Translated from the book
Elementare Feldtheorie (Elementary Field Theory)
by H. Römer and M. Forger
VCH Verlag, Weinheim 1993

7 Special Relativity

7.1 The Principle of Relativity

The origin of the theory of special relativity is associated with the difficulties to bring the behavior of electromagnetic waves, as predicted by Maxwell's theory and confirmed by experiment, into agreement with the conception of space and time that underlies Newtonian mechanics. Einstein's analysis of this situation has led to a fundamental revision of the physical concept of space and time and, as a consequence, to the formulation of a new relativistic mechanics by which the problem of consistency between mechanics and electrodynamics is solved in an elegant way. The point of departure of this analysis, and hence of special relativity as a whole, is the substitution of the Newtonian (or Galilean) principle of relativity, implicitly contained in Newtonian mechanics, by the Einsteinian (or Lorentzian) principle of relativity suggested by the laws of electrodynamics, in particular by the negative outcome of all experiments to measure motion relative to the (hypothetical) ether (Michelson-Morley experiment).

Apart from the differences, the Newtonian and the Einsteinian principle of relativity also share a number of common features. We shall begin by discussing these.

An observer in a given state of motion defines a *reference system* R. For the quantitative description of phenomena in nature, such an observer must introduce *coordinates*, i.e., a system of measuring rods at rest with respect to R and of appropriately synchronized clocks, so that the space-time position of a point particle may be determined by specifying four coordinates. In principle, these coordinates are arbitrary, as long as they allow to identify space-time positions: their choice is a matter of convenience.

An *inertial system* is a reference system in which – in terms of appropriate coordinates which will be called *natural coordinates* – every point particle that is not subject to external forces moves uniformly in a straight line. The existence of inertial systems is by no means evident. Indeed, by an appropriate choice of coordinates, we may always linearize a specific trajectory, but the possibility to simultaneously linearize the trajectories of *all* freely moving point particles, within a single coordinate system, constitutes a fundamental fact of experience. It turns out that, to an excellent approximation, a reference system R at rest with respect to the stars becomes an inertial system if we introduce spatial coordinates by choosing an oriented orthonormal frame at rest in R and a time coordinate by choosing appropriately synchronized "standard" clocks; we shall always imagine natural coordinates in an inertial system to have been introduced in this way. (The question of how to perform the synchronization will be discussed later.)

The statements on the structure of space and time that are common to the Newtonian and the Einsteinian principle of relativity can now be summarized as follows.

(R1) Space and time are homogeneous and isotropic.

(R2) All reference systems that move relative to some given inertial system uniformly in a straight line are also inertial systems and are physically equivalent among themselves: there is no physically preferred inertial system.

There is a third requirement (R3) in which the two principles differ from each other and which will be discussed below. Before that, however, we proceed to explain the meaning of the requirements (R1) and (R2).

Mathematically, the requirement (R1) is formulated by stating that physical space has the structure of a Euclidean three-dimensional affine space E^3 and physical time has the structure of a one-dimensional affine space; they can be united into physical space-time which therefore has the structure of a spatially Euclidean four-dimensional affine space E^4. The introduction of natural coordinates for an inertial system mentioned before then corresponds to the choice of a spatially Euclidean affine coordinate system, consisting of an origin in space, an oriented orthonormal 3-frame, an origin in time and a time scale.

The requirement (R2) of physical equivalence between inertial systems means that *all* physical phenomena must in *all* inertial systems obey the same laws of nature. Implementing this principle requires, however, that every observer uses the same methods to define the natural coordinates in his inertial system, that is, using the same kind of clocks, of measuring rods, of protractors, etc., and without reference to any other inertial system than his own. One says that the procedure for defining natural coordinates in inertial systems must be *universal* and *intrinsic*.

Let us now consider a single but arbitrary *event*, which is located at a *point* in space-time. In an inertial system I it will be described by natural coordinates $(t, \boldsymbol{x}) \in \mathbb{R}^4$ and in an inertial system I' by coordinates $(t', \boldsymbol{x}') \in \mathbb{R}^4$. Of course, these coordinates must be convertible into each other in a unique way, and since the event was arbitrary, this means that there must exist an invertible transformation $T_{I'I} : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ such that for the natural coordinates of any event, we have

$$T_{I'I}(t, \boldsymbol{x}) \;=\; (t', \boldsymbol{x}') \tag{7.1}$$

Obviously, composition and inversion of these transformations must obey the rules

$$T_{I''I'} \circ T_{I'I} \;=\; T_{I''I} \qquad \text{and} \qquad T_{I'I}^{-1} \;=\; T_{II'} \;. \tag{7.2}$$

Due to (R2), the transformation $T_{I'I}$ can depend only on quantities that relate I and I', not on quantities describing, e.g., the state of motion of I or I' alone (relative to a third inertial system I''). More precisely, (R2) implies that given any three inertial systems I, I' and J, there is exactly one inertial system J' with

$$T_{I'I} \;=\; T_{J'J} \;.$$

As a result, the set of all transformations $T_{I'I}$ between inertial systems forms a *group* which, for the time being, we shall denote by Γ. Indeed, given $T_{I'I} \in \Gamma$ and $T_{J'J} \in \Gamma$, we can write $T_{J'J} = T_{I''I'}$ and obtain

$$T_{J'J} \circ T_{I'I} \;=\; T_{I''I'} \circ T_{I'I} \;=\; T_{I''I} \;\in\; \Gamma \;.$$

From (R1) we obtain further restrictions on the coordinate transformations in Γ: First of all, elements of Γ must be *affine* transformations, that is, inhomogeneous linear transformations, because they transform uniform rectilinear motion into uniform rectilinear motion, without singling out particular points. The most general such transformation has the form

$$\boldsymbol{x}' \;=\; A\boldsymbol{x} \,+\, \boldsymbol{v}t \,+\, \boldsymbol{x}_0 \;, \tag{7.3}$$

$$t' \;=\; \gamma t + \boldsymbol{w}\cdot\boldsymbol{x} \,+\, t_0 \;, \tag{7.4}$$

where A is a linear transformation in $\mathbb{R}^3$, $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{x}_0$ are vectors in $\mathbb{R}^3$ and γ, t_0 are scalars in $\mathbb{R}$; the vector $-A^{-1}\boldsymbol{v}$ is to be interpreted as the relative velocity between the two inertial systems I and I'. Moreover, Γ must contain the group of rotations in three-dimensional space as a subgroup, that is, all transformations of the form given above with $A \in SO(3)$ and $\boldsymbol{v} = 0$, $\boldsymbol{w} = 0$, $\boldsymbol{x}_0 = 0$, $\gamma = 1$, $t_0 = 0$.

Newtonian mechanics, as mentioned before, implicitly contains the *Newtonian principle of relativity* which supplements the requirements (R1) and (R2) by the *postulate of absolute space and absolute time*:

(RN) Space and time are absolute, i.e., spatial distances and time differences do not depend on the reference system.

This requirement leads directly to the *Galilei invariance* of Newtonian mechanics because it forces $A \in SO(3)$ in eq. (7.3) as well as $\gamma = 1$ and $\boldsymbol{w} = 0$ in eq. (7.4). Thus the most general *Galilei transformation* is of the form

$$\boldsymbol{x}' \;=\; R\boldsymbol{x} \,+\, \boldsymbol{v}t \,+\, \boldsymbol{x}_0 \;, \tag{7.5}$$

$$t' \;=\; t \,+\, t_0 \;, \tag{7.6}$$

with $R \in SO(3)$, $\boldsymbol{v}, \boldsymbol{x}_0 \in \mathbb{R}^3$ and $t_0 \in \mathbb{R}$.

For the sake of definiteness, consider the Newtonian equations of motion for a system of point particles under the influence of pair forces, each of which is directed along the connecting line between the two particles involved and has magnitude depending only on their distance: this covers a large number of physically important cases. In the absence of external forces, all forces can be derived from a common potential $V = V(|\boldsymbol{x}_1 - \boldsymbol{x}_2|, \ldots, |\boldsymbol{x}_i - \boldsymbol{x}_j|, \ldots)$, and the Newtonian equations of motion become

$$m_i \, \frac{d^2\boldsymbol{x}_i}{dt^2} \;=\; -\, \boldsymbol{\nabla}_i V \;. \tag{7.7}$$

It is immediately verified that these equations are invariant under Galilei transformations

$$\boldsymbol{x}_i \;\mapsto\; R\boldsymbol{x}_i + \boldsymbol{v}t + \boldsymbol{x}_0 \quad , \quad t \;\mapsto\; t + t_0 \; . \tag{7.8}$$

Of course, the equations (7.7) refer to natural coordinates in an inertial system, so their invariance under the Galilei transformations (7.8) is the formal expression of the relativity principle of Newtonian mechanics. The conception of absolute time and of instantaneous interaction is already expressed through the introduction of a potential that depends only on the relative positions of the particles.

Electrodynamics (in vacuum) does *not* obey the relativity principle of Newtonian mechanics, i.e., it is *not* Galilei invariant. Indeed, Maxwell's equations (in vacuum) explicitly contain a universal velocity c and imply that light (in vacuum) propagates isotropically with this velocity.

The *ether hypothesis* interpreted this fact by giving up the principle of relativity itself, postulating that Maxwell's equations be valid only in one distinguished inertial system, that of the ether. According to this idea, light really never propagates in a complete vacuum but rather in a medium, the ether – just like sound always propagates in a medium, such as air – and the distinguished inertial system is simply that in which the ether is at rest. Indeed, some phenomena such as the Doppler effect and the aberration of light are consistent with this hypothesis. However, all experiments aiming at some kind of direct evidence for the existence of the ether, basically by trying to measure the motion of sources of light relative to the ether, among them the famous Michelson-Morley experiment, have failed.

Rather than trying to explain this failure by additional hypotheses about properties of the ether, Einstein concluded that the principle of relativity itself had to be modified. Once this point of view is adopted, one almost obviously arrives at the *Einsteinian principle of relativity* which supplements the requirements (R1) and (R2) by the *postulate of the constancy and universality of the speed of light*:

(RE) The speed of light in vacuum c is finite and universal: Relative to an arbitrary inertial system, the propagation of light in vacuum does not depend on the state of motion of the source.

This requirement leads, as we shall see, to the *Lorentz invariance* of relativistic mechanics and of electrodynamics.

Obviously, the Newtonian and the Einsteinian principle of relativity are incompatible: they contradict each other. The careful analysis of this contradiction led Einstein to his fundamental criticism of the concept of absolute time and in particular of the concept or simultaneity. Indeed, in Newtonian mechanics, the synchronization of spatially far distant clocks that are at rest in a given inertial system does not cause any problems and will be the same for any observer, independently of his state of motion: it can be carried out, for example, by simply transporting the clocks or by rigid rods or by other instantaneously propagating signals. If on the other hand one adopts the Einsteinian principle of relativity, synchronization becomes problematic and one is forced to specify a synchronization procedure. There are different possibilities for doing this, but the most natural one is

Einstein's synchronization procedure:

> In a given inertial system, consider a clock at rest and located at the origin, $U(0)$, as well as another clock of the same kind at rest but located at another point $\boldsymbol{x}$, $U(\boldsymbol{x})$. When $U(0)$ shows the time t_1, a source at rest and located at the origin emits a light pulse that propagates to the point $\boldsymbol{x}$, is reflected there and returns to the origin, arriving there when $U(0)$ shows the time t_3. $U(\boldsymbol{x})$ is said to be synchronized with $U(0)$ if, at the moment when the light pulse is reflected at the point $\boldsymbol{x}$, $U(\boldsymbol{x})$ shows the time
>
> $$t_2 \;=\; t_1 + \tfrac{1}{2}\,(t_3 - t_1) \;=\; t_1 + |\boldsymbol{x}|\,/c$$
>
> where c is the speed of light in vacuum.

