}{\partial \Phi_i}\Xi_{ik} + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_i)}[\partial_\mu \Xi_{ik} + \partial_\mu \xi_k^\nu \partial_\nu \Phi_i] - \xi_k^\mu (\partial_\mu \mathscr{L})_{\mathrm{expl}} - \mathscr{L}\partial_\mu \xi_k^\mu\right]. \tag{3.7.18}$$

Since the $\delta\alpha^k$ are independent of each other and because $\delta A = 0$ if (3.7.10) is an infinitesimal symmetry transformation for each k the bracket under the integral must be a total four-divergence of the form $-\partial_\mu \Omega_k^\mu(\Phi_j, x)$, i.e., for each k there must exist a vector field $\Omega_k^\mu(\Phi_j, x)$, such that

$$\frac{\partial \mathscr{L}}{\partial \Phi_j}\Xi_{jk} + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}[\partial_\mu \Xi_{jk} + \partial_\mu \xi_k^\nu \partial_\nu \Phi_j] - \xi_k^\mu \left(\partial_\mu^{(\mathrm{expl})} \mathscr{L}\right) - \mathscr{L}\partial_\mu \xi_k^\mu + \partial_\mu \Omega_k^\mu(\Phi_j, x) = 0. \tag{3.7.19}$$

For abbreviation we define the **canonical energy-momentum tensor**,

$$\Theta^\mu{}_\nu = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\partial_\nu \Phi_j - \mathscr{L}\delta^\mu{}_\nu. \tag{3.7.20}$$

Then (3.7.20) reads

$$\frac{\partial \mathscr{L}}{\partial \Phi_j}\Xi_{jk} + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\partial_\mu \Xi_{jk} + \partial_\mu \xi_k^\nu \Theta^\mu{}_\nu - \xi_k^\nu \left(\partial_\nu^{(\mathrm{expl})} \mathscr{L}\right) + \partial_\mu \Omega_k^\mu(\Phi_j, x) = 0. \tag{3.7.21}$$

So far this is the constraint on the transformation, defined by the Ξ_{jk} and ξ_k in (3.7.10). Now we investigate the implication for the solutions of the field equations (3.7.4).
As a first step we calculate the four-divergence of the energy-momentum tensor. Using the field equations (3.7.4) we find

$$\begin{aligned}\partial_\mu \Theta^\mu{}_\nu &= \left(\partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\right)\partial_\nu \Phi_j + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\partial_\mu \partial_\nu \Phi_j - \frac{\partial \mathscr{L}}{\partial \Phi_j}\partial_\nu \Phi_j - \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\partial_\mu \partial_\nu \Phi_j - \partial_\nu^{(\mathrm{expl})}\mathscr{L} \\ &= -\partial_\nu^{(\mathrm{expl})}\mathscr{L}.\end{aligned} \tag{3.7.22}$$

Further with the field equations we get

$$\frac{\partial \mathscr{L}}{\partial \Phi_j}\Xi_{jk} + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\partial_\mu \Xi_{jk} = \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\Xi_{jk}\right) \tag{3.7.23}$$

for the first two terms in (3.7.21). Finally using (3.7.22) in (3.7.21) we conclude that the quantities

$$j_k^\mu = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_j)}\Xi_{jk} + \Theta^\mu{}_\nu \xi_k^\nu + \Omega_k^\mu \tag{3.7.24}$$

fulfill the **equation of continuity**,

$$\partial_\mu j_k^\mu = 0. \tag{3.7.25}$$

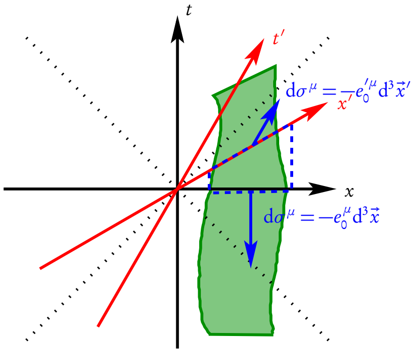
This is the first part of **Noether's theorems**: Any one-parameter symmetry implies the existence of a corresponding **Noether charge**,

$$Q_k = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\; j_k^0(x). \tag{3.7.26}$$

We note that this is a Lorentz scalar only if the continuity equation (3.7.25) is fulfilled.
To see this, we first prove that the Noether charge is conserved. This we can achieve easily by using usual 3D vector calculus. With (3.7.26) we find

$$\mathrm{d}_t Q_k = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\; \partial_t j_k^0(t,\vec{x}) = -\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\; \vec{\nabla}\cdot\vec{j}_k = 0, \tag{3.7.27}$$

where in the last step we used Gauss's theorem and the assumption that the current density vanishes at spatial infinity.



