
Special Relativity

Hendrik van Hees

March 11, 2024

Contents

Contents	3
1 Kinematics	5
1.1 Introduction	5
1.2 The special-relativistic space-time model	6
1.3 The twin “paradox”	12
1.4 General Lorentz transformations	12
1.5 Addition of velocities	15
1.6 Relative velocity	16
1.7 The Lorentz group as a Lie group	17
1.8 Fermi-Walker transport and Thomas precession	20
2 Mechanics	25
2.1 Particle dynamics	25
2.2 Motion of a particle in an electromagnetic field	29
2.2.1 Massive particle in a homogeneous electric field	30
2.2.2 Massive particle in a homogeneous magnetic field	32
2.3 Bell’s space-ship paradox	33
2.4 The action principle	37
2.4.1 The frame-dependent (1+3)-formalism	37
2.4.2 Manifestly covariant formulation	41
2.4.3 Alternative Lagrange formalism	43
2.5 Thermodynamics	44
3 Classical fields	45
3.1 The action principle for fields and Noether’s Theorem	45
3.2 Poincaré symmetry	49
3.2.1 Translations	49
3.2.2 Lorentz transformations	50
3.3 Pseudo-gauge transformations	50
3.4 The Belinfante-Rosenfeld tensor	51
3.5 Continuum mechanics	52

Contents

3.5.1	Kinematics	52
3.5.2	Dynamics of a medium	57
3.5.3	Ideal fluid	58
3.5.4	“Dust matter” and a scalar field	60
4	Classical Electromagnetism	63
4.1	Heuristic foundations	63
4.2	Manifestly covariant formulation of electrodynamics	64
4.3	The Doppler effect for light	71
4.4	The Lagrangian for the electromagnetic field	73
4.5	Gauge invariance and retarded potentials	76
4.6	The retarded fields	79
4.7	The Liénard-Wiechert potential	81
4.8	Field of a uniformly moving point charge	83
4.9	The radiation-reaction problem	85
4.9.1	The energy-momentum tensor of the Liénard-Wiechert field	85
4.9.2	The Lorentz-Abraham-Dirac equation	87
4.9.3	The free-particle solution	89
A	Tensor calculus in Minkowski space	91
A.1	Tensor algebra	91
A.2	The fundamental form	91
A.3	Generalized δ tensors in N -dimensional affine space	91
A.4	Levi-Civita tensor in (1+3)-dimensional Minkowski space	93
A.5	Differential tensor calculus	94
A.6	Integrals in Minkowski space	94
	Bibliography	99

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 century¹, 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 parameter² [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

$$c^2 t^2 - \vec{x}^2 = 0. \quad (1.2.1)$$

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

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

²Here 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.

1.2. The special-relativistic space-time model

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 \bar{t} and $\bar{\vec{x}}$ 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:

$$c^2\bar{t}^2 - \bar{\vec{x}}^2 = 0. \quad (1.2.2)$$

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 e_μ (with a lower index), $x = x^\mu e_\mu$, 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

$$x \cdot y = x^0 y^0 - \vec{x} \cdot \vec{y}, \quad (1.2.3)$$

then the Eq. (1.2.1) can be written as

$$x \cdot x =: x^2 = 0. \quad (1.2.4)$$

Obviously for the here introduced "pseudo-Cartesian" basis vectors e_μ the bilinear form is uniquely defined by

$$e_\mu \cdot e_\nu = \eta_{\mu\nu} := \begin{cases} 1 & \text{for } \mu = \nu = 0, \\ -1 & \text{for } \mu = \nu, \quad \mu, \nu \in \{1, 2, 3\}, \\ 0 & \text{for } \mu \neq \nu, \quad \mu, \nu \in \{0, 1, 2, 3\}. \end{cases} \quad (1.2.5)$$

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

$$x \cdot y = \eta_{\mu\nu} x^\mu y^\nu. \quad (1.2.6)$$

Then (1.2.2) holds true, if we assume that Bob simply also uses pseudo-Cartesian basis vectors \bar{e}_μ , 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

$$e_\nu = \bar{e}_\mu \Lambda^\mu{}_\nu \quad (1.2.7)$$

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

$$\eta_{\rho\sigma} = e_\rho \cdot e_\sigma = \Lambda^\mu{}_\rho \bar{e}_\mu \cdot \Lambda^\nu{}_\sigma \bar{e}_\nu = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta_{\mu\nu}. \quad (1.2.8)$$

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.

1. Kinematics

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 $(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})$, where \mathbf{x} and \mathbf{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 $\mathbf{x} \neq \mathbf{y}$. 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 $\mathbf{x} \neq 0$ with $\mathbf{x} \cdot \mathbf{x} = 0$. For the spacial basis vectors we have $\mathbf{e}_j \cdot \mathbf{b}_j = -1$ for $j \in \{1, 2, 3\}$. Thus we call any vector \mathbf{x} with $\mathbf{x} \cdot \mathbf{x} < 0$ a space-like vector. The time axis is described by the basis vector \mathbf{e}_0 , which fulfills $\mathbf{e}_0 \cdot \mathbf{e}_0 = 1$, and thus any vector \mathbf{x} with $\mathbf{x} \cdot \mathbf{x} > 0$ is called a time-like vector. Now any motion of a point-like object, moving along the \mathbf{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 $x_A = \lambda \mathbf{e}_0$ 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

$$\begin{pmatrix} x_B^0 \\ x_B^1 \end{pmatrix} = \begin{pmatrix} ct \\ \beta ct \end{pmatrix} \quad (1.2.9)$$

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

$$\mathbf{x}_B = \lambda \bar{\mathbf{e}}_0 = \lambda \mathbf{u}_B, \quad \lambda \in \mathbb{R}. \quad (1.2.10)$$

The unit vector

$$(\mathbf{u}_B^\mu) = \gamma \begin{pmatrix} 1 \\ \beta \end{pmatrix} \quad \text{with} \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad (1.2.11)$$

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

$$\bar{\mathbf{e}}_1 \cdot \bar{\mathbf{e}}_0 = 0, \quad \bar{\mathbf{e}}_1 \cdot \bar{\mathbf{e}}_1 = -1 \quad (1.2.12)$$

In Alice's coordinates it is obviously given by

$$\bar{\mathbf{e}}_1^\mu = \gamma \begin{pmatrix} \beta \\ 1 \end{pmatrix}. \quad (1.2.13)$$

The unit vectors $\bar{\mathbf{e}}_\mu$ 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 \mathbf{x} and \mathbf{y} happening at the same place for Alice at times t_x and t_y . The time difference according to Alice and Bob is given by

$$\Delta t_A = \frac{1}{c} \mathbf{u}_A \cdot (\mathbf{y} - \mathbf{x}) = t_y - t_x, \quad \Delta t_B = \frac{1}{c} \mathbf{u}_B \cdot (\mathbf{y} - \mathbf{x}) = \gamma \Delta t_{\text{Alice}}, \quad (1.2.14)$$

1.2. The special-relativistic space-time model

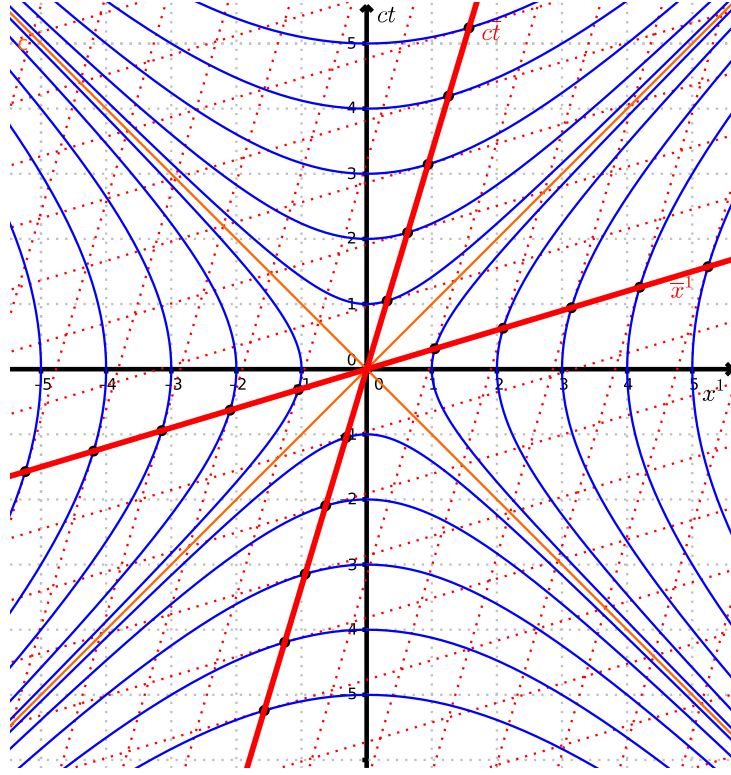


Figure 1.1: *Minkowski diagram, illustrating the construction of Bob's space-time coordinates: The thick red lines are Bob's $c\bar{t}$ and \bar{x}^1 axis. The corresponding units are given by the blue hyperbolas $(ct)^2 - (x^1)^2 = \pm n^2$ with $n \in \{1, 2, \dots\}$. We also draw the corresponding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

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

$$(\bar{y} - \bar{x})^1 = -\bar{e}_1 \cdot (\mathbf{y} - \mathbf{x}) = -\beta c \gamma \Delta t_A = -\beta c \Delta t_B. \quad (1.2.15)$$

This is indeed trivial, because during the time Δt_B between the events Bob travels the distance $\beta c \Delta t_B$ 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

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

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

$$(y - x)^\mu = \begin{pmatrix} 0 \\ y - x \end{pmatrix} \Rightarrow L_A = |y^1 - x^1|. \quad (1.2.17)$$

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

1. Kinematics

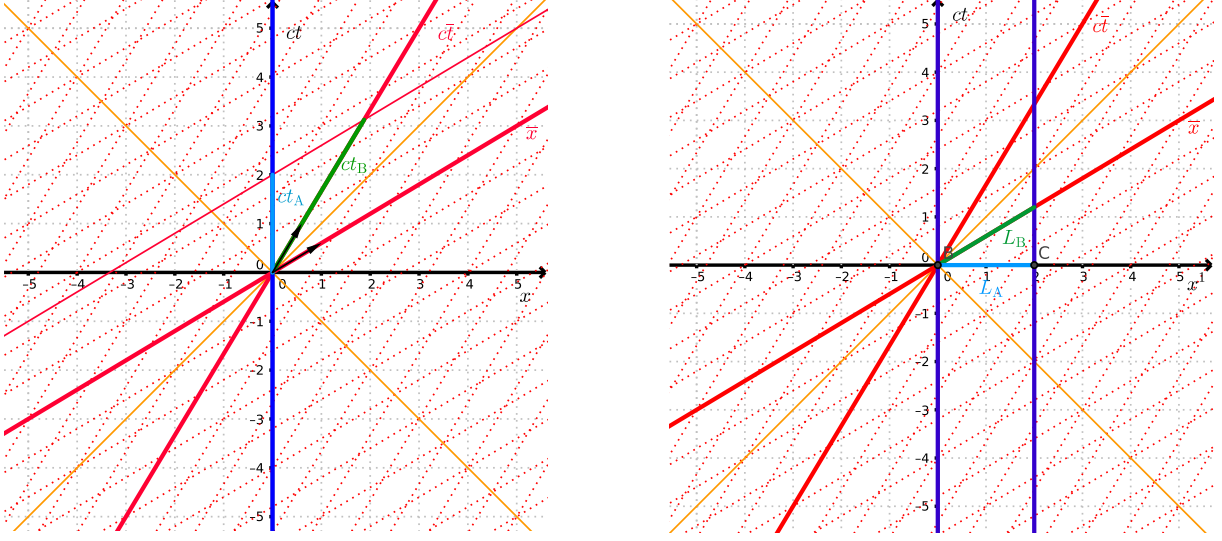


Figure 1.2: **Left:** Minkowski diagram illustrating the time-dilation effect. The time between two events at the same place $x^1 = 0$ in Alice's frame appears longer as measured by Bob, for whom the events are of course also not taking place at the same position. **Right:** Minkowski diagram illustrating 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 $L_A = 2 \text{ a.u.}$ 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 2 a.u. mark, labeled as $\bar{2}$ in his reference frame, one realizes that $L_B < 2 \text{ a.u.}$ 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.

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 $y - x = L_A > 0$):

$$\begin{aligned} \Delta t_B &= \mathbf{u}_B \cdot (\mathbf{y} - \mathbf{x}) = \gamma[(\lambda_2 - \lambda_1) - \beta L_A] \stackrel{!}{=} 0 \\ \Rightarrow \lambda_2 - \lambda_1 &= \beta L_A. \end{aligned} \quad (1.2.18)$$

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

$$L_{\text{Bob}} = -\bar{\mathbf{e}}_1 \cdot (\mathbf{y} - \mathbf{x}) = \gamma L_A (1 - \beta^2) = \frac{L_A}{\gamma} = \sqrt{1 - \beta^2} L_A. \quad (1.2.19)$$

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 \mathbf{u} , in Alice's coordinates its coordinates are obviously given by $(u^\mu) = (1, 0)^T$, 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

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

1.2. The special-relativistic space-time model

and thus the proper length of the rod by

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

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

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

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

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

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

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

and the spatial components as measured by Bob by

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

Here we have made use of the fact that we can choose as the additional two spatial vectors \bar{e}_2 and \bar{e}_3 for Bob Alice's spatial vectors e_2 and e_3 , because they are perpendicular to both \bar{e}_0 and \bar{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

$$L_B^2 = \frac{(y^1 - x^1)^2}{\gamma^2} + (y^2 - x^2)^2 + (y^3 - x^3)^2. \quad (1.2.26)$$

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.,

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

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

$$\bar{t} = \frac{1}{c} \underline{u}_B \cdot (\underline{y} - \underline{x}) = \frac{\gamma\beta}{c} (y^1 - x^1). \quad (1.2.28)$$

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,

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

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

$$\tau = \frac{1}{c} \int_0^t dt' \sqrt{\frac{dx^\mu(t')}{dt'} \frac{dx^\mu(t')}{dt'}} = \int_0^t dt' \sqrt{1 - \frac{\omega^2 R^2}{c^2}} = t \sqrt{1 - \frac{\omega^2 R^2}{c^2}}. \quad (1.3.2)$$

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 factor³.

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

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

³For references, see the Wikipedia entry on this subject:
https://en.wikipedia.org/wiki/Time_dilation_of_moving_particles

1.4. General Lorentz transformations

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 x

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

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

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

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

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

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

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

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

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

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

$$\eta_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu = \eta_{\mu\nu} \quad (1.4.7)$$

or in matrix notation

$$\hat{\Lambda}^T \hat{\eta} \hat{\Lambda} = \hat{\eta}. \quad (1.4.8)$$

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

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

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

$$\hat{R} = \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \hat{O} \end{pmatrix}, \quad (1.4.10)$$

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

1. Kinematics

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

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

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 \mathbf{u}_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 $\vec{\beta}_B = \vec{u}_B / u_B^0$ points along her new 1-Axis, leading to the new basis system, where Alice is still at rest

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

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

$$\mathbf{e}''_\mu = (B^{-1}(\beta \vec{e}'_1))^\rho{}_\mu \mathbf{e}'_\rho. \quad (1.4.13)$$

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 \mathbf{e}''_μ 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

$$\bar{\mathbf{e}}_\mu = (R_1^{-1})^\rho{}_\mu \mathbf{e}''_\rho. \quad (1.4.14)$$

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

$$\bar{\mathbf{e}}_\mu = (\hat{R}_2^{-1} \hat{B}(\beta \vec{e}'_1) \hat{R}_1^{-1})^\nu{}_\mu \mathbf{e}_\nu \Rightarrow \mathbf{e}_\mu = (\hat{R}_1 \hat{B}(\beta \vec{e}_1) \hat{R}_2)^\nu{}_\mu \bar{\mathbf{e}}_\nu = \Lambda^\nu{}_\mu \bar{\mathbf{e}}_\nu. \quad (1.4.15)$$

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

1.5. Addition of velocities

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

$$\Lambda^0_0 = \gamma = \cosh \eta, \quad \Lambda^0_1 = \Lambda^1_0 = -\beta\gamma = -\sinh \eta, \quad \Lambda^1_1 = \cosh \eta. \quad (1.4.16)$$

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

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

and the Lorentz matrix reads

$$(B^\mu_{\nu}) = \begin{pmatrix} \cosh \eta & -\sinh \eta & 0 & 0 \\ -\sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.4.18)$$

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

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

where we have made use of the identities

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

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

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

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 \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

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

1. Kinematics

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

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

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

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

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

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

which can be rewritten after some algebra as

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

For the magnitude we find, again after some algebra,

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

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 1⁴. To obtain the relative velocity we use the same idea as in the previous section. Let

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

1.7. The Lorentz group as a Lie group

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

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

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

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

Indeed, it is easily seen by direct calculation that

$$(\tilde{U}^\mu) = (\Lambda^\mu{}_\nu U^\nu) = (1, 0, 0, 0)^\top. \quad (1.6.3)$$

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

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

and from (1.5.6)

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

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

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \delta\omega^\mu{}_\nu. \quad (1.7.1)$$

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 $\delta\hat{\omega}$:

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

where we have used the usual rule for lowering indices,

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

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

$$\delta\omega_{\rho\sigma} = -\delta\omega_{\sigma\rho}. \quad (1.7.4)$$

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.

1. Kinematics

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

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

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

$$\delta \omega^0{}_j = \eta^{0\rho} \delta \omega_{\rho j} = \delta \omega_{0j}, \quad \delta \omega^j{}_0 = \eta^{j\rho} \delta \omega_{\rho 0} = \eta^{jk} \delta \omega_{k0} = -\omega_{j0} = +\omega_{0j}. \quad (1.7.6)$$

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*

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

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

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

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 $\vec{\beta} = 0$:

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

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

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

we indeed simply have

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

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

$$\delta \omega^j{}_k = \eta^{j\rho} \delta \omega_{\rho k} = -\delta \omega_{jk}, \quad \Rightarrow \delta \omega^j{}_k = -\delta \omega_{jk} = \delta \omega_{kj} = -\delta \omega^k{}_j, \quad (1.7.12)$$

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

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

1.7. The Lorentz group as a Lie group

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 $-\delta\varphi\vec{n}\cdot\hat{J}$. This we can write in terms of a differential equation for

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

Assuming that our idea is correkt this differential equation should read

$$\frac{d}{d\varphi}x' = -\vec{n}\cdot\hat{J}x. \quad (1.7.15)$$

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

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

The matrix-exponential function is defined by the corresponding series

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

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

$$(\vec{n}\cdot\hat{J})^j_k = n^i \epsilon^{ijk}, \quad (1.7.18)$$

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

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

So we have

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

and

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

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

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

Applying this to a four-vector,

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

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

1. Kinematics

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

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

with

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

This implies

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

Then the exponential series for a boost reads

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

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

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

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

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

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:

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

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

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

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

$$ds = c d\tau = \sqrt{dx^\mu dx_\mu}. \quad (1.8.1)$$

Then the four-velocity is given by

$$u^\mu = \frac{d}{ds} x^\mu. \quad (1.8.2)$$

1.8. Fermi-Walker transport and Thomas precession

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 $e'_0 = \mathbf{u}$ 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 $s + ds$ 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 $\mathbf{n}'_{(j)}(s_0) = \mathbf{u}(s_0)$ and $\mathbf{n}'_{(j)}(s_0)$ ($j \in \{1, 2, 3\}$) such that $\mathbf{n}'_{(\mu)}(s_0) \cdot \mathbf{n}'_{(\nu)}(s_0) = \eta_{\mu\nu}$ and then defines $\mathbf{n}_\mu(s)$ 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 \mathbf{u} within an infinitesimal step ds changes by

$$d\mathbf{u} = ds \mathbf{a}, \quad \mathbf{a} = \frac{d}{ds} \mathbf{u}. \quad (1.8.3)$$

Obviously it obeys

$$\mathbf{u} \cdot \mathbf{a} = 0, \quad (1.8.4)$$

because $\mathbf{u} \cdot \mathbf{u} = 1 = \text{const}$, and thus is spacelike and Minkowski orthogonal to \mathbf{u} . Now the vector $\mathbf{n}(s)$ 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 \mathbf{u} . Obviously this is achieved if we demand that \mathbf{n} is transported along the worldline from s to ds such that it undergoes an infinitesimal rotation-free Lorentz boost in the \mathbf{u} - \mathbf{a} plane, i.e., according to (1.7.7) and (1.7.8) such that

$$dn^\mu = ds(a^\mu u_\nu - u^\mu a_\nu) n^\nu \quad (1.8.5)$$

or

$$\frac{d\mathbf{n}}{ds} = \mathbf{a}(\mathbf{u} \cdot \mathbf{n}) - \mathbf{u}(\mathbf{a} \cdot \mathbf{n}). \quad (1.8.6)$$

For $\mathbf{u} \cdot \mathbf{n} = 0$ it is immediately clear that $d\mathbf{n} = -ds \mathbf{u}(\mathbf{a} \cdot \mathbf{n})$ 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 \mathbf{u} itself one gets $d\mathbf{u}/ds = \mathbf{a}$ as it must be since that is the definition of \mathbf{a} . So with Fermi-Walker transporting $\mathbf{u}(s_0)$ leads to $\mathbf{u}(s)$ when integrating the equation (1.8.6) for $\mathbf{n} = \mathbf{u}$. Further for any \mathbf{n} Fermi-Walker transport does not change the Minkowski length of this vector, because multiplying (1.8.6) with \mathbf{n} leads to

$$\mathbf{n} \cdot \frac{d\mathbf{n}}{ds} = \frac{1}{2} \frac{d}{ds} (\mathbf{n} \cdot \mathbf{n}) = 0. \quad (1.8.7)$$

1. Kinematics

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

$$\frac{d}{ds}(\mathbf{n} \cdot \mathbf{u}) = \frac{d\mathbf{n}}{ds} \cdot \mathbf{u} + \mathbf{n} \cdot \frac{d\mathbf{u}}{ds} = [\mathbf{a}(\mathbf{u} \cdot \mathbf{n}) - \mathbf{u}(\mathbf{a} \cdot \mathbf{n})] \cdot \mathbf{u} + \mathbf{n} \cdot \mathbf{a}. \quad (1.8.8)$$

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

$$\frac{d}{ds}(\mathbf{n} \cdot \mathbf{u}) = 0. \quad (1.8.9)$$

Thus, the Fermi-Walker transport defined for an arbitrary set of vectors $\mathbf{n}(s)$ 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 Fermi-Walker transported. Since by definition spin is Minkowski-orthogonal to \mathbf{u} its equation of motion is given by (1.8.6),

$$\frac{d\Sigma}{d\tau} = -\mathbf{u} \left(\frac{d\mathbf{u}}{d\tau} \cdot \Sigma \right), \quad (1.8.10)$$

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

$$\mathbf{u} \cdot \Sigma = u \cdot \Sigma = 0. \quad (1.8.11)$$

Now we want to evaluate $d\vec{\Sigma}^*/d\tau$ 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

$$\Sigma^0 = \frac{\vec{u}}{u^0} \cdot \vec{\Sigma} = \vec{\beta} \cdot \vec{\Sigma}, \quad (1.8.12)$$

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 inertial 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 :

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

We note that

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

Plugging this into (1.8.13) after some algebra one finds

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

1.8. Fermi-Walker transport and Thomas precession

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

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

For the final expression we have used

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

Thus we have with (1.8.15)

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

where we have used

$$\dot{\gamma} = \frac{d}{d\tau} \frac{\gamma}{1+\gamma} = \frac{\dot{\gamma}}{(1+\gamma)^2} \stackrel{(1.8.14)}{=} \frac{\gamma^3}{(1+\gamma)^2} \dot{\vec{\beta}} \cdot \vec{\beta}. \quad (1.8.19)$$

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

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

Further we have, again with (1.8.16)

$$\dot{\vec{\beta}} \cdot \vec{\Sigma}^* = \dot{\vec{\beta}} \cdot \vec{\Sigma} - (\vec{\beta} \cdot \dot{\vec{\beta}}) (\vec{\beta} \cdot \vec{\Sigma}) \frac{\gamma}{\gamma+1}. \quad (1.8.21)$$

Now we use this for the bracket in the first term in (1.8.18) and apply (1.8.20) in the second term, and finally we obtain

$$\dot{\Sigma}^* = \frac{\gamma^2}{\gamma+1} [(\dot{\vec{\beta}} \cdot \vec{\Sigma}^*) \vec{\beta} - (\vec{\beta} \cdot \dot{\vec{\beta}}) \vec{\Sigma}^*] = \frac{\gamma^2}{\gamma+1} (\dot{\vec{\beta}} \times \vec{\beta}) \times \vec{\Sigma}^* = \vec{\Omega}_T \times \vec{\Sigma}^*, \quad (1.8.22)$$

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

$$\vec{\Omega}_T = \frac{\gamma^2}{\gamma+1} \dot{\vec{\beta}} \times \vec{\beta} \underset{|\vec{\beta}| \ll 1}{\cong} \frac{1}{2} \dot{\vec{\beta}} \times \vec{\beta}. \quad (1.8.23)$$

The reason for this rotation is that $\vec{\beta}(\tau)$ and $\vec{\beta}(\tau + d\tau)$ need not be collinear, and thus the Lorentz boosts $\Lambda_B[\vec{\beta}(\tau)]$ can be seen built from infinitesimal Lorentz boosts in different directions, but these lead not to a pure boost again but to a boost followed by a rotation, the so-called **Wigner rotation**, and this additional rotation figures into (1.8.22) leading to the Thomas precession with angular velocity (1.8.23).

1. Kinematics

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 concept 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 19th and the beginning of the 20th 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. 4.1).