This synchronization procedure is certainly universal because it can be carried out intrinsically within any given inertial system. Moreover, it guarantees that light signals emanating from sources at rest propagate isotropically and with the universal velocity c. (Instead of light, one could use any other form of transmission of signals, provided it takes place isotropically with a known and fixed velocity of propagation v_0.) Conversely, Einstein's definition of t_2 as being simply the arithmetic mean of t_1 and t_3 is suggested by the requirement of isotropy of the propagation of light, since if more generally we set, e.g., $\tilde{t}_2 = t_1 + \epsilon(t_3 - t_1)$ with $0 \le \epsilon \le 1$, uniform rectilinear motion would still remain uniform rectilinear motion, but the propagation of light signals would for $\epsilon \neq 1/2$ take a more complicated form: the isotropy of the propagation of light would be hidden by an asymmetry in the choice of coordinates (in this case, the time coordinate).

We shall show later that Einstein's synchronization procedure coincides with the synchronization of clocks by *slow* transport.

To fix natural coordinates in inertial systems, we must not only determine the time coordinate by synchronizing clocks but also spatial coordinates: this can, as mentioned before, be done by introducing an oriented orthonormal 3-frame.

An inertial system with coordinates given by natural measuring rods and natural clocks, synchronized according to Einstein's procedure, will in the following be called a *Lorentz system.* Special relativity describes the phenomena of nature as viewed from Lorentz systems, postulating that all laws of nature have the same form in all of them. In particular, Maxwell's equations are valid in any Lorentz system.

7.2 Lorentz Transformations

The form of the coordinate transformations from one Lorentz system I to another Lorentz system I' is fixed by the Einsteinian principle of relativity. For its description it is convenient to replace t by the new time coordinate $x^0 = ct$ and to introduce the following canonical basis in $\mathbb{R}^4$:

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \quad e_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} , \quad e_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} . \tag{7.9}$$

The coordinates of an event are then given by the components of the *four-vector*

$$x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \in \mathbb{R}^4 . \tag{7.10}$$

We shall usually denote the components of x by x^μ, following the convention that greek indices $\mu, \nu, \kappa, \lambda, \ldots$ assume the values 0, 1, 2 and 3 while latin indices i, $j, k, l, \ldots$ assume the values 1, 2 and 3. Thus we have

$$x = x^\mu e_\mu = x^0 e_0 + x^1 e_1 + x^2 e_2 + x^3 e_3 \tag{7.11}$$

in analogy with

$$\boldsymbol{x} = x^i \boldsymbol{e}_i = x^1 \boldsymbol{e}_1 + x^2 \boldsymbol{e}_2 + x^3 \boldsymbol{e}_3 \tag{7.12}$$

(see Appendix). With this notation, the coordinate transformations between Lorentz systems take the form

$$x' = \Lambda x + a , \tag{7.13}$$

or written in components

$$x'^\mu = \Lambda^\mu_\nu x^\nu + a^\nu . \tag{7.14}$$

Here, the four-vector $a \in \mathbb{R}^4$ represents a shift of the origin of the space-time coordinate system. Such translations are certainly always possible but are also completely under control, so that it will be sufficient to discuss the homogeneous part. In other words, we want to determine the linear transformations Λ that give rise to coordinate transformations between Lorentz systems. Of course, these must again form a *group* which, for the time being, we shall denote by $\tilde{\Gamma}$: it is the subgroup of the group Γ introduced above consisting of those coordinate transformations that leave the origin of the space-time coordinate system invariant. It will be shown in this section that $\tilde{\Gamma}$, and hence also Γ, are completely fixed by the postulate of the constancy and universality of the speed of light.

To this end, we first define a non-degenerate symmetric bilinear form η on $\mathbb{R}^4$ which plays a fundamental role in all of relativity; it is given by

$$\eta(x, y) = \eta_{\mu\nu} x^\mu x^\nu , \tag{7.15}$$

where

$$\eta_{00} = +1 , \quad \eta_{11} = \eta_{22} = \eta_{33} = -1 , \quad \eta_{\mu\nu} = 0 \quad \text{for} \quad \mu \neq \nu . \tag{7.16}$$

One often writes $x \cdot y$ instead of $\eta(x, y)$, so that

$$x \cdot y \;=\; x^0 y^0 \;-\; \boldsymbol{x} \cdot \boldsymbol{y} \; . \tag{7.17}$$

With respect to the scalar product given by η, $\mathbb{R}^4$ becomes a four-dimensional pseudo-Euclidean vector space, called *Minkowski space*, in which the canonical basis (7.9) is an orthonormal basis (see Appendix).

The scalar product in Minkowski space is not positive definite (nor is it negative definite); in particular, there do exist four-vectors $x \neq 0$ for which $\eta(x, x) = 0$. Such four-vectors are called *lightlike* or *null*, and the set of all such vectors forms a double cone in Minkowski space, called the *light cone* (see Fig. 7.1). The term "lightlike" is motivated by the fact that a light pulse emitted at time x^0/c from the point $\boldsymbol{x}$ in the direction of the vector $\boldsymbol{y} - \boldsymbol{x}$ will arrive at time y^0/c at the point $\boldsymbol{y}$ if and only if

$$|\boldsymbol{y} - \boldsymbol{x}| \;=\; y^0 - x^0 \; ,$$

that is, if and only if the four-vector $y - x \in \mathbb{R}^4$ is lightlike (and satisfies $y^0 > x^0$), as well as by the fact that this statement, according to the postulate of the constancy and universality of the speed of light, must be valid in any Lorentz system – independently of the state of motion of the source that emits the light pulse or of the detector that receives it.

Fig. 7.1: Light cone in Minkowski space (one space dimension is suppressed)

For the homogeneous coordinate transformations $\Lambda \in \tilde{\Gamma}$ between Lorentz systems, this postulate amounts to requiring that they must map lightlike four-vectors to lightlike four-vectors:

$$\eta(x, x) \;=\; 0 \quad \Longrightarrow \quad \eta(\Lambda x, \Lambda x) \;=\; 0 \; . \tag{7.18}$$

But this implies

$$\eta(\Lambda x, \Lambda y) \;=\; a(\Lambda)\,\eta(x,y) \tag{7.19}$$

with a scalar factor $a(\Lambda)$ depending on the transformation Λ.

In order to prove this statement, we split Minkowski space into two subspaces V_+ and V_- which are orthogonal with respect to η, where V_+ consists of all scalar multiples of e_0 and V_- consists of all linear combinations of e_1, e_2 and e_3; then η is positive definite on V_+ and negative definite on V_-. Next, we set

$$a(\Lambda) \;=\; \eta(\Lambda e_0, \Lambda e_0)\;.$$

Then to begin with, we have for $x_+ \in V_+$ with $x_+ = x^0 e_0$

$$\eta(x_+, x_+) \;=\; (x^0)^2 \quad , \quad \eta(\Lambda x_+, \Lambda x_+) \;=\; (x^0)^2\,\eta(\Lambda e_0, \Lambda e_0)\;,$$

so

$$\eta(\Lambda x_+, \Lambda x_+) \;=\; a(\Lambda)\,\eta(x_+, x_+)\;. \tag{7.20-a}$$

On the other hand, we have for $x_+ \in V_+$ and $x_- \in V_-$

$$\eta(\Lambda x_+, \Lambda x_-) \;=\; 0\;, \tag{7.20-b}$$

$$\eta(\Lambda x_-, \Lambda x_-) \;=\; a(\Lambda)\,\eta(x_-, x_-)\;. \tag{7.20-c}$$

For the proof, we may assume without loss of generality that $x_+ \neq 0$, $x_- \neq 0$ and may replace the vector x_+ in eqn (7.20-b) by the rescaled vector

$$y_+ \;=\; \sqrt{-\,\frac{\eta(x_-, x_-)}{\eta(x_+, x_+)}}\;x_+ \;=\; \sqrt{-\,\eta(x_-, x_-)}\;e_0$$

The normalization of y_+ has been chosen in such a way that the vectors $x_- + y_+$ and $x_- - y_+$ are both lightlike:

$$\eta(x_- \pm y_+, x_- \pm y_+) \;=\; 0\;.$$

This implies

$$\eta(\Lambda(x_- \pm y_+), \Lambda(x_- \pm y_+)) \;=\; 0\;,$$

i.e.,

$$\eta(\Lambda x_-, \Lambda x_-) \,+\, \eta(\Lambda y_+, \Lambda y_+) \,\pm\, 2\eta(\Lambda x_-, \Lambda y_+) \;=\; 0\;,$$

so we get

$$\eta(\Lambda x_-, \Lambda y_+) \;=\; 0$$

and, using eqn (7.20-a),

$$\eta(\Lambda x_-, \Lambda x_-) \;=\; -\,\eta(\Lambda y_+, \Lambda y_+) \;=\; -\,a(\Lambda)\,\eta(y_+, y_+) \;=\; a(\Lambda)\,\eta(x_-, x_-)\;.$$

Combining eqns (7.20-a)–(7.20-c) gives, for any four-vector $x \in \mathbb{R}^4$,

$$\eta(\Lambda x, \Lambda x) \;=\; a(\Lambda)\,\eta(x, x)\;.$$

Replacing x in this formula first by $x + y$ and then by $x - y$ and taking the difference between the two resulting equations, we arrive at eqn (7.19).

In index notation and in matrix notation, eqn (7.19) takes the form

$$\eta_{\kappa\lambda}\, \Lambda^{\kappa}_{\mu}\, \Lambda^{\lambda}_{\nu} \;=\; a(\Lambda)\; \eta_{\mu\nu} \tag{7.21}$$

and

$$\Lambda^{\mathrm{T}} \eta \Lambda \;=\; a(\Lambda)\; \eta \tag{7.22}$$

respectively, where the factor $a(\Lambda)$ is always positive: $a(\Lambda) > 0$.

The possibility $a(\Lambda) = 0$ can be excluded since η is non-degenerate and Λ has an inverse. Indeed, suppose $a(\Lambda) < 0$; then η would be negative definite on the three-dimensional subspace V_- but positive definite on the three-dimensional subspace ΛV_-, i.e., would be simultaneously positive and negative definite on their intersection $V_- \cap \Lambda V_-$, which would be possible only if this intersection were trivial: $V_- \cap \Lambda V_- = \{0\}$. This however can be excluded on purely dimensional grounds: according to the formula

$$\dim\left(V_1 + V_2\right) \,+\, \dim\left(V_1 \cap V_2\right) \;=\; \dim\left(V_1\right) \,+\, \dim\left(V_2\right)$$

which is valid for any two subspaces V_1 and V_2 of an arbitrary finite-dimensional vector space W, we have in the pesent case $\dim\left(V_- \cap \Lambda V_-\right) \geq 2$.

Moreover, the factor $a(\Lambda)$ is obviously multiplicative with respect to the composition of transformations:

$$a(\Lambda_1 \Lambda_2) \;=\; a(\Lambda_1)\; a(\Lambda_2)\ . \tag{7.23}$$

Therefore, the subset of linear transformations Λ on Minkowski space $\mathbb{R}^4$ that satisfy the equations (7.19), (7.21) and (7.22) is a group $\hat{\Gamma}$ which is essentially built of two subgroups:

(a) *Scale transformations* are linear maps $D_\lambda : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ of the form

$$D_\lambda x \;=\; \lambda x \qquad \text{with} \quad \lambda > 0\ , \tag{7.24}$$

thus satisfying the relation $a(D_\lambda) = \lambda^2$. They form a subgroup of $\hat{\Gamma}$ which is isomorphic to the group $\mathbb{R}^+$ of positive real numbers, equipped with the standard multiplication.

(b) *Lorentz transformations* are linear maps $\Lambda : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ with the property that

$$\eta(\Lambda x, \Lambda y) \;=\; \eta(x, y)\ , \tag{7.25}$$

or written in index notation,

$$\eta_{\kappa\lambda}\, \Lambda^{\kappa}_{\mu}\, \Lambda^{\lambda}_{\nu} \;=\; \eta_{\mu\nu}\ , \tag{7.26}$$

or written in matrix notation

$$\Lambda^{\mathrm{T}} \eta \Lambda \;=\; \eta\ , \tag{7.27}$$

thus satisfying the relation $a(\Lambda) = 1$. They form a subgroup of $\hat{\Gamma}$ which is identical with the pseudo-orthogonal group $O(1,3)$; it is also called the *Lorentz group* and is often denoted by L.