To see that Q_k is also a Lorentz scalar if (3.7.25) is fulfilled we use the 4D version of Gauss's theorem (A.7.16) with the 4D volume M depicted in the figure, where the shaded region indicates the region, where the charge and current densities are non-zero. Then one has

$$\begin{aligned}
0 &= \int_M \mathrm{d}^4x \partial_\mu j_k^\mu \\
&= \int_{\partial M} \mathrm{d}^3\sigma_\mu j_k^\mu \\
&= \int_{t'=0} \mathrm{d}^3\vec{x}'\, j_k'^0(t=0,\vec{x}') - \int_{t=0} \mathrm{d}^3\vec{x}\, j_k^0(t=0,\vec{x}),
\end{aligned} \tag{3.7.28}$$

which indeed implies that

$$Q_k' - Q_k = 0 \Rightarrow Q_k' = Q_k. \tag{3.7.29}$$

3.7.2 The free electromagnetic field

3.7.3 Systems of charged point particles

3.7.4 Continuum-mechanical formulation

Appendix A

Tensor calculus in Minkowski space

A.1 Tensor algebra

A.2 Affine spaces

A.3 The fundamental form

A.4 Generalized δ tensors in N-dimensional affine space

The alternating tensors ("k=forms") play a special role, particularly in tensor calculus, as we shall see in the following sections.

We start with the invariant generalized δ tensors, whose components $\delta^{\mu_1\cdots\mu_k}_{\nu_1\ldots\nu_k}$ with respect to an arbitry Carstesian basis are defined as the determinant

$$\delta^{\mu_1\cdots\mu_k}_{\nu_1\ldots\nu_k} = \det\begin{pmatrix} \delta^{\mu_1}_{\nu_1} & \delta^{\mu_1}_{\nu_2} & \ldots & \delta^{\mu_1}_{\nu_k} \\ \delta^{\mu_2}_{\nu_1} & \delta^{\mu_2}_{\nu_2} & \ldots & \delta^{\mu_2}_{\nu_k} \\ \vdots & \vdots & \vdots & \vdots \\ \delta^{\mu_k}_{\nu_1} & \delta^{\mu_k}_{\nu_2} & \ldots & \delta^{\mu_k}_{\nu_k} \end{pmatrix}. \tag{A.4.1}$$

This immediately implies that in N-dimensional space we have $k \in \{1,2,\ldots,N\}$ because for $k > N$ all the δ symbols must be 0, because there are only N different values, $\{1,\ldots,N\}$ for the indices[1] since the δ symbol is obviously antisymmetric under exchange of any pair of lower or any pair of upper indices. Also obviously $\delta^{\mu_1\cdots\mu_k}_{\nu_1\ldots\nu_k} \neq 0$ only if $\{\mu_1,\ldots,\mu_k\} = \{\nu_1,\ldots,\nu_k\}$, because otherwise in each of the products of k Kronecker δ's summed over to calculate the determinant vanishes.

Obviously the δ symbols for $k = N$ are invariant under arbitrary basis transformations: Writing $\bar{x}^{\mu} = T^{\mu}{}_{\nu}\, x^{\nu}$ and $\bar{x}_{\mu} = U^{\nu}{}_{\mu}\, x_{\nu}$, which implies $U^{\mu}{}_{\nu}\, T^{\nu}{}_{\rho} = T^{\mu}{}_{\rho}\, U^{\rho}{}_{\nu} = \delta^{\mu}_{\nu}$, we find

$$\bar{\delta}^{\rho_1\cdots\rho_N}_{\sigma_1\ldots\sigma_N} = T^{\rho_1}{}_{\mu_1} \cdots T^{\rho_N}{}_{\mu_N}\, U^{\nu_1}{}_{\sigma_1} \cdots U^{\nu_N}{}_{\sigma_N}\, \delta^{\mu_1\cdots\mu_N}_{\nu_1\ldots\nu_N} = \det(\hat{T})\det(\hat{U})\delta^{\mu_1,\ldots,\mu_N}_{\nu_1\ldots\nu_N} = \delta^{\mu_1,\ldots,\mu_N}_{\nu_1\ldots\nu_N}. \tag{A.4.2}$$

Next we show that contracting the generalized δ-symbol of rank $2k$ over one index pair, gives the δ-symbol of rank $2k-2$. To evaluate the contraction $\delta^{\mu_1\cdots\mu_{k-1}\mu_k}_{\nu_1\ldots\nu_{k-1}\mu_k}$ we use Laplace's formula to the last

[1] In this section we count the tensor components with indices running over $\{1,\ldots,N\}$ for simplicity of the argument.