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

2. Mechanics

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

$$ds = c d\tau = dt \sqrt{\eta_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}}, \quad (2.1.1)$$

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

$$\frac{dx}{d\tau} = c \frac{dx}{ds} = c \mathbf{u}, \quad (2.1.2)$$

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

$$\frac{dx}{d\tau} \cdot \frac{dx}{d\tau} = c^2, \quad (2.1.3)$$

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 2nd law in a covariant way. To this end we define the **four-momentum** of the particle as

$$\mathbf{p} = m c \mathbf{u}, \quad (2.1.4)$$

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

$$(p^\mu) = m c \gamma \begin{pmatrix} 1 \\ \vec{\beta} \end{pmatrix}, \quad (2.1.5)$$

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

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

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):

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

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

2.1. Particle dynamics

or defining

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

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

$$E_0 = mc^2 \quad (2.1.9)$$

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

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

Since $\gamma \rightarrow \infty$ for $\beta \rightarrow 1$ 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 2nd Law for the motion of a particle in a four-dimensionally covariant way as

$$\frac{d}{d\tau}p^\mu = mc \frac{d}{d\tau}u^\mu = m \frac{d^2}{d\tau^2}x^\mu = K^\mu, \quad (2.1.11)$$

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

$$\frac{d\mathbf{u}}{d\tau} \cdot \mathbf{u} = 0, \quad (2.1.12)$$

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

$$\mathbf{K} \cdot \mathbf{u} = \eta_{\mu\nu}K^\mu u^\nu = 0. \quad (2.1.13)$$

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

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

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

$$mc \frac{du^0}{d\tau} = K^0 = \frac{\vec{u}}{u^0} \cdot \vec{K} = \vec{\beta} \cdot \vec{K}. \quad (2.1.15)$$

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

$$mc \frac{du^0}{dt} = \vec{\beta} \cdot \frac{\vec{K}}{u^0} \quad (2.1.16)$$

2. Mechanics

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)

$$\frac{dE}{dt} = \frac{dE_{\text{kin}}}{dt} = \vec{v} \cdot \vec{F}. \quad (2.1.17)$$

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

$$\frac{d}{dt} \vec{p} = \frac{1}{u_0} \frac{d}{d\tau} \vec{p} = \frac{1}{u_0} \vec{K} = \vec{F}, \quad (2.1.18)$$

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

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

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

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

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

$$ds^2 = \eta_{\mu\nu} dt^2 \frac{dx^\mu}{dt} \frac{dx^\nu}{dt} = dt^2 (c^2 - \vec{v} \cdot \vec{v}) = 0 \Rightarrow d\tau = 0, \quad (2.1.21)$$

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

$$\frac{dp^\mu}{d\lambda} = K^\mu. \quad (2.1.22)$$

2.2. Motion of a particle in an electromagnetic field

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

$$p_\mu K^\mu = 0. \quad (2.1.23)$$

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

$$\frac{d\vec{p}}{dt} = \vec{K} \quad (2.1.24)$$

and close the equations of motion simply via (2.1.20)

$$\frac{d\vec{x}}{dt} = \vec{v} = c \frac{\vec{p}}{|\vec{p}|}. \quad (2.1.25)$$

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. 4.1 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

$$\bar{F}^{\mu\nu}(\bar{x}^\rho) = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}(x^\gamma) \quad \text{with} \quad x^\gamma = (\Lambda^{-1})^\gamma_\rho \bar{x}^\rho. \quad (2.2.1)$$

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

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

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} :

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

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

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

2. Mechanics

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

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

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

$$\frac{1}{u_0} \frac{d}{d\tau} \vec{p} = \frac{d}{dt} \vec{p} = q \vec{E} + \frac{\vec{v}}{c} \times \vec{B}, \quad (2.2.6)$$

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):

$$\vec{u} = \frac{1}{mc} \vec{p} = \frac{1}{c} \frac{d}{d\tau} \vec{x}. \quad (2.2.7)$$

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

$$\frac{d}{dt} \vec{x} = c \frac{\vec{p}}{p^0}. \quad (2.2.8)$$

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

$$\vec{v} \cdot \frac{d}{dt} \vec{p} = q \vec{v} \cdot \vec{E}, \quad (2.2.9)$$

but this is precisely the temporal component of the first equation in (2.2.5) divided by u^0 . So the ‘‘on-shell’’ condition $p_\mu p^\mu = m^2 c^2 = \text{const}$ is compatible with the equations of motion by construction, and thus after solving (2.2.6) we can evaluate $p^0 = \sqrt{m^2 c^2 + \vec{p}^2}$ 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 $\vec{E} = (E, 0, 0)^T$ 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)

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

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

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

2.2. Motion of a particle in an electromagnetic field

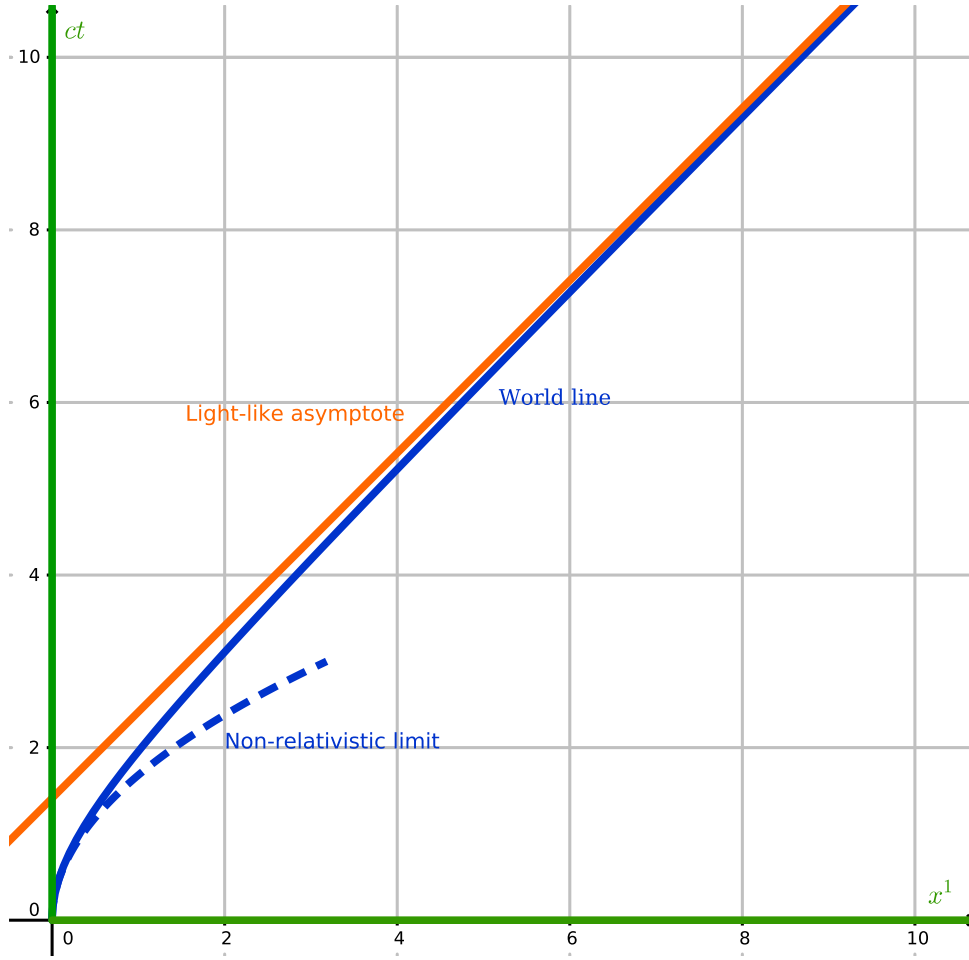


Figure 2.1: Minkowski diagram for the “hyperbolic motion” of a charged particle due to a homogeneous electric field. The corresponding world line has a light-like asymptote, which indicates that even a light signal sent from x^1 at a time $t >$ cannot reach an observer being accelerated with the charged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

and thus

$$v^1(t) = \frac{dx^1}{dt} = \frac{c p^1}{p^0} = \frac{q c E t}{\sqrt{m^2 c^2 + q^2 E^2 t^2}}. \quad (2.2.12)$$

This is easily integrated to

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

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 $|q E t| \ll m c$ or $t \ll m c / (q E)$. Then we can expand the square root in (2.2.13), leading to

$$x^1(t) = \frac{m c^2}{q E} \left(\frac{q^2 E^2 t^2}{2 m^2 c^2} + \mathcal{O} \left[\left(\frac{q E t}{m c} \right)^4 \right] \right) = \frac{q E t^2}{2 m} + \dots \quad (2.2.14)$$

2. Mechanics

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

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

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 $t_H = \frac{mc}{qE}$ (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

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

The 3-component is immediately integrated

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

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

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

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

$$\xi(\tau) = \xi_0 \exp(-i\omega\tau). \quad (2.2.19)$$

Separating back into real and imaginary part gives

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

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

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

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

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

2.3. Bell's space-ship paradox

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

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

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

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

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

$$\frac{dt}{d\tau} = u^0 \Rightarrow t = \tau \sqrt{1 + \vec{u}_0^2} = \text{const.} \quad (2.2.25)$$

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

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

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

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

and substitution of (2.2.27) in (2.2.22)

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

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

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

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

²I 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.

2. Mechanics

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 were 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. 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 $p^\mu = mc u^\mu$ between momentum and four-velocity That means we can write (2.2.5) as

$$\frac{du^0}{d\tau} = \frac{\alpha}{c} u^1, \quad \frac{du^1}{d\tau} = \frac{\alpha}{c} u^0. \quad (2.3.1)$$

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 2nd equation (2.4.1) and use the first one

$$\frac{d^2 u^1}{d\tau^2} = \frac{\alpha}{c} \frac{du^0}{d\tau} = \frac{\alpha^2}{c^2} u^1. \quad (2.3.2)$$

2.3. Bell's space-ship paradox

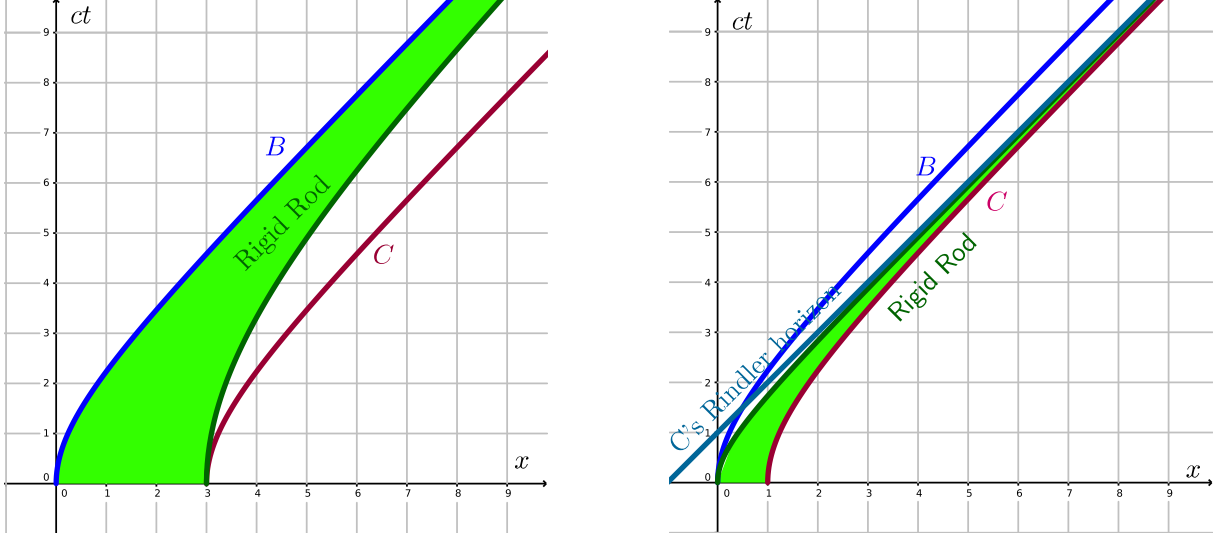


Figure 2.2: Minkowski diagram illustrating Bell's space-ship paradox. **Left:** Comparing the world tube of a rigid rod of proper length L_A attached to the front of spaceship B to that of spaceship C , one sees that the front end of the rod always stays behind spaceship C , i.e., a real elastic string 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

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

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

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

Thus our solution reads

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

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

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

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

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

2. Mechanics

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.,

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

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

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

Indeed doing the matrix multiplication explicitly, one immediately finds

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

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 instantaneous restframe, i.e.,

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

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

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

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

$$\beta_{\text{rod}}(\tau_0) = \frac{\dot{x}_{\text{rod}}^1(\tau_0)}{\dot{x}_{\text{rod}}^0(\tau_0)} = \tanh\left(\frac{\alpha\tau_0}{c}\right) \quad (2.3.13)$$

and thus

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

Finally the proper acceleration is given by

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

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

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

2.4. The action principle

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

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

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 proper acceleration $\alpha_{\text{rod}} \rightarrow \infty$, 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 asymptotic 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

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

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**

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

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

2. Mechanics

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

$$\frac{d}{dt} \delta q = \delta \dot{q}. \quad (2.4.2)$$

We note that in the following taking *partial* derivatives of functions of q , \dot{q} , and t means that we treat the variables t , q , and \dot{q} 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

$$\delta A = \int_{t_1}^{t_2} dt \left[\delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} \right]. \quad (2.4.3)$$

Now using (2.4.2) we get by integration by parts

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

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

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

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:

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

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.

2.4. The action principle

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 $dt f(\dot{\vec{x}})$ that is invariant. Obviously the only covariant expression $\propto dt$ we can build from the time derivatives of \vec{x} is

$$dt \sqrt{\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} = dt c \sqrt{1 - \vec{\beta}^2}, \quad \vec{\beta} = \frac{1}{c} \dot{\vec{x}}. \quad (2.4.7)$$

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

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

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

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

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

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

This gives

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

Since then also

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

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

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

With the full Lagrangian

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

2. Mechanics

we find the equations of motion via

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

where $\vec{A} = (A^i)^3$. Further we get

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

and

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

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

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

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

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

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

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

we find

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

and thus

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

Also setting

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

finally the equations of motion (2.4.18) read

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

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).

³Here 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

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

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

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

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**

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} = \frac{\partial L}{\partial x^\mu}. \quad (2.4.27)$$

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

$$mc \frac{d}{d\lambda} \frac{\dot{x}_\mu}{\sqrt{\dot{x}^\alpha \dot{x}_\alpha}} = \frac{q}{c} F_{\alpha\mu} \dot{x}^\alpha \quad \text{with} \quad F_{\alpha\mu} = \partial_\alpha A_\mu - \partial_\mu A_\alpha, \quad (2.4.28)$$

where we have introduced the partial derivative,

$$\partial_\mu = \frac{\partial}{\partial x^\mu}. \quad (2.4.29)$$

Here it is important to realize that the partial derivative wrt. contravariant vector components (with upper indices) leads to covariant tensor-field components (with lower indices). For a more detailed explanation, see App. A.

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

$$\dot{x}^\mu \frac{\partial L}{\partial \dot{x}^\mu} = L, \quad (2.4.30)$$

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

$$\ddot{x}^\mu \frac{\partial L}{\partial \dot{x}^\mu} + \dot{x}^\mu \frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} = \dot{x}^\mu \frac{\partial L}{\partial x^\mu} + \ddot{x}^\mu \frac{\partial L}{\partial \dot{x}^\mu} \Rightarrow \dot{x}^\mu \frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} = \dot{x}^\mu \frac{\partial L}{\partial x^\mu}. \quad (2.4.31)$$

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

$$\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = \dot{x}_\mu \dot{x}^\mu = c^2, \quad (2.4.32)$$

2. Mechanics

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

$$m \ddot{x}_\alpha = \frac{q}{c} F_{\alpha\mu} \dot{x}^\mu. \quad (2.4.33)$$

This formalism can be generalized to find the equation of motion in other kinds of external fields by making the ansatz

$$L = -mc \sqrt{\eta_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma} + L_{\text{int}}(\underline{x}, \underline{\dot{x}}) \quad (2.4.34)$$

with L_{int} being also a homogeneous function in the $\underline{\dot{x}}$ of degree 1, i.e.,

$$\dot{x}^\rho \frac{\partial L_{\text{int}}}{\partial \dot{x}^\rho} = L_{\text{int}}. \quad (2.4.35)$$

As shown above in (2.4.31) this implies

$$\dot{x}^\rho \frac{d}{d\lambda} \frac{\partial L_{\text{int}}}{\partial \dot{x}^\rho} = \dot{x}^\rho \frac{\partial L_{\text{int}}}{\partial x^\rho}. \quad (2.4.36)$$

Now the Euler-Lagrange equations of the extended Hamiltonian principle reads

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} = -mc \frac{d}{d\lambda} \frac{\dot{x}_\mu}{\sqrt{\eta_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma}} + \frac{d}{d\lambda} \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} = \frac{\partial L}{\partial x^\mu} = \frac{\partial L_{\text{int}}}{\partial x^\mu}. \quad (2.4.37)$$

Now we have

$$\sqrt{\eta_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma} = c \dot{\tau} \quad (2.4.38)$$

and thus

$$mc \frac{d}{d\lambda} \frac{\dot{x}_\mu}{\sqrt{\eta_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma}} = \frac{dp_\mu}{d\lambda} \stackrel{(2.4.38)}{=} \frac{d}{d\lambda} \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} - \frac{\partial L_{\text{int}}}{\partial x^\mu}. \quad (2.4.39)$$

Multiplying this with \dot{x}^μ and using (2.4.38) we have

$$\dot{x}^\mu \frac{dp_\mu}{d\lambda} = 0. \quad (2.4.40)$$

Dividing this with $\dot{\tau}^2$ we find this implies that

$$\frac{dx^\mu}{d\tau} \frac{dp_\mu}{d\tau} = \frac{1}{m} p^\mu \frac{dp_\mu}{d\tau} = 0, \quad (2.4.41)$$

which makes the equation of motion (2.4.39) automatically consistent with the on-shell condition

$$p_\mu p^\mu = m^2 c^2, \quad (2.4.42)$$

and we can use τ as the parameter of the worldline in (2.4.39).

As another example for a covariant equation of motion we consider the motion of a particle in an external **scalar field**. For the interaction Lagrangian we have two choices to be a homogeneous function of order 1 of \dot{x}^μ ,

$$L_1 = -g \sqrt{\dot{x}_\mu \dot{x}^\mu} \Phi(x), \quad L_2 = -g' \dot{x}^\mu \partial_\mu \Phi = \frac{d}{d\lambda} \Phi, \quad (2.4.43)$$

because we want the Lagrangian to be a Lorentz scalar itself and the action parametrization invariant.

2.4. The action principle

The second choice, however does not lead to any consequence in the equation of motion, because it is a total derivative of a function of the coordinates only wrt. λ , and this part of the Lagrangian leads to a contribution of the action

$$A_2[x^\mu] = -g' \{ \Phi[x(\lambda_2)] - \Phi[x(\lambda_1)] \} \Rightarrow \delta A_2 = -g' \left[\delta x^\mu \partial_\mu \Phi(x) \right]_{\lambda_1}^{\lambda_2} \equiv 0. \quad (2.4.44)$$

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

$$L = -\sqrt{\dot{x}_\mu \dot{x}^\mu} [mc + g\Phi(x)]. \quad (2.4.45)$$

The evaluation of the Euler-Lagrange equations lead, after setting $\lambda = \tau$, to the equations of motion

$$(m + g\Phi)\dot{x}^\mu = \frac{g}{c}(c^2\eta^{\mu\nu} - \dot{x}^\mu \dot{x}^\nu)\partial_\nu \Phi, \quad (2.4.46)$$

where now the dot means the derivative wrt. τ . Obviously by contraction with $\dot{x}_\mu = c u_\mu$ one indeed finds that the constraint $u_\mu u^\mu = 1 = \text{const}$ is compatible with this equation of motion since

$$\dot{x}_\mu (c^2\eta^{\mu\nu} - \dot{x}^\mu \dot{x}^\nu) = 0, \quad (2.4.47)$$

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 $\propto \sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}$), which simplifies the analysis greatly.

To that end we write as an alternative Lagrangian instead of (2.4.34):

$$L = -\frac{m}{2}\eta_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma + L_{\text{int}}(\underline{x}, \underline{\dot{x}}) \quad (2.4.48)$$

with the same interaction Lagrangian, L_{int} , as in (2.4.34), which is a homogeneous function in the $\underline{\dot{x}}$ of degree 1, i.e., fulfilling (2.4.35).

Then the Euler-Lagrange equation reads

$$\frac{d}{d\lambda} \frac{\partial L}{\partial \dot{x}^\mu} = -m\ddot{x}_\mu + \frac{d}{d\lambda} \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} = \frac{\partial L_{\text{int}}}{\partial x^\mu}. \quad (2.4.49)$$

Comparing this with the equation of motion for the original Lagrangian (2.4.37) we see that we only get the same equation of motion, if

$$\eta_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma = \text{const}. \quad (2.4.50)$$

But this constraint is always fulfilled, because since L does not depend explicitly on the world-line parameter λ , the Hamilton-function-like quantity

$$\tilde{H} = P_\mu \dot{x}^\mu - L = \text{const.}, \quad (2.4.51)$$

for the solutions of the equations of motion, where

$$P_\mu = \frac{\partial L}{\partial \dot{x}^\mu} = -m\dot{x}_\mu + \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} \quad (2.4.52)$$

2. Mechanics

is the canonical momentum. Because of (2.4.35), (2.4.52) indeed reads

$$\tilde{H} = -m\dot{x}_\mu\dot{x}^\mu + \frac{m}{2}\dot{x}_\mu\dot{x}^\mu + \dot{x}^\mu \frac{\partial L_{\text{int}}}{\partial \dot{x}^\mu} - L_{\text{int}} = -\frac{m}{2}\dot{x}_\mu\dot{x}^\mu = -\frac{m}{2}\eta_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma = \text{const.} \quad (2.4.53)$$

Now we can choose the value of this quantity as we like, and if we choose

$$\eta_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma = c^2, \quad (2.4.54)$$

then we have again chosen $\lambda = \tau$.

2.5 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 fields

3.1 The action principle for fields and Noether's Theorem

In this section we shall establish **Hamilton's least-action principle for fields** in its formulation by **Lagrange**, which has the advantage to be manifestly covariant under Poincaré transformations. 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.

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

$$L = \int_{\mathbb{R}^3} d^3x \mathcal{L}(\Phi_i, \partial_\mu \Phi_i; t, \vec{x}). \quad (3.1.1)$$

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

$$A[\Phi_i] = \int_{t_1}^{t_2} dt L[\Phi_i, t] = \frac{1}{c} \int_{V^{(4)}} d^4x \mathcal{L}(\Phi_i, \partial_\mu \Phi_i; t, \vec{x}). \quad (3.1.2)$$

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.

3. Classical fields

The variation is then easily calculated as

$$\begin{aligned} c\delta A[\Phi_i] &= \int_{V^{(4)}} d^4x \left[\delta\Phi_i \frac{\partial \mathcal{L}}{\partial \Phi_i} + \delta(\partial_\mu \Phi_i) \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_i)} \right] \\ &= \int_{V^{(4)}} d^4x \delta\Phi_i \left[\frac{\partial \mathcal{L}}{\partial \Phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_i)} \right]. \end{aligned} \quad (3.1.3)$$

Here the Einstein summation relation applies to both the indices of space-time components 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**

$$\frac{\partial \mathcal{L}}{\partial \Phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_i)} = 0. \quad (3.1.4)$$

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

$$\mathcal{L}' = \mathcal{L} + \partial_\mu \Omega^\mu(\Phi_i, x) \quad (3.1.5)$$

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

$$\Delta \mathcal{L} = \partial_\mu \Omega^\mu = \partial_\mu \Phi_i \frac{\partial \Omega^\mu}{\partial \Phi_i} + [\partial_\mu \Omega^\mu(\Phi_i, x)]_{\text{expl}}. \quad (3.1.6)$$

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

$$\frac{\partial \Delta \mathcal{L}}{\partial (\partial_\mu \Phi_i)} = \frac{\partial \Omega^\mu}{\partial \Phi_i} \Rightarrow \partial_\mu \frac{\partial \Delta \mathcal{L}}{\partial (\partial_\mu \Phi_i)} = \partial_\mu \frac{\partial \Omega^\mu}{\partial \Phi_i}. \quad (3.1.7)$$

On the other hand

$$\frac{\partial \Delta \mathcal{L}}{\partial \Phi_i} = \frac{\partial}{\partial \Phi_i} \partial_\mu \Omega^\mu = \partial_\mu \frac{\partial \Omega^\mu}{\partial \Phi_i}. \quad (3.1.8)$$

Eqs. (3.1.7) and (3.1.8) thus imply that

$$\frac{\partial \Delta \mathcal{L}}{\partial \Phi_i} - \partial_\mu \frac{\partial \Delta \mathcal{L}}{\partial (\partial_\mu \Phi_i)} \equiv 0, \quad (3.1.9)$$

i.e., the addition of $\Delta \mathcal{L}$ does not contribute to the variation of the action, and thus the Lagrangians \mathcal{L} and \mathcal{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 that it allows to formulate **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

3.1. The action principle for fields and Noether's Theorem

proper orthochronous Poincaré group. Let the group depend on parameters α^k . Then we can write an infinitesimal transformation of this kind in the form

$$x^\mu \rightarrow x'^\mu = x^\mu - \delta\alpha^k \xi_k^\mu(x), \quad \Phi_i \rightarrow \Phi'_i(x') = \Phi_i(x) + \delta\alpha^k \Xi_{ik}(\Phi_i, x). \quad (3.1.10)$$

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

$$\delta A = \int_{\mathbb{R}^4} d^4x' \mathcal{L}(\Phi'_i, \partial'_\mu \Phi'_i, x') - \int_{\mathbb{R}^4} d^4x \mathcal{L}(\Phi_i, \partial_\mu \Phi_i, x), \quad (3.1.11)$$

we note that

$$\delta(\partial_\mu \Phi_i) = \partial'_\mu \Phi'_i - \partial_\mu \Phi_i = \frac{\partial x^\nu}{\partial x'^\mu} \partial_\nu [\Phi_i + \delta\alpha^k \Xi_{ik}] - \partial_\mu \Phi_i. \quad (3.1.12)$$

Now we need the inverse matrix of

$$J^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} = \delta^\mu_\nu - \delta\alpha^k \partial_\nu \xi_k^\mu. \quad (3.1.13)$$

It is easy to see that up to linear order in the $\delta\alpha^k$

$$\frac{\partial x^\nu}{\partial x'^\mu} = \delta^\nu_\mu + \delta\alpha^k \partial_\mu \xi_k^\nu + \mathcal{O}(\delta\alpha^2). \quad (3.1.14)$$

To prove this we calculate, using (3.1.13)

$$J^\mu{}_\rho \frac{\partial x^\rho}{\partial x'^\sigma} = (\delta^\mu_\rho - \delta\alpha^k \partial_\rho \xi_k^\mu) (\delta^\rho_\sigma + \delta\alpha^k \partial_\sigma \xi_k^\rho) = \delta^\mu_\sigma + \mathcal{O}(\delta\alpha^2). \quad (3.1.15)$$

Thus (3.1.14) is verified. Plugging this into (3.1.12) we find after some algebra

$$\delta(\partial_\mu \Phi_i) = \delta\alpha^k [\partial_\mu \Xi_{ik} + \partial_\mu \xi_k^\nu \partial_\nu \Phi_i]. \quad (3.1.16)$$

Further the four-volume element transforms as

$$d^4x' = d^4x \det\left(\frac{\partial x'}{\partial x}\right) = d^4x [1 - \delta\alpha^k \partial_\mu \xi_k^\mu + \mathcal{O}(\delta\alpha^2)]. \quad (3.1.17)$$

The latter follows from taking the determinant of (3.1.12) and only keeping terms up to linear order in the $\delta\alpha^k$ which only occur in the product of the diagonal elements of the matrix. All other terms in the determinant obviously are at least of order $\mathcal{O}(\delta\alpha^2)$.

Plugging all this into (3.1.11) we find after some algebra

$$\delta A = \delta\alpha^k \int_{\mathbb{R}^4} d^4x \left[\frac{\partial \mathcal{L}}{\partial \Phi_i} \Xi_{ik} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_i)} [\partial_\mu \Xi_{ik} + \partial_\mu \xi_k^\nu \partial_\nu \Phi_i] - \xi_k^\mu (\partial_\mu \mathcal{L})_{\text{expl}} - \mathcal{L} \partial_\mu \xi_k^\mu \right]. \quad (3.1.18)$$

Since the $\delta\alpha^k$ are independent of each other and because $\delta A = 0$ if (3.1.10) 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

$$\frac{\partial \mathcal{L}}{\partial \Phi_j} \Xi_{jk} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} [\partial_\mu \Xi_{jk} + \partial_\mu \xi_k^\nu \partial_\nu \Phi_j] - \xi_k^\mu (\partial_\mu^{\text{expl}} \mathcal{L}) - \mathcal{L} \partial_\mu \xi_k^\mu + \partial_\mu \Omega_k^\mu(\Phi_j, x) = 0. \quad (3.1.19)$$

3. Classical fields

For abbreviation we define the **canonical energy-momentum tensor**,

$$\Theta^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \partial_\nu \Phi_j - \mathcal{L} \delta^\mu{}_\nu. \quad (3.1.20)$$

Then (3.1.20) reads

$$\frac{\partial \mathcal{L}}{\partial \Phi_j} \Xi_{jk} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \partial_\mu \Xi_{jk} + \partial_\mu \xi_k^\nu \Theta^\mu{}_\nu - \xi_k^\nu (\partial_\nu^{(\text{expl})} \mathcal{L}) + \partial_\mu \Omega_k^\mu(\Phi_j, x) = 0. \quad (3.1.21)$$

So far this is the constraint on the transformation, defined by the Ξ_{jk} and ξ_k^ν in (3.1.10) to be a symmetry transformation. Now we investigate the implication for the solutions of the field equations (3.1.4).