We see immediately that every transformation in $\hat{\Gamma}$ can be written uniquely as the product of a scale transformation and a Lorentz transformation, where the order of the factors in the product is irrelevant because scale transformations commute with Lorentz transformations (even with general linear transformations). Thus $\hat{\Gamma}$ has the structure of a direct product

$$\hat{\Gamma} \;\simeq\; \mathbb{R}^+ \times L \;. \tag{7.28}$$

Our next step will be to investigate the structure of the Lorentz group L more closely. We begin with the statement that L is not connected but rather splits into four different connected components:

$$L \;=\; L_+^\uparrow \cup L_+^\downarrow \cup L_-^\uparrow \cup L_-^\downarrow \;. \tag{7.29}$$

For the proof, consider first the determinant and second the sign of the 00 diagonal element: For $\Lambda \in L$, we have

$$\det(\eta) \;=\; \det(\Lambda^{\mathrm{T}}\,\eta\,\Lambda) \;=\; \det(\Lambda^{\mathrm{T}})\,\det(\eta)\,\det(\Lambda)$$

and therefore

$$\det(\Lambda) \;=\; \pm\,1\;, \tag{7.30}$$

as well as

$$1 \;=\; \eta_{00} \;=\; \eta_{\kappa\lambda}\,\Lambda^\kappa_{\;0}\,\Lambda^\lambda_{\;0} \;=\; (\Lambda^0_{\;0})^2 - (\Lambda^1_{\;0})^2 - (\Lambda^2_{\;0})^2 - (\Lambda^3_{\;0})^2$$

and therefore

$$\left|\Lambda^0_{\;0}\right| \geq 1\;. \tag{7.31}$$

Then by definition,

$$\begin{aligned} L_+^\uparrow &= \{\,\Lambda \in L \;/\; \det(\Lambda) = +1\,,\; \Lambda^0_{\;0} \geq +1\,\}\;, \\ L_+^\downarrow &= \{\,\Lambda \in L \;/\; \det(\Lambda) = +1\,,\; \Lambda^0_{\;0} \leq -1\,\}\;, \\ L_-^\uparrow &= \{\,\Lambda \in L \;/\; \det(\Lambda) = -1\,,\; \Lambda^0_{\;0} \geq +1\,\}\;, \\ L_-^\downarrow &= \{\,\Lambda \in L \;/\; \det(\Lambda) = -1\,,\; \Lambda^0_{\;0} \leq -1\,\}\;. \end{aligned} \tag{7.32}$$

These subsets may be combined into subgroups by forming the group L_+ of *proper Lorentz transformations,*

$$L_+ \;=\; L_+^\uparrow \cup L_+^\downarrow \;=\; \{\,\Lambda \in L \;/\; \det(\Lambda) = 1\,\}\;, \tag{7.33}$$

the group $L^\uparrow$ of *orthochronous Lorentz transformations,*

$$L^\uparrow \;=\; L_+^\uparrow \cup L_-^\uparrow \;=\; \{\,\Lambda \in L \;/\; \Lambda^0_{\;0} \geq 1\,\}\;, \tag{7.34}$$

and their intersection, the group $L_+^\uparrow$ of *proper orthochronous Lorentz transformations,* as specified in eqn (7.32).

Within each of the subsets $L_-^\uparrow$, $L_-^\downarrow$ and $L_+^\downarrow$ specified in eqn (7.32) above, we can identify a distinguished element, namely the *spatial reflection* or *parity transformation* P, the *time reversal transformation* T and the *total reflection* PT, respectively, which are defined by

$$P\begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \;=\; \begin{pmatrix} x^0 \\ -\,\boldsymbol{x} \end{pmatrix} , \qquad \text{i.e.,} \qquad P \;=\; \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ,$$

$$T\begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \;=\; \begin{pmatrix} -\,x^0 \\ \boldsymbol{x} \end{pmatrix} , \qquad \text{i.e.,} \qquad T \;=\; \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} , \tag{7.35}$$

$$PT(x) \;=\; -\,x\ , \qquad \text{i.e.,} \qquad PT \;=\; \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ,$$

It is immediately verified that every Lorentz transformation $\Lambda \in L$ can be uniquely represented in the form

$$\Lambda \;=\; P^n\, T^m\, \Lambda_0 \qquad \text{with} \quad n, m \in \{0,1\}\ ,\ \Lambda_0 \in L_+^\uparrow\ . \tag{7.36}$$

Therefore, the problem is reduced to the task of clarifying the structure of the proper orthochronous Lorentz group; in particular, we still have to show that $L_+^\uparrow$ is indeed connected.

To this end, we remark first that $L_+^\uparrow$ contains, in a natural way, the rotation group $SO(3)$:

$$\tilde{R} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix} \in L_+^\uparrow \qquad \text{for} \quad R \in SO(3)\ . \tag{7.37}$$

Indeed, we obviously have

$$\tilde{R}\begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \;=\; \begin{pmatrix} x^0 \\ R\boldsymbol{x} \end{pmatrix} \tag{7.38}$$

and hence

$$\eta(\tilde{R}x, \tilde{R}y) \;=\; x^0 y^0 \,-\, (R\boldsymbol{x})\cdot(R\boldsymbol{y}) \;=\; x^0 y^0 \,-\, \boldsymbol{x}\cdot\boldsymbol{y} \;=\; \eta(x,y)$$

as well as $\det(\tilde{R}) = 1$ and $\tilde{R}^0_0 = 1$. Normally, the $\tilde{R}$ are identified with the R and are called the *rotations* in Minkowski space; they form a subgroup of the (proper orthochronous) Lorentz group isomorphic to the rotation group $SO(3)$. For later use, note also that a (proper orthochronous) Lorentz transformation $\Lambda \in L_+^\uparrow$ is a rotation if and only if

$$\Lambda e_0 \;=\; e_0\ . \tag{7.39}$$

Next we shall determine those (proper orthochronous) Lorentz transformations $\Lambda \in L_+^\uparrow$ that leave, say, the coordinates x^2 and x^3 fixed, i.e., are of the form

$$\Lambda = \begin{pmatrix} \Lambda^0_0 & \Lambda^0_1 & 0 & 0 \\ \Lambda^1_0 & \Lambda^1_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

For these, we must have

$$\begin{aligned} (\Lambda^0_0)^2 - (\Lambda^1_0)^2 &= \eta(\Lambda e_0, \Lambda e_0) = \eta(e_0, e_0) = +1 , \\ (\Lambda^0_1)^2 - (\Lambda^1_1)^2 &= \eta(\Lambda e_1, \Lambda e_1) = \eta(e_1, e_1) = -1 , \\ \Lambda^0_0 \Lambda^0_1 - \Lambda^1_0 \Lambda^1_1 &= \eta(\Lambda e_0, \Lambda e_1) = \eta(e_0, e_1) = 0 , \end{aligned}$$

as well as

$$\Lambda^0_0 \geq 1$$

and

$$\det(\Lambda) = \Lambda^0_0 \Lambda^1_1 - \Lambda^1_0 \Lambda^0_1 = 1 .$$

The most general solution of this system of equations can be written in the form

$$\Lambda^0_0 = \Lambda^1_1 = \cosh\theta \quad , \quad \Lambda^1_0 = \Lambda^0_1 = -\sinh\theta$$

with some $\theta \in \mathbb{R}$ which is arbitrary; we shall denote this transformation by $\Lambda_1(\theta)$:

$$\Lambda_1(\theta) = \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{7.40}$$

Lorentz transformations of the form $\Lambda_1(\theta)$ are called *boosts* in the 1-direction and the parameter θ is called the *rapidity.* Its interpretation can be derived by considering the action of $\Lambda_1(\theta)$ on four-vectors $x \in \mathbb{R}^4$:

$$\begin{aligned} x'^0 &= \cosh\theta\, x^0 - \sinh\theta\, x^1 , \\ x'^1 &= \cosh\theta\, x^1 - \sinh\theta\, x^0 , \\ x'^2 &= x^2 \quad , \quad x'^3 = x^3 . \end{aligned} \tag{7.41}$$

In particular, we have $x'^1 = 0$ if and only if $x^1 = ct\tanh\theta = vt$. Thus $\Lambda_1(\theta)$ describes a coordinate transformation in which the origin of I', as viewed from I, moves with velocity

$$v = c\tanh\theta \tag{7.42}$$

in the 1-direction. Expressing the rapidity θ in terms of the velocity v, the transformations (7.41) become

$$\begin{aligned} x'^0 &= \gamma\,(x^0 - \beta x^1) , \\ x'^1 &= \gamma\,(x^1 - \beta x^0) , \\ x'^2 &= x^2 \quad , \quad x'^3 = x^3 , \end{aligned} \tag{7.43}$$

with

$$\beta \;=\; \frac{v}{c} \qquad \text{and} \qquad \gamma \;=\; \frac{1}{\sqrt{1-\beta^2}} \;=\; \frac{1}{\sqrt{1-v^2/c^2}} \tag{7.44}$$

or even more explicitly,

$$\begin{aligned} t' &= \frac{t - v x^1/c^2}{\sqrt{1-v^2/c^2}} \; , \\ x'^1 &= \frac{x^1 - vt}{\sqrt{1-v^2/c^2}} \; , \\ x'^2 &= x^2 \quad , \quad x'^3 \;=\; x^3 \; . \end{aligned} \tag{7.45}$$

The definition of boosts $\Lambda_2(\theta)$ in the 2-direction and boosts $\Lambda_3(\theta)$ in the 3-direction, as well as, more generally, of boosts $\Lambda(\boldsymbol{n},\theta)$ along an arbitrary unit vector $\boldsymbol{n}$ in 3-space, is analogous. For example, eqn (7.40) is extended to

$$\begin{aligned} \Lambda_1(\theta) &= \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \\ \Lambda_2(\theta) &= \begin{pmatrix} \cosh\theta & 0 & -\sinh\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sinh\theta & 0 & \cosh\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \\ \Lambda_3(\theta) &= \begin{pmatrix} \cosh\theta & 0 & 0 & -\sinh\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\theta & 0 & 0 & \cosh\theta \end{pmatrix} , \end{aligned} \tag{7.46}$$

whereas, in analogy with eqns (7.41) and (7.43), the action of $\Lambda(\boldsymbol{n},\theta)$ on four-vectors in Minkowski space is given by

$$\begin{aligned} x'^0 &= \cosh\theta\, x^0 \;-\; \sinh\theta\, (\boldsymbol{n}\cdot\boldsymbol{x}) \; , \\ \boldsymbol{x}'_{\parallel} &= \cosh\theta\, \boldsymbol{x}_{\parallel} \;-\; \sinh\theta\, x^0\, \boldsymbol{n} \; , \\ \boldsymbol{x}'_{\perp} &= \boldsymbol{x}_{\perp} \; , \end{aligned} \tag{7.47}$$

with

$$\boldsymbol{v} \;=\; c\, \tanh\theta\, \boldsymbol{n} \; , \tag{7.48}$$

or

$$\begin{aligned} x'^0 &= \gamma\, (x^0 \;-\; \beta\, \boldsymbol{n}\cdot\boldsymbol{x}) \; , \\ \boldsymbol{x}'_{\parallel} &= \gamma\, (\boldsymbol{x}_{\parallel} \;-\; \beta \boldsymbol{n} x^0) \; , \\ \boldsymbol{x}'_{\perp} &= \boldsymbol{x}_{\perp} \; , \end{aligned} \tag{7.49}$$

with

$$\beta \;=\; \frac{v}{c} \qquad \text{and} \qquad \gamma \;=\; \frac{1}{\sqrt{1-\beta^2}} \;=\; \frac{1}{\sqrt{1-v^2/c^2}} \tag{7.50}$$

or even more explicitly,

$$\begin{aligned} t' &= \frac{t - (\boldsymbol{v}\cdot\boldsymbol{x})/c^2}{\sqrt{1 - v^2/c^2}} \;, \\ \boldsymbol{x}'_\parallel &= \frac{\boldsymbol{x}_\parallel - \boldsymbol{v}t}{\sqrt{1 - v^2/c^2}} \;, \\ \boldsymbol{x}'_\perp &= \boldsymbol{x}_\perp \;, \end{aligned} \tag{7.51}$$

where $\boldsymbol{x}_\parallel$ ($\boldsymbol{x}'_\parallel$) and $\boldsymbol{x}_\perp$ ($\boldsymbol{x}'_\perp$) denote the components of $\boldsymbol{x}$ ($\boldsymbol{x}'$) parallel and orthogonal to $\boldsymbol{n}$, respectively.