row,

$$\delta^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_k} = \sum_{j=1}^{k} (-1)^{j+k} \delta^{\mu_k}_{\nu_j} \delta^{\mu_1\cdots\mu_{k-1}}_{\nu_1\ldots\hat{\nu}_j\ldots\nu_k}, \tag{A.4.3}$$

where the $\hat{\nu}_j$ index in the last expression means to leave this one index out. Now we just calculate the contraction by setting $\nu_k = \mu_k$, which implies summation from 1 to N. Then we get

$$\delta^{\mu_1\cdots\mu_{k-1}\mu_k}_{\nu_1\cdots\nu_{k-1}\mu_k} = N\delta^{\mu_1,\ldots,\mu_{k-1}}_{\nu_1,\ldots,\nu_{k-1}} + \sum_{j=1}^{k-1} (-1)^{k+j} \delta^{\mu_1,\ldots,\mu_{k-1}}_{\nu_1\ldots\hat{\nu}_j\ldots\nu_j} \tag{A.4.4}$$

So in the sum we sum $k-1$-times over the same $\delta^{\mu_1\cdots\mu_k}_{\nu_1\ldots\nu_{k-1}}$ symbol, only the index ν_j is put to the $(k-1)^{(\mathrm{th})}$ position. We just have to bring it back to the original position, which needs $k-1-j$ exchanges, i.e., we finaly get

$$\delta^{\mu_1\cdots\mu_k}_{\nu_1\cdots\mu_k} = (N-k+1)\delta^{\mu_1,\ldots,\mu_{k-1}}_{\nu_1,\ldots,\nu_{k-1}}. \tag{A.4.5}$$

Iterating this equation leads to the general contraction

$$\delta^{\mu_1\cdots\mu_j\mu_{j+1}\cdots\mu_k}_{\nu_1\cdots\nu_j\mu_{j+1}\cdots\mu_k} = (N-k+1)(N-k+2)\cdots(N-j)\delta^{\mu_1\cdots\mu_j}_{\nu_1\cdots\nu_j} = \frac{(N-j)!}{(N-k)!}\delta^{\mu_1\cdots\mu_j}_{\nu_1\cdots\nu_j}. \tag{A.4.6}$$

Since now for $k=N$ the generalized δ symbol is an invariant tensor according to (A.4.2), so are all $(2k)$-rank δ symbols since we can get them as contractions over the $(N-k)$ other indices.

A.5 Levi-Civita tensor in (1+3)-dimensional Minkowski space

Now we turn back to (1+3)-dimensional Minkowski space with all Greek indices running over $\{0,1,2,3\}$. Of course, for the generalized δ symbol nothing changes. Now we introduce the **Levi-Civita tensor** by its pseudo-Cartesian contravariant components as

$$\epsilon^{\mu\nu\rho\sigma} = \delta^{\mu\nu\rho\sigma}_{0\,123}, \tag{A.5.1}$$

i.e., it is +1 (−1) if the ordered quadrupel $(\mu,\nu\rho,\sigma)$ is an even (odd) permutation of the lexically ordered quadrupel $(0,1,2,3)$[2]. Then the covariant components are given by the usual rule of lowering contravariant indices,

$$\epsilon_{\alpha\beta\gamma\delta} = \eta_{\alpha\mu}\eta_{\beta\nu}\eta_{\gamma\rho}\eta_{\gamma\sigma}\epsilon^{\mu\nu\rho\sigma} = \epsilon^{\mu\nu\rho\sigma}\det\hat{\eta} = -\epsilon^{\alpha\beta\gamma\delta}. \tag{A.5.2}$$

Since we have

$$\epsilon^{\mu\nu\rho\sigma}\epsilon_{\alpha\beta\gamma\delta} = -\delta^{\mu\nu\rho\sigma}_{\alpha\beta\gamma\delta}, \tag{A.5.3}$$

because for $(\alpha,\beta,\gamma,\delta) = (\mu\nu\rho\sigma) = (0,1,2,3)$ the values of both sides of the equation are −1 and the symbol is antisymmetric under exchange of any pair of upper and of any pair of lower indices. From (A.4.6) we get the important formulas

$$\epsilon^{\mu\nu\rho\delta}\epsilon_{\alpha\beta\gamma\delta} = -\delta^{\mu\nu\rho}_{\alpha\beta\gamma}, \tag{A.5.4}$$

$$\epsilon^{\mu\nu\gamma\delta}\epsilon_{\alpha\beta\gamma\delta} = -2\delta^{\mu\nu}_{\alpha\beta}, \tag{A.5.5}$$

$$\epsilon^{\mu\beta\gamma\delta}\epsilon_{\alpha\beta\gamma\delta} = -6\delta^{\mu}_{\alpha}. \tag{A.5.6}$$