As a first step we calculate the four-divergence of the energy-momentum tensor. Using the field equations (3.1.4) we find

$$\begin{aligned} \partial_\mu \Theta^\mu{}_\nu &= \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \right) \partial_\nu \Phi_j + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \partial_\mu \partial_\nu \Phi_j - \frac{\partial \mathcal{L}}{\partial \Phi_j} \partial_\nu \Phi_j - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \partial_\mu \partial_\nu \Phi_j - \partial_\nu^{(\text{expl})} \mathcal{L} \\ &= -\partial_\nu^{(\text{expl})} \mathcal{L}. \end{aligned} \quad (3.1.22)$$

Further with the field equations we get

$$\frac{\partial \mathcal{L}}{\partial \Phi_j} \Xi_{jk} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \partial_\mu \Xi_{jk} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \Xi_{jk} \right) \quad (3.1.23)$$

for the first two terms in (3.1.21). Finally using (3.1.22) and (3.1.23) in (3.1.21) we conclude that the quantities

$$j_k^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_j)} \Xi_{jk} + \Theta^\mu{}_\nu \xi_k^\nu + \Omega_k^\mu \quad (3.1.24)$$

fulfill the **equation of continuity**,

$$\partial_\mu j_k^\mu = 0. \quad (3.1.25)$$

This is the first part of **Noether's theorems**: Any one-parameter symmetry implies the existence of a corresponding conserved **Noether charge**,

$$Q_k = \int_{\mathbb{R}^3} d^3 \vec{x} j_k^0(x). \quad (3.1.26)$$

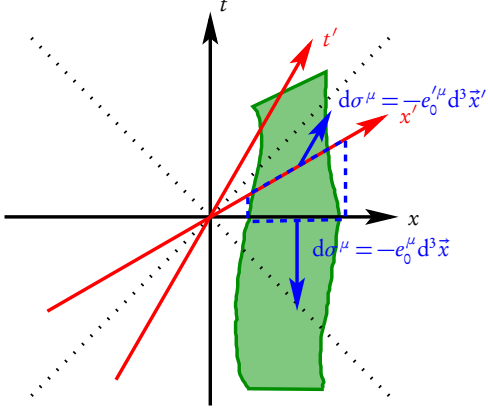
We note that this is a Lorentz scalar only if the continuity equation (3.1.25) is fulfilled.

To see this, we first prove that the Noether charge is conserved. This we can achieve easily by using usual 3D vector calculus. With (3.1.26) we find

$$d_t Q_k = \int_{\mathbb{R}^3} d^3 \vec{x} \partial_t j_k^0(t, \vec{x}) = - \int_{\mathbb{R}^3} d^3 \vec{x} \vec{\nabla} \cdot \vec{j}_k = 0, \quad (3.1.27)$$

where in the last step we used Gauss's theorem and the assumption that the current density vanishes at spatial infinity.

3.2. Poincaré symmetry



To see that Q_k is also a Lorentz scalar if (3.1.25) is fulfilled we use the 4D version of Gauss's theorem (A.6.16) 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

$$\begin{aligned}
 0 &= \int_M d^4x \partial_\mu j_k^\mu \\
 &= \int_{\partial M} d^3\sigma_\mu j_k^\mu \\
 &= \int_{t'=0} d^3\vec{x}' j_k^0(0, \vec{x}') - \int_{t=0} d^3\vec{x} j_k^0(0, \vec{x}),
 \end{aligned} \tag{3.1.28}$$

which indeed implies that

$$Q'_k - Q_k = 0 \Rightarrow Q'_k = Q_k. \tag{3.1.29}$$

3.2 Poincaré symmetry

We can investigate the symmetry of a quite general class of relativistic field theory using the above formalism. For all applications needed in this manuscript, we consider **linear representations** of the proper orthochronous Poincaré group. To this end we use the results of Sect. 1.7, according to which an infinitesimal Poincaré transformation acts on the space-time four-vector as

$$\delta x^\mu = \delta \omega_{\alpha\beta} g^{\alpha\mu} x^\beta + \delta a^\mu \tag{3.2.1}$$

where $\delta \omega_{\alpha\beta} = -\delta \omega_{\beta\alpha} = \text{const.}$ and $\delta a^\mu = \text{const.}$ It is convenient to treat the translation and Lorentz subgroups separately. In the following we bring the transformations to the general form (3.1.10). Here the index k of course becomes a four-vector index which we denote by α for translations and a pair $\alpha\beta$ for Lorentz transformations.

3.2.1 Translations

The translations are given by (3.2.1) with $\delta \omega_{\alpha\beta} = 0$:

$$\delta x^\mu = \delta a^\mu = \delta a_\alpha g^{\mu\alpha} \Rightarrow \xi^{\mu\alpha} = -\eta^{\mu\alpha}. \tag{3.2.2}$$

The fields usually transform as “scalar fields” under translations, i.e.,

$$\delta \phi'_i = 0 \Rightarrow \Xi_i^\alpha = 0. \tag{3.2.3}$$

The Lagrangian usually does not explicitly depend on x , and according to the symmetry condition (3.1.21) this implies that the translations are a symmetry with $\Omega_i^\alpha = 0$, and from (3.1.24) it follows that the Noether current is the **canonical energy-momentum tensor**, Θ^μ_α .

3. Classical fields

3.2.2 Lorentz transformations

Here we have

$$\delta x^\mu = \delta \omega_{\alpha\beta} \eta^{\mu\alpha} x^\beta = \frac{1}{2} \delta \omega_{\alpha\beta} (\eta^{\mu\alpha} x^\beta - \eta^{\mu\beta} x^\alpha) \Rightarrow \xi^{\mu\alpha\beta} = \frac{1}{2} (x^\alpha \eta^{\mu\beta} - x^\beta \eta^{\mu\alpha}). \quad (3.2.4)$$

The fields usually transform linearly, i.e., according to

$$\delta \phi_i = \frac{1}{2} \delta \omega_{\alpha\beta} \Sigma_{ij}^{\alpha\beta} \phi_j \Rightarrow \Xi_i^{\alpha\beta} = \frac{1}{2} \Sigma_{ij}^{\alpha\beta} \phi_j, \quad (3.2.5)$$

where $\Sigma_{ij}^{\alpha\beta} = -\Sigma_{ij}^{\beta\alpha} = \text{const}$. The Lagrangian is usually a **scalar field** built from the fields ϕ_j and their derivatives $\partial_\mu \phi_j$. As already assumed for the invariance under space-time translations, we again assume the Lagrangian to be not explicit dependent on x .

From (3.1.24) then the conserved currents define the **angular-momentum tensor**,

$$J^{\mu\alpha\beta} = -J^{\mu\beta\alpha} = x^\alpha \Theta^{\mu\beta} - x^\beta \Theta^{\mu\alpha} + S^{\mu\alpha\beta} \quad \text{with} \quad S^{\mu\alpha\beta} = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \Sigma_{ij}^{\alpha\beta} \phi_j. \quad (3.2.6)$$

3.3 Pseudo-gauge transformations

It is clear that the conserved currents from our symmetry analysis above are not uniquely determined: If j^μ is a conserved current, i.e., fulfilling $\partial_\mu j^\mu = 0$ we can always use an arbitrary antisymmetric tensor field, $\Omega^{\mu\nu} = -\Omega^{\nu\mu}$ to define a new current by

$$\tilde{j}^\mu = j^\mu + \partial_\nu \Omega^{\mu\nu}. \quad (3.3.1)$$

Because of the antisymmetry first of all the local conservation law, $\partial_\mu \tilde{j}^\mu = \partial_\mu j^\mu$ holds true identically, but also the conserved charge,

$$Q = \int_{\mathbb{R}^3} d^3x j^0 = \int_{\mathbb{R}^3} d^3x \tilde{j}^0 \quad (3.3.2)$$

does not change, because using the usual three-dimensional Gauss's Integral Theorem shows that

$$\int_{\mathbb{R}^3} d^3x \partial_\nu \Omega^{0\nu} = \int_{\mathbb{R}^3} d^3x \partial_j \Omega^{0j} = 0, \quad (3.3.3)$$

assuming that $\Omega^{0\nu}$ vanishes sufficiently quickly at spatial infinity. In the first step we have used that $\Omega^{00} = 0$ due to the antisymmetry of $\Omega^{\mu\nu}$.

This shows that \tilde{j}^μ describes the same physical content as the original Noether current derived from the above symmetry analysis, i.e., a local conservation law as well as the conserved total charge associated with the symmetry. In this sense the Noether theorem provides a unique conserved quantity but does not uniquely specify in any physical way between different distribution of the corresponding density and current density.

A redefinition of a conserved current via an arbitrary antisymmetric tensor file $\Omega^{\mu\nu}$, as described above, is sometimes called a **pseudo-gauge transformation**.

3.4 The Belinfante-Rosenfeld tensor

As we shall see, in the case of the energy-momentum and the angular-momentum tensors one often encounters difficulties with the canonical energy-momentum tensor. E.g., as we shall show later, in the case of electrodynamics in an naive application of the transformation laws for the electrodynamic potential A^μ as a vector field, the canonical energy-momentum tensor turns out to be not gauge invariant, though this is the case for the corresponding conserved “Noether charges”, i.e., the total energy and momentum of the (free) electromagnetic field.

A hint of, how to cure such kinds of “deficiencies” provides the General Theory of Relativity (GR), where the energy-momentum tensor of all fields not describing gravitation turns out to provide the sources of the gravitational field, which is described by the pseudo-metric of the pseudo-Riemannian space-time manifold¹, and this energy-momentum tensor is necessarily symmetric. In this sense general relativity provides a physical meaning at least to the local distribution of energy in and momentum and their current densities although in GR there are in general no globally conserved energy and momentum.

As has been shown by Belinfante and Rosenfeld [Bel40, Ros40], within SR we can always define a symmetric energy-momentum tensor from the canonical one by a pseudo-gauge transformation making use of the spin part of angular-momentum tensor (3.2.6). Indeed, the local conservation law for angular momentum gives

$$\partial_\mu J^{\mu\rho\nu} = \Theta^{\rho\nu} - \Theta^{\nu\rho} + \partial_\mu S^{\mu\rho\nu} = 0. \quad (3.4.1)$$

Now we look for a pseudo-gauge transformation

$$\tilde{\Theta}^{\mu\nu} =: T^{\mu\nu} = \Theta^{\mu\nu} + \partial_\rho \Omega^{\mu\rho\nu}, \quad \Omega^{\mu\rho\nu} = -\Omega^{\rho\mu\nu} \quad (3.4.2)$$

such that $T^{\mu\nu} = T^{\nu\mu}$. Now (3.4.1) gives a hint that we can use $S^{\mu\rho\nu}$ to achieve this. Indeed, since $S^{\mu\rho\nu} = -S^{\mu\nu\rho}$ we can define the tensor

$$\Omega^{\mu\rho\nu} = \frac{1}{2}(S^{\mu\nu\rho} - S^{\rho\nu\mu} + S^{\nu\mu\rho}), \quad (3.4.3)$$

which is indeed antisymmetric under exchange of the indices μ and ρ . Further, using (3.4.1) gives

$$T^{\mu\nu} = \Theta^{\mu\nu} + \frac{1}{2}\partial_\rho(S^{\mu\nu\rho} + S^{\nu\mu\rho}) - \frac{1}{2}(\Theta^{\mu\nu} - \Theta^{\nu\mu}) = \frac{1}{2}(\Theta^{\mu\nu} + \Theta^{\nu\mu}) + \frac{1}{2}\partial_\rho(S^{\mu\nu\rho} + S^{\nu\mu\rho}), \quad (3.4.4)$$

which is indeed symmetric under exchange of the indices μ and ν . It is clear that the total conserved energy-momentum vector

$$P^\nu = \int_{\mathbb{R}^3} d^3x T^{0\nu} = \int_{\mathbb{R}^3} d^3x \Theta^{0\nu} \quad (3.4.5)$$

does not change under this pseudo-gauge transformation.

We also note that now we can define the total angular-momentum density tensor by

$$\tilde{j}^{\mu\alpha\beta} = x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha}, \quad (3.4.6)$$

which looks as if there is no more “spin” involved. However, indeed, because of $T^{\mu\nu} = T^{\nu\mu}$ and $\partial_\mu T^{\mu\nu} = 0$, first of all we have the local conservation law,

$$\partial_\mu \tilde{j}^{\mu\alpha\beta} = 0. \quad (3.4.7)$$

¹See, e.g., [LL96] or [MTW73] for an introduction to General Relativity.

3. Classical fields

Further

$$\begin{aligned}\tilde{j}^{\mu\alpha\beta} &= x^\alpha \Theta^{\mu\beta} - x^\beta \Theta^{\mu\alpha} + x^\alpha \partial_\rho \Omega^{\mu\rho\beta} - x^\beta \partial_\rho \Omega^{\mu\rho\alpha} \\ &= J^{\mu\alpha\beta} - S^{\mu\alpha\beta} + \partial_\rho (x^\alpha \Omega^{\mu\rho\beta} - x^\beta \Omega^{\mu\rho\alpha}) - (\Omega^{\mu\alpha\beta} - \Omega^{\mu\beta\alpha})\end{aligned}\quad (3.4.8)$$

Plugging in the definition for $\Omega^{\mu\alpha\beta}$ (3.4.3) one finds that the last term in (3.4.8) cancels the second, such that the relation between $\tilde{j}^{\mu\alpha\beta}$ and $J^{\mu\alpha\beta}$ is again a pseudo-gauge transformation,

$$\tilde{j}^{\mu\alpha\beta} = J^{\mu\alpha\beta} + \partial_\nu \Omega^{\mu\nu\alpha\beta}, \quad \Omega^{\mu\nu\alpha\beta} = x^\alpha \Omega^{\mu\nu\beta} - x^\beta \Omega^{\mu\nu\alpha}, \quad (3.4.9)$$

where $\Omega^{\mu\nu\alpha\beta}$ indeed fulfills the necessary anti-symmetry condition $\Omega^{\mu\nu\alpha\beta} = -\Omega^{\nu\mu\alpha\beta}$.

3.5 Continuum mechanics

As we shall see later, when treating the electromagnetic interaction between charged matter and the electromagnetic field as the paradigmatic example for a concrete relativistic dynamical model, the point-particle model to describe macroscopic bodies is highly problematic in the theory of relativity and that a **field description** of dynamical systems is more natural.

As an example we treat the ideal fluid, which is also often useful as a classical (i.e., non-quantum) model of matter in applications of the theory of relativity as relativistic heavy-ion collisions and for simple classical models for the constitutive relations of macroscopic electrodynamics in a fully consistent relativistic form. We follow here [Sop08] and for the kinematical part [RZ13] adapted to our special-relativistic framework.

3.5.1 Kinematics

As in non-relativistic continuum mechanics, i.e., the mechanics of fluids (gases and liquids) and elastic solids, there are two kinds of description, usually referred to as the **Lagrangian** and the **Eulerian** formulation. From the point of view of point-particle mechanics the Lagrange formulation is most natural. Instead of considering a discrete set of interacting point particles, described by functions of their position vectors in an inertial reference frame as function of time or, equivalently, in the relativistic context, the space-time four-vector as a function of proper time to get a manifestly covariant description in terms of “world lines”, one introduces the idea of “material elements” (or more naturally “fluid elements”, though the here developed formalism for the kinematics also holds for solids). These are “macroscopically small”, i.e., you consider volume elements (in some arbitrary inertial frame) that are small in comparison to the typical length scale across which the macroscopic properties of the continuous matter change considerable, so that one can treat them formally as “infinitesimally small” volume elements d^3x but at the same time “microscopically large”, such that they contain many atoms or molecules, over which fast changing “thermal motion” one can average as far as only the macroscopic properties of the medium are of interest.

The kinematical description in the Lagrange formalism then is to define first a “standard configuration” by cartesian coordinates ξ^A ($A \in \{1, 2, 3\}$) in some arbitrary inertial frame and a density $\tilde{n}(\vec{\xi})$ giving the number of particles per volume within this material frame. Now the kinematics of the continuous medium can be described by the space-time four vectors $x^\mu(t, \vec{\xi})$ or, more conveniently in the manifestly covariant description by $x^\mu(s, \vec{\xi})$, where $s = c\tau$ with τ the proper time of the fluid element labelled by $\vec{\xi}$.

3.5. Continuum mechanics

This defines the fluid, as seen by an inertial observer, as a **time-like world-line congruence**. In the following we define **material derivatives** with respect to coordinate time, D_t , and s or proper time τ , D_s and D_τ by $D_t x = \partial_t x^\mu(t, \vec{\xi})$ and $D_\tau x^\mu = c \partial_s x^\mu(s, \vec{\xi}) = c u^\mu(s, \vec{\xi})$. Clearly $u_\mu u^\mu = 1$.

We also assume that $x^\mu(s, \vec{\xi})$ as well as $x^\mu(t, \vec{\xi})$ are invertible in the sense that we can uniquely derive $\vec{\xi}(s, \vec{\xi})$ as well as $\vec{\xi}(t, \vec{\xi})$. Also since $D_t x = D_\tau x / u^0$ one can always uniquely express t as a function $t(s, \vec{\xi})$ and s as a function $s(t, \vec{\xi})$, which is all in complete analogy to point-particle mechanics.

The other point of view is the more intuitive Eulerian description, which describes the fluid as seen by an observer located at a given space-time point x^μ and observing the flow of the medium just at \vec{x} at time t . This leads to a field-theoretical description with the three scalar fields $\xi^A(x)$, which just tells us which fluid element, given by the material coordinates $\vec{\xi}$ if in standard configuration, is located at the observer's position \vec{x} at time t . The normalized four-velocity field of the fluid is then given by $u^\mu(x) \equiv u^\mu[t, \vec{\xi}(t, \vec{x})]$.

This also allows us to calculate material time and proper-time derivatives of arbitrary scalar, vector, or tensor components in this Eulerian description:

$$D_t A(x) = \partial_t A[t, \vec{x}(t, \vec{\xi})] = \partial_t A + \vec{v} \cdot \vec{\nabla} A(x). \quad (3.5.1)$$

Here A stands for any scalar-like field quantity (including vector- or tensor-field components) Multiplying with u_0 one gets the manifestly covariant form

$$D_\tau A(x) = c D_s A(x) = c u^\mu(x) \partial_\mu A(x). \quad (3.5.2)$$

E.g., for the proper-acceleration field we get

$$a^\mu(x) = D_\tau u^\mu(x) = c u^\nu \partial_\nu u^\mu. \quad (3.5.3)$$

We note that, because $u_\mu u^\mu = 1 = \text{const}$ from this we immediately get

$$a^\mu u_\mu = c u_\mu u^\nu \partial_\nu u^\mu = \frac{c}{2} u^\nu \partial_\nu (u_\mu u^\mu) = 0, \quad (3.5.4)$$

as to be expected from point-particle mechanics in the Lagrangian description.

The next quantity we can calculate is the **particle-number four-flow**. We assume here that the particle number is conserved². The particle-number four-current $j_N^\mu(x)$ is given by $(j_N^\mu) = (c \rho_N, \rho_N \vec{v})$, where $\vec{v} = D_t \vec{x}$ is the three-velocity field of the medium. As we shall see in a moment this is a vector field. It can be expressed in manifestly covariant form as $j_N^\mu = n(x) c u^\mu$, where $n(x) = \rho_N / u^0$ is a scalar field. Its physical meaning is that it gives the particle-number density as measured in the instantaneous rest frame of the medium element. Indeed the particle number in the volume element $d^3 x$ is $\rho_N d^3 x = n(x) d^3 x^*$, and the volume element as measured in the observational frame, $d^3 x$, is Lorentz contracted compared to its measurement in the rest frame $d^3 x^*$, i.e., $\rho_N d^3 x = \rho_N / u_0 d^3 x^* = n d^3 x^*$.

Now we express j_N^μ in terms of the scalar fields $\xi^A(x)$. To that end we note that at time t the fluid element labelled by $\xi^A(x)$ in the standard configuration is located at \vec{x} . The corresponding number of

²This is not always generally valid in the relativistic realm. E.g., if you have some fluid at temperatures higher than the rest energy mc^2 of some particles, e.g., a gas electrons and positrons there can be annihilation and pair-creation processes like $e^+ + e^- \leftrightarrow \gamma + \gamma$. Then in the fluid-dynamics description we have to choose some conserved quantity to count the "amount of matter" like, in this example, electric charge instead of the particle numbers themselves.

3. Classical fields

particles is $dN = \tilde{n}(\xi)d^3\xi$, but $d^3\xi = \det(\partial_j \xi^A)d^3x$. This means that

$$\begin{aligned} \rho_N(x) &= \tilde{n}(\vec{\xi})\Delta^{jkl}(\partial_j \xi^1)(\partial_k \xi^2)(\partial_l \xi^3) \\ &= \tilde{n}(\vec{\xi})\epsilon^{0jkl}(\partial_j \xi^1)(\partial_k \xi^2)(\partial_l \xi^3) \\ &= \tilde{n}(\vec{\xi})\epsilon^{0\alpha\beta\gamma}(\partial_\alpha \xi^1)(\partial_\beta \xi^2)(\partial_\gamma \xi^3) \\ &= \frac{1}{3!}\tilde{n}(\vec{\xi})\epsilon^{0\alpha\beta\gamma}\Delta_{ABC}(\partial_\alpha \xi^A)(\partial_\beta \xi^B)(\partial_\gamma \xi^C), \end{aligned} \quad (3.5.5)$$

where here and in the following we define as the Levi-Civita symbol in d dimensions as the symbol $\Delta^{a_1 a_2 \dots a_d} = \Delta_{a_1 \dots a_d}$ that is totally antisymmetric under any permutations of two of its indices and $\Delta^{12 \dots d} = \Delta_{12 \dots d} = 1$. In the context of Minkowski four-tensor indices we have $\Delta^{0123} = \epsilon^{0123} = 1$ and $\epsilon_{0123} = -\Delta_{0123}$.

Now (3.5.5) suggests that

$$j_N^\mu(x) = \frac{c}{3!}\tilde{n}(\vec{\xi})\epsilon^{\mu\alpha\beta\gamma}\Delta_{ABC}(\partial_\alpha \xi^A)(\partial_\beta \xi^B)(\partial_\gamma \xi^C). \quad (3.5.6)$$

For $\mu = 0$ this immediately follows from (3.5.5). For $\mu = j$ exactly one of the indices α, β and γ must be 0. Then the other two of these indices take values different from 0 and j . In our totally symmetrized notation in (3.5.6) this simply boils down to

$$j_N^j(x) = \frac{c}{2!}\tilde{n}(\vec{\xi})\epsilon^{j0bc}\Delta_{ABC}(\partial_0 \xi^A)(\partial_b \xi^B)(\partial_c \xi^C). \quad (3.5.7)$$

Now we have

$$x^\mu = x^\mu[t, \vec{\xi}(x)]. \quad (3.5.8)$$

Taking the derivative wrt. t at fixed \vec{x} for a spatial component x^a this leads to

$$0 = \partial_t x^a = \partial_t x^a(t, \vec{\xi})|_{\vec{\xi}=\vec{\xi}(x)} + (\partial_t \xi^A)\partial_A x^a(t, \vec{\xi})|_{\vec{\xi}=\vec{\xi}(x)} = v^a(x) + c(\partial_0 \xi^A)\partial_A x^a. \quad (3.5.9)$$

From this we find

$$c\partial_0 \xi^A = -v^a \partial_a \xi^A. \quad (3.5.10)$$

Plugging this in (3.5.7) after using $\epsilon^{j0bc} = -\epsilon^{0jbc}$ indeed leads to the expected result:

$$j_N^j = +\frac{1}{2!}\tilde{n}(\vec{\xi})\epsilon^{0jbc}\Delta_{ABC}v^a(\partial_a \xi^A)(\partial_b \xi^B)(\partial_c \xi^C) = \frac{1}{2}\rho_N\Delta^{jbc}\Delta_{abc}v^a = \rho_N v^j. \quad (3.5.11)$$

From the derivation of the current (3.5.6) it is also clear that it must obey the continuity equation. Formally this is easy to prove, using (3.5.6):

$$\partial_\mu J_N^\mu = \frac{c}{3!}\epsilon^{\mu\alpha\beta\gamma}\Delta_{ABC}\left[(\partial_D \tilde{n})(\partial_\mu \xi^D)(\partial_\alpha \xi^A)(\partial_\beta \xi^B)(\partial_\gamma \xi^C) + 3\tilde{n}(\partial_\mu \partial_\alpha \xi^A)(\partial_\beta \xi^B)(\partial_\gamma \xi^C)\right]. \quad (3.5.12)$$

Both terms in the square bracket vanish due to the contraction with $\epsilon^{\mu\alpha\beta\gamma}$: The first term, because any non-vanishing contribution must have $\{A, B, C\} = \{1, 2, 3\}$, and thus D is necessarily equal with one of these indices (A, B, C), e.g., A . Then $(\partial_\mu \xi^D)(\partial_\alpha^A)$ is symmetric under exchange of μ and α and this cancels any such contribution when contracted with $\epsilon^{\mu\alpha\beta\gamma}$. The second term in the square bracket

3.5. Continuum mechanics

is explicitly symmetric under exchange of the indices μ and α and thus its contraction with $\epsilon^{\mu\alpha\beta\gamma}$ vanishes too. Thus finally indeed

$$\partial_\mu J_N^\mu = 0. \quad (3.5.13)$$

Next we consider the **deformations of the medium** as a function of (proper) time. By using proper time rather than coordinate time to describe the motion of the infinitesimal volume elements of the medium we already use an invariant description of time by defining it in the local rest frame of the medium. Further we also have to distinguish between true physical deformations of one such volume element relative to some other neighboring element and the effect of length contraction. This can also be achieved by looking at these deformations in the local rest frame of the medium. Finally, as in non-relativistic fluid dynamics, which is locally valid in the local rest frame, we have to also separate the rigid rotations of the medium element as a whole from the true intrinsic deformations. To this end we consider the volume element at $x^\mu = x^\mu(\tau, \xi^A)$ and an infinitesimally close volume element at $x^\mu + \delta\epsilon s^\mu$, where $s^\mu = s^\mu(x)$ is an arbitrary space-like ‘‘displacement vector’’. For the proper velocity of the displacement vector we find on one hand

$$D_\tau s_\mu = c u^\nu \partial_\nu s_\mu, \quad (3.5.14)$$

as for any vector field. On the other hand

$$c u^\mu(x + \delta\epsilon s) = c u^\mu(x) + c \delta\epsilon s^\nu \partial_\nu u^\mu(x) = c u^\mu(x) + c \delta\epsilon u^\nu \partial_\nu s^\mu. \quad (3.5.15)$$

From this we finally get

$$D_\tau s_\mu = c u^\nu \partial_\nu s_\mu = c s^\nu \partial_\nu u_\mu. \quad (3.5.16)$$

The deformation is thus given as a linear mapping with the tensor $\partial_\nu u_\mu$, and we have to decompose this deformation in the above described covariant way. To that end we introduce the projection operator to the local rest frame,

$$\Delta_{\mu\nu} = \eta_{\mu\nu} - u_\mu u_\nu. \quad (3.5.17)$$

First we note that

$$\partial_\nu u_\mu = (\Delta_\nu^\sigma + u^\sigma u_\nu)(\Delta_\mu^\rho + u^\rho u_\mu) \partial_\sigma u_\rho = \Delta_\nu^\sigma \Delta_\mu^\rho \partial_\sigma u_\rho + a_\mu u_\nu, \quad (3.5.18)$$

where we have used (3.5.4).

The first term now can be decomposed first in its symmetric and antisymmetric part. Since the projectors Δ_α^β project to the space-like hypersurface Minkowski-orthogonal to u_μ , i.e., to the local rest frame of the medium, this is merely a tensor in 3D space, and thus the antisymmetric part is already irreducible under the symmetry group of rotations of this hypersurface. The symmetric part can further be decomposed in two irreducible pieces, namely a scalar part transforming trivially under rotations in the hypersurface and the remainder. The scalar part is given by the trace, i.e., we have

$$\Theta_{\nu\mu} = \frac{1}{2} \Delta_\nu^\sigma \Delta_\mu^\rho (\partial_\sigma u_\rho + \partial_\rho u_\sigma) = \sigma_{\nu\mu} + \frac{1}{3} \Theta \Delta_{\nu\mu}, \quad (3.5.19)$$

where

$$\Theta = \Theta_\mu^\mu = \partial_\rho u^\rho, \quad (3.5.20)$$

where in the final step we have used (3.5.4) again. From (3.5.19) we immediately find

$$\sigma_\mu^\mu = 0. \quad (3.5.21)$$

3. Classical fields

So, finally we have the displacement (3.5.16) decomposed in the irreducible parts

$$D_\tau s_\mu = s^\nu \left(a_\mu u_\nu + c \omega_{\nu\mu} + c \sigma_{\nu\mu} + \frac{c}{3} \Theta \Delta_{\nu\mu} \right). \quad (3.5.22)$$

The first two terms describe the rigid displacement of the medium cell in the local rest frame³, acceleration and rotation, where $\omega_{\nu\mu} = -\omega_{\mu\nu}$ is the **kinematic vorticity** of the fluid cell, and the 3rd and 4th terms the intrinsic deformations of the cell. The term with $\sigma_{\nu\mu}$ describes the **shear deformation** of the cell, perserving its proper volume, while Θ describes the change of the proper volume by **compression or expansion**.