Thus $\Lambda(\boldsymbol{n},\theta)$ is a coordinate transformation between two Lorentz systems which move with velocity $\boldsymbol{v}$ in relation to each other, where $\boldsymbol{n}$, θ and $\boldsymbol{v}$ are related by eqn (7.48); we shall also write $B(\boldsymbol{v})$ rather than $\Lambda(\boldsymbol{n},\theta)$. The advantage of using the rapidity θ instead of the velocity v (or the dimensionless velocity $\beta = v/c$) lies in the simple composition rule: successive application of two boosts in the same direction produces another boost, also in the same direction, where the rapidities simply add:

$$\Lambda(\boldsymbol{n},\theta_1)\,\Lambda(\boldsymbol{n},\theta_2) \;=\; \Lambda(\boldsymbol{n},\theta_1+\theta_2)\;. \tag{7.52}$$

Translating this rule from the rapidities to the velocities gives

$$B(\boldsymbol{v}_1)\,B(\boldsymbol{v}_2) \;=\; B(\boldsymbol{v}) \tag{7.53}$$

with

$$\boldsymbol{v}_1 \;=\; v_1\boldsymbol{n} \quad,\quad \boldsymbol{v}_2 \;=\; v_2\boldsymbol{n} \quad,\quad \boldsymbol{v} \;=\; v\boldsymbol{n} \tag{7.54}$$

and

$$v \;=\; \frac{v_1 + v_2}{1 + v_1 v_2/c^2}\;. \tag{7.55}$$

The last equation is known as the *relativistic addition theorem of velocities.* Moreover, we see from equations (7.47)–(7.51) that the speed of light c constitutes an upper bound for the possible relative velocity between any two inertial systems, since the quantity γ becomes infinite for $v = c$ and imaginary for $v > c$, whereas eqns (7.52) and (7.55) show that even the trick of combining several successive boosts in the same direction will not allow us to bypass this upper bound. In passing, we note that successive application of two boosts in different directions will no longer produce a boost, i.e., the set of boosts does *not* provide a subgroup of the (proper orthochronous) Lorentz group. However, boosts do behave naturally under rotations:

$$\begin{aligned} \tilde{R}\,\Lambda(\boldsymbol{n},\theta)\,\tilde{R}^{-1} &= \Lambda(R\boldsymbol{n},\theta) \\ &\text{or} \\ \tilde{R}\,B(\boldsymbol{v})\,\tilde{R}^{-1} &= B(R\boldsymbol{v})\;. \end{aligned} \tag{7.56}$$

Finally, every (proper orthochronous) Lorentz transformation Λ can be written uniquely as the product of a boost $B(\Lambda)$ with a rotation $\tilde{R}(\Lambda)$:

$$\Lambda \;=\; B(\Lambda)\,\tilde{R}(\Lambda)\;. \tag{7.57}$$

Such a decomposition is known in mathematics as a *Cartan decomposition* – in this case of the (proper orthochronous) Lorentz group.

For the proof, assume that $\Lambda \in L_+^\uparrow$ is given, and write

$$\Lambda e_0 \;=\; x_\Lambda \;=\; \begin{pmatrix} x_\Lambda^0 \\ \boldsymbol{x}_\Lambda \end{pmatrix} .$$

Due to

$$(x_\Lambda^0)^2 - (\boldsymbol{x}_\Lambda)^2 \;=\; \eta(e_0, e_0) \;=\; 1 \quad , \quad x_\Lambda^0 \geq 1$$

we may put

$$x_\Lambda^0 \;=\; \cosh\theta \quad , \quad |\boldsymbol{x}_\Lambda| \;=\; \sinh\theta \quad , \quad \boldsymbol{n} \;=\; \boldsymbol{x}_\Lambda / |\boldsymbol{x}_\Lambda|$$

and obtain

$$\Lambda(\boldsymbol{n}, -\theta)\, \Lambda\, e_0 \;=\; e_0 \; ,$$

i.e., $\Lambda(\boldsymbol{n}, -\theta)\, \Lambda$ is a rotation, thus proving eqn (7.57) with $B(\Lambda) = \Lambda(\boldsymbol{n}, \theta)$ and $\tilde{R}(\Lambda) = \Lambda(\boldsymbol{n}, -\theta)\, \Lambda$.

The uniqueness of the decomposition is simply based on the fact that a (proper orthochronous) Lorentz transformation which is simultaneously a boost and a rotation must be the identity.

Using eqn (7.56), we may even show that every (proper orthochronous) Lorentz transformation $\Lambda \in L_+^\uparrow$ can be written in the form

$$\Lambda \;=\; \tilde{R}(\Lambda)\, B_0(\Lambda)\, \tilde{R}'(\Lambda) \; , \tag{7.58}$$

where $\tilde{R}$ and $\tilde{R}'$ are rotations and B_0 is a boost in some fixed direction $\boldsymbol{n}_0$; such a decomposition is however no longer unique. The decomposition (7.57) also implies that the proper orthochronous Lorentz group must be connected because the rotation group $SO(3)$ is connected: the fact that every element $R \in SO(3)$ can be connected to the identity element through a continuous curve (which will be taken for granted here) implies that the same holds true for every element $\Lambda \in L_+^\uparrow$. Therefore, as has been stated before, the Lorentz group as a whole really does have precisely four connected components.

Concluding, we are now in a position to prove that the homogeneous coordinate transformations between Lorentz systems are exactly the proper orthochronous Lorentz transformations:

$$\tilde{\Gamma} \;=\; L_+^\uparrow \; . \tag{7.59}$$

A first hint in this direction is provided by the observation that $\tilde{\Gamma}$ must be contained in the subgroup L of $\hat{\Gamma}$: Indeed, one may argue that, according to the principle of relativity and the isotropy of physical space, the factor $a(\Lambda)$ in eqns (7.19), (7.21) and (7.22) can only depend on the modulus of the relative velocity of the inertial systems involved:

$$a(\Lambda) \;=\; a(|\boldsymbol{v}|) \; .$$

But then we also get

$$a(\Lambda^{-1}) \;=\; a(|\boldsymbol{v}|) \; .$$

Applzing the product rule (7.23) gives

$$a(|\boldsymbol{v}|)^2 \;=\; a(0) \;=\; 1\;,$$

and hence $a(|\boldsymbol{v}|) = 1$. In what follows, we shall however use a different procedure, which allows to handle the discrete parts as well.

To this end, we start out from the observation that, due to isotropy of physical space and by continuity, $\tilde{\Gamma}$ should contain

(a) arbitrary spatial rotations $\tilde{R}$,

(b) for every velocity $\boldsymbol{v}$ satisfying the condition $v < c$, precisely one boost $\tilde{B}(\boldsymbol{v})$ describing the transition from a given Lorentz system to a new one whose origin moves relatively to the old one with velocity $\boldsymbol{v}$.

On the other hand, $\tilde{\Gamma}$ should not contain

(c) the scale transformations D_λ,

(d) the reflections P and T,

(e) combinations of these.

Indeed, the presence of such transformations within $\tilde{\Gamma}$ would violate the principle of universality of the procedure for fixing coordinates in each Lorentz system, since they would represent transformations between coordinate systems without relative motion and with parallel coordinate axes. Natural coordinates in two such systems must however be identical.

Now according to the equations (7.19), (7.21) and (7.22), $\tilde{\Gamma}$ is a subgroup of $\hat{\Gamma}$. But we know from eqns (7.28) and (7.36) that every linear transformation $\Lambda \in \hat{\Gamma}$ can be written uniquely in the form

$$\Lambda \;=\; D_\lambda\, P^n\, T^m\, \Lambda_0 \qquad \text{with} \quad \lambda > 0\;,\;\; n, m \in \{0,1\}\;,\;\; \Lambda_0 \in L_+^\uparrow\;. \tag{7.60}$$

In particular, every $\tilde{\Lambda} \in \tilde{\Gamma}$ is of the form

$$\tilde{\Lambda} \;=\; D_\lambda\, P^n\, T^m\, \Lambda \qquad \text{with} \quad \lambda > 0\;,\;\; n, m \in \{0,1\}\;,\;\; \Lambda \in L_+^\uparrow\;, \tag{7.61}$$

and according to (a) and (b), all $\Lambda \in L_+^\uparrow$ will appear in this way. But every $\Lambda \in L_+^\uparrow$ can be associated with only *one* value for λ, for n and for m, since otherwise the subgroup property of $\tilde{\Gamma}$ would force at least one of the forbidden transformations in (c)–(e) to belong to $\tilde{\Gamma}$; thus in eqn (7.61), the numbers λ, n and m are in fact functions of Λ, with

$$\lambda(\Lambda_1\Lambda_2) \;=\; \lambda(\Lambda_1)\,\lambda(\Lambda_2) \tag{7.62}$$

and

$$\begin{aligned} n(\Lambda_1\Lambda_2) \;&=\; n(\Lambda_1) + n(\Lambda_2) \qquad \bmod 2\;, \\ m(\Lambda_1\Lambda_2) \;&=\; m(\Lambda_1) + m(\Lambda_2) \qquad \bmod 2\;. \end{aligned} \tag{7.63}$$

Moreover, we have $n(1) = m(1) = 0$, and as $n(\Lambda)$ and $m(\Lambda)$ must depend continuously on Λ and $L_+^\uparrow$ is connected, we get

$$n(\Lambda) \;=\; m(\Lambda) \;=\; 0$$

for all $\Lambda \in L_+^\uparrow$. It only remains to be shown that, similarly,

$$\lambda(\Lambda) \;=\; 1$$

for all $\Lambda \in L_+^\uparrow$; this is based on the multiplicativity expressed by eqn (7.62).

For the proof, we use the decomposition (7.57): $\Lambda = B\tilde{R}$ implies $\lambda(\Lambda) = \lambda(B)\,\lambda(\tilde{R})$, so it suffices to show that $\lambda(\tilde{R}) = 1$ and $\lambda(B) = 1$. Now every rotation $\tilde{R}$ and every boost B can be written in the form

$$\tilde{R} \;=\; \tilde{S}\,\tilde{R}_3(\varphi)\,\tilde{S}^{-1} \qquad \text{and} \qquad B \;=\; \tilde{S}\,\Lambda_3(\theta)\,\tilde{S}^{-1}\;,$$

respectively, where $\tilde{S}$ is an appropriate rotation while $\tilde{R}_3(\varphi)$ is a rotation about the 3-axis with angle φ and $\Lambda_3(\theta)$ is a boost in the 3-direction with velocity $v = c\tanh\theta$. Due to the multiplicativity expressed by eqn (7.62), we have

$$\lambda(\tilde{R}) \;=\; \lambda(\tilde{S})\,\lambda(\tilde{R}_3(\varphi))\,\lambda(\tilde{S}^{-1}) \;=\; \lambda(\tilde{R}_3(\varphi))$$

and

$$\lambda(B) \;=\; \lambda(\tilde{S})\,\lambda(\Lambda_3(\theta))\,\lambda(\tilde{S}^{-1}) \;=\; \lambda(\Lambda_3(\theta))\;,$$

as well as

$$\tilde{R}_3(\varphi_1 + \varphi_2) \;=\; \tilde{R}_3(\varphi_1)\,\tilde{R}_3(\varphi_2)$$

and

$$\Lambda_3(\theta_1 + \theta_2) \;=\; \Lambda_3(\theta_1)\,\Lambda_3(\theta_2)\;,$$

so that

$$\lambda(\tilde{R}) \;=\; \lambda(\tilde{R}_3(\varphi)) \;=\; \exp(\text{const.}\,\varphi)$$

and

$$\lambda(B) \;=\; \lambda(\Lambda_3(\theta)) \;=\; \exp(\text{const.}\,\theta)\;.$$

However, when combining rotations about different axes or boosts in different directions, the angles or rapidities do not simply add, so that the only possibility to satisfy these conditions is to set $\text{const.} \equiv 0$, i.e., $\lambda \equiv 1$, as was to be shown.