[2]Note that there are different conventions concerning the sign are found in the literature. Some authors define in this way the covariant Levi-Civita tensor components, and this leads to a sign change

Now a k"-rank tensor is called **totally antisymmetric** if its components $T^{\mu_1 \cdots \mu_k}$ are antisymmetric under exchange of any pair of indices. It is clear that only totally antisymmetric tensors of rank $k \leq 4$ exist since each index can only take the four different values 0, 1, 2, or 3. Then one defines the **Hodge dual** of this totally antisymmetric tensor as the $(4-k)$-rank tensor by its components

$$\begin{aligned} {}^\dagger T_{\nu_{k+1} \cdots \nu_4} &= \frac{1}{k!} \epsilon_{\nu_{k+1} \cdots \nu_4 \mu_1 \cdots \mu_k} T^{\mu_1 \cdots \mu_k}, \\ {}^\dagger T^{\nu_{k+1} \cdots \nu_4} &= \frac{1}{k!} \epsilon^{\mu_1 \cdots \mu_k \nu_{k+1} \cdots \nu_4} T_{\mu_1 \cdots \mu_k}. \end{aligned} \tag{A.5.7}$$

E.g., a totally antisymmetric Tensor $\boldsymbol{T}$ of rank 3 is mapped to a vector ${}^\dagger \boldsymbol{T}$ by

$$({}^\dagger T)_\mu = \frac{1}{3!} \epsilon_{\mu\nu\rho\sigma} T^{\nu\rho\sigma}. \tag{A.5.8}$$

Taking the Hodge-dual twice leads back to the original tensor up to a sign, because of (A.5.4-A.5.6). E.g., for our 3rd-rank-tensor example we have

$$({}^{\dagger\dagger} T)^{\alpha\beta\gamma} = \epsilon^{\mu\alpha\beta\gamma} ({}^\dagger T)_\mu = \frac{1}{3!} \epsilon^{\mu\alpha\beta\gamma} \epsilon_{\mu\nu\rho\sigma} T^{\nu\rho\sigma} = -\frac{1}{3!} \delta^{\alpha\beta\gamma}_{\nu\rho\sigma} T^{\nu\rho\sigma} = -T^{\alpha\beta\gamma}. \tag{A.5.9}$$

In the last step we have used the total antisymmetry of the tensor components under permutation of its indices. Thus Hodge dualization is an invertible mapping between totally antisymmetric tensors.

A.6 Differential tensor calculus

Now we consider **tensor fields**. These provide tensor-valued observables $\boldsymbol{T}(\boldsymbol{x})$ as a function of the spacetime point given by the spacetime fourvector $\boldsymbol{x}$. The transformation properties under proper orthochronous Lorentz transformations for its components is obviously

$$\bar{T}^{\mu\nu\cdots}(\bar{x}^\alpha) = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \cdots T^{\rho\sigma\cdots}(x^\beta), \quad \bar{x}^\alpha = \Lambda^\alpha{}_\beta x^\beta. \tag{A.6.1}$$

It is now easy to show that the partial derivative of the components of a tensor field of rank k with respect to x^α leads to a covariant (lower) tensor index, i.e., it defines a new tensor of rank $(k+1)$.
To show this we remember the properties of Lorentz-transformation matrices:

$$\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}. \tag{A.6.2}$$

Contracting with $\eta^{\sigma\alpha}$ gives

$$\eta_{\mu\nu} \eta^{\sigma\alpha} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \delta^\alpha_\rho, \tag{A.6.3}$$

i.e.,

$$(\Lambda^{-1})^\alpha{}_\mu = \eta_{\mu\nu} \eta^{\sigma\alpha} \Lambda^\nu{}_\sigma = \Lambda_\mu{}^\alpha. \tag{A.6.4}$$

The covariant component of a fourvector thus transforms as

$$\bar{x}_\mu = \eta_{\mu\alpha} \bar{x}^\alpha = \eta_{\mu\alpha} \Lambda^\alpha{}_\beta x^\beta = \eta_{\mu\alpha} \eta^{\beta\gamma} \Lambda^\alpha{}_\beta x_\gamma = \Lambda_\mu{}^\gamma x_\gamma = (\Lambda^{-1})^\gamma{}_\mu x_\gamma. \tag{A.6.5}$$

Now we have

$$\begin{aligned} \bar{\partial}_\alpha \bar{T}^{\mu\nu\ldots} &= \frac{\partial}{\partial \bar{x}^\alpha} \bar{T}^{\mu\nu\ldots} = \frac{\partial x^\beta}{\partial \bar{x}^\alpha} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \cdots \partial_\beta T^{\rho\sigma\ldots} \\ &= (\Lambda^{-1})^\beta{}_\alpha \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \cdots \partial_\beta T^{\rho\sigma\ldots} \\ &= \Lambda_\alpha{}^\beta \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \cdots \partial_\beta T^{\rho\sigma\ldots}, \end{aligned} \tag{A.6.6}$$

and this is precisely the transformation law for tensor components with a covariant indices α and β, as claimed.