After having some intuitive picture of the kinematics of the fluid cell we can look at this kinematics again using our field-theoretical description with the three scalar fields $\xi^A(x)$ introduced above and which we can use in our Lagrangian framework to derive equations of motion for (perfect) fluids and (nondissipative) elastic bodies. Since the Lagrange density should be a Lorentz scalar, it is clear that the basic building block are the six scalars,

$$G^{AB} = -(\partial_\mu \xi^A)(\partial^\mu \xi^B), \quad (3.5.23)$$

the **deformation matrix**.

Its meaning becomes clear in a **local rest frame of the medium cell**, where $\partial_0^* \xi^A = 0^4$. There the deformation matrix is obviously given by

$$G^{AB} = (\partial_a^* \xi^A)(\partial_a^* \xi^B). \quad (3.5.24)$$

This matrix encodes information on the displacement of the fluid cell in the same sense as s^μ above, but now looked at from the “laboratory frame” at a given time t in terms of the material coordinates:

$$\xi^A(t, \vec{x} + \delta \vec{x}) = \xi^A(t, \vec{x}) + \delta x^a \partial_a \xi^A = \xi^A(t, \vec{x}) + \delta x^a M_a^A. \quad (3.5.25)$$

Obviously we have

$$G^{AB} = M_a^A M_a^B \Leftrightarrow \hat{G} = \hat{M} \hat{M}^T, \quad (3.5.26)$$

and \hat{G} is a positive definite symmetric matrix, from which we can get \hat{M} up to a rotation. Indeed, if we write $\hat{M} = \hat{M}' \hat{R}$, where $\hat{R} \in \text{SO}(3)$ we find $\hat{G} = \hat{M} \hat{M}^T = \hat{M}' \hat{R} \hat{R}^T \hat{M}'^T = \hat{M}' \hat{M}'^T$. So we can write

$$\hat{M} = \sqrt{\hat{G}} \hat{R}. \quad (3.5.27)$$

So \hat{G} describes the deformation of a medium cell at a given time t modulo a rigid rotation. So it is perfectly suited to describe, e.g., an elastic body, whose elastic energy depends precisely on the deformations of different parts relative to its “relaxed state” but not on rigid rotations of the body as a whole.

It is also easy to find the relation of the deformation matrix to the above discussed particle-number current. Using (3.5.5) in the local rest frame we have

$$J_N^{*\mu} = (nc, 0, 0, 0) \quad (3.5.28)$$

³This term arises, because we consider neighboring medium cells at constant proper time. In the local rest frame we have $u^\mu = (1, 0, 0, 0)$ and thus $a^\mu = (0, \vec{a})$, i.e., the term describes the change of the displacement’s velocity during the corresponding time $\delta t = \delta \tau = \delta \epsilon s^0 / c$.

⁴Here and in the following we write $x^{*\mu}$ for the space-time coordinates wrt. the local rest frame of the fluid cell.

with

$$n = \tilde{n}(\vec{\xi}) \det \hat{M} = \tilde{n}(\vec{\xi}) \sqrt{\det \hat{G}^*}. \quad (3.5.29)$$

In a general frame we obviously have

$$J_N^\mu = c n u^\mu \Rightarrow n = \sqrt{J_N^\mu J_{N\mu}} = \tilde{n}(\xi) \sqrt{\det \hat{G}}. \quad (3.5.30)$$

Here we have used that $\hat{G} = (G^{ab})$, given in the manifestly covariant form (3.5.23), are Lorentz scalars, and thus also $\det \hat{G}$ is a Lorentz scalar.

3.5.2 Dynamics of a medium

Now we can write down the Lagrange density for a general continuous medium. According to the above kinematic considerations we write the Lagrange density as a function of the three ‘‘material coordinates’’ $\xi^A(x)$ and the deformation matrix G^{AB} , which is a function of the derivatives. For the following it is most convenient to also separate a factor n ,

$$\mathcal{L}(\xi^A, G^{AB}) = -n U(\xi^A, G^{AB}). \quad (3.5.31)$$

To interpret this ansatz and find out the physical meaning of U we calculate the **canonical energy-momentum tensor** using (3.1.20) for this Lagrange density:

$$\Theta^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \xi^A)} \partial_\nu \xi^A - \mathcal{L} \delta_\nu^\mu. \quad (3.5.32)$$

First we determine the derivative of n :

$$\begin{aligned} \frac{\partial n}{\partial (\partial_\mu \xi^A)} &= \frac{1}{2n} \frac{\partial (n)^2}{\partial (\partial_\mu \xi^A)} \\ &= \frac{1}{2c^2 n} \frac{\partial}{\partial (\partial_\mu \xi^A)} J_N^\alpha J_{N\alpha} \\ &= \frac{J_{N\alpha}}{c^2 n} \frac{\partial J_N^\alpha}{\partial (\partial_\mu \xi^A)} \\ &= \frac{\tilde{n} J_{N\alpha}}{2cn} \epsilon_{ABC} \epsilon^{\alpha\mu\rho\sigma} \partial_\rho \xi^B \partial_\sigma \xi^C. \end{aligned} \quad (3.5.33)$$

Here we have used (3.5.6) for the current. Within (3.5.32) we need the combination

$$\frac{\partial n}{\partial (\partial_\mu \xi^A)} \partial_\nu \xi^A = \frac{J_{N\alpha} \tilde{n}}{cn} \underbrace{\epsilon_{ABC} \epsilon^{\alpha\mu\rho\sigma} \partial_\nu \xi^A \partial_\rho \xi^B \partial_\sigma \xi^C}_{F^{\alpha\mu}{}_\nu}. \quad (3.5.34)$$

Now $F^{\alpha\mu}{}_\nu = -F^{\mu\alpha}{}_\nu$, i.e., this expression is only nonvanishing if $\mu \neq \alpha$. For a nonvanishing contribution we must also have either $\nu = \mu$, for which case $F^{\alpha\mu}{}_\nu = J_N^\alpha/c$, or $\nu = \alpha$, for which case $F^{\alpha\mu}{}_\nu = -J_N^\mu/c$. Thus we have

$$F^{\alpha\mu}{}_\nu = \frac{1}{c} (J_N^\alpha \delta_\nu^\mu - \delta_\nu^\alpha J_N^\mu). \quad (3.5.35)$$

3. Classical fields

Using this in (3.5.34) with $J_N^\alpha = n c u^\alpha$ leads finally to

$$\frac{\partial n}{\partial(\partial_\mu \xi^A)} \partial_\nu \xi^A = n(\delta_\nu^\mu - u^\mu u_\nu) = n \Delta_\nu^\mu. \quad (3.5.36)$$

Using this in (3.5.32) with the Lagrangian (3.5.31) after some simple calculations, using

$$\frac{\partial G^{CD}}{\partial(\partial_\mu \xi^A)} = -(\partial^\mu \xi^C \delta^{AD} + \partial^\mu \xi^D \delta^{AC}), \quad (3.5.37)$$

we finally get

$$\Theta_{\mu\nu} = n U u_\mu u_\nu + 2n \frac{\partial U}{\partial G^{CD}} (\partial_\mu \xi^C) (\partial_\nu \xi^D) = T_{\mu\nu}, \quad (3.5.38)$$

were in the final step have we used the symmetry $\Theta_{\mu\nu} = \Theta_{\nu\mu}$, i.e., as usual for field theories involving only scalar fields the canonical energy-momentum tensor is symmetric and thus identical with the symmetric Belinfante energy-momentum tensor.

In the local restframe of the medium we have $(u_\mu^*) = (1, 0, 0, 0)$ and thus we can identify $\epsilon^* = nU$ as the energy density measured in this local rest frame, and thus U is the energy per particle, and $T^{*ab} = -S^{*ab}$, where S^{*ab} is the **stress tensor** in the local rest frame. The latter follows from momentum conservation $\partial_\mu T^{\mu a} = 0$, and it is characteristic for the continuum-mechanics equations of motion that the three equations for energy and momentum are equivalent to the Euler-Lagrange equations of motion, because

$$\begin{aligned} \partial_\mu T^{\mu\nu} &= \partial_\mu \left[(\partial^\nu \xi_A) \frac{\partial \mathcal{L}}{\partial(\partial_\mu \xi^A)} - \eta^{\mu\nu} \mathcal{L} \right] \\ &= (\partial_\mu \partial^\nu \xi^A) \frac{\partial \mathcal{L}}{\partial(\partial_\mu \xi^A)} + (\partial^\nu \xi_A) \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \xi^A)} - (\partial^\nu \xi^A) \frac{\partial \mathcal{L}}{\partial \xi^A} - (\partial^\nu \partial_\mu \xi^A) \frac{\partial \mathcal{L}}{\partial(\partial_\mu \xi^A)} \\ &= (\partial^\nu \xi^A) \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \xi^A)} - \frac{\partial \mathcal{L}}{\partial \xi^A} \right). \end{aligned} \quad (3.5.39)$$

Since by assumption the \vec{x} are diffeomorphisms of the $\xi^A(x^0, \vec{x})$ for all $x^0 = ct$, the momentum conservation equations, i.e., (3.5.39) for $\nu = a \in \{1, 2, 3\}$ are indeed equivalent to the Euler-Lagrange equations for the fields ξ^A , and as expected from Noether's theorem also energy is conserved, i.e., (3.5.39) for $\nu = 0$.

3.5.3 Ideal fluid

To apply our formalism of the previous Sect. to the case of an **ideal fluid** we have to use some thermodynamics, which in the relativistic context we will treat in detail in the next Chapter. Here, it is sufficient to know that the usual formalism of thermodynamical quantities apply in the (local) rest-frame of the fluid independent of whether we work within Newtonian or special-relativistic physics. For the internal energy of a fluid for a (reversible) change of state we have

$$d\mathcal{U} = T dS - P dV + \mu dN, \quad (3.5.40)$$

where T is the temperature, S the entropy, p the pressure, V the volume, N the particle number, and μ the chemical potential of the fluid, and $\mathcal{U} = UN$ the internal energy of the matter in the fluid cell under consideration with all these quantities defined in its local rest frame. Thus also $n = N/V$.

3.5. Continuum mechanics

Thus the “natural” choice of independent variables for the internal energy are the **extensive** quantities S , V , and N , of which \mathcal{U} , which is an extensive quantity itself, must be a homogeneous function of order 1, i.e.,

$$\lambda\mathcal{U}(S, V, N) = \mathcal{U}(\lambda S, \lambda V, \lambda N). \quad (3.5.41)$$

Taking the derivative of this equation with respect to λ and then setting $\lambda = 1$ leads to

$$\mathcal{U}(S, V, N) = S\partial_S\mathcal{U} + V\partial_V\mathcal{U} + N\partial_N\mathcal{U}. \quad (3.5.42)$$

From (3.5.40) this yields the **Gibbs-Duham relation**

$$\mathcal{U} = TS - PV + \mu N. \quad (3.5.43)$$

Taking the total differential and using again (3.5.40) gives

$$SdT - VdP + Nd\mu = 0. \quad (3.5.44)$$

Deviding this by N and defining $s = S/N$, the entropy per particle,

$$sdT - \frac{1}{n}dP + d\mu = 0. \quad (3.5.45)$$

Deviding the Gibbs-Duham relation (3.5.43) by N and taking the total differential yields

$$dU = sdT + Tds - \frac{1}{n}dP + \frac{1}{n^2}Pdn + d\mu, \quad (3.5.46)$$

and using (3.5.45) leads to

$$dU = Tds + \frac{P}{n^2}dn \Rightarrow U = U(n, s), \quad \partial_s U = T, \quad \partial_n U = \frac{P}{n^2}. \quad (3.5.47)$$

In the following we also need the corresponding relation for the **enthalpy per particle**, $h = H/N = (\mathcal{U} + PV)/N = U + P/n$:

$$dh = dU + \frac{dP}{n} - \frac{P}{n^2}dn \stackrel{(3.5.47)}{=} Tds + \frac{dP}{n}. \quad (3.5.48)$$

For an ideal fluid we thus have

$$\mathcal{L} = -nU(n, s). \quad (3.5.49)$$

Using (3.5.36) we find for the energy-momentum tensor

$$T_{\mu\nu} = \frac{\partial \mathcal{L}}{\partial n} \frac{\partial n}{\partial(\partial^\mu \xi^A)} \partial_\nu \xi^A - \mathcal{L} \eta_{\mu\nu} = (nU + P)u_\mu u_\nu - P\eta_{\mu\nu}. \quad (3.5.50)$$

Usually one also uses $\varepsilon = nU$, the energy density in the local rest frame of the fluid cell, i.e., a *scalar field* [Oll08]:

$$T_{\mu\nu} = (\varepsilon + P)u_\mu u_\nu - P\eta_{\mu\nu}. \quad (3.5.51)$$

According to the above derivation, the **equations of motion** now are

$$\partial_\mu T^{\mu\nu} = 0, \quad \partial_\mu J_N^\mu = \partial_\mu (nu^\mu) = 0. \quad (3.5.52)$$

3. Classical fields

Since the energy-conservation equation, i.e., the first equation for $\nu = 0$ follows from the spatial equations for $\nu \in \{1, 2, 3\}$ these are four independent equations of motion for the 5 independent fields u^μ (only \vec{u} are independent, because $u_\mu u^\mu = 1$) and two independent thermodynamic quantities. So in addition one also needs an **equation of state** like $P = P(\varepsilon, n)$.

One important consequence of the ideal-fluid equations of motion is that the fluid motion is **isentropic**. To derive this, we contract the first equation in (3.5.52) with u_ν , and plug in (3.5.50), using the 2nd equation (3.5.52), $u_\nu \partial_\mu u^\nu = \partial_\mu (u_\nu u^\nu)/2 = 0$, and (3.5.48)

$$u_\nu \partial_\mu T^{\mu\nu} = \partial_\mu \underbrace{[(U + P/n) n u^\mu]}_h - u^\nu \partial_\nu P = T n u^\mu \partial_\mu s = 0. \quad (3.5.53)$$

This means that the **entropy is conserved**. To see this we can rewrite the latter equation as

$$n u^\mu \partial_\mu s = \partial_\mu (n s u^\mu) = 0. \quad (3.5.54)$$

Here ns is the entropy density as measured in the local rest frame of the fluid and thus, since $J_S^\mu = n s u^\mu$ it the entropy four-current, (3.5.54) has the meaning of (local) entropy conservation.

3.5.4 “Dust matter” and a scalar field

From a more fundamental point of view, one can treat matter as interacting with fields like the electromagnetic field, which we shall treat in Chpt. 4 in great detail. In this Sect. we restrict ourselves to the interaction of “free particles” (often called “dust matter” in relation with general-relativistic astrophysics). Since for non-interacting particles the energy per particle in its rest frame simply is $U = mc^2$, where m is the invariant mass of the particles, the corresponding contribution to the Lagrange density (3.5.31) is

$$\mathcal{L}_0^{(\text{mech})} = -mc^2 n. \quad (3.5.55)$$

Now we like to couple the particles to a real scalar field. The Lagrange density for the free field reads

$$\mathcal{L}_0^{(\text{field})} = \frac{1}{2} (\partial_\mu \Phi) (\partial^\mu \Phi) - \frac{M^2}{2} \Phi^2, \quad (3.5.56)$$

where M is a parameter⁵. The simplest interaction term between the dust medium and the field is

$$\mathcal{L}^{(\text{int})} = -gn\Phi, \quad (3.5.57)$$

where g is another constant.

The equation of motion of the field is directly derived from the corresponding Euler-Lagrange equation (3.1.4) for the single scalar field Φ :

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} = \square \Phi = \frac{\partial \mathcal{L}}{\partial \Phi} = -M^2 \Phi - gn. \quad (3.5.58)$$

This implies the **inhomogeneous Klein-Gordon equation**

$$(\square + M^2)\Phi = -gn. \quad (3.5.59)$$

⁵In the quantized version, $M\hbar/c$ is the mass of the corresponding “scalar particles”, described by the (quantized) Klein-Gordon equation.

3.5. Continuum mechanics

The equations of motion for the medium is again most simply derived from the energy-momentum conservation law, cf. (3.5.39). To this end we evaluate the energy-momentum tensor. For the three contributions (3.5.55-3.5.57) we get, using (3.5.32) and (3.5.38) for the mechanical and interaction part

$$\Theta_{(\text{mech})}^{\mu\nu} = mc^2 n u_\mu u^\nu, \quad (3.5.60)$$

$$\Theta_{(\text{field})}^{\mu\nu} = (\partial^\mu \Phi)(\partial^\nu \Phi) - \frac{1}{2} [(\partial_\alpha \Phi)(\partial^\alpha \Phi) - M^2 \Phi^2] \eta^{\mu\nu} \quad (3.5.61)$$

$$\Theta_{(\text{int})}^{\mu\nu} = g \Phi n u^\mu u^\nu. \quad (3.5.62)$$

The equation of motion for the medium then follows from energy-momentum conservation $\partial_\mu \Theta^{\mu\nu} = 0$, using $\partial_\mu (n u^\mu) = 0$,

$$n(mc^2 + g\Phi)u^\mu \partial_\mu u^\nu = g n (\eta^{\mu\nu} - u^\mu u^\nu) \partial_\mu \Phi := K^\nu. \quad (3.5.63)$$

The scalar field thus provides a contribution to an “effective mass” of the particle, corresponding to the left-hand side of (3.5.63), as well as a Minkowski force given by the right-hand side, fulfilling the constraint equation $u_\nu K^\nu$ as it is enforced by the identity $u_\nu u^\nu = 0$, from which $\partial_\mu u_\nu u^\nu = 2u_\nu \partial_\mu u^\nu = 0$ follows.

One should note that one gets back to the point-particle limit by writing the particle-number current for a single particle as

$$J_N^\mu(x) = c \int d\lambda \frac{dy^\mu(\lambda)}{d\lambda} \delta^{(4)}[x - y(\lambda)], \quad (3.5.64)$$

where λ is an arbitrary world-line parameter. Written in this form the current is written in a parameter-independent and manifestly covariant way. Here $y^\mu(\lambda)$ is the necessarily time-like world line of the particle.

One common physically motivated choice for λ is the coordinate-time $t_y = y^0/c$,

$$J_N^\mu(x) = c \int dt_y \left(\frac{c}{d_t \vec{y}(t_y)} \right) \delta[c(t - t_y)] \delta^{(3)}[\vec{x} - \vec{y}(t_y)] = \left(\frac{c}{d_t \vec{y}(t)} \right) \delta^{(3)}[\vec{x} - \vec{y}(t)]. \quad (3.5.65)$$

Form this expression it is most easy to prove the particle-number conservation,

$$\partial_\mu J_N^\mu = (\partial_t + d_t \vec{y} \cdot \vec{\nabla}) \delta^{(3)}[\vec{x} - \vec{y}(t)] = 0. \quad (3.5.66)$$

To derive the point-particle equation of motion we note that from (3.5.65) we get, using $u^\mu = \gamma(1, d_t \vec{y}/c)$ and $\gamma = u^0 = 1/\sqrt{1 - (d_t \vec{y}/c)^2}$,

$$J_N^\mu(x) = \frac{c}{\gamma} u^\mu \delta^{(3)}[\vec{x} - \vec{y}(t)] \Rightarrow n = \frac{1}{\gamma} \delta^{(3)}[\vec{x} - \vec{y}(t)]. \quad (3.5.67)$$

To derive the point-particle equation of motion, we note that from (3.5.55) and (3.5.57)

$$L_0^{(\text{mech})} + L^{(\text{int})} = \int_{\mathbb{R}^3} d^3x [-mc^2 - g\Phi] n = -\sqrt{1 - (d_t \vec{y}/c)^2} [mc^2 + g\Phi(y)]. \quad (3.5.68)$$

The Euler-Lagrange equations for variations of this Lagrangian with respect to \vec{y} and rewriting the time derivatives in terms of proper-time derivatives yields

$$\left[m + \frac{g}{c^2} \Phi(y) \right] d_\tau^2 x^\mu = g \left[\eta^{\mu\nu} - \frac{1}{c^2} (d_\tau y^\mu)(d_\tau y^\nu) \right] \partial_\mu \Phi. \quad (3.5.69)$$

3. Classical fields

This is a well-defined equation of motion, if one takes the field Φ as an external given field, but one should note however, that in this point-like limit a self-consistent solution of (3.5.69) and the field equation (3.5.59) is not a well-defined problem. For the analogous case of the electromagnetic field this is known as the **radiation reaction problem**. This will be discussed in some detail in Chpt. 4.

Chapter 4

Classical Electromagnetism

4.1 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

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (4.1.1)$$

$$\vec{\nabla} \times \vec{E} + \frac{1}{c} \partial_t \vec{B} = 0, \quad (4.1.2)$$

$$\vec{\nabla} \cdot \vec{E} = \rho, \quad (4.1.3)$$

$$\vec{\nabla} \times \vec{B} - \frac{1}{c} \partial_t \vec{E} = \frac{1}{c} \vec{j}. \quad (4.1.4)$$

Eqs. (4.1.1) and (4.1.2) 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** (4.1.1) 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** (4.1.2) 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** (4.1.3) and the **Maxwell-Ampere Law** (4.1.4), 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 $d^2\vec{f}$ around position \vec{x} , $dI = d^2\vec{f} \cdot \vec{j}(t, \vec{x})$ 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 “mediating” 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**

4. Classical Electromagnetism

electromagnetic field (\vec{E}, \vec{B}) as

$$\frac{d}{dt} \left(\frac{m\dot{\vec{x}}}{\sqrt{1-\dot{\vec{x}}^2/c^2}} \right) = q\vec{E} + \frac{q}{c} \vec{v} \times \vec{B} \quad (4.1.5)$$

or, in manifestly covariant form (2.4.33)

$$m \frac{d^2}{d\tau^2} x_\alpha = \frac{q}{c} F_{\alpha\mu} \frac{d}{d\tau} x^\mu. \quad (4.1.6)$$

Here, τ is the proper time of the particle along his world line. The spatial part is also readily derived by using $d\tau = dt \sqrt{1-\dot{\vec{x}}^2/c^2}$ from (4.1.5) by deviding this equation by this square-root factor. Using $u^\mu = 1/c dx^\mu/d\tau$ then gives

$$m d_\tau^2 \vec{x} = q(u^0 \vec{E} + \vec{u} \times \vec{B}). \quad (4.1.7)$$

In any way, (4.1.5) 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 (4.1.1-4.1.4) 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

$$\vec{j}(t, \vec{x}) = \rho(t, \vec{x}) \vec{v}(t, \vec{x}). \quad (4.1.8)$$

Then the **force density** is given by

$$\vec{f}(t, \vec{x}) = \rho(t, \vec{x}) \vec{E}(t, \vec{x}) + \frac{1}{c} \vec{j} \times \vec{B}, \quad (4.1.9)$$

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

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

4.2 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 (4.1.1-4.1.4) 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 (4.1.4) and using (4.1.3) we find

$$\frac{1}{c} \vec{\nabla} \cdot \vec{j} = -\frac{1}{c} \vec{\nabla} \cdot \partial_t \vec{E} = -\frac{1}{c} \partial_t \vec{\nabla} \cdot \vec{E} = -\frac{1}{c} \partial_t \rho \Rightarrow \partial_t \rho + \vec{\nabla} \cdot \vec{j} = 0. \quad (4.2.1)$$

4.2. Manifestly covariant formulation of electrodynamics

The latter equation is the **continuity equation**. Its physical meaning becomes clear when integrating (4.2.1) over an arbitrary volume at rest. Then we find, using Gauss's theorem

$$\int_V d^3x \partial_t \rho(t, \vec{x}) = \frac{d}{dt} \int_V d^3x \rho(t, \vec{x}) = \frac{d}{dt} Q_V(t) = - \int_V d^3x \vec{\nabla} \cdot \vec{j}(t, \vec{x}) = - \int_{\partial V} d^2\vec{f} \cdot \vec{j}(t, \vec{x}). \quad (4.2.2)$$

Here ∂V is the (closed) boundary surface of V with the convention to orient the surface-normal vectors $d^2\vec{f}$ 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 (4.2.1) 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

$$\partial_\mu = \frac{\partial}{\partial x^\mu}. \quad (4.2.3)$$

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 (4.2.1) as

$$\partial_\mu j^\mu(t, \vec{x}) = 0, \quad (4.2.4)$$

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

$$j^\mu(x) = \begin{pmatrix} c\rho(t, \vec{x}) \\ \vec{j}(t, \vec{x}) \end{pmatrix}. \quad (4.2.5)$$

To confirm that (4.2.3) 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, $T^{\mu\nu\rho\dots}(x)$.

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

$$\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu, \quad \eta_{\mu\rho} \Lambda^\mu{}_\nu \Lambda^\rho{}_\sigma = \eta_{\nu\sigma} \Rightarrow \eta_{\nu\sigma} (\Lambda^{-1})^\sigma{}_\beta = \eta_{\mu\beta} \Lambda^\mu{}_\nu \Rightarrow (\Lambda^{-1})^\alpha{}_\beta = \eta_{\mu\beta} \eta^{\nu\alpha} \Lambda^\mu{}_\nu. \quad (4.2.6)$$

From this the transformation of the co-variant components read

$$\bar{x}_\rho = \eta_{\mu\rho} \Lambda^\mu{}_\nu x^\nu = \eta_{\mu\rho} \eta^{\nu\sigma} \Lambda^\mu{}_\nu x_\sigma = (\Lambda^{-1})^\sigma{}_\rho x_\sigma. \quad (4.2.7)$$

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

$$\bar{T}^{\mu\nu\rho\dots}(\bar{x}) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \dots T^{\alpha\beta\gamma\dots}(x). \quad (4.2.8)$$

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

$$\bar{\partial}_\sigma \bar{T}^{\mu\nu\rho\dots}(\bar{x}) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \dots \frac{\partial x^\delta}{\partial \bar{x}^\sigma} \partial_\delta T^{\alpha\beta\gamma\dots}(x) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\rho{}_\gamma \dots (\Lambda^{-1})^\delta{}_\sigma \partial_\delta T^{\alpha\beta\gamma\dots}(x), \quad (4.2.9)$$

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

4. Classical Electromagnetism

This implies that the left-hand side of (4.2.5), 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 (4.1.1) and (4.1.2) to introduce the **electromagnetic potentials**. From (4.1.1) we know that \vec{B} has a **vector potential** \vec{A} such that

$$\vec{B}(t, \vec{x}) = \vec{\nabla} \times \vec{A}(t, \vec{x}). \quad (4.2.10)$$

Using this in (4.1.2) we find

$$\vec{\nabla} \times \left(\vec{E} + \frac{1}{c} \partial_t \vec{A} \right) = 0, \quad (4.2.11)$$

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

$$\vec{E} + \frac{1}{c} \partial_t \vec{A} = -\vec{\nabla} \Phi \Rightarrow \vec{E}(t, \vec{x}) = -\vec{\nabla} \Phi(t, \vec{x}) - \frac{1}{c} \partial_t \vec{A} = -\vec{\nabla} \Phi(t, \vec{x}) - \partial_0 \vec{A}. \quad (4.2.12)$$

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

$$[A^\mu(x)] = \begin{pmatrix} \Phi(x) \\ \vec{A}(x) \end{pmatrix}, \quad (4.2.13)$$

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**,

$$F_{\alpha\beta}(x) = \partial_\alpha A_\beta(x) - \partial_\beta A_\alpha(x). \quad (4.2.14)$$

It is clear that the observable phenomena are described by this Faraday tensor, which can be expressed with the electric and magnetic field components (\vec{E}, \vec{B}), as we shall see next. This implies that the same physical situation is described also by the four-potential,

$$A'_\mu = A_\mu + \partial_\mu \chi \quad (4.2.15)$$

with an arbitrary scalar field χ . This is known as the **gauge invariance** of electromagnetism. We come back to this very important property in Sect. 4.5.