Thus we have definitely identified, on the basis of Einstein's principle of relativity alone, the group of all linear homogeneneous coordinate transformations between Lorentz systems as being the proper orthochronous Lorentz group.

What remains to be added are the translations. One defines the *Poincaré group* P, the *proper Poincaré group* P_+, the *orthochronous Poincaré group* $P^\uparrow$ and the *proper orthochronous Poincaré group* $P_+^\uparrow$ as the groups of affine transformations in Minkowski space whose homogeneous parts are the corresponding versions of the Lorentz group:

$$P \;=\; \{\,(a,\Lambda)\;/\;a \in \mathbb{R}^4\;,\;\Lambda \in L\;\}\;, \tag{7.64}$$

$$P_+ \;=\; \{\,(a,\Lambda)\;/\;a \in \mathbb{R}^4\;,\;\Lambda \in L_+\,\}\;, \tag{7.65}$$

$$P^\uparrow \;=\; \{\,(a,\Lambda)\;/\;a \in \mathbb{R}^4\;,\;\Lambda \in L^\uparrow\;\}\;, \tag{7.66}$$

$$P_+^\uparrow \;=\; \{\,(a,\Lambda)\;/\;a \in \mathbb{R}^4\;,\;\Lambda \in L_+^\uparrow\,\}\;. \tag{7.67}$$

The multiplication law results from the interpretation of the pair (a, Λ) as a Lorentz transformation Λ followed by a translation by a:

$$(a_1, \Lambda_1)\,(a_2, \Lambda_2) \;=\; (a_1 + \Lambda_1 a_2, \Lambda_1 \Lambda_2)\;. \tag{7.68}$$

We have thus proved completely, on the basis of Einstein's principle of relativity alone, that *the invariance group of special relativity*, i.e., the group of all coordinate transformations between Lorentz systems, *is the proper orthochronous Poincaré group.*

7.3 On the Geometry of Minkowski Space

In this section we shall investigate the question which four-vectors $x \in \mathbb{R}^4$ can be transformed into each other by proper orthochronous Lorentz transformations. Two vectors x and x' in Minkowski space will be called *equivalent* if there exists a proper orthochronous Lorenty transformation $\Lambda \in L_+^\uparrow$ such that $x' = \Lambda x$; in this case we write $x \approx x'$. Obviously, "$\approx$" does satisfy the defining properties of an equivalence relation, namely

$$\begin{array}{rc} \text{Reflexivity:} & x \approx x\;, \\ \text{Symmetry:} & x \approx x' \implies x' \approx x\;, \\ \text{Transitivity:} & x \approx x' \text{ and } x' \approx x'' \implies x \approx x''\;. \end{array}$$

For the rotation group and vectors in three-dimensional Euclidean space $\mathbb{R}^3$, the solution of the corresponding equivalence problem is well known: Vectors can be transformed into each other by a rotation if and only if they have the same length.

For the Lorentz group and vectors in four-dimensional Minkowski space $\mathbb{R}^4$, the second part of this statement remains valid: two four-vectors x and x' can certainly be transformed into each other by a Lorentz transformation only if their invariant "length" with respect to the Lorentz metric η is the same:

$$x^2 \;=\; x'^2\;.$$

Definition: A non-zero four-vector x in Minkowski space is called

$$\begin{array}{ccc} \textit{timelike} & \text{if} & x^2 > 0\;, \\ \textit{lightlike} \text{ or } \textit{null} & \text{if} & x^2 = 0\;, \\ \textit{spacelike} & \text{if} & x^2 < 0\;. \end{array}$$

This gives four different classes of four-vectors, the fourth class being trivial, consisting of the zero four-vector alone. Obviously, four-vectors in different classes are inequivalent. Moreover, each class of timelike four-vectors of a given invariant "length" and the class of lightlike four-vectors decompose into two distinct subclasses, according to the sign of their time component x^0 (note that timelike and lightlike vectors always satisfy $x^0 \neq 0$).

In order to show this, we consider an arbitrary four-vector

$$x = \begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix} \in \mathbb{R}^4$$

and its image

$$\Lambda x = x' = \begin{pmatrix} x'^0 \\ \boldsymbol{x}' \end{pmatrix} \in \mathbb{R}^4$$

under a proper orthochronous Lorentz transformation $\Lambda \in L_+^\uparrow$.

If x is timelike or lightlike, so is x', and we have $|x^0| > |\boldsymbol{x}| > 0$ and $|x'^0| > |\boldsymbol{x}'| > 0$ or $|x^0| = |\boldsymbol{x}| > 0$ and $|x'^0| = |\boldsymbol{x}'| > 0$, respectively. In this case, $\Lambda^0_0 \geq 1$ implies that x'^0 and x^0 have the same sign:

$$\begin{aligned}
x^0 \geq 0 \quad \Longrightarrow \quad x'^0 \;&=\; \Lambda^0_0 x^0 + \Lambda^0_1 x^1 + \Lambda^0_2 x^2 + \Lambda^0_3 x^3 \\
&\geq\; \Lambda^0_0 x^0 - \sqrt{(\Lambda^0_1)^2 + (\Lambda^0_2)^2 + (\Lambda^0_3)^2}\,|\boldsymbol{x}| \\
&\geq\; \Lambda^0_0\,(x^0 - |\boldsymbol{x}|) \;\geq\; 0\;, \\
x^0 \leq 0 \quad \Longrightarrow \quad x'^0 \;&=\; \Lambda^0_0 x^0 + \Lambda^0_1 x^1 + \Lambda^0_2 x^2 + \Lambda^0_3 x^3 \\
&\leq\; \Lambda^0_0 x^0 + \sqrt{(\Lambda^0_1)^2 + (\Lambda^0_2)^2 + (\Lambda^0_3)^2}\,|\boldsymbol{x}| \\
&\leq\; \Lambda^0_0\,(x^0 + |\boldsymbol{x}|) \;\leq\; 0\;.
\end{aligned}$$

Moreover, fixing a unit vector $\boldsymbol{e} \in \mathbb{R}^3$, we can verify the following statements:
If x is timelike, then by applying a boost $\Lambda = B(\boldsymbol{v})$ with velocity $\boldsymbol{v} = c\boldsymbol{x}/x^0$, x can be brought to the normal form x_{N}, where $x^0_{\mathrm{N}} = (\operatorname{sgn} x^0)\sqrt{x^2}$ and $\boldsymbol{x}_{\mathrm{N}} = 0$.
If x is lightlike, then by applying a boost $\Lambda = B(\boldsymbol{v})$ with velocity

$$\boldsymbol{v} = -(\operatorname{sgn} x^0)\, c\, \frac{1 - |\boldsymbol{x}|^2}{1 + |\boldsymbol{x}|^2}\, \frac{\boldsymbol{x}}{|\boldsymbol{x}|}\;,$$

x can be brought to the form x', where $x'^0 = \operatorname{sgn} x^0$ and $\boldsymbol{x}' = \boldsymbol{x}/|\boldsymbol{x}|$. By a subsequent rotation, $\boldsymbol{x}'$ may then be transformed into $\boldsymbol{e}$.

If x is spacelike, then by applying a boost $\Lambda = B(\boldsymbol{v})$ with velocity $\boldsymbol{v} = cx^0\boldsymbol{x}/|\boldsymbol{x}|^2$, x can be brought to the form x', where $x'^0 = 0$ and $\boldsymbol{x}' = \sqrt{-x^2}\,\boldsymbol{x}/|\boldsymbol{x}|$. By a subsequent rotation, $\boldsymbol{x}'$ may then be transformed into $\sqrt{-x^2}\,\boldsymbol{e}$.

In this way, we arrive at the following subdivision of Minkowski space into $L_+^\uparrow$-invariant subsets (see Fig. 7.2):

$$\mathbb{R}^4 = V^+ \cup V^- \cup V_0^+ \cup V_0^- \cup R \cup \{0\}\;. \tag{7.69}$$

Fig. 7.2: Subdivision of Minkowski space into $L_+^\uparrow$-invariant subsets (two space dimensions are suppressed)

For these subsets, the following terminology is used:

1. Elements of
$$V^+ \;=\; \{\, x \in \mathbb{R}^4 \;/\; x^2 > 0 \;,\; x^0 > 0 \,\} \tag{7.70}$$
and of
$$V^- \;=\; \{\, x \in \mathbb{R}^4 \;/\; x^2 > 0 \;,\; x^0 < 0 \,\} \tag{7.71}$$
are called *future oriented* and *past oriented timelike vectors*, respectively. By applying a proper orthochronous Lorentz transformation, such vectors may be brought to the *normal form*
$$\begin{pmatrix} +\sqrt{x^2} \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -\sqrt{x^2} \\ 0 \end{pmatrix} \;, \tag{7.72}$$
respectively.

2. Elements of
$$V_0^+ \;=\; \{\, x \in \mathbb{R}^4 \;/\; x^2 = 0 \;,\; x^0 > 0 \,\} \tag{7.73}$$
and of
$$V_0^- \;=\; \{\, x \in \mathbb{R}^4 \;/\; x^2 = 0 \;,\; x^0 < 0 \,\} \tag{7.74}$$
are called *future oriented* and *past oriented, lightlike vectors*, respectively. By applying a proper orthochronous Lorenty transformation, such vectors may be brought to the *normal form*
$$\begin{pmatrix} +1 \\ \boldsymbol{e} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ \boldsymbol{e} \end{pmatrix} \;, \tag{7.75}$$
respectively, where $\boldsymbol{e}$ denotes a fixed unit vector in $\mathbb{R}^3$.

3. Elements of
$$R \;=\; \{\, x \in \mathbb{R}^4 \;/\; x^2 < 0 \,\} \tag{7.76}$$
are called *spacelike vectors*. By applying a proper orthochronous Lorentz transformation, such vectors may be brought to the *normal form*
$$\begin{pmatrix} 0 \\ \sqrt{-x^2}\,\boldsymbol{e} \end{pmatrix} , \tag{7.77}$$
where $\boldsymbol{e}$ denotes a fixed unit vector in $\mathbb{R}^3$. In particular, proper orthochronous Lorentz transformations can change the sign of the time component of a spacelike vector.

The sets V_0^+ and V_0^- are also called the *forward* or *future light cone* and the *backward* or *past light cone*, respectively, whereas the sets
$$\bar{V}^+ \;=\; V^+ \cup V_0^+ \cup \{0\} \tag{7.78}$$
and
$$\bar{V}^- \;=\; V^- \cup V_0^- \cup \{0\} \tag{7.79}$$
are sometimes called the *forward* or *future cone* and the *backward* or *past cone*, respectively; then V^+ is the interior and $V_0^+ \cup \{0\}$ is the boundary of the forward cone whereas V^- is the interior and $V_0^- \cup \{0\}$ is the boundary of the past cone.

Finally we agree to call two events represented by the four-vectors x and y in $\mathbb{R}^4$ (with $x \neq y$) *timelike separated* or *lightlike separated* or *spacelike separated* whenever we have $(y-x)^2 > 0$ or $(y-x)^2 = 0$ or $(y-x)^2 < 0$, respectively. In the case of timelike or lightlike separation, we can also define their temporal order: if $(y-x)^2 \geq 0$, we say x is *before* y and y is *after* x if $y^0 > x^0$.