A.7 Integrals in Minkowski space

As in usual 3D Euclidean tensor calculus we can define various types of covariant integrals in Minkowski space and derive integral theorems of the Stokes and Gauss type.

The most basic ones can be defined without taking recourse to any notion of a scalar product for Euclidean spaces or a fundamental form ("pseudoscalar product") for Minkowski space. Here a bare differential manifold without any notions of a metric or pseudometric is sufficient. As for special relativity we only need integration on affine spaces we restrict ourselves to this case, but the notion of integrals on general differential manifolds is very similar.

The key are the generalized Kronecker symbols introduced in Sect. A.4. For Minkowski space we have $N = 4$ and thus three types of integrals, because we can define $1 \ldots 4$-dimensional sub-manifolds. Here we only consider the most simple case that such a $d \leq 4$-dimensional manifold is defined by a set of parameters q^k ($k \in \{1, \ldots, d\}$) with domain in a cuboid $Q_d(a_1, b_1) \otimes (a_2, b_2) \otimes \cdots \otimes (a_d, b_d)$ by a function $x^\mu(q^k)$ with x^μ the spacetime four-vector. Then we can define integration measures by

$$\begin{aligned} \mathrm{d}^d \sigma^{\mu_1 \mu_2 \cdots \mu_d} &= \mathrm{d}q^1 \mathrm{d}q^2 \cdots \mathrm{d}q^d \, \delta^{\mu_1 \mu_2 \cdots \mu_d}_{\nu_1 \nu_2 \ldots \nu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \frac{\partial x^{\nu_2}}{\partial q^2} \cdots \frac{\partial x^{\nu_d}}{\partial q^d} \\ &= \mathrm{d}q^1 \mathrm{d}q^2 \cdots \mathrm{d}q^d \, \epsilon^{\mu_1 \cdots \mu_d} \det\left(\frac{\partial x^\nu}{\partial q^k} \right). \end{aligned} \tag{A.7.1}$$

The main reason is two-fold: First of all it generalizes the corresponding notions in 3D Euclidean space, where for, e.g., $d = 2$ the above construction defines a 2D surface element $\mathrm{d}^2 \sigma^{ab}$, which usually is mapped by a Hodge dualization to the surface-element normal vector $\mathrm{d}^2 \vec{f}$, but first we do not consider the Hodge dualization but work with the integration measures of the type (A.7.1) only.

The second reason to use (A.7.1) is the independence of this description of the choice of parameters. Using other parameters $\tilde{q}^k$ the (A.7.2) gets the same form again, because then we have

$$\begin{aligned} \mathrm{d}^d \sigma^{\mu_1 \mu_2 \cdots \mu_d} &= \mathrm{d}\tilde{q}^1 \mathrm{d}\tilde{q}^2 \cdots \mathrm{d}\tilde{q}^d \, \epsilon^{\mu_1 \mu_2 \cdots \mu_d} \det\left(\frac{\partial x^\nu}{\partial q^k} \right) \det\left(\frac{\partial q^k}{\partial \tilde{q}^l} \right) \\ &= \mathrm{d}\tilde{q}^1 \mathrm{d}\tilde{q}^2 \cdots \mathrm{d}\tilde{q}^d \, \epsilon^{\mu_1 \mu_2 \cdots \mu_d} \det\left(\frac{\partial x^\nu}{\partial q^k} \frac{\partial q^k}{\partial \tilde{q}^l} \right) \\ &= \mathrm{d}\tilde{q}^1 \mathrm{d}\tilde{q}^2 \cdots \mathrm{d}\tilde{q}^d \, \epsilon^{\mu_1 \mu_2 \cdots \mu_d} \det\left(\frac{\partial x^\nu}{\partial \tilde{q}^l} \right) \\ &= \mathrm{d}\tilde{q}^1 \mathrm{d}\tilde{q}^2 \cdots \mathrm{d}\tilde{q}^d \, \delta^{\mu_1 \mu_2 \cdots \mu_d}_{\nu_1 \nu_2 \ldots \nu_d} \frac{\partial x^{\nu_1}}{\partial \tilde{q}^1} \frac{\partial x^{\nu_2}}{\partial \tilde{q}^2} \cdots \frac{\partial x^{\nu_d}}{\partial \tilde{q}^d} \end{aligned} \tag{A.7.2}$$