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 also 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 (4.2.13). Since $F_{\alpha\beta} = -F_{\beta\alpha}$ we just need to calculate the six components F_{a0} and F_{ab} with $a, b \in \{1, 2, 3\}$:

$$\begin{aligned} F_{0a} &= \partial_0 A_a - \partial_a A_0 = -\partial_0 A^a - \partial_a A^0 = -\frac{1}{c} \partial_t A^a - \partial_a \Phi = E_a, \\ F_{ab} &= \partial_a A_b - \partial_b A_a = -\partial_a A^b + \partial_b A^a = \epsilon_{bac} (\vec{\nabla} \times \vec{A})_c = -\epsilon_{abc} B_c. \end{aligned} \quad (4.2.16)$$

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

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}. \quad (4.2.17)$$

4.2. Manifestly covariant formulation of electrodynamics

Raising both indices we get

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix} \quad (4.2.18)$$

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.,

$$\partial_\mu F^{\mu\nu} = \begin{pmatrix} \vec{\nabla} \cdot \vec{E} \\ -\partial_0 \vec{E} + \vec{\nabla} \times \vec{B} \end{pmatrix} = \begin{pmatrix} \rho \\ \vec{j}/c \end{pmatrix} = \frac{1}{c}(j^\nu). \quad (4.2.19)$$

In the second step we have used the inhomogeneous Maxwell equations (4.1.3-4.1.4) 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

$$\tilde{\epsilon}^{\mu\nu\rho\sigma} = \Lambda^\mu_\alpha \Lambda^\nu_\beta \Lambda^\rho_\gamma \Lambda^\sigma_\delta \epsilon^{\alpha\beta\gamma\delta} = \epsilon^{\mu\nu\rho\sigma} \det(\hat{\Lambda}) = \epsilon^{\mu\nu\rho\sigma}. \quad (4.2.20)$$

In the same way we find for the covariant components

$$\epsilon_{\mu\nu\rho\sigma} = \epsilon^{\mu\nu\rho\sigma} \det(\hat{\eta}) = -\epsilon^{\mu\nu\rho\sigma}. \quad (4.2.21)$$

Finally we note that

$$\epsilon_{\alpha\beta\gamma\delta} \epsilon^{\mu\nu\rho\sigma} = -\delta_{\alpha\beta\gamma\delta}^{\rho\mu\nu\sigma}, \quad (4.2.22)$$

where

$$\delta_{\alpha\beta\gamma\delta}^{\mu\nu\rho\sigma} = \det \begin{pmatrix} \delta_\alpha^\mu & \delta_\beta^\mu & \delta_\gamma^\mu & \delta_\delta^\mu \\ \delta_\alpha^\nu & \delta_\beta^\nu & \delta_\gamma^\nu & \delta_\delta^\nu \\ \delta_\alpha^\rho & \delta_\beta^\rho & \delta_\gamma^\rho & \delta_\delta^\rho \\ \delta_\alpha^\sigma & \delta_\beta^\sigma & \delta_\gamma^\sigma & \delta_\delta^\sigma \end{pmatrix}. \quad (4.2.23)$$

This means that $\delta_{\alpha\beta\gamma\delta}^{\mu\nu\rho\sigma}$ is antisymmetric 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 (4.2.22) and the properties of the determinant (4.2.23) it is easy to evaluate all possible contractions of two Levi-Civita tensors,

$$\epsilon_{\alpha\beta\gamma\delta} \epsilon^{\mu\nu\rho\delta} = -\det \begin{pmatrix} \delta_\alpha^\mu & \delta_\beta^\mu & \delta_\gamma^\mu \\ \delta_\alpha^\nu & \delta_\beta^\nu & \delta_\gamma^\nu \\ \delta_\alpha^\rho & \delta_\beta^\rho & \delta_\gamma^\rho \end{pmatrix} = -\delta_{\alpha\beta\gamma}^{\mu\nu\rho}, \quad (4.2.24)$$

$$\epsilon_{\alpha\beta\gamma\delta} \epsilon^{\mu\nu\gamma\delta} = -2 \det \begin{pmatrix} \delta_\alpha^\mu & \delta_\beta^\mu \\ \delta_\alpha^\nu & \delta_\beta^\nu \end{pmatrix} = -2\delta_{\alpha\beta}^{\mu\nu}, \quad (4.2.25)$$

$$\epsilon_{\alpha\beta\gamma\delta} \epsilon^{\mu\beta\gamma\delta} = -6\delta_\alpha^\mu. \quad (4.2.26)$$

4. Classical Electromagnetism

The Levi-Civita 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

$$\dagger F^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}. \quad (4.2.27)$$

The tensor defined by the components $\dagger F^{\mu\nu}$ are called the **Hodge dual** of the Faraday tensor. Using (4.2.24-4.2.26) 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

$$\begin{aligned} \dagger(\dagger F)_{\alpha\beta} &= \frac{1}{2} \epsilon_{\alpha\beta\mu\nu} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \\ &= +\frac{1}{4} \epsilon_{\alpha\beta\mu\nu} \epsilon^{\rho\sigma\mu\nu} F_{\rho\sigma} \\ &= -\frac{1}{2} (\delta_{\alpha}^{\rho} \delta_{\beta}^{\sigma} - \delta_{\alpha}^{\sigma} \delta_{\beta}^{\rho}) F_{\rho\sigma} \\ &= -\delta_{\alpha}^{\rho} \delta_{\beta}^{\sigma} F_{\rho\sigma} = -F_{\alpha\beta}. \end{aligned} \quad (4.2.28)$$

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

$$\dagger F^{0\nu} = \frac{1}{2} \epsilon^{0\nu\alpha\beta} F_{\alpha\beta} = -B_{\nu}, \quad \nu \in \{1, 2, 3\} \quad (4.2.29)$$

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

$$\dagger F^{mn} = \frac{1}{2} \epsilon^{mn\alpha\beta} F_{\alpha\beta} = \epsilon^{mn0b} F_{0b} = \epsilon^{mnb} E_b. \quad (4.2.30)$$

In matrix notation this reads

$$\dagger F^{\mu\nu} = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{pmatrix}. \quad (4.2.31)$$

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 (4.2.19)

$$\partial_{\mu} \dagger F^{\mu\nu} = \begin{pmatrix} \vec{\nabla} \cdot \vec{B} \\ -\partial_0 \vec{B} - \vec{\nabla} \times \vec{E} \end{pmatrix} = 0, \quad (4.2.32)$$

where in the last step we have used the homogeneous Maxwell equations (4.1.1-4.1.2). So the Maxwell equations are equivalent to the manifestly covariant relativistic field equations for the Faraday tensor and its dual (4.2.19) and (4.2.32):

$$\partial_{\mu} F^{\mu\nu} = \frac{1}{c} j^{\nu}, \quad \partial_{\mu} \dagger F_{\mu\nu} = 0. \quad (4.2.33)$$

4.2. Manifestly covariant formulation of electrodynamics

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,

$$E^\mu = F^{\mu\nu} u_\nu = \begin{pmatrix} \vec{u} \cdot \vec{E} \\ u^0 \vec{E} + \vec{u} \times \vec{B} \end{pmatrix}, \quad B^\mu = \dagger F^{\mu\nu} u_\nu = \begin{pmatrix} \vec{u} \cdot \vec{B} \\ u^0 \vec{B} - \vec{u} \times \vec{E} \end{pmatrix}. \quad (4.2.34)$$

Indeed, for the observer at rest in the reference frame at hand, we have $u = (1, 0, 0, 0)^T$ and for this observer we have $(E^\mu) = (0, \vec{E})$ and $(B^\mu) = (0, \vec{B})$.

(4.2.34) 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 applying the appropriate boost matrix

$$\hat{\Lambda} = \begin{pmatrix} u^0 & -\vec{u}^T \\ -\vec{u} & \mathbb{1}_3 + (u^0 - 1)\vec{u}\vec{u}^T/\vec{u}^2 \end{pmatrix} \quad (4.2.35)$$

to the vectors given in (4.2.34),

$$(\vec{E}^\mu) = (\Lambda^\mu{}_\nu E^\nu) = \begin{pmatrix} 0 \\ \vec{E} \end{pmatrix}, \quad (\vec{B}^\mu) = (\Lambda^\mu{}_\nu B^\nu) = \begin{pmatrix} 0 \\ \vec{B} \end{pmatrix}, \quad (4.2.36)$$

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

$$\vec{E} = \vec{E}_\parallel + \gamma(\vec{E}_\perp + \vec{\beta} \times \vec{B}) = \gamma(\vec{E} + \vec{\beta} \times \vec{B}) - \frac{\gamma^2}{1+\gamma}\vec{\beta}(\vec{\beta} \cdot \vec{E}), \quad (4.2.37)$$

$$\vec{B} = \vec{B}_\parallel + \gamma(\vec{B}_\perp - \vec{\beta} \times \vec{E}) = \gamma(\vec{B} - \vec{\beta} \times \vec{E}) - \frac{\gamma^2}{1+\gamma}\vec{\beta}(\vec{\beta} \cdot \vec{B}). \quad (4.2.38)$$

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

$$\vec{V}_\parallel = \frac{\vec{\beta} \cdot \vec{V}}{\beta^2} \vec{\beta}, \quad \vec{V}_\perp = \vec{V} - \vec{V}_\parallel, \quad (4.2.39)$$

and

$$\vec{E}_\parallel = \gamma \vec{E} + \frac{1-\gamma}{\beta^2} \vec{\beta}(\vec{\beta} \cdot \vec{E}) = \gamma \vec{E} + \frac{1-\gamma}{1-1/\gamma^2} \vec{\beta}(\vec{\beta} \cdot \vec{E}) = \gamma \vec{E} = \gamma \vec{E} - \frac{\gamma^2}{1+\gamma} \vec{\beta}(\vec{\beta} \cdot \vec{E}), \quad (4.2.40)$$

and the analogous equation for \vec{B}_\parallel .

Of course, one obtains the same result by applying the boost transformation to the Faraday tensor and then reads off \vec{E} and \vec{B} , using (4.2.17),

$$\vec{F}^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma}. \quad (4.2.41)$$

From (4.2.34) with some algebra one can derive that

$$F_{\mu\nu} = E_\mu u_\nu - E_\nu u_\mu - \epsilon_{\mu\nu\rho\sigma} B^\rho u^\sigma. \quad (4.2.42)$$

4. Classical Electromagnetism

Now the homogeneous Maxwell equations (4.2.19) read

$$\partial_\mu F^{\mu\nu} = \partial_\mu E^\mu u^\nu - u^\mu \partial_\mu E^\nu - \epsilon^{\mu\nu\rho\sigma} (\partial_\mu B_\rho u_\sigma) = \frac{1}{c} j^\nu. \quad (4.2.43)$$

The dual tensor is

$$\dagger F^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} = \epsilon^{\mu\nu\rho\sigma} E_\rho u_\sigma + B^\mu u^\nu - B^\nu u^\mu, \quad (4.2.44)$$

and the inhomogeneous Maxwell equations (4.2.32) become

$$\partial_\mu \dagger F^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} \partial_\mu E_\rho u_\sigma + \partial_\mu B^\mu u^\nu - u^\mu \partial_\mu B^\nu. \quad (4.2.45)$$

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

$$F_{\mu\nu} F^{\mu\nu} = -\dagger F_{\mu\nu} \dagger F^{\mu\nu} = -2(\vec{E}^2 - \vec{B}^2), \quad F_{\mu\nu} \dagger F^{\mu\nu} = -4\vec{E} \cdot \vec{B}. \quad (4.2.46)$$

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

$$\vec{F} = \vec{E} + i\vec{B}. \quad (4.2.47)$$

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

$$\vec{F}^2 = \vec{E}^2 - \vec{B}^2 - 2i\vec{E} \cdot \vec{B}. \quad (4.2.48)$$

This shows that the invariants of the electromagnetic field (4.2.46) 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 (4.2.37-4.2.38) can be expressed as a complexified rotation matrices $\hat{O} \in \text{SO}(3, \mathbb{C})$, 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 $\text{SO}(3)$, 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 (4.2.37-4.2.38) we find

$$\vec{\tilde{F}} = \hat{\beta}(\hat{\beta} \cdot \vec{F}) - i\beta\gamma \hat{\beta} \times \vec{F} + \gamma[\hat{\beta} \times (\vec{F} \times \vec{\beta})] = \hat{D}_{\hat{\beta}}(i\eta)\vec{F}, \quad (4.2.49)$$

where $\hat{\beta} = \vec{\beta}/\beta$ and $\eta = \text{artanh } \beta$. 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 $\text{SO}(3, \mathbb{C})$. The usual rotations are represented by the subgroup $\text{SO}(3)$ and rotation-free boosts by those with imaginary rotation angles $i\eta$, 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 (4.1.1-4.1.4). Taking the divergence reads

$$\vec{\nabla} \cdot \vec{F} = \rho \quad (4.2.50)$$

and the curl

$$\vec{\nabla} \times \vec{F} = \vec{\nabla} \times (\vec{E} + i\vec{B}) = -\frac{1}{c} \partial_t \vec{B} + \frac{i}{c} \partial_t \vec{E} + \frac{i}{c} \vec{j} = \frac{i}{c} (\partial_t \vec{F} + \vec{j}) \quad (4.2.51)$$

or

$$\vec{\nabla} \times \vec{F} - \frac{i}{c} \partial_t \vec{F} = \frac{i}{c} \vec{j}. \quad (4.2.52)$$

4.3 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 $\vec{k} = k(\cos \alpha, \sin \alpha, 0) = k\vec{n}$. Then also in the frame Σ' we have $\vec{k}' = k'(\cos \alpha', \sin \alpha', 0) = k'\vec{n}'$. As we shall see, the parameters $k' = c\omega'$ and α' are different from the corresponding values $k = \omega c$ and α , where the source is moving with $\vec{v} = \beta c \vec{e}_1$, i.e., both the frequency of the electromagnetic wave changes when the source is moving compared to the frequency when the source is at rest (**Doppler effect** of light) and also the propagation direction (**aberration** of light).

In the frame, where the light source is at rest a plane-wave solution for the electric field reads

$$\vec{E}'(\underline{x}') = \vec{E}'_0 \cos(\underline{k}' \cdot \underline{x}'). \quad (4.3.1)$$

We assume that in Σ' the axes are chosen such that $\vec{n}' = (\cos \alpha', \sin \alpha', 0)^T$ and $\vec{E}'_0 = E'_{01} \vec{e}'_1 + E'_{02} \vec{e}'_2$. Here

$$\vec{e}'_1 = \begin{pmatrix} -\sin \alpha' \\ \cos \alpha' \\ 0 \end{pmatrix}, \quad \vec{e}'_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (4.3.2)$$

are two vectors perpendicular to \vec{n}' such that $\vec{e}'_1 \times \vec{e}'_2 = \vec{n}'$. The magnetic field is uniquely determined by $\vec{B}'_0 = \vec{n}' \times \vec{E}'_0 = (E'_{01} \vec{e}'_2 - E'_{02} \vec{e}'_1)$ and

$$\vec{E}'(\underline{x}') = \vec{E}'_0 \cos(\underline{k}' \cdot \underline{x}'), \quad \vec{B}'(\underline{x}') = \vec{B}'_0 \cos(\underline{k}' \cdot \underline{x}'). \quad (4.3.3)$$

Further due to Lorentz invariance we have $\underline{k}' \cdot \underline{x}' = \underline{k} \cdot \underline{x}$ and thus because of (4.2.37)

$$\vec{E}(\underline{x}) = \vec{E}_0 \cos(\underline{k} \cdot \underline{x}), \quad \vec{B}(\underline{x}) = \vec{B}_0 \cos(\underline{k} \cdot \underline{x}) \quad (4.3.4)$$

with

$$\vec{E}_0 = \begin{pmatrix} -E'_{01} \sin \alpha' \\ \gamma E'_{01} (\beta + \cos \alpha') \\ E'_{02} \gamma (1 + \beta \cos \alpha') \end{pmatrix}, \quad \vec{B}_0 = \frac{1}{c} \begin{pmatrix} E'_{02} \sin \alpha' \\ -E'_{02} \gamma (\beta + \cos \alpha') \\ E'_{01} \gamma (1 + \beta \cos \alpha') \end{pmatrix}. \quad (4.3.5)$$

Now

$$\underline{k} = \hat{\Lambda}^{-1} \underline{k}' = \hat{\Lambda}^{-1} \begin{pmatrix} k' \\ k' \cos \alpha' \\ k' \sin \alpha' \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k' (1 + \beta \cos \alpha') \\ \gamma k' (\beta + \cos \alpha') \\ k' \sin \alpha' \\ 0 \end{pmatrix} = \begin{pmatrix} k \\ k \vec{n} \end{pmatrix} \quad (4.3.6)$$

and thus

$$k = \gamma k' (1 + \beta \cos \alpha'), \quad \vec{n} = \begin{pmatrix} (\beta + \cos \alpha') / (1 + \beta \cos \alpha') \\ \sqrt{1 - \beta^2} \sin \alpha' / (1 + \beta \cos \alpha') \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{pmatrix}. \quad (4.3.7)$$

As expected the fields (4.3.5) in Σ are of the same form as in Σ' , i.e.,

$$\vec{e}_1 = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{E} = (E_{01} \vec{e}_1 + E_{02} \vec{e}_2) \cos(\underline{k} \cdot \underline{x}), \quad \vec{B} = \frac{1}{c} \vec{n} \times \vec{E}. \quad (4.3.8)$$

4. Classical Electromagnetism

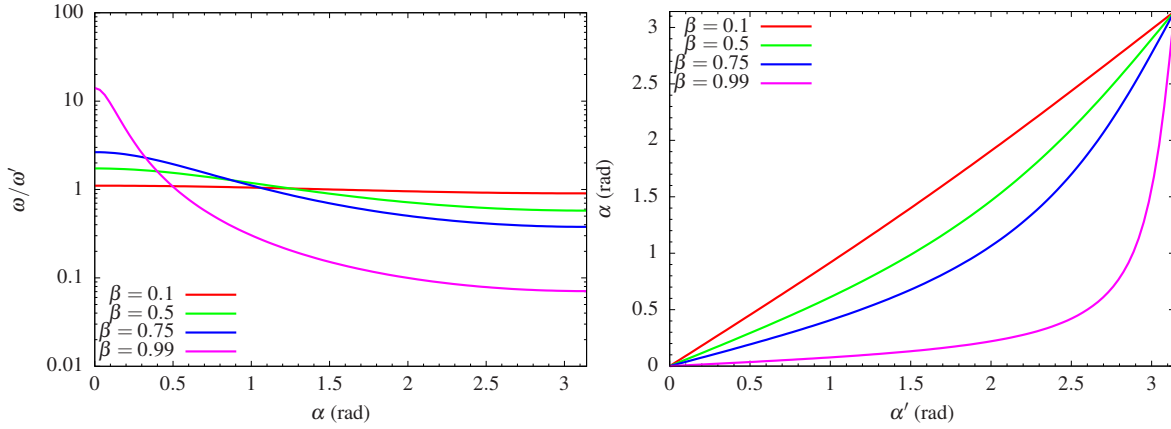


Figure 4.1: *Left Doppler effect for light according to (4.3.14); right: aberration of light according to (4.3.12).*

Also the observer in Σ thus sees a plane wave as in the restframe of the source, Σ' but with different parameters. This must be so, because the Maxwell equations are a system of relativistically covariant field equations, as discussed above, and the plane-wave solutions of the (free) Maxwell equations are uniquely determined.

According to (4.3.6) the frequency changes from its rest-frame value to

$$\omega = ck = \gamma\omega'(1 + \beta \cos\alpha'). \quad (4.3.9)$$

This is the **Doppler effect** for light, and analogously to the Doppler effect for sound the frequency depends on the speed of the light source and the direction of observation, \vec{n} (parametrized with the angle α), relative to the direction of the velocity of the light source, $\vec{v} = \beta c\vec{e}_1$. The effect gets largest, if the light source moves directly towards the observer ($\alpha = 0$) or away from him ($\alpha = \pi$). Then from (4.3.7) also $\alpha' = 0$ or $\alpha' = \pi$, respectively. According to (4.3.9) we get for these cases (upper sign is for $\alpha = \alpha' = 0$ and lower sign is for $\alpha = \alpha' = \pi$)

$$\omega_{\pm} = \gamma\omega'(1 \pm \beta) = \omega' \sqrt{\frac{1 \pm \beta}{1 \mp \beta}}. \quad (4.3.10)$$

If the light source is moving towards the observer $\omega = \omega_+ > \omega'$. If the source is moving away from the observer we get $\omega = \omega_- < \omega'$. That means a light-source moving towards the observer leads to a blue-shift of the light, one that is moving away from the observer to a red-shift.

Another interesting case is if the light is observed in Σ in the direction perpendicular to the source, i.e., for $\alpha = \pi/2$. According to (4.3.7) this implies $\cos\alpha' = -\beta$ und $\sin\alpha' = \sqrt{1 - \beta^2}$. Using again (4.3.9) we find

$$\omega = \gamma\omega'(1 - \beta^2) = \omega' \sqrt{1 - \beta^2}, \quad (4.3.11)$$

i.e., contrary to the case of the (non-relativistic) acoustic Doppler effect one also has a Doppler shift of the frequency when the light is observed in a direction transverse to the velocity of the source, i.e., a **transverse Doppler effect**, and it's a red-shift by the inverse Lorentz factor $1/\gamma$. This can be understood as an effect purely due to time dilation: Considering the light wave as a clock, which ticks at a rate $T' = 2\pi/\omega'$ in the frame it moves with $v = \beta c$ it ticks at rate $T = 2\pi/\omega = \gamma T' = \gamma 2\pi/\omega'$ and thus $\omega = \omega'/\gamma$.

4.4. The Lagrangian for the electromagnetic field

Further the angle α is changed against the angle α' in the restframe of the source. This is known as the **aberration of light**. Defining $\alpha, \alpha' \in (-\pi, \pi]$ we get from (4.3.7)

$$\alpha = \text{sign}(\sin \alpha') \arccos\left(\frac{\beta + \cos \alpha'}{1 + \beta \cos \alpha'}\right). \quad (4.3.12)$$

As can be seen in Fig. 4.1, that usually $\alpha < \alpha'$ (except for the special cases, discussed above, $\alpha = \alpha' = 0$ and $\alpha = \alpha' = \pi$). The direction of the one polarization vector changes correspondingly according to (4.3.8). The other polarization vector perpendicular to the plane spanned by the velocity of the light-source and the direction of the wave vector (given in the frame Σ and in our choice of coordinates the (12)-plane), does not change.

Of course we can express the Doppler effect also directly with the angle α , defined in Σ , because obviously

$$\underline{k}' = \hat{\Lambda} \underline{k} = \hat{\Lambda} \begin{pmatrix} k \\ k \cos \alpha \\ k \sin \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k (1 - \beta \cos \alpha) \\ \gamma k (-\beta + \cos \alpha) \\ k \sin \alpha \\ 0 \end{pmatrix}. \quad (4.3.13)$$

As to be expected this deviates from (4.3.6) only by substituting $(k, \alpha) \rightarrow (k', \alpha')$ and $\beta \rightarrow -\beta$. From the temporal component of (4.3.13) we get (see the right plot in 4.1)

$$\omega' = \omega \frac{1 - \beta \cos \alpha}{\sqrt{1 - \beta^2}} \Rightarrow \omega = \omega' \frac{\sqrt{1 - \beta^2}}{1 - \beta \cos \alpha}. \quad (4.3.14)$$

Finally we compare the intensity of the light wave as observed in the two frames. The intensity is given by

$$\varepsilon = \frac{1}{2} \vec{E}_0^2 = \frac{1}{2} \vec{E}'_0^2 \gamma^2 (1 + \beta \cos \alpha')^2 = \varepsilon' \left(\frac{\omega}{\omega'}\right)^2. \quad (4.3.15)$$

The light intensity changes in comparison to that measured in the rest-frame of the source with the square of the ‘‘Doppler factor’’ ω/ω' .

4.4 The Lagrangian for the electromagnetic field

In this Section we analyze first the free electromagnetic field and then the ‘‘dust medium’’ as the most simple continuum-mechanical model for charged particles interacting through the electromagnetic interaction in terms of the field-theoretical action principle discussed in 3.1.

For the free fields we can find the Lagrangian for the electromagnetic four-potential A^μ quite easily from the fact that the (inhomogeneous) Maxwell equations are *linear*, i.e., the Lagrangian must be a quadratic form of the A^μ and its derivatives $\partial_\nu A^\mu$. Further the Maxwell equations are **gauge invariant**, i.e., also the action should not change under a gauge transformation (4.5.2). Thus we expect the Lagrangian to be a bilinear form of the gauge-invariant field-strength tensor $F_{\mu\nu}$. The two possible Poincaré invariants are

$$\mathcal{L}_1 = \alpha F_{\mu\nu} F^{\mu\nu}, \quad \mathcal{L}_2 = \beta^\dagger F^{\mu\nu} F_{\mu\nu} = \beta \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} F_{\mu\nu} \quad (4.4.1)$$

with some constants α and β . The second candidate for a Lagrangian is however ruled out by the empirical fact that the electromagnetic interaction is **invariant under spatial reflections**, which is a Lorentz transformation with determinant (-1) . Since the Levi-Civita tensor changes its sign under such

4. Classical Electromagnetism

Lorentz transformations, \mathcal{L}_2 is a pseudoscalar rather than a scalar quantity under spatial reflections. Thus we set $\beta = 0$. So up to the factor α the Lagrangian for the free electromagnetic field is determined to be \mathcal{L}_1 . The usual convention for the choice of the coefficient is $\alpha = -1/4$. Thus we have

$$\mathcal{L}_0^{(\text{em})} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} = -\frac{1}{2}F^{\mu\nu}\partial_\mu A_\nu. \quad (4.4.2)$$

To see that this leads to the correct source-free Maxwell equations, we calculate the variation of the corresponding action,

$$\delta S_0^{(\text{em})} = -\frac{1}{2} \int d^4x F^{\mu\nu} \delta F_{\mu\nu} = - \int d^4x F^{\mu\nu} \partial_\mu \delta A_\nu = + \int d^4x \delta A_\nu \partial_\mu F^{\mu\nu} \stackrel{!}{=} 0. \quad (4.4.3)$$

Since this should hold for all δA_ν , the equations of motion indeed are the source-free inhomogeneous Maxwell equations, i.e., (4.2.33) with $j^\nu = 0$,

$$\partial_\mu F^{\mu\nu} = 0. \quad (4.4.4)$$

Next we evaluate the **canonical energy-momentum tensor**. From the general formula (3.1.20) we get

$$\Theta_{\mu\nu}^{(\text{em})} = \frac{\partial \mathcal{L}_0^{(\text{em})}}{\partial(\partial^\mu A_\rho)} \partial_\nu A_\rho - \mathcal{L} \eta_{\mu\nu} = F_{\rho\mu} \partial_\nu A^\rho + \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta_{\mu\nu}. \quad (4.4.5)$$

We note that this cannot be the observable energy-momentum tensor since it is not a gauge-invariant expression. As discussed in Sect. 3.3, Noether's theorem does not uniquely determine the "conserved currents" of a symmetry (here the symmetry under space-time translations) but only up to a "pseudogauge transformation", i.e., one can introduce an alternative energy-momentum tensor of the form,

$$T_{\mu\nu}^{(\text{em})} = \Theta_{\mu\nu}^{(\text{em})} + \partial_\rho \Omega^{\mu\rho\nu} \quad (4.4.6)$$

with an arbitrary field $\Omega^{\mu\rho\nu}$ obeying $\Omega^{\mu\rho\nu} = -\Omega^{\nu\rho\mu}$. For (4.4.6) the same local conservation law holds as for (4.4.5),

$$\partial^\mu T_{\mu\nu}^{(\text{em})} = \partial^\mu \Theta_{\mu\nu}^{(\text{em})} = 0, \quad (4.4.7)$$

and the total energy-momentum vector,

$$P^\nu = \int_{\mathbb{R}^3} d^3x T_{\text{em}}^{0\nu} = \int_{\mathbb{R}^3} d^3x \Theta_{\text{em}}^{0\nu} \quad (4.4.8)$$

is also the same for both energy-momentum tensors. To make (4.4.6) gauge invariant, we should have

$$\partial_\rho \Omega^{\mu\rho\nu} = -F^{\rho\mu} \partial_\rho A^\nu \stackrel{(4.4.4)}{=} -\partial_\rho (F^{\rho\mu} A^\nu) \Rightarrow \Omega^{\mu\rho\nu} = -F^{\rho\mu} A^\nu, \quad (4.4.9)$$

which is indeed antisymmetric under exchange of the indices ρ and μ . With this choice we get the **symmetric energy-momentum tensor** of the electromagnetic field,

$$T_{\mu\nu}^{(\text{em})} = F_{\rho\mu} F_\nu{}^\rho + \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta_{\mu\nu} = \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta_{\mu\nu} - F_{\rho\mu} F^\rho{}_\nu, \quad (4.4.10)$$

which is also obviously gauge invariant.