This subdivision has profound physical significance:

Suppose first that x and y are timelike or lightlike separated, i.e.,
$$\left|y^0 - x^0\right| \geq |\boldsymbol{y} - \boldsymbol{x}| \;\; .$$
Obviously, such events can, at least in principle, be connected by a signal or a causal influence whose velocity of propagation does not exceed the speed of light. If on the other hand x and y are spacelike separated, i.e.,
$$\left|y^0 - x^0\right| < |\boldsymbol{y} - \boldsymbol{x}| \;\; ,$$
then a signal to be exchanged between two such events would have to propagate with a medium velocity
$$v \;=\; c\,\frac{|\boldsymbol{y} - \boldsymbol{x}|}{|y^0 - x^0|} \;>\; c \; .$$
If such signals existed, then by passing to another inertial system we could switch the temporal order between the event of its emission and the event of its reception. The consequences of this possibility are absurd: a phenomenon of this kind would jeopardize the possibility to distinguish in an invariant way between cause and

effect, and we would be forced to give up either the principle of relativity or the principle of causality. This fundamental problem is the real reason behind the special role of the speed of light as the maximal possible velocity of propagation for signals and for any kind of causal influence. In particular, the velocity of point particles is limited by the speed of light – at least as far as they may be used to transmit information or any other kind of causal influence. If on the other hand one wants to transmit information by using wave phenomena, one has to take into account that the phase velocity

$$v_{\mathrm{ph}} \; = \; \omega(\boldsymbol{k})/k$$

of a plane wave may very well exceed the speed of light because such a wave is inadequate for transmitting signals. Transmitting information requires using pulses, that is, wave packets, which propagate with the group velocity

$$v_{\mathrm{gr}} \; = \; |\boldsymbol{\nabla}_{\boldsymbol{k}}\, \omega(\boldsymbol{k})| \; .$$

For example, we have for a wave guide

$$\omega(\boldsymbol{k}) \; = \; \sqrt{\omega_0^2 + k^2 c^2} \; ,$$

so that

$$v_{\mathrm{ph}} \; = \; \frac{\sqrt{\omega_0^2 + k^2 c^2}}{k} \; > \; c \quad , \quad v_{\mathrm{gr}} \; = \; \frac{k c^2}{\sqrt{\omega_0^2 + k^2 c^2}} \; < \; c \; .$$

A more precise analysis reveals that the really relevant quantity is the front end velocity

$$v_{\mathrm{fr}} \; = \; \lim_{k \to \infty} \frac{\omega(k)}{k} \; ;$$

this is the velocity of propagation in space of the first signal of a wave whose source has been turned on instantaneously.

Quite generally and as a matter of principle it may be stated that only timelike or lightlike separated events can be causally related and that their causal relationship (the order of cause and effect) is independent of the chosen inertial system. For spacelike separated events, on the other hand, one can always find an inertial system I in which they are simultaneous, but they will then no longer be simultaneous in an inertial system I' that is moving with respect I: this is the *relativity of simultaneity.*

To conclude this section, we report two other interesting results on the characterization of Lorentz transformations as transformations between inertial systems.

1. As early as 1911, P. Frank and H. Rothe have investigated to what extent the general requirements (R1) and (R2) of the principle of relativity alone – that is, homogeneity and isotropy of space and time together with physical equivalence between all inertial systems – determine the transformations between inertial systems. We have already seen that these transformations must be affine and must form a group. It turns out that for the homogeneneous transformations there remain only three possibilities:

a) homogeneous Galilei transformations,

b) Lorentz transformations with a finite limit velocity c_∞,

c) four-dimensional rotations.

Possibility c) can be immediately discarded because in this case one could by an appropriate rotation invert the sign of any coordinate, so that the temporal order between any two events would always depend on the inertial system. Homogeneous Galilei transformations result from Lorentz transformations by considering the limit $c_\infty \to \infty$. Which of the two possibilities a) or b) is realized therefore depends on whether there exists a finite limit velocity c_∞ for signals or not; if it does, it is also a limit for the relative velocity between inertial systems.

2. The orthochronous Lorentz transformations can be characterized solely by their property of preserving the temporal order between events. More precisely, E.C. Zeeman has in 1964 proved the following theorem:

   Let $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ be an invertible *causal* mapping, i.e., for any two events $x, y \in \mathbb{R}^4$, $f(y)$ is timelike separated and after $f(x)$ if and only if y is timelike separated and after x. Then f is of the form

   $$f(x) \;=\; \lambda \varLambda x \,+\, a \;,$$

   where $\lambda > 0$, $\varLambda \in L^\uparrow$ and $a \in \mathbb{R}^4$. Observe that no additional assumptions on f (such as differentiability or even continuity) are needed: all these properties already follow from the hypothesis of causality.

7.4 Behavior under Lorentz Transformations

7.4.1 Time Dilatation

Consider a clock at rest in an inertial system I. Two subsequent time beats of the clock correspond to events represented by four-vectors x and y with timelike difference vector $\Delta x = y - x$. In I, we have $x = (ct, \boldsymbol{x})$, $y = (c(t + \Delta t), \boldsymbol{x})$ and $\Delta x = (c\Delta t, 0)$ where $c\Delta t = \sqrt{(\Delta x)^2}$. By applying a boost with velocity $-\boldsymbol{v}$, we pass to an inertial system I' in which the clock moves with velocity $\boldsymbol{v} = c\beta\boldsymbol{n}$ and find

$$\Delta t' \;=\; \Delta t \, \frac{1}{\sqrt{1 - v^2/c^2}} \;. \tag{7.80}$$

The effect, called relativistic *time dilatation*, is that moving clocks run more slowly, since $\Delta t \leq \Delta t'$. The shortest possible time difference is the time difference Δt measured in the inertial system I in which the clock is at rest and is called the *proper time difference*.

The effect of time dilatation can be verified experimentally, e.g., by the increased lifetime of rapidly moving unstable particles, such as pions and muons in cosmic radiation or the great majority of particles produced in accelerator experiments.

It is also the reason why one has to be careful when synchronizing clocks by physical transport from one place to another. In a given inertial system, moving a clock by a distance L with velocity v requires the time $T = L/v$. According to the clock itself, however, the time elapsed is only

$$T_0 \;=\; (L/v)\,\sqrt{1-v^2/c^2}\;.$$

For $v \ll c$, the difference

$$\Delta T \;=\; T - T_0 \;=\; (L/v)\,(1-\sqrt{1-v^2/c^2})$$

is to first order given by

$$\Delta T \;=\; \frac{Lv}{c^2}\;.$$

Thus as announced before, the difference can be made arbitrarily small by sufficiently slow transport.

7.4.2 Lorentz Contraction, Relativity of Simultaneity

Consider a measuring rod at rest in an inertial system I and oriented along some given unit vector $\boldsymbol{e}$. At a given time t, the two ends of the measuring rod correspond to events represented by four-vectors x and y with spacelike difference vector $\Delta x = y - x$. In I, we have $x = (ct, \boldsymbol{x})$, $y = (ct, \boldsymbol{y})$ and $\Delta x = (0, \Delta l\,\boldsymbol{e})$ where $\Delta l = \sqrt{-(\Delta x)^2}$. By applying a boost with velocity $-\boldsymbol{v}$, we pass to an inertial system I' in which the measuring rod moves with velocity $\boldsymbol{v} = c\beta\boldsymbol{n}$ and find

$$\Delta t' \;=\; \frac{\beta\,\boldsymbol{n}\cdot\boldsymbol{e}}{c}\,\frac{\Delta l}{\sqrt{1-v^2/c^2}}\;. \tag{7.81}$$

In particular, we see that $\Delta t' \neq 0$ but $\Delta t = 0$ (relativity of simultaneity). Moreover,

$$\Delta\boldsymbol{x}'_{\|} \;=\; \frac{\Delta l}{\sqrt{1-v^2/c^2}}\,\boldsymbol{e}_{\|} \quad , \quad \Delta\boldsymbol{x}'_{\perp} \;=\; \Delta l\;\boldsymbol{e}_{\perp}\;.$$

However, because of $\Delta t' \neq 0$, the length of the measuring rod in I' is *not* equal to the absolute value of $\Delta\boldsymbol{x}'$ but rather to the spatial distance between the world points of the rod's beginning and end point when these are simultaneous in I':

$$\Delta l' \;=\; |\boldsymbol{y}' - \boldsymbol{x}'| \qquad \text{when} \qquad y'^0 \;=\; x'^0\;.$$

Now we have

$$\begin{aligned} x'^0 \;&=\; \gamma\,(x^0 + \beta\,\boldsymbol{n}\cdot\boldsymbol{x}) \quad &,\quad y'^0 \;&=\; \gamma\,(y^0 + \beta\,\boldsymbol{n}\cdot\boldsymbol{y}) \\ \boldsymbol{x}'_{\|} \;&=\; \gamma\,(\boldsymbol{x}_{\|} + \beta\boldsymbol{n}x^0) \quad &,\quad \boldsymbol{y}'_{\|} \;&=\; \gamma\,(\boldsymbol{y}_{\|} + \beta\boldsymbol{n}y^0) \\ \boldsymbol{x}'_{\perp} \;&=\; \boldsymbol{x}_{\perp} \quad &,\quad \boldsymbol{y}'_{\perp} \;&=\; \boldsymbol{y}_{\perp} \end{aligned}$$

and therefore

$$\begin{aligned} y'^0 \;=\; x'^0 \qquad &\Longrightarrow \qquad y^0 - x^0 \;=\; -\,\beta\,\boldsymbol{n}\cdot(\boldsymbol{y}-\boldsymbol{x}) \;=\; -\,\beta\,\boldsymbol{n}\cdot\boldsymbol{e}\,\Delta l\;, \\ &\Longrightarrow \qquad \boldsymbol{y}'_{\|} - \boldsymbol{x}'_{\|} \;=\; \gamma\Big((\boldsymbol{y}_{\|} - \boldsymbol{x}_{\|}) + \beta\boldsymbol{n}\,(y^0 - x^0)\Big) \\ &\qquad\qquad\qquad\qquad\;=\; \gamma\,(1-\beta^2)\,\Delta l\;\,(\boldsymbol{n}\cdot\boldsymbol{e})\,\boldsymbol{n}\;, \end{aligned}$$

so that

$$\Delta l' \;=\; \Delta l\,\sqrt{1-v^2/c^2} \quad \text{when} \quad \boldsymbol{n}\parallel\boldsymbol{e} \quad , \quad \Delta l' \;=\; \Delta l \quad \text{when} \quad \boldsymbol{n}\perp\boldsymbol{e}\;. \tag{7.82}$$

The effect, called *Lorentz contraction*, is that moving bodies are contracted in the direction of motion, since $\Delta l' \leq \Delta l$, whereas there is no contraction orthogonal to the direction of motion; thus volumes are also reduced by the same factor:

$$V' \;=\; V\sqrt{1-v^2/c^2}\;. \tag{7.83}$$

The greatest possible length of a measuring rod is the length Δl measured in the inertial system I in which it is at rest and is called its *proper length*.

The effect of Lorentz contraction differs from that of time dilatation in that it cannot be observed directly. One reason is that the image produced by a rapidly flying object on a detector, such as the retina of an observer or a photographic plate, is determined by the distribution of photons emitted or reflected by the object at different times so that they may arrive at the detector at a given instant, that is, simultaneously. A detailed analysis shows that the resulting image is not contracted in one direction (flattened) but rather rotated by an angle $\varphi = \arctan(v/c)$.

7.4.3 Addition Theorem of Velocities

Consider a body moving uniformly in a straight line with respect to an inertial system I, with velocity $\boldsymbol{v}$; it will then also be moving uniformly in a straight line with respect to any other inertial system I', with velocity $\boldsymbol{v}'$. Assuming that I' moves with velocity $-\boldsymbol{w}$ as measured from I, we may divide the spatial components of eq. (7.49) by its time component to obtain

$$\boldsymbol{v}'_\parallel \;=\; \frac{\boldsymbol{v}_\parallel + \boldsymbol{w}}{1+\boldsymbol{v}\cdot\boldsymbol{w}/c^2} \quad , \quad \boldsymbol{v}'_\perp \;=\; \boldsymbol{v}_\perp/\gamma\;. \tag{7.84}$$

In particular, when $\boldsymbol{v}\parallel\boldsymbol{w}$, we recover the previously derived relativistic addition theorem of velocities, eq. (7.55). Again, the speed of light c appears as an upper bound for velocities.