Since under proper orthochronous Lorentz transformations $\mathrm{d}^d \sigma^{\mu_1 \mu_2 \cdots \mu_d}$ transform like contravariant tensor components we can integrate antisymmetric rank-d tensor fields over such a d-dimensional hypersurface with a scalar as a result, i.e., we can define invariant integrals of the form

$$\int_M \mathrm{d}^d \sigma^{\mu_1 \cdots \mu_d} T_{\mu_1 \cdots \mu_d}(\underline{x}) = \int_Q \mathrm{d}q^1 \cdots \mathrm{d}q^d \, \delta^{\mu_1 \mu_2 \cdots \mu_d}_{\nu_1 \nu_2 \ldots \nu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \cdots \frac{\partial x^{\nu_d}}{\partial q^d} T_{\mu_1 \cdots \mu_d}[\underline{x}(q)]. \tag{A.7.3}$$

Now the so defined submanifolds M have a boundary, which is defined as all points described by $q^j = b_j$ and $q^j = a_j$ for one $j \in \{1,\dots,d\}$ and all the other q^k's running over their intervals. This manifold we denote as ∂M. Concerning the orientation of the corresponding hypersurface elements we use the definition for the parts with $q^j = b_j = \text{const}$

$$\mathrm{d}^{d-1}\sigma^{\mu_1\cdots\hat{\mu}_j\cdots\mu_d} = \delta^{\mu_1\cdots\hat{\mu}_j\cdots\mu_d}_{\nu_1\cdots\hat{\nu}_j\cdots\nu_d}\frac{\partial x^{\nu_1}}{\partial q^1}\cdots\frac{\partial x^{\nu_{j-1}}}{\partial q^{j-1}}\frac{\partial x^{\nu_{j+1}}}{\partial q^{j+1}}\cdots\frac{\partial x^{\nu_d}}{\partial q^d}\bigg|_{q^j=b_j} \tag{A.7.4}$$

and for the parts with $q^j = a_j = \text{const}$

$$\mathrm{d}^{d-1}\sigma^{\mu_1\cdots\hat{\mu}_j\cdots\mu_d} = -\,\delta^{\mu_1\cdots\hat{\mu}_j\cdots\mu_d}_{\nu_1\cdots\hat{\nu}_j\cdots\nu_d}\frac{\partial x^{\nu_1}}{\partial q^1}\cdots\frac{\partial x^{\nu_{j-1}}}{\partial q^{j-1}}\frac{\partial x^{\nu_{j+1}}}{\partial q^{j+1}}\cdots\frac{\partial x^{\nu_d}}{\partial q^d}\bigg|_{q^j=a_j}. \tag{A.7.5}$$

The hat over the indices indicates the ones to be left out in the expression, including an appropriate sign convention,

$$\delta^{\mu_1\cdots\hat{\mu}_j\cdots\mu_d}_{\nu_1\cdots\hat{\nu}_j\cdots\nu_d} = (-1)^{j+1}\delta^{\mu_1\cdots\mu_{j-1}\mu_{j+1}\cdots\mu_d}_{\nu_1\cdots\nu_{j-1}\nu_{j+1}\cdots\nu_d}. \tag{A.7.6}$$

This defines the orientation of the boundary ∂M relative to M. The orientation of the latter is defined by the parametrization. This definition is chosen in a way to make the statement of the **generalized Stokes theorem of integration** most simple:

$$\int_M \mathrm{d}^d\sigma^{\mu_1\cdots\mu_d}\partial_{\mu_1}T_{\mu_2\cdots\mu_d} = \int_{\partial M}\mathrm{d}^{d-1}\sigma^{\mu_2\cdots\mu_d}T_{\mu_2\cdots\mu_d}. \tag{A.7.7}$$