4.4. The Lagrangian for the electromagnetic field

Now we consider a charged dust medium as introduced in Sect. 3.5.4 as the most simple example of a charged medium interacting via the electromagnetic interaction. The Lagrangian for the free dust particles is again given by (3.5.56),

$$\mathcal{L}_0^{(\text{mech})} = -mc^2 n. \quad (4.4.11)$$

The interaction with the electromagnetic field is described by

$$\mathcal{L}^{(\text{int})} = -\frac{q}{c} J_N^\mu A_\mu, \quad (4.4.12)$$

where q is the charge of the dust particles. We first note that this provides a gauge-invariant contribution to the action, because for $A'_\mu = A_\mu + \partial_\mu \chi$ the gauge-transformed Lagrangian reads

$$\mathcal{L}^{(\text{int})'} = \mathcal{L}^{(\text{int})} - \frac{q}{c} J_N^\mu \partial_\mu \chi = \mathcal{L}^{(\text{int})} - \frac{q}{c} \partial_\mu (J_N^\mu \chi), \quad (4.4.13)$$

due to particle-number conservation. Thus the gauge-transformed interaction Lagrangian differs from the original one only by a total divergence and thus the corresponding action functionals are the same. We also obtain the correct inhomogeneous Maxwell equations since the variation of the interaction part of the action wrt. the A^μ is

$$\delta S_{\text{int}} = - \int d^4x \frac{q}{c} J_N^\mu \delta A_\mu. \quad (4.4.14)$$

Adding this to the variation of the action for the free electromagnetic field (4.4.3) indeed leads to the correct inhomogeneous Maxwell equations with non-vanishing sources,

$$\partial_\mu F^{\mu\nu} = \frac{q}{c} J_N^\nu. \quad (4.4.15)$$

As to be expected the electromagnetic four-current density is

$$J^\nu = q J_N^\nu = q c n u^\nu. \quad (4.4.16)$$

To find also the equations of motion for the medium we use energy-momentum conservation. First we work with the canonical energy-momentum tensor. In addition to (4.4.5) we have the energy-momentum tensor for the free dust particles (3.5.60),

$$\Theta_{\mu\nu}^{(\text{mech})} = mc^2 n u^\mu u^\nu, \quad (4.4.17)$$

and with (3.5.32-3.5.34) we find for the contribution of the interaction Lagrangian

$$\Theta_{\mu\nu}^{(\text{int})} = \frac{q}{c} J_{N\mu} A_\nu = \rho n u_\mu A_\nu. \quad (4.4.18)$$

The equations of motion for the medium follow from the local energy-momentum conservation after some algebra,

$$\partial_\mu \Theta_{\text{mech}}^{\mu\nu} = mc^2 n u^\mu \partial_\mu u^\nu = -\partial_\mu \Theta_{\text{em}}^{\mu\nu} - \partial_\mu \Theta_{\text{int}}^{\mu\nu} = q n F^{\nu\rho} u_\rho = \frac{1}{c} F^{\nu\rho} J_\rho. \quad (4.4.19)$$

We note that despite the fact that the canonical energy-momentum tensors for the fields (4.4.5) as well as from the interaction term (4.4.18) are *not* gauge invariant, we obtain the expected gauge-invariant

4. Classical Electromagnetism

equations of motion, since obviously on the right-hand side of (4.4.20) the expected gauge-invariant Lorentz-force density occurs.

This is, of course, expected, because the equations of motion for the medium also follow from the variation of the action, which is gauge invariant (cf. Sect. 3.5.2).

Again, we can use the same pseudo-gauge transformation as for the free field (4.4.9) to obtain a gauge-invariant total energy-momentum tensor for the electromagnetic field and the interaction, because with (4.4.15) we get

$$\partial_\rho \Omega^{\mu\rho\nu} = -\partial_\rho (F^{\rho\mu} A^\nu) = -nq u^\mu A^\nu - F^{\rho\mu} \partial_\rho A^\nu. \quad (4.4.20)$$

This leads to

$$T_{\text{em}}^{\mu\nu} = \Theta_{\text{em}}^{\mu\nu} + \Theta_{\text{int}}^{\mu\nu} + \partial_\rho \Omega^{\mu\rho\nu} = \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} \eta^{\mu\nu} - F_\rho{}^\mu F^{\rho\nu} = F^\mu{}_\rho F^{\rho\nu} + \frac{1}{4} \eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma}, \quad (4.4.21)$$

which is formally the same gauge-invariant expression as for the free electromagnetic field. The interaction term is now “hidden” in the field expressions.

Of course the same equation of motion for the medium (4.4.19) can be derived from this, using (4.4.19), because

$$\partial_\mu \Theta_{\text{mech}}^{\mu\nu} = mc^2 n u^\mu \partial_\mu u^\nu = -\partial_\mu T_{\text{em}}^{\mu\nu} = -\partial_\mu (\Theta_{\text{em}}^{\mu\nu} + \Theta_{\text{int}}^{\mu\nu} + \partial_\rho \Omega^{\mu\rho\nu}) \equiv -\partial_\mu (\Theta_{\text{em}}^{\mu\nu} + \Theta_{\text{int}}^{\mu\nu}). \quad (4.4.22)$$

The expressions for the various quantities defined by the components of the symmetric and gauge invariant energy-momentum tensor of the electromagnetic field (4.4.21) in terms of the 3D formalism of electrodynamics can be simply read off the definition (4.4.21) by expressing the Faraday tensor in terms of (\vec{E}, \vec{B}) . The result is

$$\varepsilon_{\text{em}} = T_{\text{em}}^{00} = \frac{1}{2} (\vec{E}^2 + \vec{B}^2), \quad (4.4.23)$$

$$c\Pi_{\text{em}}^j = \frac{1}{c} S^j = T^{0j} = \vec{E} \times \vec{B}, \quad (4.4.24)$$

$$T_{\text{Max}}^{jk} = -T_{\text{em}}^{(jk)} = \frac{1}{2} (\vec{E}^2 + \vec{B}^2) \delta^{jk} - E^j E^k - B^j B^k. \quad (4.4.25)$$

Here, ε_{em} denotes the **energy density** of the electromagnetic field, $\vec{\Pi}$ its **momentum density**, \vec{S} the **energy-current density** (or **Poynting vector**), and T_{Max}^{jk} the **Maxwell stress tensor**, expressing the stress tensor of the medium due to the electromagnetic interaction.

4.5 Gauge invariance and retarded potentials

As we have already mentioned in Sect. 4.2 the homogeneous Maxwell equations guarantee the existence of the four-vector potential A_μ with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (4.5.1)$$

and for any scalar field χ

$$A'_\mu = A_\mu + \partial_\mu \chi \quad (4.5.2)$$

we get the same Faraday tensor (4.5.1) since $\partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi = 0$. This is called **gauge invariance** of the electromagnetic potential.

4.5. Gauge invariance and retarded potentials

Plugging (4.5.1) into the inhomogeneous Maxwell equation (4.2.19) we get the equation of motion for the four-potential,

$$\partial^\mu F_{\mu\nu} = \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = \square A_\nu - \partial_\nu \partial_\mu A^\mu = \frac{1}{c} j^\nu. \quad (4.5.3)$$

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

$$\square = \partial_\mu \partial^\mu = \frac{1}{c^2} \partial_t^2 - \vec{\nabla}^2 = \frac{1}{c^2} \partial_t^2 - \Delta. \quad (4.5.4)$$

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 (4.5.3).

Obviously the equations (4.5.3) become separated for the four components of A^μ if one chooses the **Lorenz-gauge condition**,

$$\partial_\mu A^\mu = 0, \quad (4.5.5)$$

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 (4.5.5) this condition is also fulfilled in any other inertial frame.

Thus with the Lorenz-gauge condition (4.5.5) the equation of motion for the vector potential (4.5.3) simplifies to

$$\square A^\nu = \frac{1}{c} j^\nu. \quad (4.5.6)$$

This equation is consistent with the gauge condition, because

$$0 = \square \partial_\nu A^\nu = \frac{1}{c} \partial_\nu j^\nu, \quad (4.5.7)$$

and this is fulfilled according to the conservation of electric charge, which leads to the continuity equation (4.2.4). This condition of charge conservation has to be fulfilled thus not only as an integrability condition for the Maxwell equations but also to fulfill gauge invariance.

Now we investigate, how to solve (4.5.3) 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 is 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

$$\square \Phi(x) = J(x), \quad (4.5.8)$$

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.,

$$\square_x D(x, x') = \delta^{(4)}(x - x'). \quad (4.5.9)$$

4. Classical Electromagnetism

Having found such a solution, a particular solution of the inhomogeneous equation (4.5.8) is given by

$$\Phi(x) = \int_{\mathbb{R}^4} d^4x' D(x, x') J(x'). \quad (4.5.10)$$

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**, it 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 evoking the argument with the causality principle, defining the **causal time arrow**. Thus, we look for the so-called **retarded Green's function**, subject to the constraint

$$D(x, x') \equiv D(x - x') = \Theta(x^0 - x'^0) d(x - x'). \quad (4.5.11)$$

In this ansatz we have already taken into account that (4.5.9) 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

$$\square D(y) = \delta^{(4)}(y). \quad (4.5.12)$$

To this end we express D as a Fourier transform wrt. y^0

$$D(y) = \int_{\mathbb{R}} \frac{dk^0}{2\pi} \exp(-ik^0 y^0) \tilde{D}(k^0, \vec{y}). \quad (4.5.13)$$

Then we get from (4.5.12)

$$\delta^{(4)}(y) = \delta^{(3)}(\vec{y}) \int_{\mathbb{R}} \frac{dk^0}{2\pi} \exp(-ik^0 y^0) = \int_{\mathbb{R}} \frac{dk^0}{2\pi} (-k^{02} - \Delta) \tilde{D}(k^0, \vec{y}) \quad (4.5.14)$$

or

$$-(k^{02} + \Delta) \tilde{D}(k^0, \vec{y}) = \delta^{(3)}(\vec{y}). \quad (4.5.15)$$

Due to the invariance under rotations, this suggests the ansatz $\tilde{D}(k^0, \vec{y}) = F(k^0, r)/r$ with $r = |\vec{y}|$. Then (4.5.15) for $r \neq 0$ simplifies to

$$\begin{aligned} \partial_r^2 F(k^0, r) &= -k^{02} F(k^0, r) \Rightarrow F(k^0, r) = A \exp(ik_0 r) + B \exp(-ik_0 r) \\ \Rightarrow D(k^0, r) &= \frac{A}{r} \exp(ik_0 r) + \frac{B}{r} \exp(-ik_0 r). \end{aligned} \quad (4.5.16)$$

Plugging this into (4.5.13) we get

$$D(y) = \frac{A}{r} \delta(y^0 - r) + \frac{B}{r} \delta(y^0 + r). \quad (4.5.17)$$

Due to the above introduced retardation condition that $D(y) = 0$ for $y^0 < 0$ we must have $B = 0$. Further, for $k^{02} = 0$ and $B = 0$ (4.5.15) must reduce to the known Green's function for $-\Delta$ which is $1/(4\pi r)$, leading to $A = 1/(4\pi)$. So we finally obtain

$$D(y) = \frac{1}{4\pi|\vec{y}|} \delta(y^0 - |\vec{y}|). \quad (4.5.18)$$

4.6. The retarded fields

As we shall see below, for many applications it is more convenient to rewrite this in a manifestly covariant way:

$$D(y) = \frac{1}{2\pi} \Theta(y^0) \delta(y^{02} - \vec{y}^2) = \frac{1}{2\pi} \Theta(y^0) \delta(y \cdot y) \equiv \frac{1}{2\pi} \Theta(y^0) \delta(y^2). \quad (4.5.19)$$

We note that this function is invariant under orthochronous Lorentz transformations, which do not change the sign of the time component of any four-vector. This again reflects our particular choice of the Green's function of the D'Alembert operator, i.e., the *retarded* one.

To solve (4.5.6) it is most convenient to use the form (4.5.18) of the Green's function, leading to the **retarded four-potential**,

$$A^\mu(x) = \int_{\mathbb{R}^4} d^4 x' \frac{\delta(x^0 - x'^0 - |\vec{x} - \vec{x}'|)}{4\pi c |\vec{x} - \vec{x}'|} j^\mu(x') = \int_{\mathbb{R}^3} d^3 x' \frac{j^\mu(x^0 - |\vec{x} - \vec{x}'|, \vec{x}')}{4\pi c |\vec{x} - \vec{x}'|}. \quad (4.5.20)$$

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

$$A^\mu(x) = \int_{\mathbb{R}^3} d^3 x' \frac{j^\mu(t - |\vec{x} - \vec{x}'|/c, \vec{x}')}{4\pi c |\vec{x} - \vec{x}'|}. \quad (4.5.21)$$

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} at time t 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 (4.5.21) is a valid solution for potentials, leading to a Faraday tensor that solves all Maxwell equations, we must check that the Lorenz-gauge condition (4.6.5) is fulfilled, because only then the solution of (4.6.6) leads to a valid solution of the Maxwell equations. To prove this, we write the solution in unintegrated form,

$$A^\mu(x) = \int_{\mathbb{R}^4} d^4 x' D(x - x') \frac{j^\mu(x')}{c}, \quad (4.5.22)$$

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

$$\begin{aligned} \partial_\mu A^\mu(x) &= \int_{\mathbb{R}^4} d^4 x' \frac{\partial}{\partial x^\mu} D(x - x') \frac{j^\mu(x')}{c} \\ &= - \int_{\mathbb{R}^4} d^4 x' \frac{\partial}{\partial x'^\mu} D(x - x') \frac{j^\mu(x')}{c} \\ &= \int_{\mathbb{R}^4} d^4 x' D(x - x') \frac{\partial}{\partial x'^\mu} \frac{j^\mu(x')}{c} = 0, \end{aligned} \quad (4.5.23)$$

where in the last step we have made use of the continuity equation (4.2.4) again. Thus the retarded solution (4.5.21) indeed fulfills the Lorenz-gauge condition (4.6.5) and the wave equations (4.6.6) and thus provides a valid electromagnetic potential.

4.6 The retarded fields

With (4.5.21) 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,

4. Classical Electromagnetism

$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. It is, however, simpler to use the non-covariant notation,

$$\vec{E}(t, \vec{x}) = -\vec{\nabla}A^0(t, \vec{x}) - \frac{1}{c}\partial_t \vec{A}(t, \vec{x}), \quad \vec{B}(t, \vec{x}) = \vec{\nabla} \times \vec{A}(t, \vec{x}). \quad (4.6.1)$$

To evaluate the derivatives, it is generally more convenient to work with the four-dimensional integral (4.5.23), taking the derivatives and only afterwards perform the x'^0 (or t' integral) to get rid of the δ distribution:

$$A^\mu(t, \vec{x}) = \int_{\mathbb{R}} dt' \int_{\mathbb{R}^3} d^3x' \frac{\delta(t-t'-|\vec{x}-\vec{x}'|/c)}{4\pi c|\vec{x}-\vec{x}'|} j^\mu(t', \vec{x}'). \quad (4.6.2)$$

Now we evaluate the derivatives we need in (4.6.1). Introducing the abbreviation $\vec{R} = \vec{r} - \vec{r}'$ and $R = |\vec{r} - \vec{r}'| = |\vec{R}|$, we find

$$\begin{aligned} \partial_t \vec{A}(t, \vec{x}) &= \int_{\mathbb{R}} dt' \int_{\mathbb{R}^3} d^3x' \frac{\partial_t \delta(t-t'-R/c)}{4\pi cR} \vec{j}(t', \vec{x}') \\ &= - \int_{\mathbb{R}} dt' \int_{\mathbb{R}^3} d^3x' \frac{\partial'_t \delta(t-t'-R/c)}{4\pi cR} \vec{j}(t', \vec{x}') \\ &= \int_{\mathbb{R}} dt' \int_{\mathbb{R}^3} d^3x' \frac{\delta(t-t'-R/c)}{4\pi cR} \partial'_t \vec{j}(t', \vec{x}') \\ &= \int_{\mathbb{R}^3} d^3x' \frac{1}{4\pi cR} \left[\partial'_t \vec{j}(t', \vec{x}') \right]_{t'=t-R/c}. \end{aligned} \quad (4.6.3)$$

Further we have

$$\begin{aligned} \vec{\nabla}A^0 &= \int_{\mathbb{R}} dt' \int_{\mathbb{R}^3} d^3x' \rho(t', \vec{x}') \vec{\nabla} \frac{\delta(t-t'-R/c)}{4\pi R} \\ &= \int_{\mathbb{R}^3} dt' \int_{\mathbb{R}^3} d^3x' \rho(t', \vec{x}') \left[-\partial_t \delta(t-t'-R/c) \frac{\vec{R}}{4\pi cR^2} - \delta(t-t'-R/c) \frac{\vec{R}}{4\pi R^3} \right] \\ &= \int_{\mathbb{R}^3} dt' \int_{\mathbb{R}^3} d^3x' \rho(t', \vec{x}') \left[\partial'_t \delta(t-t'-R/c) \frac{\vec{R}}{4\pi cR^2} - \delta(t-t'-R/c) \frac{\vec{R}}{4\pi R^3} \right] \\ &= - \int_{\mathbb{R}^3} d^3x' \left[\frac{\partial'_t \rho(t', \vec{x}') \vec{R}}{4\pi cR^2} + \frac{\rho(t', \vec{x}') \vec{R}}{4\pi R^3} \right]_{t'=t-R/c} \end{aligned} \quad (4.6.4)$$

and

$$\begin{aligned} \vec{\nabla} \times \vec{A}(t, \vec{x}) &= \int_{\mathbb{R}} dt' \int_{\mathbb{R}^3} d^3x' \left[\partial'_t \delta(t-t'-R/c) \frac{\vec{R}}{4\pi c^2 R^2} - \delta(t-t'-R/c) \frac{\vec{R}}{4\pi c R^3} \right] \vec{j}(t', \vec{x}') \\ &= + \int_{\mathbb{R}^3} d^3x' \left[\frac{\partial'_t \vec{j}(t', \vec{x}') \times \vec{R}}{4\pi c^2 R^2} + \frac{\vec{j}(t', \vec{x}') \times \vec{R}}{4\pi c R^3} \right]_{t'=t-R/c}. \end{aligned} \quad (4.6.5)$$

Plugging (4.6.3-4.6.5) into (4.6.1) we find the **Jefimenko equations**,

$$\vec{E}(t, \vec{x}) = \frac{1}{4\pi} \int_{\mathbb{R}^3} d^3x' \left[\frac{\rho(t', \vec{x}') \vec{R}}{R^3} + \frac{\partial'_t \rho(t', \vec{x}') \vec{R}}{cR^2} - \frac{\partial'_t \vec{j}(t', \vec{x}')}{c^2 R} \right]_{t'=t-R/c}, \quad (4.6.6)$$

$$\vec{B}(t, \vec{x}) = \frac{1}{4\pi} \int_{\mathbb{R}^3} d^3x' \left[\frac{\vec{j}(t', \vec{x}') \times \vec{R}}{cR^3} + \frac{\partial'_t \vec{j}(t', \vec{x}') \times \vec{R}}{c^2 R^2} \right]_{t'=t-R/c}. \quad (4.6.7)$$

4.7 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 $u_\mu u^\mu = 1$ with $u^0 = \gamma > 0$ 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. Then, defining $\rho^* = \text{sign } \rho \sqrt{j_\mu j^\mu}$, we have

$$(j^\mu) = c\rho^* \gamma \begin{pmatrix} 1 \\ \vec{\beta} \end{pmatrix} = c\rho^* u^\mu. \quad (4.7.1)$$

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

$$\rho^* = u_\mu j^\mu. \quad (4.7.2)$$

Now for a point charge q we have

$$j^\mu(x) = q \frac{d}{dt} y^\mu(t) \delta^{(3)}[\vec{x} - \vec{y}(t)] = qc \int_{\mathbb{R}} dt' \frac{d}{dt'} y^\mu(t') \delta^{(4)}[x - y(t')]. \quad (4.7.3)$$

Since this is invariant under reparametrization of the world line, we can rewrite this immediately in manifestly covariant form as

$$j^\mu(x) = qc \int_{\mathbb{R}} d\tau \dot{y}^\mu(\tau) \delta^{(4)}[x - y(\tau)] = qc^2 \int_{\mathbb{R}} d\tau u^\mu(\tau) \delta^{(4)}[x - y(\tau)]. \quad (4.7.4)$$

Here and in the following we write a dot for a derivative with respect to *proper* time of the particle. Now we check the continuity equation. To that end we note that

$$\frac{d}{d\tau} \delta^{(4)}[x - y(\tau)] = -\dot{y}^\mu(\tau) \partial_\mu \delta^{(4)}[x - y(\tau)] = -c u^\mu(\tau) \partial_\mu \delta^{(4)}[x - y(\tau)]. \quad (4.7.5)$$

From this we immediately find through partial integration,

$$\partial_\mu j^\mu(x) = qc^2 \int_{\mathbb{R}} d\tau u^\mu(\tau) \partial_\mu \delta^{(4)}[x - y(\tau)] = -qc \int_{\mathbb{R}} d\tau \frac{d}{d\tau} \delta^{(4)}[x - y(\tau)] = 0. \quad (4.7.6)$$

One can of course also verify the validity of the continuity equation, using the (1+3)-form in (4.7.3), according to which

$$\begin{aligned} \vec{\nabla} \cdot \vec{j} &= q \frac{d}{dt} \vec{y}(t) \cdot \vec{\nabla} \delta^{(3)}[\vec{x} - \vec{y}(t)], \\ \partial_0 j^0 &= q \frac{d}{dt} \delta^{(3)}[\vec{x} - \vec{y}(t)] = -q \frac{d}{dt} \vec{y}(t) \cdot \vec{\nabla} \delta^{(3)}[\vec{x} - \vec{y}(t)] \\ &\Rightarrow \partial_\mu j^\mu = 0. \end{aligned} \quad (4.7.7)$$

¹As we shall see, some problems occur, if $|\vec{\beta}| = |\vec{u}|/u^0 = |d\vec{y}/dt|/c$ approaches asymptotically 1 for $\tau \rightarrow \pm\infty$, 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 $|\vec{\beta}| \rightarrow 1$ -asymptotic behavior.

4. Classical Electromagnetism

Now the four-potential is easily evaluated using (4.5.22) and the manifestly covariant expression (4.5.19) for the retarded Green's function of the D'Alembert operator. With the covariant form for the four-current density (4.7.4) we integrate out the δ distribution, using $\delta[f(\tau)] = \sum_{\tau_0} \delta(\tau - \tau_0)/|\dot{f}(\tau_0)|$ to get

$$A^\mu(x) = \frac{q}{2\pi} \int_{-\infty}^{\infty} d\tau u^\mu(\tau) \theta[x_0 - y_0(\tau)] \delta\{[x - y(\tau)]^2\} \quad (4.7.8)$$

$$= \frac{q}{4\pi} \int_{-\infty}^{\infty} d\tau \frac{u^\mu(\tau)}{u(\tau) \cdot [x - y(\tau)]} \delta(\tau - \tau_{\text{ret}}) \quad (4.7.9)$$

$$= \frac{q}{4\pi} \frac{u^\mu(\tau_{\text{ret}})}{u(\tau_{\text{ret}}) \cdot [x - y(\tau_{\text{ret}})]}, \quad (4.7.10)$$

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

$$x^0 - y^0(\tau_{\text{ret}}) = |\vec{x} - \vec{y}(\tau_{\text{ret}})|. \quad (4.7.11)$$

Of course, the four-vector

$$z(\tau_{\text{ret}}, x) = x - y(\tau_{\text{ret}}) \quad (4.7.12)$$

is **light-like**, i.e., $z \cdot z = 0$ and $z^0 > 0$. One should note that in (4.7.8) 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 (4.7.9), one has to take this x -dependence of τ_{ret} into account. As we shall demonstrate below, it is again more advantageous to use the unintegrated expression (4.7.8).

This means that we can determine the retarded proper time in a Minkowski diagram by drawing the particle's world line $y(\tau)$ and then construct the light cone around the space-time point, where the fields are observed, x . The intersection of the past light-cone with the world line of the particle then determines τ_{ret} . Since z^0 is a monotonously decreasing function of τ , for each x the retarded proper time τ_{ret} is uniquely determined.

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 (4.7.10) it is more convenient to use the form (4.7.8). 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 anyway we exclude this contribution, because we can calculate the field only for x not on the particle's world line.

Thus we only need to consider

$$\begin{aligned} \partial_\mu \delta\{[x - y(\tau)]^2\} &= 2[x_\mu - y_\mu(\tau)] \delta'\{[x - y(\tau)]^2\} \\ &= 2[x_\mu - y_\mu(\tau)] \left(\frac{d}{d\tau} [x - y(\tau)]^2 \right)^{-1} \frac{d}{d\tau} \delta\{[x - y(\tau)]^2\} \\ &= -\frac{x_\mu - y_\mu(\tau)}{\dot{y}(\tau) \cdot [x - y(\tau)]} \frac{d}{d\tau} \delta\{[x - y(\tau)]^2\}. \end{aligned} \quad (4.7.13)$$

Inserting this into the corresponding derivative of (4.7.8), and integration by parts, ignoring the singular

4.8. Field of a uniformly moving point charge

distribution through the derivative of the Heaviside unit-step function as argued above, we finally get

$$\begin{aligned} \partial_\mu A_\nu(x) &= \frac{q}{2\pi} \int_{\mathbb{R}} d\tau \frac{d}{d\tau} \left\{ \frac{[x_\mu - y_\mu(\tau)]u_\nu(\tau)}{u \cdot [x - y(\tau)]} \right\} \Theta[x^0 - y^0(\tau)] \delta\{[x - y(\tau)]^2\} \\ &= \left[\frac{q}{4\pi c u(\tau) \cdot z} \frac{d}{d\tau} \left[\frac{z_\mu u_\nu(\tau)}{u(\tau) \cdot z} \right] \right]_{\tau=\tau_{\text{ret}}}, \end{aligned} \quad (4.7.14)$$

where in the final step we performed the same evaluation of the integral over τ which lead from (4.7.8) to (4.7.10). Antisymmetrization of this expression wrt. the indices μ and ν finally yields

$$F_{\mu\nu}(x) = \left[\frac{q}{4\pi c u(\tau) \cdot z} \frac{d}{d\tau} \left[\frac{z_\mu u_\nu - z_\nu u_\mu}{u \cdot z} \right] \right]_{\tau=\tau_{\text{ret}}}. \quad (4.7.15)$$

From now on we write

$$R = u(\tau) \cdot z = \frac{1}{c} \dot{y} \cdot z, \quad Q = \frac{1}{c} \dot{u}(\tau) \cdot z = \frac{1}{c^2} \ddot{y} \cdot z. \quad (4.7.16)$$

One should note that, for $\tau = \tau_{\text{ret}}$, R is the distance of the particle at proper time τ_{ret} from the observation point \vec{x} as measured in the momentary rest frame of the particle, i.e., $R > 0$. Working out the proper-time derivative in (4.7.15) we find

$$F_{\mu\nu} = \frac{q}{4\pi R^3} \left\{ \frac{R}{c^2} (z_\mu \ddot{y}_\nu - z_\nu \ddot{y}_\mu) - Q (z_\mu u_\nu - z_\nu u_\mu) + (z_\mu u_\nu - z_\nu u_\mu) \right\}. \quad (4.7.17)$$

One should note that the first two terms $\propto \ddot{y}$ parametrically fall off with $1/R$, while the remaining part falls off like $1/R^2$. As we shall see, this first part is the **radiative part**, i.e., the part of the field that describes the propagation of **energy and momentum** to “spatial infinity”, while the second part describes the particle’s electrostatic Coulomb field as observed from a reference frame, where the particle moves with four-velocity u .