7.4.4 Doppler Effect and Aberration of Light

Consider a plane wave

$$u(t,\boldsymbol{x}) \;=\; u_0\,\exp\left(-\,i(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})\right) \;=\; u_0\,\exp\left(-\,i\varphi(t,\boldsymbol{x})\right)\;.$$

Putting

$$k \;=\; \begin{pmatrix}\omega/c\\ \boldsymbol{k}\end{pmatrix} \in \mathbb{R}^4\;, \tag{7.85}$$

we can write its phase as $\varphi(t,\boldsymbol{x}) = k\cdot x$. This phase cannot depend on the reference frame, which is possible if and only if under Lorentz transformations, k transforms in the same way as x, or in other words, if k is a four-vector. Thus under a boost $\Lambda(\boldsymbol{n},\theta) = B(\boldsymbol{v})$, k transforms according to

$$\begin{aligned} k'^0 &= \gamma\,(k^0 - \beta\,\boldsymbol{n}\cdot\boldsymbol{k}) \;, \\ \boldsymbol{k}'_\parallel &= \gamma\,(\boldsymbol{k}_\parallel - \beta\boldsymbol{n}k^0) \;, \\ \boldsymbol{k}'_\perp &= \boldsymbol{k}_\perp \;, \end{aligned} \tag{7.86}$$

with $k^0 = \omega/c$, $k'^0 = \omega'/c$ (cf. eq. (7.49)). In particular, for light waves, we have $\omega^2/c^2 = \boldsymbol{k}^2$, i.e., $k^2 = 0$: k is lightlike. In this case, the first equation in (7.86) becomes

$$\omega' = \gamma\,\omega\,(1 - \beta\cos\theta) \;, \tag{7.87}$$

where θ denotes the angle between $\boldsymbol{n}$ and $\boldsymbol{k}$: this frequency shift is the relativistic *Doppler effect*. For $\beta \ll 1$, one finds the classical Doppler effect, which agrees with the one that would result from the ether hypothesis. In addition, there is a transversal Doppler effect, occuring even when $\theta = \pi/2$; it is however of order β^2. Finally, using $\boldsymbol{k}'_\perp = \boldsymbol{k}_\perp$ and $|\boldsymbol{k}_\perp| = |\boldsymbol{k}|\sin\theta$, $|\boldsymbol{k}'_\perp| = |\boldsymbol{k}'|\sin\theta'$, we get

$$\sin\theta' = \frac{|\boldsymbol{k}|}{|\boldsymbol{k}'|}\sin\theta = \frac{\omega}{\omega'}\sin\theta \;,$$

so using the previous equation,

$$\sin\theta' = \sqrt{1-\beta^2}\,\frac{\sin\theta}{1 - \beta\cos\theta} \;. \tag{7.88}$$

Thus the apparent direction of a light ray depends on the reference frame: this is the relativistic *aberration* of light. For $\beta \ll 1$, one finds the classical aberration effect, which agrees with the one that would result from the ether hypothesis; it is observed as as shift of the position of fixed stars due to the orbital motion of the earth around the sun.

7.5 Relativistic Cinematics of a Point Particle

In non-relativistic mechanics, the motion of a point particle is described by its trajectory $t \mapsto \boldsymbol{x}(t)$ which gives its position as a function of time. A relativistic theory requires an analogous treatment which must however be more symmetric in t and $\boldsymbol{x}$. To this end, one introduces a new parameter σ and describes the motion in terms of a *world line* $\sigma \mapsto x(\sigma)$ where

$$x(\sigma) = \begin{pmatrix} x^0(\sigma) \\ \boldsymbol{x}(\sigma) \end{pmatrix} \;. \tag{7.89}$$

The parameter σ is arbitrary and has no physical significance, but it will be required to increase monotonically with t:

$$\frac{dx^0}{d\sigma} > 0 \;. \tag{7.90}$$

In addition, physically realizable motions will satisfy

$$\left|\frac{d\boldsymbol{x}}{d\sigma}\right| \Big/ \frac{dx^0}{d\sigma} \;=\; \frac{1}{c}\left|\frac{d\boldsymbol{x}}{dt}\right| \;\leq\; 1 \;,$$

i.e.,

$$\left(\frac{dx}{d\sigma}\right)^2 \;\geq\; 0 \;. \tag{7.91}$$

This condition is Lorentz invariant and does not depend on the parametrization. In a space-time diagram, the world line will be the graph of a function over the time axis (with respect to any inertial system) and its inclination will never be less than 1 (see Fig. 7.3).

Fig. 7.3: Space-time diagram with world line of a point particle (two space dimensions are suppressed)

The Lorentz invariant quantity

$$\tau_{12} \;=\; \frac{1}{c}\int_{\sigma_1}^{\sigma_2} d\sigma\,\sqrt{\left(\frac{dx}{d\sigma}\right)^2} \;=\; \int_{t_1}^{t_2} dt\,\sqrt{1-\frac{v^2(t)}{c^2}} \tag{7.92}$$

does not depend on the parametrization and is the *proper time* measured by a comoving observer, using a set of clocks each of which is in uniform rectilinear motion and momentarily at rest relative to the particle. Such a set of clocks can be considered as the idealization of a real clock carried along the trajectory and insensitive to acceleration. More realistically, one can measure the acceleration along the trajectory and use the result to correct for the deviation in the clock's pace, as compared to a non-accelerated clock of the same design. For example, a quartz clock is to a good approximation insensitive to acceleration, in contrast to a pendulum clock. In the same sense, the "biological clock" of a comoving observer is insensitive to acceleration, at least when the acceleration is not too great. Therefore, τ_{12} is also the time interval measured and experienced by a comoving observer.

Proper time τ is a particulary natural and preferred Lorentz invariant parameter for *timelike world lines*, i.e., world lines $\sigma \mapsto x(\sigma)$ satisfying

$$\left(\frac{dx}{d\sigma}\right)^2 \; > \; 0 \; . \tag{7.93}$$

They are the world lines along which the speed of light is never reached. Obviously,

$$\frac{d\tau}{d\sigma} \;=\; \frac{1}{c}\,\sqrt{\left(\frac{dx}{d\sigma}\right)^2} \;=\; \frac{dt}{d\sigma}\,\sqrt{1-\frac{v^2(t)}{c^2}} \; . \tag{7.94}$$

Moreover, eq. (7.92) leads to the inequality

$$\tau_{12} \;\leq\; t_2 - t_1 \; , \tag{7.95}$$

i.e., proper time differences are always smaller than or at most equal to coordinate time differences, as measured in an arbitrary inertial system.

A classical thought experiment will help to further substantiate this result. Imagine two twin brothers, one of whom remains at home in an inertial system while the other embarks on a spaceship to undertake a high speed round trip; then when he gets back and both meet again, the travelling twin will be younger than the twin who has stayed back at home. To clarify this assertion, we show in Fig. 7.4 the world lines of the two brothers, as viewed from the inertial system of the one who has stayed back at home.

Fig. 7.4: Twin experiment: world line of twin A staying at home and general world line of travelling twin B, as viewed from twin A's rest frame (two space dimensions are suppressed)

This clear prediction of special relativity is very often questioned in the name of "common sense" – so much that the situation has come to be referred to as the *twin paradox*. Briefly, the main objection raised and used to classify the result as paradoxical goes as follows.

Special relativity states that a clock runs most slowly in its rest frame. But of course, just as twin B move in relation to twin A, so does twin A move in relation to twin B. Therefore, it is argued, just as twin A sees twin B age more slowly, so should twin B see twin A age more slowly, and this is paradoxical: the only way out would be that both twins have aged by the same amount.

Of course, this argument is fallacious because it is based on the tacit assumption that the situation of the twins is symmetrical, which is not the case. Indeed, the world line of A is straight, whereas the world line of B must be curved, at least somewhere, if B and A are to meet again. This difference holds in *any* inertial system. It can be observed physically, for example, by noting that B, in contrast to A, will undergo acceleration, at least somewhere along his journey. One might therefore be tempted to consider this acceleration as being the cause of the age difference between A and B, shifting the effect into the realm of general relativity.

As we shall see, however, this view is not correct either. Indeed, the world line of B can be arranged in such a way that B is almost always in uniform rectilinear motion. Fig. 7.5 shows such a situation, in which the world line of B is piecewise straight; acceleration is restricted to short (i.e., in principle, arbitrarily short) intervals, in order to change the direction of the velocity of B. It would be very strange if the entire age difference were accumulated during the short periods of acceleration. Moreover, even if these are supposed to be spread out over longer intervals, in order to avoid physically unbearable accelerations, the possibility of considering the acceleration phase as the cause of the age difference between A and B can be ruled out by comparing two possible journeys of B, B_1 and B_2, carried out with identical periods of acceleration but different periods of uniform rectilinear motion, since the age differences accumulated in the course of these two journeys – that between A and B_1 and that between A and B_2 – are *not* equal. (Otherwise, one would be led to the conclusion that one and the same cause can have two different effects, and in fact arbitrarily many different effects, which is absurd.)

Sceptics use the limiting case of a piecewise straight world line to argue that the situation of the twins should be symmetrical after all and that there should thus be no age difference, since the period of acceleration, being arbitrarily short, should have no effect.

Of course, special relativity does predict an age difference, even for piecewise straight world lines. In fact, the effect is in this case particularly easy to compute. Assuming that the velocities of B on the two straight pieces of B's world line in Fig. 7.5, as measured in A's rest frame, are $\pm v$, we get

$$\tau_{12} \;=\; \sqrt{1 - v^2/c^2}\,(t_2 - t_1) \;<\; t_2 - t_1 \;.$$

The world lines of A and B are fundamentally different: that of A is straight whereas that of B is only piecewise straight. *The twin effect is not of dynamical but of purely geometrical nature.*

In a triangle with three positive timelike sides x, y and $z = x + y$, the side z is, measured in the Lorentz metric η, *longer* than the two other sides x and y taken together. Indeed,

$$z^2 \;=\; x^2 + 2x{\cdot}y + y^2 \;.$$

Fig. 7.5: Twin experiment: world line of twin A staying at home and piecewise straight world line of travelling twin B, as viewed from twin A's rest frame (two space dimensions are suppressed)

The quantity $2x{\cdot}y$ is best evaluated in the rest frame of x:

$$2x{\cdot}y \;=\; 2x^0y^0 \;=\; 2\,\sqrt{x^2}\,\sqrt{y^2+\boldsymbol{y}^2} \;\geq\; 2\,\sqrt{x^2}\,\sqrt{y^2}\;.$$

Thus

$$z^2 \;\geq\; x^2+2\,\sqrt{x^2}\,\sqrt{y^2}+y^2 \;=\; \left(\sqrt{x^2}+\sqrt{y^2}\right)^2\;,$$

i.e.,

$$\sqrt{z^2} \;\geq\; \sqrt{x^2}+\sqrt{y^2}\;. \tag{7.96}$$

The twin effect is based on the geometry of timelike triangles in the Lorentz metric. It is geometric in the sense that a "detour" in space-time leads to a shorter proper time. In this respect, the Lorentz metric differs from the Euclidean metric, for which any two sides of a triangle taken together are longer than the third. It is true that the world line of the travelling twin cannot be realized without acceleration, but it is inappropriate to consider this acceleration as the cause of the age difference between the two, just as it would be absurd to consider the bends in the corners of a triangle as the cause of the fact that the triangle is closed or that any of its two sides taken together are longer than the third.

We want to study more closely how the age difference does come about by analyzing a simple example (see Fig. 7.6). Assume that A is at rest in an inertial system I whereas B, as seen from I, moves away from A with velocity v during the time interval $0 \leq t \leq T/2$ and then moves back towards A with velocity $-v$ during the time interval $T/2 \leq t \leq T$, so that both meet again at time $t = T$. Both of them agree to emit light signals with frequency ν_0, in order to keep the other informed about their aging process. We want to calculate at what time B receives the light signals emitted by A and at what time A receives the light signals emitted by B.

To this end, we first write down the world lines of A and B, parametrized by the respective proper time (which for A is simply t, the time coordinate in I, but for B is $\tau = t/\gamma + \text{const.}$), as well as the world lines of the light signals emitted by A and by B.