It is also very simple to prove since one has just to take each of the integrals on the left-hand side in the sum over μ_1 separately. To this end one uses the definition (A.4.1) of the δ-tensor components and expand the determinant wrt. the 1st row,

$$\begin{aligned}
\int_M \mathrm{d}^d\sigma^{\mu_1\cdots\mu_d}\partial_{\mu_1}T_{\mu_2\cdots\mu_d} &= \sum_{k=1}^d\int_Q \mathrm{d}q^1\cdots\mathrm{d}q^d(-1)^{k+1}\delta^{\mu_1}_{\nu_k}\delta^{\mu_2\cdots\cdots\cdots\mu_d}_{\nu_2\cdots\nu_{k-1}\nu_{k+1}\nu_d}T_{\mu_2\cdots\mu_d}\frac{\partial x^{\nu_1}}{\partial q^1}\cdots\frac{\partial x^{\nu_d}}{\partial q^d}\\
&= \sum_{k=1}^d\int_Q \mathrm{d}q^1\cdots\mathrm{d}q^d(-1)^{k+1}\delta^{\mu_2\cdots\cdots\cdots\mu_d}_{\nu_2\cdots\nu_{k-1}\nu_{k+1}\nu_d}\frac{\partial}{\partial q^k}T_{\mu_2\cdots\mu_d}\\
&\qquad\frac{\partial x^{\nu_1}}{\partial q^1}\cdots\frac{\partial x^{\nu_{k-1}}}{\partial q^{k-1}}\frac{\partial x^{\nu_{k+1}}}{\partial q^{k+1}}\cdots\frac{\partial x^{\nu_d}}{\partial q^d}\\
&= \sum_{k=1}^d\int_{Q_k} \mathrm{d}q^1\cdots\mathrm{d}q^{k-1}\mathrm{d}q^{k+1}\cdots\mathrm{d}q^d(-1)^{k+1}\delta^{\mu_2\cdots\cdots\cdots\mu_d}_{\nu_2\cdots\nu_{k-1}\nu_{k+1}\nu_d}\left[T_{\mu_2\cdots\mu_d}\right]_{q_k=a_k}^{q_k=b_k}\\
&\qquad\frac{\partial x^{\nu_1}}{\partial q^1}\cdots\frac{\partial x^{\nu_{k-1}}}{\partial q^{k-1}}\frac{\partial x^{\nu_{k+1}}}{\partial q^{k+1}}\cdots\frac{\partial x^{\nu_d}}{\partial q^d}\\
&= \int_{\partial M}\mathrm{d}^{d-1}\sigma^{\mu_2\cdots\mu_d}T_{\mu_2\cdots\mu_d}.
\end{aligned} \tag{A.7.8}$$

This is the general **Stokes's integral theorem**.

Now we can also use the Levi"=Civita symbols to consider the duals of the possible integration measures $\mathrm{d}^d\sigma^{\mu_1\cdots\mu_d}$, where $d \in \{1,2,3,4\}$, and the corresponding Stokes's integral theorem translates into the general **Gauss's integral theorem**.

To that end we define

$$\mathrm{d}^4\sigma^* = -\frac{1}{4!}\epsilon_{\mu_1\cdots\mu_4}\mathrm{d}^4\sigma^{\mu_1\cdots\mu_4} = \mathrm{d}^4 x, \tag{A.7.9}$$

$$\mathrm{d}^3\sigma^*_{\mu_1} = -\frac{1}{3!}\epsilon_{\mu_1\mu_2\mu_3\mu_4}\mathrm{d}^3\sigma^{\mu_2\cdots\mu_4}, \tag{A.7.10}$$

$$\mathrm{d}^2\sigma^*_{\mu_1\mu_2} = -\frac{1}{2!}\epsilon_{\mu_1\mu_2\mu_3\mu_4}\mathrm{d}^2\sigma^{\mu_3\mu_4}, \tag{A.7.11}$$

$$\mathrm{d}^3\sigma^*_{\mu_1\mu_2\mu_3} = -\epsilon_{\mu_1\mu_2\mu_3\mu_4}\mathrm{d}\sigma^{\mu_4}. \tag{A.7.12}$$

This leads to scalar integrals of the type

$$\int_M \mathrm{d}^k\sigma_{\mu_1\cdots\mu_k} V^{\mu_1\cdots\mu_k} \tag{A.7.13}$$

over totally antisymmetric tensor vields of rank k.

To state **Gauss's integral laws**, it is more lucid to discuss the three different cases separately. We start with

$$\int_M \mathrm{d}^4\sigma^* \partial_{\mu_1} V^{\mu_1} = \int_{\partial M} \mathrm{d}^3\sigma^*_{\mu_1} V^{\mu_1}. \tag{A.7.14}$$