4.8 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

$$y(\tau) = c u \tau, \quad u = \text{const}. \quad (4.8.1)$$

For given x the retarded time is determined by the equation

$$z^2(\tau) = [x - y(\tau)]^2 = 0, \quad x^0 > y^0(\tau). \quad (4.8.2)$$

Inserting (4.8.1), evaluating the square and simplifying a bit, we find the quadratic equation for τ

$$\tau^2 - \frac{2u \cdot x}{c} \tau + \frac{x^2}{c^2} = 0 \quad (4.8.3)$$

with the two solutions

$$\tau_\pm = \frac{u \cdot x}{c} \pm \frac{1}{c} \sqrt{(u \cdot x)^2 - x^2}. \quad (4.8.4)$$

4. Classical Electromagnetism

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

$$\tau_{\text{ret}} = \tau_- = \frac{\mathbf{u} \cdot \mathbf{x}}{c} - \frac{1}{c} \sqrt{(\mathbf{u} \cdot \mathbf{x})^2 - x^2}. \quad (4.8.5)$$

The denominator in (4.7.10) then is

$$R = \mathbf{u} \cdot \mathbf{z} = \mathbf{u} \cdot [x - \gamma(\tau_{\text{ret}})] = \mathbf{u} \cdot \mathbf{x} - [\mathbf{u} \cdot \mathbf{x} - \sqrt{(\mathbf{u} \cdot \mathbf{x})^2 - x^2}] = \sqrt{(\mathbf{u} \cdot \mathbf{x})^2 - x^2}. \quad (4.8.6)$$

Thus the four-potential is

$$A^\mu(x) = \frac{q u^\mu}{4\pi \sqrt{(\mathbf{u} \cdot \mathbf{x})^2 - x^2}}. \quad (4.8.7)$$

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

$$A^0(x) = \frac{q}{4\pi r}, \quad \vec{A}(x) = 0. \quad (4.8.8)$$

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 (4.8.7) is the only covariant generalization of (4.8.8) which can be built with the only four-vectors \mathbf{u} and \mathbf{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 (4.8.8).

The fields are given by (4.7.17). The radiative part $\propto \ddot{\mathbf{j}}$ vanishes and thus

$$F_{\mu\nu} = \frac{q}{4\pi [(\mathbf{u} \cdot \mathbf{x}) - x^2]^{3/2}} (z_\mu u_\nu - z_\nu u_\mu). \quad (4.8.9)$$

In (1+3)-notation, using (4.2.17) we can read off the electric and magnetic field components as

$$\vec{E} = \frac{q}{4\pi} \frac{\gamma(\vec{x} - \vec{v}t)}{[(\mathbf{u} \cdot \mathbf{x})^2 - x^2]^{3/2}}, \quad \vec{B} = \frac{q}{4\pi} \frac{\vec{u} \times \vec{x}}{[(\mathbf{u} \cdot \mathbf{x})^2 - x^2]^{3/2}} = \vec{\beta} \times \vec{E}. \quad (4.8.10)$$

In the final step we have used $\vec{u} = \gamma\vec{\beta} = \gamma\vec{v}/c$. The same result can of course be obtained by taking calculating $\vec{E} = -\partial_t \vec{A}/c - \nabla A^0$ and $\vec{B} = \nabla \times \vec{A}$ for the potentials given by (4.8.7).

To put this in a more familiar form, note that

$$(\mathbf{u} \cdot \mathbf{x})^2 - x^2 = \gamma^2 (ct - \vec{\beta} \cdot \vec{x})^2 - c^2 t^2 + \vec{x}^2 = \gamma^2 (\vec{x}_\parallel - \vec{v}t)^2 + \vec{x}_\perp^2, \quad (4.8.11)$$

where

$$\vec{x}_\parallel = \frac{\vec{\beta} \cdot \vec{x}}{\beta^2} \vec{\beta}, \quad \vec{x}_\perp = \vec{x} - \vec{x}_\parallel \quad (4.8.12)$$

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

$$\vec{E} = \frac{q}{4\pi} \frac{\gamma(\vec{x} - \vec{v}t)}{[\gamma^2 (\vec{x}_\parallel - \vec{v}t)^2 + \vec{x}_\perp^2]^{3/2}}. \quad (4.8.13)$$

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., $\vec{E} \propto \vec{x} - \vec{v}t$.

4.9 The radiation-reaction problem

While the treatment of the interaction between a charged medium² and the electromagnetic field is well-defined, i.e., it is, at least in principle, possible to solve the coupled equations for the electromagnetic field and the continuum-mechanics equations of motion for the charged medium self-consistently.

As we shall show now, the task to formulate a self-consistent theory for **point particles** is plagued by notorious problem, well-known since the first formulations by H. A. Lorentz and M. Abraham. A relativistic generalization has been derived by Dirac. For a nice review of these and other “classical-electron models”, see [Erb61]. Here we shall follow a modern derivation in the spirit of regularization and renormalization known from modern quantum field theory [Lec18, LM07].

The problem is to describe the influence of the electromagnetic field of the point article on its motion. It is clear that this “self field” of the particle is singular along its world line. Of course, a free electron should simply move with constant velocity. As we shall see in the next section, of course a free point particle does not radiate and thus there should also be no self-interaction of it with its own electromagnetic field, which is nothing than the “boosted Coulomb field” as calculated in Sect. 4.8. On the other hand if we want to calculate the proper total energy of such a particle (i.e., the total energy of the electron and its Coulomb field in the particle’s rest frame), we find a diverging result, because of the singularity of the electrostatic Coulomb potential at the position of the particle. Now, on the other hand, we can interpret this proper energy of the particle simply as described by its observed physical mass, m , which implies that already for the most simple case of a free charged particle we somehow must handle the unphysical infinite Coulomb-field energy and then take the unobservable “bare mass”, i.e., the mass of a fictitious particle neglecting the energy of the Coulomb field, such that it cancels the diverging field energy of this Coulomb field, leading to the observed mass of the particle.

Now, if the particle is accelerated by some external force (which can well be by an additional electromagnetic field as, e.g., in a particle accelerator), there are electromagnetic waves emitted carrying energy and momentum away from the particle. Thus we expect that the equation of motion for the particle should contain some **radiation-reaction force** or **self-force** acting like a friction force due to the loss of energy and momentum due to the radiation field.

Formally we shall handle this by applying a covariant regularization of the Liénard-Wiechert fields, making the parts that diverge along the worldline of the particle finite and enabling us to formally subtract the divergent self-energy of the particle and also obtain an expression for the radiation-reaction force, which will turn out to be formally finite. Then we shall investigate the notorious problems of the resulting **Lorentz-Abraham-Dirac equation** as **runaway solutions** and **preacceleration**, which indicates a violation of causality of this classical theory of a charged point particle. It should, however, be kept in mind that **quantum effects** become important already at much larger spacetime scales than the preacceleration time scale, which is closely related to a typical length scale, called the **classical electron radius**, which is defined as the typical radius of a particle of finite extent for which the entire mass is given by its (then of course finite!) electromagnetic self-energy. In this sense one can say that the classical theory of charged point particles itself hints at its invalidity!

4.9.1 The energy-momentum tensor of the Liénard-Wiechert field

We use the electromagnetic field of an arbitrarily moving massive point charge, given by the Liénard-Wiechert field (4.7.17) in 4.4.10 to evaluate the energy-momentum tensor of the electromagnetic field.

²Of course the medium model can be changed to an ideal fluid, leading to **magnetohydrodynamics** or an elastic medium as a simple model for a dielectric.

4. Classical Electromagnetism

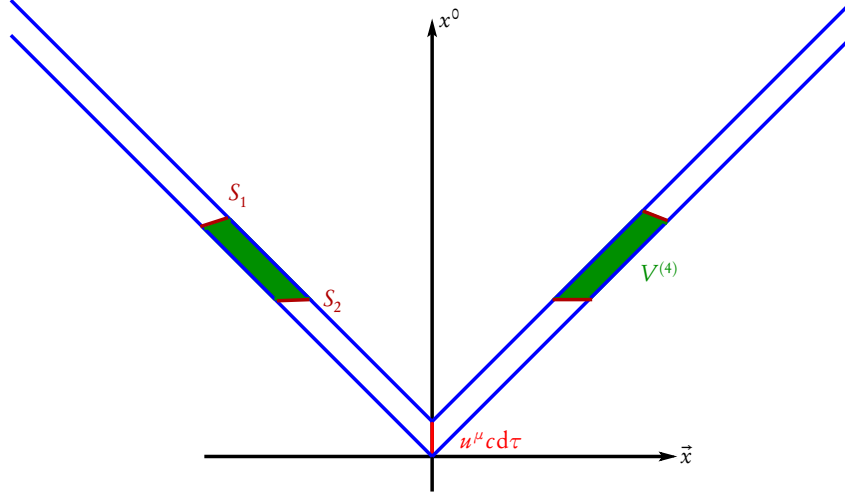


Figure 4.2: Radiation originating from an infinitesimal piece of the particle's worldline.

After a longer but straight-forward calculation we find

$$T_{\mu\nu}(x) = -\frac{q^2}{16\pi^2 R^4} \left(\frac{\ddot{y}^2}{c^4} + \frac{Q^2}{R^2} \right) z_\mu z_\nu \quad (4.9.1)$$

$$+ \frac{q^2}{16\pi^2 R^4} \left[\frac{2Q}{R^2} z_\mu z_\nu + \frac{1}{c^2} (z_\mu \ddot{y}_\nu + \ddot{y}_\mu z_\nu) - QR(z_\mu u_\nu + u_\mu z_\nu) \right] \quad (4.9.2)$$

$$+ \frac{q^2}{16\pi^2 R^6} \left[R(z_\mu u_\nu + u_\mu z_\nu) - z_\mu z_\nu - \frac{R^2}{2} \eta^{\mu\nu} \right]. \quad (4.9.3)$$

Since parametrically $Q, z_\mu \sim R$ the contribution (4.9.1) goes like $1/R^2$, (4.9.2) like $1/R^3$, and (4.9.3) like $1/R^4$. One should note again that for any x the argument for τ in the above quantities has to be taken $\tau = \tau_{\text{ret}} \equiv \tau_{\text{ret}}(x)$, cf. the solution of (4.4.11).

To evaluate the energy and momentum [Sch60] transferred from the accelerated particle to the electromagnetic radiation field, we note that the field due to an infinitesimal piece of the particle's worldline between proper times τ and $d\tau$ is located between the corresponding two light-cones cf. Fig. 4.2, where we have chosen the inertial reference frame such that the particle at the considered moment τ is at rest. Since in the green four-volume $V^{(4)}$ we have $\partial_\mu T^{\mu\nu} = 0$ we have

$$\int_{V^{(4)}} d^4x \partial_\mu T^{\mu\nu} = \int_{\partial V^{(4)}} d^3\Sigma_\mu T^{\mu\nu} = 0.$$

To evaluate the irradiated total momentum we can push the integration volume, $V^{(4)}$, very far away from the considered piece of the particle's world-line to get only the part of the energy and momentum which is carried to spatial infinity by the electromagnetic field. Then only the piece (4.9.1) contributes, and the parts of the 3D boundary $\partial V^{(4)}$ along the light cones do not contribute, because the corresponding normal vectors $d\Sigma^\mu \propto z^\mu$. Then it is clear that the total irradiated momentum from the considered infinitesimal piece of the particle's trajectory can be evaluated using any of the space-like hyper-surfaces like S_1 and S_2 . For convenience we choose S_2 :

$$P^\nu = \frac{1}{c} \int_{S_2} d^3\Sigma_\mu T^{\mu\nu} \quad (4.9.4)$$

4.9. The radiation-reaction problem

is a four-vector, which becomes clear from using the 4D Gauss's integral Theorem to the shaded region in Fig. 4.2 [Sch60]. If we integrate between two regions very far away from the infinitesimal piece of the particle's worldline under consideration, i.e., in the limit $R \rightarrow \infty$ the only piece of the energy-momentum tensor we have to consider in evaluating (4.9.4) is the "radiation piece" in the far-field approximation (4.9.1). Since the result is a four-vector we can also choose an arbitrary inertial reference frame. The most simple is the rest frame of the particle at τ , i.e., $u^\mu = \dot{y}^\mu/c = (1, 0, 0, 0)$. Then we can choose the spatial base vectors such that $\ddot{y} = (0, 0, 0, a)$. We also choose the origin of the frame such that $t = \tau_{\text{ret}}$ and $\vec{y}(\tau_{\text{ret}}) = 0$. As an integration region we choose the spherical shell with $R \rightarrow \infty$ around $y(\tau_{\text{ret}})$, i.e.,

$$S: \quad x = \begin{pmatrix} 0 \\ R \cos \varphi \sin \vartheta \\ R \sin \varphi \sin \vartheta \\ R \cos \vartheta \end{pmatrix}, \quad R \in (0, c d\tau), \quad \vartheta \in [0, \pi], \varphi \in [0, 2\pi). \quad (4.9.5)$$

Then

$$d\sigma = c d\tau d\vartheta d\varphi \sin \vartheta R^2 (1, 0, 0, 0) \quad (4.9.6)$$

and thus

$$dP^\mu = -\frac{q^2 d\tau}{(4\pi)^2 R c^4} \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi (-a^2)(1 - \cos^2 \vartheta) \begin{pmatrix} R \\ \vec{x} \end{pmatrix} = \frac{q^2 a^2}{6\pi c^4} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (4.9.7)$$

i.e., in covariant form

$$\dot{P}^\mu = -\frac{q^2}{6\pi c^4} (\ddot{y} \cdot \ddot{y}) u^\mu, \quad (4.9.8)$$

where the proper-time arguments are τ_{ret} . One should note that this is only the radiation energy and momentum carried to spatial infinity by the electromagnetic wave. The total change four-momentum due to radiation is of course given by

$$\Delta P^\mu = -\frac{q^2}{6\pi c^4} \int_{\mathbb{R}} d\tau (\ddot{y} \cdot \ddot{y}) u^\mu. \quad (4.9.9)$$

We stress again that in order for having a well-defined finite total change of the four-momentum we have to assume that $\ddot{y} \neq 0$ only in a finite spatial region or we must have a sufficiently strong convergence $\ddot{y} \rightarrow 0$ for $\tau \rightarrow \pm\infty$. Note that this assumptions are *not* fulfilled for hyperbolic motion, i.e., the motion of the charged particle in an everywhere homogeneous constant electric field (which is anyway an only academic example).

4.9.2 The Lorentz-Abraham-Dirac equation

The above consideration shows that there must be a backreaction on the particle due to the energy and momentum radiated off as calculated by (4.9.8), i.e., there must a kind of additional "friction force" for a particle in addition to the external force (e.g., due to an external electromagnetic field) accelerating the particle, caused by the interaction of the particle with its own electromagnetic field. Of course, this field is dominated by the near-field part, i.e., the boosted Coulomb field part, of the retarded field due to the term in (4.7.17) proportional to $z_\mu u_\nu - z_\nu u_\mu$. Obviously this Coulomb field diverges along the worldline of the particle, and thus we have to regularize it in some way to be able to make mathematical

4. Classical Electromagnetism

sense of the self-force on the particle. A very convenient regularization scheme, provided in [Lec18, LM07, Nak13], is to introduce a small spatial scale, ϵ and write as a **regularized retarded propagator**

$$D_\epsilon(x) = \frac{\Theta(x^0)}{2\pi} \delta(x^2 - \epsilon^2) = \frac{1}{4\pi} \frac{\delta(x^0 - \sqrt{\vec{x}^2 + \epsilon^2})}{\sqrt{\vec{x}^2 + \epsilon^2}} \quad (4.9.10)$$

instead of (4.5.19). Using this instead of the original retarded propagator for the D'Alembert operator in the calculation of the fields, instead of (4.7.14) one arrives at

$$\partial_\mu A_\nu^{(\epsilon)}(x) = -\frac{q_0}{2\pi} \int_{\mathbb{R}} d\tau \Theta(x^0 - z^0) \frac{z_\mu u_\nu}{z \cdot u} \frac{d}{d\tau} \delta(z^2 - \epsilon^2), \quad (4.9.11)$$

where again $z = x - y(\tau)$.

Now we want to evaluate this expression at $x = y(\tau_0)$, which thanks to our regularization is now well-defined for any given time-like worldline of the particle. To this end we write

$$\frac{d}{d\tau} \delta(z^2 - \epsilon^2) = -2z \cdot \dot{y} \delta'(z^2 - \epsilon^2) = -\frac{1}{\epsilon} z \cdot \dot{y} \frac{d}{d\epsilon} \delta(z^2 - \epsilon^2). \quad (4.9.12)$$

Plugging this into (4.9.11) leads to

$$\partial_\mu A_\nu^{(\epsilon)}[y(\tau_0)] = \frac{q}{2\pi\epsilon} \frac{d}{d\epsilon} \int_{\mathbb{R}} d\tau \Theta(y_0^0 - y^0) \delta[(y - y_0)^2 - \epsilon^2] (y_0 - y)_\mu u_\nu. \quad (4.9.13)$$

Since finally we want the expression for $\epsilon \rightarrow 0^+$, it is sufficient to expand the expression in powers of ϵ , which due to the δ distribution is equivalent to an expansion of the trajectory around τ_0 :

$$z(\tau) - z(\tau_0) = \dot{y}(\tau_0)(\tau - \tau_0) + \frac{1}{2} \ddot{y}(\tau_0)(\tau - \tau_0)^2 + \mathcal{O}(\tau^3), \quad (4.9.14)$$

from which

$$\delta[(y - y_0)^2 - \epsilon^2] = \delta[c^2 \tau^2 - \epsilon^2 + \mathcal{O}(\tau^3)]. \quad (4.9.15)$$

Using this in (4.9.13) yields

$$\begin{aligned} \partial_\mu A_\nu^{(\epsilon)}(y_0) &= \frac{q}{2\epsilon\pi} \frac{d}{d\epsilon} \left[\frac{1}{2\epsilon} \left(-u_0 \epsilon + \frac{\dot{y}}{2c^2} \epsilon^2 - \frac{\ddot{y}}{3!c^3} \epsilon^3 \right)_\mu u_\nu(\tau_0 - \epsilon) + \mathcal{O}(\epsilon^4) \right] \\ &= \frac{q}{4\pi\epsilon} \left[\frac{1}{c^2} u_{0\mu} \ddot{y}_{0\nu} + \frac{1}{2c^2} \ddot{y}_{0\mu} u_{0\nu} \right. \\ &\quad \left. - \epsilon \left(u_{0\mu} \ddot{y}_{\nu} / c^3 + \ddot{y}_{0\mu} \ddot{y}_{0\nu} / c^4 + \frac{1}{3} \ddot{y}_{0\nu} u_{0\mu} / c^2 \right) + \mathcal{O}(\epsilon^2) \right]. \end{aligned} \quad (4.9.16)$$

Finally antisymmetrizing in the indices μ and ν leads to the regularized field-strength tensor at the point $z_0 = z(\tau_0)$,

$$F_{\mu\nu}^{(\epsilon)}(z_0) = \frac{q}{4\pi} \left[\frac{1}{2\epsilon c^2} (u_{0\mu} \ddot{y}_{0\nu} - u_{0\nu} \ddot{y}_{0\mu}) - \frac{2}{3c^3} (u_{0\mu} \ddot{y}_{0\nu} - u_{0\nu} \ddot{y}_{0\mu}) \right] + \mathcal{O}(\epsilon). \quad (4.9.17)$$

Thus the regularized self-force reads

$$K_{\text{self},\mu}^{(\epsilon)} = q F_{\mu\nu}^{(\epsilon)}(z_0) u_0^\nu = -\frac{q^2}{8\pi\epsilon c^2} \ddot{y}_{0\mu} + \frac{q^2}{6\pi c^3} \left[\ddot{y}_{0\mu} - (u_0 \cdot \ddot{y}_0) u_{0\mu} \right] + \mathcal{O}(\epsilon). \quad (4.9.18)$$

4.9. The radiation-reaction problem

As expected this expression diverges for $\epsilon \rightarrow 0$, but writing down the equation of motion leads to

$$m_0 \ddot{y}_{0\mu} = K_\mu^{(\text{ext})} - \frac{q^2}{8\pi c^2 \epsilon} \ddot{y}_{0\mu} + \frac{q^2}{6\pi c^3} \left[\ddot{y}_{0\mu} - (u_0 \cdot \ddot{y}_0) u_{0\mu} \right] + \mathcal{O}(\epsilon). \quad (4.9.19)$$

To make this finite, we argue that the 2nd term on the right-hand side is of the form of the left-hand side with a diverging mass. Bringing this term to the right-hand side and interpreting the resulting coefficient of $\ddot{y}_{0\mu}$ as the observable finite mass m of the particle we find, now dropping the subscript 0 since obviously this equation holds for any choice of τ_0 , the **Lorentz-Abraham-Dirac** (LAD) equation of motion for a charged particle taking into account the interaction of the particle with its own radiation field including the **radiation-reaction force**, considering that some amount of energy and momentum is carried away from the particle to spatial infinity by the electromagnetic field,

$$\dot{p}_\mu = m \ddot{y}_\mu = K_\mu^{(\text{ext})} + \frac{q^2}{6\pi c^3} \left[\ddot{y}_\mu - u_\mu (u \cdot \ddot{y}) \right]. \quad (4.9.20)$$

Since

$$\dot{y}^2 = c^2 u^2 = c^2 = \text{const} \Rightarrow \dot{y} \cdot \dot{y} = 0 \Rightarrow \ddot{y} \cdot \dot{y} + \dot{y} \cdot \ddot{y} = c \ddot{y} \cdot u + \dot{y} \cdot \ddot{y} = 0 \quad (4.9.21)$$

we can rewrite the last term in (4.9.21),

$$\dot{p}_\mu = m \ddot{y}_\mu = K_\mu^{(\text{ext})} + \frac{q^2}{6\pi c^3} \left[\ddot{y}_\mu + \frac{1}{c} u_\mu (\dot{y} \cdot \ddot{y}) \right]. \quad (4.9.22)$$

The radiation-reaction term contains a term $\propto \ddot{y}$, the so-called Schott term. As we shall see in a moment, this gives rise to some characteristic problems with the LAD equation, since in contradistinction to the usual point-particle equations of motion in Newtonian as well as special-relativistic mechanics we have a third-order differential equation, and thus we need to provide initial conditions not only for y and \dot{y} but also for \ddot{y} .

Comparing the self-interaction contribution to the force with the momentum per unit time, which is carried away by the electromagnetic field (4.9.8) to spatial infinity, we see that the Schott term describes a piece of energy and momentum that is “reabsorbed” again by the particle at later times. The total momentum balance is nevertheless correctly taken into account under the assumptions made above after Eq. (4.9.9), i.e., the last term in the LAD equation as written in (4.9.22) balances precisely the total energy and momentum carried away by the electromagnetic field, i.e., it describes the expected **friction force** due to **radiation damping**. To qualitatively understand this, one should remember that $\dot{y} \cdot \ddot{y} < 0$, because $u \cdot \ddot{y} = 0$, i.e., \ddot{y} is Minkowski-orthogonal to the time-like dimensionless four-velocity u and thus necessarily spacelike.

4.9.3 The free-particle solution

The LAD equation (4.9.20) is very unusual compared to the standard equations of motion in point-particle mechanics, which usually are of 2nd order in the time derivatives and thus usually have unique solutions when considering the initial-value problem, where $y(\tau = \tau_0) = y_0$ and $\dot{y}(\tau = \tau_0) = c u_0$ are given. Since the LAD equation is of 3rd order, when considering the description of the motion as usual as an initial-value problem, we need in addition an initial value for the proper acceleration, $\ddot{y}(\tau = \tau_0) = a_0$. This is a priori not a problem. On the other hand, as discussed above, in order to solve the radiation-reaction problem to have a proper balance of energy-momentum conservation, we rather have to impose additional “asymptotic conditions” $\ddot{y}(\tau) \rightarrow 0$ for $\tau \rightarrow \pm\infty$.

4. Classical Electromagnetism

We start our discussion of examples with the most simple case, where no external force is acting, and we expect then that the particle should simply move with constant velocity, and the only electromagnetic field thus should be the corresponding (boosted) Coulomb field. Of course, setting $K_\mu^{(\text{ext})} = 0$ in (4.9.22), indeed $\dot{y} = c u = \text{const}$ is a solution of the equation of motion. However, we wish to show that this is the unique solution, given the asymptotic conditions. To that end it is convenient to define the time scale $\tau_q = q^2 / (6m\pi c^3)$ and write the LAD equation for the free charged particle in the form

$$\ddot{y}_\mu = \tau_q \left(\ddot{y}_\mu + \frac{u_\mu}{c} \ddot{y}^2 \right) \Rightarrow \dot{u}_\mu = \tau_q (\ddot{u}_\mu + u_\mu \dot{u}^2). \quad (4.9.23)$$

As we shall see, it is useful to introduce the new independent variable λ by [Lec18]

$$\lambda = \exp(\tau / \tau_q). \quad (4.9.24)$$

Then we have

$$\dot{f} = \frac{d\lambda}{d\tau} \frac{df}{d\lambda} = \frac{\lambda}{\tau_q} f'(\lambda), \quad (4.9.25)$$

where the prime now denotes a derivative wrt. λ . To express (4.9.24) in terms of derivatives wrt. λ we need

$$\dot{y} = \frac{\lambda}{\tau_q} y', \quad \ddot{y} = \frac{\lambda}{\tau_q^2} y' + \frac{\lambda^2}{\tau_q^2} y'' \quad (4.9.26)$$

Plugging this into (4.9.23) the equations of motion read

$$u''_\mu + u_\mu u'^2 = 0. \quad (4.9.27)$$

Since further $u \cdot u = 1$, we we have $u' \cdot u = 0$. Thus by multiplying this equation with u' we get

$$u \cdot u'' = 0 \Rightarrow u'^2 = -K^2 \quad (4.9.28)$$

Here we can assume $K \geq 0$ since u' is necessarily time-like. With this (4.9.27) reads

$$u'' = +K^2 u \Rightarrow u = C_1 \exp(K\lambda) + C_2 \exp(-K\lambda). \quad (4.9.29)$$

From $u \cdot u = 1$ we find

$$C_1^2 = C_2^2 = 0, \quad C_1 \cdot C_2 = \frac{1}{2}. \quad (4.9.30)$$

Now we investigate the asymptotic conditions. Since $\lambda \rightarrow 0$ for $\tau \rightarrow -\infty$ and $\lambda \rightarrow \infty$ for $\tau \rightarrow \infty$ we find that indeed these conditions $\ddot{u} \rightarrow 0$ for $\tau \rightarrow \pm\infty$ can only be fulfilled for $K = 0$ and thus $u = C_1 + C_2 = \text{const}$.

We note that for $K \neq 0$, the solutions (4.9.29) are indeed unphysical since (a) the particle gets spontaneously accelerated although there is no other particle with which it interacts (“self-acceleration”) and (b) the energy of the particle diverges for $\tau \rightarrow \infty$. Indeed for $\tau \rightarrow \infty$ we have $u \simeq C_1 \exp(K\lambda)$ and thus

$$E = mc^2 u^0 = mc^2 C_1^0 \exp(K\lambda) = mc^2 |\vec{C}_1| \exp(K\lambda). \quad (4.9.31)$$

These run-away solutions are excluded by the asymptotic conditions, as expected.

Appendix A

Tensor calculus in Minkowski space

In this Appendix we briefly summarize the basic formalism of tensor calculus in Minkowski space.

A.1 Tensor algebra

We start with the general definition of a finite-dimensional real vector space. It consists of a set V of **vectors**, which we denote with bold-faced symbols \mathbf{a} , \mathbf{b} , \dots together with an algebraic operation $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} + \mathbf{b}$, which forms an Abelian group, i.e.,

$$\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in V : (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (\text{A.1.1})$$

$$\forall \mathbf{a}, \mathbf{b} \in V : \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (\text{A.1.2})$$

$$\exists \mathbf{0} \in V \forall \mathbf{a} \in V : \mathbf{a} + \mathbf{0} = \mathbf{a}, \quad (\text{A.1.3})$$

$$\forall \mathbf{a} \in V \exists (-\mathbf{a}) \in V : \mathbf{a} + (-\mathbf{a}) = \mathbf{0}. \quad (\text{A.1.4})$$

Further for each $\mathbf{a} \in V$ and $\lambda \in \mathbb{R}$ there is an algebraic operation $(\mathbf{a}, \lambda) \mapsto \lambda \mathbf{a}$ that fulfills

$$\forall \lambda_1, \lambda_2 \in \mathbb{R}, \mathbf{a} \in V : \lambda_1(\lambda_2 \mathbf{a}) = (\lambda_1 \lambda_2) \mathbf{a}, \quad (\text{A.1.5})$$

$$\forall \mathbf{a} \in V : 0 \mathbf{a} = \mathbf{0}, \quad (\text{A.1.6})$$

$$\forall \lambda \in \mathbb{R}, \mathbf{a}, \mathbf{b} \in V : \lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}, \quad (\text{A.1.7})$$

$$\forall \lambda_1, \lambda_2 \in \mathbb{R}, \mathbf{a} \in V : (\lambda_1 + \lambda_2) \mathbf{a} = \lambda_1 \mathbf{a} + \lambda_2 \mathbf{a}. \quad (\text{A.1.8})$$

Any finite set of vectors \mathbf{e}_i is called **linearly independent** if

$$\sum_j \lambda_j \mathbf{e}_j = \mathbf{0} \Rightarrow \lambda_j = 0. \quad (\text{A.1.9})$$

A.2 The fundamental form

A.3 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.

A. Tensor calculus in Minkowski space

We start with the invariant generalized δ tensors, whose components $\delta_{\nu_1 \dots \nu_k}^{\mu_1 \dots \mu_k}$ with respect to an arbitrary Cartesian basis are defined as the determinant

$$\delta_{\nu_1 \dots \nu_k}^{\mu_1 \dots \mu_k} = \det \begin{pmatrix} \delta_{\nu_1}^{\mu_1} & \delta_{\nu_2}^{\mu_1} & \dots & \delta_{\nu_k}^{\mu_1} \\ \delta_{\nu_1}^{\mu_2} & \delta_{\nu_2}^{\mu_2} & \dots & \delta_{\nu_k}^{\mu_2} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{\nu_1}^{\mu_k} & \delta_{\nu_2}^{\mu_k} & \dots & \delta_{\nu_k}^{\mu_k} \end{pmatrix}. \quad (\text{A.3.1})$$

This immediately implies that in N -dimensional space we have $k \in \{1, 2, \dots, N\}$ because for $k > N$ all the δ symbols must be 0, because there are only N different values, $\{1, \dots, N\}$ for the indices¹ since the δ symbol is obviously antisymmetric under exchange of any pair of lower or any pair of upper indices. Also obviously $\delta_{\nu_1 \dots \nu_k}^{\mu_1 \dots \mu_k} \neq 0$ only if $\{\mu_1, \dots, \mu_k\} = \{\nu_1, \dots, \nu_k\}$, 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 $U^\mu_\nu T^\rho_\mu = T^\rho_\nu U^\mu_\nu = \delta^\rho_\nu$, we find

$$\bar{\delta}_{\sigma_1 \dots \sigma_N}^{\rho_1 \dots \rho_N} = T^{\rho_1}_{\mu_1} \dots T^{\rho_N}_{\mu_N} U^{\nu_1}_{\sigma_1} \dots U^{\nu_N}_{\sigma_N} \delta_{\nu_1 \dots \nu_N}^{\mu_1 \dots \mu_N} = \det(\hat{T}) \det(\hat{U}) \delta_{\nu_1 \dots \nu_N}^{\mu_1 \dots \mu_N} = \delta_{\nu_1 \dots \nu_N}^{\mu_1 \dots \mu_N}. \quad (\text{A.3.2})$$

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 $\delta_{\nu_1 \dots \nu_{k-1} \mu_k}^{\mu_1 \dots \mu_{k-1} \mu_k}$ we use Laplace's formula to the last row,

$$\delta_{\nu_1 \dots \nu_k}^{\mu_1 \dots \mu_k} = \sum_{j=1}^k (-1)^{j+k} \delta_{\nu_j}^{\mu_k} \delta_{\nu_1 \dots \hat{\nu}_j \dots \nu_k}^{\mu_1 \dots \mu_{k-1}}, \quad (\text{A.3.3})$$

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

$$\delta_{\nu_1 \dots \nu_{k-1} \mu_k}^{\mu_1 \dots \mu_{k-1} \mu_k} = N \delta_{\nu_1 \dots \nu_{k-1}}^{\mu_1 \dots \mu_{k-1}} + \sum_{j=1}^{k-1} (-1)^{k+j} \delta_{\nu_1 \dots \hat{\nu}_j \dots \nu_j}^{\mu_1 \dots \mu_{k-1}} \quad (\text{A.3.4})$$

So in the sum we sum $k-1$ -times over the same $\delta_{\nu_1 \dots \nu_{k-1}}^{\mu_1 \dots \mu_{k-1}}$ symbol, only the index ν_j is put to the $(k-1)$ th position. We just have to bring it back to the original position, which needs $k-1-j$ exchanges, i.e., we finally get

$$\delta_{\nu_1 \dots \mu_k}^{\mu_1 \dots \mu_k} = (N - k + 1) \delta_{\nu_1 \dots \nu_{k-1}}^{\mu_1 \dots \mu_{k-1}}. \quad (\text{A.3.5})$$

Iterating this equation leads to the general contraction

$$\delta_{\nu_1 \dots \nu_j \mu_{j+1} \dots \mu_k}^{\mu_1 \dots \mu_j \mu_{j+1} \dots \mu_k} = (N - k + 1)(N - k + 2) \dots (N - j) \delta_{\nu_1 \dots \nu_j}^{\mu_1 \dots \mu_j} = \frac{(N - j)!}{(N - k)!} \delta_{\nu_1 \dots \nu_j}^{\mu_1 \dots \mu_j}. \quad (\text{A.3.6})$$

Since now for $k = N$ the generalized δ symbol is an invariant tensor according to (A.3.2), so are all $(2k)$ -rank δ symbols since we can get them as contractions over the $(N - k)$ other indices.

¹In this section we count the tensor components with indices running over $\{1, \dots, N\}$ for simplicity of the argument.

A.4 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

$$\epsilon^{\mu\nu\rho\sigma} = \delta_{0123}^{\mu\nu\rho\sigma}, \quad (\text{A.4.1})$$

i.e., it is +1 (−1) if the ordered quadrupel (μ, ν, ρ, σ) 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,

$$\epsilon_{\alpha\beta\gamma\delta} = \eta_{\alpha\mu}\eta_{\beta\nu}\eta_{\gamma\rho}\eta_{\delta\sigma}\epsilon^{\mu\nu\rho\sigma} = \epsilon^{\mu\nu\rho\sigma} \det \hat{\eta} = -\epsilon^{\alpha\beta\gamma\delta}. \quad (\text{A.4.2})$$

Since we have

$$\epsilon^{\mu\nu\rho\sigma}\epsilon_{\alpha\beta\gamma\delta} = -\delta_{\alpha\beta\gamma\delta}^{\mu\nu\rho\sigma}, \quad (\text{A.4.3})$$

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.3.6) we get the important formulas

$$\epsilon^{\mu\nu\rho\delta}\epsilon_{\alpha\beta\gamma\delta} = -\delta_{\alpha\beta\gamma}^{\mu\nu\rho}, \quad (\text{A.4.4})$$

$$\epsilon^{\mu\nu\gamma\delta}\epsilon_{\alpha\beta\gamma\delta} = -2\delta_{\alpha\beta}^{\mu\nu}, \quad (\text{A.4.5})$$

$$\epsilon^{\mu\beta\gamma\delta}\epsilon_{\alpha\beta\gamma\delta} = -6\delta_{\alpha}^{\mu}. \quad (\text{A.4.6})$$

Now a k "-rank tensor is called **totally antisymmetric** if its components $T^{\mu_1 \dots \mu_k}$ 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

$$\dagger T_{\nu_{k+1} \dots \nu_4} = \frac{1}{k!} \epsilon_{\nu_{k+1} \dots \nu_4 \mu_1 \dots \mu_k} T^{\mu_1 \dots \mu_k}, \quad (\text{A.4.7})$$

$$\dagger T^{\nu_{k+1} \dots \nu_4} = \frac{1}{k!} \epsilon^{\mu_1 \dots \mu_k \nu_{k+1} \dots \nu_4} T_{\mu_1 \dots \mu_k}.$$

E.g., a totally antisymmetric Tensor T of rank 3 is mapped to a vector $\dagger T$ by

$$(\dagger T)_{\mu} = \frac{1}{3!} \epsilon_{\mu\nu\rho\sigma} T^{\nu\rho\sigma}. \quad (\text{A.4.8})$$

Taking the Hodge-dual twice leads back to the original tensor up to a sign, because of (A.4.4-A.4.6). E.g., for our 3rd-rank-tensor example we have

$$(\dagger\dagger T)^{\alpha\beta\gamma} = \epsilon^{\mu\alpha\beta\gamma} (\dagger T)_{\mu} = \frac{1}{3!} \epsilon^{\mu\alpha\beta\gamma} \epsilon_{\mu\nu\rho\sigma} T^{\nu\rho\sigma} = -\frac{1}{3!} \delta_{\nu\rho\sigma}^{\alpha\beta\gamma} T^{\nu\rho\sigma} = -T^{\alpha\beta\gamma}. \quad (\text{A.4.9})$$

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.

²Note 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

A.5 Differential tensor calculus

Now we consider **tensor fields**. These provide tensor-valued observables $T(x)$ as a function of the spacetime point given by the spacetime fourvector x . The transformation properties under proper orthochronous Lorentz transformations for its components is obviously

$$\bar{T}^{\mu\nu\dots}(\bar{x}^\alpha) = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \dots T^{\rho\sigma\dots}(x^\beta), \quad \bar{x}^\alpha = \Lambda^\alpha{}_\beta x^\beta. \quad (\text{A.5.1})$$

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:

$$\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}. \quad (\text{A.5.2})$$

Contracting with $\eta^{\sigma\alpha}$ gives

$$\eta_{\mu\nu} \eta^{\sigma\alpha} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \delta_\rho^\alpha, \quad (\text{A.5.3})$$

i.e.,

$$(\Lambda^{-1})^\alpha{}_\mu = \eta_{\mu\nu} \eta^{\sigma\alpha} \Lambda^\nu{}_\sigma = \Lambda_\mu{}^\alpha. \quad (\text{A.5.4})$$

The covariant component of a fourvector thus transforms as

$$\bar{x}_\mu = \eta_{\mu\alpha} \bar{x}^\alpha = \eta_{\mu\alpha} \Lambda^\alpha{}_\beta x^\beta = \eta_{\mu\alpha} \eta^{\beta\gamma} \Lambda^\alpha{}_\beta x_\gamma = \Lambda_\mu{}^\gamma x_\gamma = (\Lambda^{-1})^\gamma{}_\mu x_\gamma. \quad (\text{A.5.5})$$

Now we have

$$\begin{aligned} \bar{\partial}_\alpha \bar{T}^{\mu\nu\dots} &= \frac{\partial}{\partial \bar{x}^\alpha} \bar{T}^{\mu\nu\dots} = \frac{\partial x^\beta}{\partial \bar{x}^\alpha} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \dots \partial_\beta T^{\rho\sigma\dots} \\ &= (\Lambda^{-1})^\beta{}_\alpha \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \dots \partial_\beta T^{\rho\sigma\dots} \\ &= \Lambda_\alpha{}^\beta \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \dots \partial_\beta T^{\rho\sigma\dots}, \end{aligned} \quad (\text{A.5.6})$$

and this is precisely the transformation law for tensor components with a covariant indices α and β , as claimed.

A.6 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.3. For Minkowski space we have $N = 4$ and thus three types of integrals, because we can define 1...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

A.6. Integrals in Minkowski space

parameters q^k ($k \in \{1, \dots, d\}$) with domain in a cuboid $Q_d(a_1, b_1) \otimes (a_2, b_2) \otimes \dots \otimes (a_d, b_d)$ by a function $x^\mu(q^k)$ with x^μ the spacetime four-vector. Then we can define integration measures by

$$\begin{aligned} d^d \sigma^{\mu_1 \mu_2 \dots \mu_d} &= dq^1 dq^2 \dots dq^d \delta_{\nu_1 \nu_2 \dots \nu_d}^{\mu_1 \mu_2 \dots \mu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \frac{\partial x^{\nu_2}}{\partial q^2} \dots \frac{\partial x^{\nu_d}}{\partial q^d} \\ &= dq^1 dq^2 \dots dq^d \epsilon^{\mu_1 \dots \mu_d} \det \left(\frac{\partial x^\nu}{\partial q^k} \right). \end{aligned} \quad (\text{A.6.1})$$

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 $d^2 \sigma^{ab}$, which usually is mapped by a Hodge dualization to the surface-element normal vector $d^2 \vec{f}$, but first we do not consider the Hodge dualization but work with the integration measures of the type (A.6.1) only.

The second reason to use (A.6.1) is the independence of this description of the choice of parameters. Using other parameters \tilde{q}^k the (A.6.2) gets the same form again, because then we have

$$\begin{aligned} d^d \sigma^{\mu_1 \mu_2 \dots \mu_d} &= d\tilde{q}^1 d\tilde{q}^2 \dots d\tilde{q}^d \epsilon^{\mu_1 \mu_2 \dots \mu_d} \det \left(\frac{\partial x^\nu}{\partial q^k} \right) \det \left(\frac{\partial q^k}{\partial \tilde{q}^l} \right) \\ &= d\tilde{q}^1 d\tilde{q}^2 \dots d\tilde{q}^d \epsilon^{\mu_1 \mu_2 \dots \mu_d} \det \left(\frac{\partial x^\nu}{\partial q^k} \frac{\partial q^k}{\partial \tilde{q}^l} \right) \\ &= d\tilde{q}^1 d\tilde{q}^2 \dots d\tilde{q}^d \epsilon^{\mu_1 \mu_2 \dots \mu_d} \det \left(\frac{\partial x^\nu}{\partial \tilde{q}^l} \right) \\ &= d\tilde{q}^1 d\tilde{q}^2 \dots d\tilde{q}^d \delta_{\nu_1 \nu_2 \dots \nu_d}^{\mu_1 \mu_2 \dots \mu_d} \frac{\partial x^{\nu_1}}{\partial \tilde{q}^1} \frac{\partial x^{\nu_2}}{\partial \tilde{q}^2} \dots \frac{\partial x^{\nu_d}}{\partial \tilde{q}^d} \end{aligned} \quad (\text{A.6.2})$$

Since under proper orthochronous Lorentz transformations $d^d \sigma^{\mu_1 \mu_2 \dots \mu_d}$ 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

$$\int_M d^d \sigma^{\mu_1 \dots \mu_d} T_{\mu_1 \dots \mu_d}(\underline{x}) = \int_Q dq^1 \dots dq^d \delta_{\nu_1 \nu_2 \dots \nu_d}^{\mu_1 \mu_2 \dots \mu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \dots \frac{\partial x^{\nu_d}}{\partial q^d} T_{\mu_1 \dots \mu_d}[\underline{x}(q)]. \quad (\text{A.6.3})$$

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 $q^j = b_j = \text{const}$

$$d^{d-1} \sigma^{\mu_1 \dots \hat{\mu}_j \dots \mu_d} = \delta_{\nu_1 \dots \hat{\nu}_j \dots \nu_d}^{\mu_1 \dots \hat{\mu}_j \dots \mu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \dots \frac{\partial x^{\nu_{j-1}}}{\partial q^{j-1}} \frac{\partial x^{\nu_{j+1}}}{\partial q^{j+1}} \dots \frac{\partial x^{\nu_d}}{\partial q^d} \Bigg|_{q^j=b_j} \quad (\text{A.6.4})$$

and for the parts with $q^j = a_j = \text{const}$

$$d^{d-1} \sigma^{\mu_1 \dots \hat{\mu}_j \dots \mu_d} = - \delta_{\nu_1 \dots \hat{\nu}_j \dots \nu_d}^{\mu_1 \dots \hat{\mu}_j \dots \mu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \dots \frac{\partial x^{\nu_{j-1}}}{\partial q^{j-1}} \frac{\partial x^{\nu_{j+1}}}{\partial q^{j+1}} \dots \frac{\partial x^{\nu_d}}{\partial q^d} \Bigg|_{q^j=a_j}. \quad (\text{A.6.5})$$

The hat over the indices indicates the ones to be left out in the expression, including an appropriate sign convention,

$$\delta_{\nu_1 \dots \hat{\nu}_j \dots \nu_d}^{\mu_1 \dots \hat{\mu}_j \dots \mu_d} = (-1)^{j+1} \delta_{\nu_1 \dots \nu_{j-1} \nu_{j+1} \dots \nu_d}^{\mu_1 \dots \mu_{j-1} \mu_{j+1} \dots \mu_d}. \quad (\text{A.6.6})$$

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:

$$\int_M d^d \sigma^{\mu_1 \dots \mu_d} \partial_{\mu_1} T_{\mu_2 \dots \mu_d} = \int_{\partial M} d^{d-1} \sigma^{\mu_2 \dots \mu_d} T_{\mu_2 \dots \mu_d}. \quad (\text{A.6.7})$$

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.3.1) of the δ -tensor components and expand the determinant wrt. the 1st row,

$$\begin{aligned} \int_M d^d \sigma^{\mu_1 \dots \mu_d} \partial_{\mu_1} T_{\mu_2 \dots \mu_d} &= \sum_{k=1}^d \int_Q dq^1 \dots dq^d (-1)^{k+1} \delta_{\nu_k}^{\mu_1} \delta_{\nu_2 \dots \nu_{k-1} \nu_{k+1} \nu_d}^{\mu_2 \dots \mu_d} T_{\mu_2 \dots \mu_d} \frac{\partial x^{\nu_1}}{\partial q^1} \dots \frac{\partial x^{\nu_d}}{\partial q^d} \\ &= \sum_{k=1}^d \int_Q dq^1 \dots dq^d (-1)^{k+1} \delta_{\nu_2 \dots \nu_{k-1} \nu_{k+1} \nu_d}^{\mu_2 \dots \mu_d} \frac{\partial}{\partial q^k} T_{\mu_2 \dots \mu_d} \\ &\quad \frac{\partial x^{\nu_1}}{\partial q^1} \dots \frac{\partial x^{\nu_{k-1}}}{\partial q^{k-1}} \frac{\partial x^{\nu_{k+1}}}{\partial q^{k+1}} \dots \frac{\partial x^{\nu_d}}{\partial q^d} \\ &= \sum_{k=1}^d \int_{Q_k} dq^1 \dots dq^{k-1} dq^{k+1} \dots dq^d (-1)^{k+1} \delta_{\nu_2 \dots \nu_{k-1} \nu_{k+1} \nu_d}^{\mu_2 \dots \mu_d} [T_{\mu_2 \dots \mu_d}]_{q_k=a_k}^{q_k=b_k} \\ &\quad \frac{\partial x^{\nu_1}}{\partial q^1} \dots \frac{\partial x^{\nu_{k-1}}}{\partial q^{k-1}} \frac{\partial x^{\nu_{k+1}}}{\partial q^{k+1}} \dots \frac{\partial x^{\nu_d}}{\partial q^d} \\ &= \int_{\partial M} d^{d-1} \sigma^{\mu_2 \dots \mu_d} T_{\mu_2 \dots \mu_d}. \end{aligned} \quad (\text{A.6.8})$$

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 $d^d \sigma^{\mu_1 \dots \mu_d}$, 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

$$d^4 \sigma^* = -\frac{1}{4!} \epsilon_{\mu_1 \dots \mu_4} d^4 \sigma^{\mu_1 \dots \mu_4} = d^4 x, \quad (\text{A.6.9})$$

$$d^3 \sigma_{\mu_1}^* = -\frac{1}{3!} \epsilon_{\mu_1 \mu_2 \mu_3 \mu_4} d^3 \sigma^{\mu_2 \dots \mu_4}, \quad (\text{A.6.10})$$

$$d^2 \sigma_{\mu_1 \mu_2}^* = -\frac{1}{2!} \epsilon_{\mu_1 \mu_2 \mu_3 \mu_4} d^2 \sigma^{\mu_3 \mu_4}, \quad (\text{A.6.11})$$

$$d^3 \sigma_{\mu_1 \mu_2 \mu_3}^* = -\epsilon_{\mu_1 \mu_2 \mu_3 \mu_4} d\sigma^{\mu_4}. \quad (\text{A.6.12})$$

This leads to scalar integrals of the type

$$\int_M d^k \sigma_{\mu_1 \dots \mu_k} V^{\mu_1 \dots \mu_k} \quad (\text{A.6.13})$$

over totally antisymmetric tensor fields of rank k .

To state **Gauss's integral laws**, it is more lucid to discuss the three different cases separately. We start with

$$\int_M d^4 \sigma^* \partial_{\mu_1} V^{\mu_1} = \int_{\partial M} d^3 \sigma_{\mu_1}^* V^{\mu_1}. \quad (\text{A.6.14})$$

A.6. Integrals in Minkowski space

To prove it, we simply have to write out the integral on the left-hand side of the equation according to the definition (A.6.9) and use the properties of the Hodge dualization operation discussed in Sect. A.4,

$$\begin{aligned}
\int_M d^4 \sigma^* \partial_{\mu_1} V^{\mu_1} &= -\frac{1}{4!} \int_M d^4 \sigma^{\nu_1 \mu_2 \mu_3 \mu_4} \epsilon_{\nu_1 \mu_2 \mu_3 \mu_4} \partial_{\mu_1} V^{\mu_1} \\
&= \frac{1}{4!} \frac{1}{3!} \int_M d^4 \sigma^{\nu_1 \mu_2 \mu_3 \mu_4} \epsilon_{\nu_1 \mu_2 \mu_3 \mu_4} \epsilon^{\nu_2 \nu_3 \nu_4 \mu_1} \partial_{\mu_1} \dagger V_{\nu_2 \nu_3 \nu_4} \\
&= +\frac{1}{4!} \frac{1}{3!} \int_M d^4 \sigma^{\nu_1 \mu_2 \mu_3 \mu_4} \delta_{\nu_1 \mu_2 \mu_3 \mu_4}^{\mu_1 \nu_2 \nu_3 \nu_4} \partial_{\mu_1} \dagger V_{\nu_2 \nu_3 \nu_4} \\
&= +\frac{1}{3!} \int_M d^4 \sigma^{\mu_1 \mu_2 \mu_3 \mu_4} \partial_{\mu_1} \dagger V_{\mu_2 \mu_3 \mu_4}.
\end{aligned} \tag{A.6.15}$$

With Stokes's theorem (A.6.8) we get

$$\begin{aligned}
\int_M d^4 \sigma^* \partial_{\mu_1} V^{\mu_1} &= \frac{1}{3!} \int_{\partial M} d^3 \sigma^{\mu_2 \mu_3 \mu_4} \dagger V_{\mu_2 \mu_3 \mu_4} \\
&= -\frac{1}{3!} \int_{\partial M} d^3 \sigma^{\mu_2 \mu_3 \mu_4} \epsilon_{\mu_1 \mu_2 \mu_3 \mu_4} V^{\mu_1} \\
&= \int_{\partial M} d^3 \sigma_{\mu_1}^* V^{\mu_1},
\end{aligned} \tag{A.6.16}$$

which is (A.6.14).

In a similar way we get for an antisymmetric 2nd-rank tensor field

$$\int_M d^3 \sigma_{\mu_1 \mu_2}^* \partial_{\mu_2} V^{\mu_1 \mu_2} = \frac{1}{2!} \int_{\partial M} d^2 \sigma_{\mu_1 \mu_2}^* V^{\mu_1 \mu_2}, \tag{A.6.17}$$

for a totally antisymmetric 3rd-rank tensor field

$$\frac{1}{2!} \int_M d^2 \sigma_{\mu_1 \mu_2}^* \partial_{\mu_3} V^{\mu_1 \mu_2 \mu_3} = \frac{1}{3!} \int_{\partial M} d\sigma_{\mu_1 \mu_2 \mu_3}^* V^{\mu_1 \mu_2 \mu_3}. \tag{A.6.18}$$

A. Tensor calculus in Minkowski space

Bibliography

- [Bel40] F. J. Belinfante, On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields, *Physica* **7**, 449 (1940), [https://doi.org/10.1016/S0031-8914\(40\)90091-X](https://doi.org/10.1016/S0031-8914(40)90091-X).
- [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>.
- [Ein05] A. Einstein, On the electrodynamics of moving bodies, *Annalen Phys.* **17**, 891 (1905), [Annalen Phys.14,194(2005)], <https://doi.org/10.1002/andp.200590006>.
- [Erb61] T. Erber, The classical theories of radiation reaction, *Fortschritte der Physik* **9** (1961), <https://doi.org/10.1002/prop.19610090702>.
- [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).
- [Lec18] K. Lechner, *Classical Electrodynamics*, Springer International Publishing AG, Cham (2018), <https://doi.org/10.1007/978-3-319-91809-9>.
- [LL96] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics 2, The Classical Theory of Fields*, Butterworth Heinemann, Oxford, 4 edn. (1996).
- [LM07] K. Lechner and P. A. Marchetti, Variational principle and energy-momentum tensor for relativistic electrodynamics of point charges, *Ann. Phys. (NY)* **322**, 1162 (2007), <https://dx.doi.org/10.1016/j.aop.2006.07.002>.
- [Min09] H. Minkowski, Raum und Zeit, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 18, p. 75-88 **18**, 75 (1909).
- [MTW73] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation*, W. H. Freeman, San Francisco (1973).
- [Nak13] C. Nakhleh, The Lorentz-Dirac and Landau-Lifshitz equations from the perspective of modern renormalization theory, *Am. J. Phys* **81**, 180 (2013), <https://dx.doi.org/10.1119/1.4773292>.

Bibliography

- [Noe10] F. Noether, Zur Kinematik des starren Körpers in der Relativtheorie, *Annalen der Physik* **336**, 919 (1910).
- [Oll08] J.-Y. Ollitrault, Relativistic hydrodynamics for heavy-ion collisions, *Eur. J. Phys.* **29**, 275 (2008), <https://doi.org/10.1088/0143-0807/29/2/010>.
- [Ros40] L. J. H. C. Rosenfeld, Sur le tenseur d'impulsion-énergie, *Roy. Belg. Memoirs de Classes de Science* **18**, 063504 (1940).
- [RZ13] L. Rezzolla and O. Zanotti, *Relativistic hydrodynamics*, Oxford University Press, Oxford (2013), <https://dx.doi.org/10.1093/acprof:oso/9780198528906.001.0001>.
- [Sch60] A. Schild, On the radiation emitted by an accelerated point charge, *J. Math. Anal. and Appls.* **1**, 127 (1960), [https://doi.org/10.1016/0022-247X\(60\)90033-0](https://doi.org/10.1016/0022-247X(60)90033-0).
- [Sil14] L. Silberstein, *The Theory of Relativity*, MacMillan&Co., London (1914), <https://archive.org/details/theoryofrelativi00silbrich>.
- [Sop08] D. E. Soper, *Classical field theory*, Dover Publications, Minneola, New York (2008).
- [Tho26] L. H. Thomas, The motion of the spinning electron, *Nature* **117** (1926), <https://doi.org/10.1038/2F117514a0>.