To prove it, we simply have to write out the integral on the left-hand side of the equation according to the definition (A.7.9) and use the properties of the Hodge dualization operation discussed in Sect. A.5,

$$\begin{aligned}
\int_M \mathrm{d}^4\sigma^* \partial_{\mu_1} V^{\mu_1} &= -\frac{1}{4!}\int_M \mathrm{d}^4\sigma^{\nu_1\mu_2\mu_3\mu_4}\epsilon_{\nu_1\mu_2\mu_3\mu_4}\partial_{\mu_1}V^{\mu_1} \\
&= \frac{1}{4!}\frac{1}{3!}\int_M \mathrm{d}^4\sigma^{\nu_1\mu_2\mu_3\mu_4}\epsilon_{\nu_1\mu_2\mu_3\mu_4}\epsilon^{\nu_2\nu_3\nu_4\mu_1}\partial_{\mu_1}{}^{\dagger}V_{\nu_2\nu_3\nu_4} \\
&= +\frac{1}{4!}\frac{1}{3!}\int_M \mathrm{d}^4\sigma^{\nu_1\mu_2\mu_3\mu_4}\delta^{\mu_1\nu_2\ \nu_3\ \nu_4}_{\nu_1\ \mu_2\mu_3\mu_4}\partial_{\mu_1}{}^{\dagger}V_{\nu_2\nu_3\nu_4} \\
&= +\frac{1}{3!}\int_M \mathrm{d}^4\sigma^{\mu_1\mu_2\mu_3\mu_4}\partial_{\mu_1}{}^{\dagger}V_{\mu_2\mu_3\nu_4}.
\end{aligned} \tag{A.7.15}$$

With Stokes's theorem (A.7.8) we get

$$\begin{aligned}
\int_M \mathrm{d}^4\sigma^* \partial_{\mu_1} V^{\mu_1} &= \frac{1}{3!}\int_{\partial M} \mathrm{d}^3\sigma^{\mu_2\mu_3\mu_4}{}^{\dagger}V_{\mu_2\mu_3\mu_4} \\
&= -\frac{1}{3!}\int_{\partial M} \mathrm{d}^3\sigma^{\mu_2\mu_3\mu_4}\epsilon_{\mu_1\mu_2\mu_3\mu_4}V^{\mu_1} \\
&= \int_{\partial M} \mathrm{d}^3\sigma^*_{\mu_1} V^{\mu_1},
\end{aligned} \tag{A.7.16}$$

which is (A.7.14).

In a similar way we get for an antisymmetric 2nd-rank tensor field

$$\int_M \mathrm{d}^3\sigma^*_{\mu_1}\partial_{\mu_2}V^{\mu_1\mu_2} = \frac{1}{2!}\int_{\partial_M}\mathrm{d}^2\sigma^*_{\mu_1\mu_2}V^{\mu_1\mu_2}, \tag{A.7.17}$$

for a totally antisymmetric 3rd-rank tensor field

$$\frac{1}{2!}\int_M \mathrm{d}^2\sigma^*_{\mu_1\mu_2}\partial_{\mu_3}V^{\mu_1\mu_2\mu_3} = \frac{1}{3!}\int \mathrm{d}\sigma^*_{\mu_1\mu_2\mu_3}V^{\mu_1\mu_2\mu_3}. \tag{A.7.18}$$

Bibliography

[Bel87] J. S. Bell, *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge, New York, Melbourne (1987).

[Bor09] M. Born, Die Theorie des starren Elektrons in der Kinematik des Relativitätsprinzips, Ann. Phys. (Leipzig) **335**, 1 (1909), `https://doi.org/10.1002/andp.19093351102`.

[Cro15] B. Crowell, *Special Relativity* (2015), `https://www.lightandmatter.com`.

[DeW11] B. DeWitt, *Bryce DeWitt's Lectures on Gravitation*, Springer, Heidelberg, Dordrecht, London, New York (2011), `https://dx.doi.org/10.1007/978-3-540-36911-0`.

[Ein05] A. Einstein, On the electrodynamics of moving bodies, Annalen Phys. **17**, 891 (1905), [Annalen Phys.14,194(2005)], `https://dx.doi.org/10.1002/andp.200590006`.

[Her10] G. Herglotz, Über den vom Standpunkt des Relativitätsprinzips aus als "starr" zu bezeichnenden Körper, Annalen der Physik **336**, 393 (1910).

[Jac98] J. D. Jackson, *Classical electrodynamics*, John Wiley&Sons, Hoboken, NJ, 3rd edn. (1998).

[Min09] H. Minkowski, Raum und Zeit, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 18, p. 75-88 **18**, 75 (1909).

[Noe10] F. Noether, Zur Kinematik des starren Körpers in der Relativtheorie, Annalen der Physik **336**, 919 (1910).

[Sil14] L. Silberstein, *The Theory of Relativity*, MacMillan&Co., London (1914), `https://archive.org/details/theoryofrelativi00silbrich`.

[Tho26] L. H. Thomas, The motion of the spinning electron, Nature **117** (1926), `https://dx.doi.org/10.1038%2F117514a0`.