Fig. 7.6: Twin experiment: the origin of the age difference; see text for more details (two space dimensions are suppressed)

World line of A: $x_{\mathrm{A}}(t) = (ct, 0)$

World line of B: $$x_{\mathrm{B}}(\tau) = \begin{cases} \gamma\tau(c, v) & \text{for} \quad 0 \le \tau \le T/2\gamma \\ T(0, v) + \gamma\tau(c, -v) & \text{for} \quad T/2\gamma \le \tau \le T/\gamma \end{cases}$$

World lines of signals of A: $\xi_n(\sigma_1) = x_{\mathrm{A}}(t_n) + \sigma_1(c, +c)\ ; \quad t_n = n/\nu_0$

World lines of signals of B: $\eta_n(\sigma_2) = x_{\mathrm{B}}(\tau_n) + \sigma_2(c, -c)\ ; \quad \tau_n = n/\nu_0$

The arrival times of the signals are simply found from the intersection points of the world lines. Putting $\beta = v/c$, we find

Arrival of signals from A at B:

$$\tilde{\tau}_n = \begin{cases} \dfrac{n}{\nu_0}\sqrt{\dfrac{1+\beta}{1-\beta}} & \text{for} \quad 0 \le \tilde{\tau}_n \le T/2\gamma \\ \left(\dfrac{n}{\nu_0} + T\beta\right)\sqrt{\dfrac{1-\beta}{1+\beta}} & \text{for} \quad T/2\gamma \le \tilde{\tau}_n \le T/\gamma \end{cases}$$

Arrival of signals from B at A:

$$\tilde{t}_n = \begin{cases} \dfrac{n}{\nu_0}\sqrt{\dfrac{1+\beta}{1-\beta}} & \text{for} \quad 0 \le n/\nu_0 \le T/2\gamma \\ & \text{i.e.} \quad 0 \le \tilde{t}_n \le (T/2)(1+\beta) \\ \dfrac{n}{\nu_0}\sqrt{\dfrac{1-\beta}{1+\beta}} + T\beta & \text{for} \quad T/2\gamma \le n/\nu_0 \le T/\gamma \\ & \text{i.e.} \quad (T/2)(1+\beta) \le \tilde{t}_n \le T \end{cases}$$

From this result, we may conclude the following.

a) Both A and B distinguish two phases: in the first phase they see signals coming from the receding partner, in the second phase they see signals coming from the approaching partner.

b) Both for A and for B, the arrival frequency of the signals, as compared to the emission frequency ν_0, is reduced to

$$\nu_1 = \nu_0 \sqrt{\frac{1-\beta}{1+\beta}}$$

during the first phase and enhanced to

$$\nu_2 = \nu_0 \sqrt{\frac{1+\beta}{1-\beta}}$$

during the second phase, by factors

$$\sqrt{\frac{1-\beta}{1+\beta}} \qquad \text{and} \qquad \sqrt{\frac{1+\beta}{1-\beta}}\,,$$

respectively, which are precisely the relative frequency shifts of the Doppler effect. They are the same for A and B, so in this respect, we do have symmetry between A and B. If the world line of B were straight, there would be only the first phase, and the symmetry between A and B would be complete.

c) An asymmetry of geometric nature arises from the fact that for B, both phases last for exactly half of the total proper time of the journey, whereas for A, the first phase corresponds to the proper time interval $0 \le t \le (T/2)(1+\beta)$ and the second to the proper time interval $(T/2)(1+\beta) \le t \le T$: for A, phase 1 lasts longer than phase 2.

d) The total number of signals received is

for B:

$$\begin{aligned} N_\mathrm{B} &= \frac{T\nu_0}{2\gamma}\sqrt{\frac{1-\beta}{1+\beta}} + \frac{T\nu_0}{2\gamma}\sqrt{\frac{1+\beta}{1-\beta}} \\ &= \frac{T\nu_0}{2}\,(1-\beta+1+\beta) = T\nu_0\,. \end{aligned}$$

for A:

$$N_A = \frac{T\nu_0}{2}(1+\beta)\sqrt{\frac{1-\beta}{1+\beta}} + \frac{T\nu_0}{2}(1-\beta)\sqrt{\frac{1+\beta}{1-\beta}}$$
$$= T\nu_0\sqrt{1-\beta^2}\ .$$

Thus,

$$N_A = \sqrt{1-\beta^2}\, N_B < N_B\ ,$$

that is, A receives less signals than B: B has aged less than A. The asymmetry arises from the fact that for A the phase of higher signal frequency (blue shift) is shorter than the phase of lower signal frequency (red shift). In other words, the age difference is accumulated during the period in which A and B are in different phases, rather than during the (arbitrarily short) period of acceleration.

The last doubts as to the real existence of the twin phenomenon should have been eliminated by the chronometer experiment of Hafele and Keating which consisted in transporting a high precision clock in a standard airplane around the world and measuring the time difference to a clock of the same design maintained at rest.

After this excursion, let us return to the subject of relativistic cinematics. The four-vector

$$u = \frac{dx}{d\tau} \tag{7.97}$$

is called the *four-velocity*; it satisfies

$$u = \frac{dx}{dt}\frac{dt}{d\tau} = \gamma\begin{pmatrix} c \\ \boldsymbol{v}\end{pmatrix}\ , \tag{7.98}$$

where as before $\boldsymbol{v} = d\boldsymbol{x}/dt$ is the ordinary velocity. Obviously,

$$u^2 = \left(\frac{dx}{d\tau}\right)^2 = c^2\ . \tag{7.99}$$

Next, we define the *four-acceleration*

$$a = \frac{du}{d\tau} = \frac{d^2x}{d\tau^2} \tag{7.100}$$

and obtain

$$a = \frac{du}{dt}\frac{dt}{d\tau} = \gamma^2\begin{pmatrix} 0 \\ \boldsymbol{a}\end{pmatrix} + \frac{\gamma^4\,\boldsymbol{v}\cdot\boldsymbol{a}}{c^2}\begin{pmatrix} c \\ \boldsymbol{v}\end{pmatrix}\ , \tag{7.101}$$

where $\boldsymbol{a} = d\boldsymbol{v}/dt = d^2\boldsymbol{x}/dt^2$ is the ordinary acceleration.[1]. Due to eq. (7.99), we have

$$2u\cdot\frac{du}{d\tau} = \frac{d}{d\tau}(u^2) = 0\ ,$$

[1] Thus in the case of the four-acceleration, we deviate from our standard convention, in that the three-dimensional vector formed by the spatial components of the four-acceleration a is not the ordinary acceleration $\boldsymbol{a}$; according to eq. (7.101) and eq. (7.103), this is only true in the rest frame of the particle.

i.e.,

$$u \cdot a \;=\; 0 \;. \tag{7.102}$$

This implies that a must be a spacelike four-vector, since u is a timelike four-vector. In the momentary rest frame of the particle, we have

$$u \;=\; \begin{pmatrix} c \\ 0 \end{pmatrix} \quad , \quad a \;=\; \begin{pmatrix} 0 \\ \boldsymbol{a} \end{pmatrix} . \tag{7.103}$$

As a simple example of a problem of relativistic cinematics, we discuss the case of uniformly accelerated motion, restricting ourselves for the sake of simplicity to linear motion in one space dimension. (Thus during the remainder of this section, $x(\tau)$ and $v(\tau)$ will represent the non-trivial spatial component of the four-vector elsewhere denoted by $x(\tau)$ and $u(\tau)$, respectively.) This kind of motion can be characterized, independently of the inertial system employed, by the Lorentz invariant condition

$$\left(\frac{du}{d\tau}\right)^2 \;=\; -\,a^2 \;=\; \text{const.}\;. \tag{7.104}$$

As we shall see, it is particularly convenient to use the rapidity θ of the momentary rest frame of the particle, relative to the rest frame of the particle at the beginning of its motion, as a function of the proper time τ. Due to

$$\beta \;=\; \tanh\theta \quad , \quad \gamma \;=\; \cosh\theta \quad , \quad \beta\gamma \;=\; \sinh\theta$$

(cf. eq. (7.42) and eq. (7.44)), we have

$$u \;=\; c\begin{pmatrix} \cosh\theta \\ \sinh\theta \end{pmatrix} \quad , \quad a \;=\; c\,\frac{d\theta}{d\tau}\begin{pmatrix} \sinh\theta \\ \cosh\theta \end{pmatrix}$$

and thus

$$a^2 \;=\; -\,c^2\left(\frac{d\theta}{d\tau}\right)^2 \;.$$

Comparing with eq. (7.104), this yields

$$\frac{d\theta}{d\tau} \;=\; \frac{a}{c}$$

with the solution

$$\theta(\tau) \;=\; a\tau/c \tag{7.105}$$

corresponding to the initial condition $\theta(0) = 0$. Thus we get

$$\beta(\tau) \;=\; \tanh\left(a\tau/c\right) \;, \tag{7.106}$$

$$\frac{dt}{d\tau} \;=\; \gamma(\tau) \;=\; \cosh\left(a\tau/c\right) \;, \tag{7.107}$$

$$\frac{1}{c}\frac{dx}{d\tau} \;=\; \beta(\tau)\,\gamma(\tau) \;=\; \sinh\left(a\tau/c\right) \;, \tag{7.108}$$

and by integration in τ with initial conditions $t(0) = 0$, $x(0) = x_0$,

$$t(\tau) \;=\; \frac{c}{a}\,\sinh\left(a\tau/c\right)\;, \tag{7.109}$$

$$x(\tau) \;=\; \frac{c^2}{a}\,\cosh\left(a\tau/c\right)\,+\,x_0\;. \tag{7.110}$$

Expressing τ in terms of t, we finally arrive at

$$x(t) \;=\; \frac{c^2}{a}\,\sqrt{1+(at/c)^2}\,+\,x_0\;, \tag{7.111}$$

$$v(t) \;=\; \frac{at}{\sqrt{1+(at/c)^2}}\;. \tag{7.112}$$

In particular, we see that $t \to \infty$ is equivalent to $\tau \to \infty$ and that for large times

$$x(t) \;\sim\; ct \quad,\quad v(t) \;\sim\; c \qquad \text{for} \quad t \to \infty\;, \tag{7.113}$$

whereas for small times

$$x(t) \;\sim\; \tfrac{1}{2}\,at^2\,+\,x_0 \quad,\quad v(t) \;\sim\; at \qquad \text{for} \quad at \ll c\;, \tag{7.114}$$

which is the Newtonian limit.

7.6 Covariant Formalism

In the course of our discussion of the theory of special relativity, we have so far identified the proper orthochronous Poincaré transformations as the admissible coordinate transformations between Lorentz systems (that is, inertial systems with natural coordinates), have studied the structure of these transformations and have analyzed some of the resulting physical consequences. It is now time to proceed to the core of Einstein's principle of relativity, which in its full generality states that all laws of physics must have the same form in all Lorentz systems, or expressed differently, must preserve their form under proper orthochronous Poincaré transformations. Mathematically, this means that all laws of physics must be written in relativistically *covariant* form, that is, as equations between quantities that transform in the same way under the proper orthochronous Lorentz group. In order to verify compatibility of a given physical theory with Einstein's principle of relativity, it is therefore necessary to first of all organize the physical quantities appearing in that theory into scalars, vectors or, more generally, tensors over Minkowski space $\mathbb{R}^4$; one also speaks of world or four-scalars, world or four-vectors and world or four-tensors, respectively. We have already encountered examples of this procedure in the previous section, where in the course of our discussion of the relativistic cinematics of point particles we introduced proper time τ as a world scalar, as well as the position four-vector x, the four-velocity u and the four-acceleration a as world vectors. Other important examples can be constructed from differential operators, such as the ordinary differential operator $d/d\tau$, a world scalar, the partial diffential operator

$$\partial \;=\; (\partial_0\,,\boldsymbol{\nabla}) \;=\; \left(\frac{1}{c}\frac{\partial}{\partial t}\,,\boldsymbol{\nabla}\right)~, \tag{7.115}$$

a world covector, and its square, the wave operator or d'Alembertian

$$\Box \;\equiv\; \partial^2 \;=\; \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta~, \tag{7.116}$$

a world scalar.

The natural formalism for handling relativistically covariant theories is thus the tensor algebra and tensor analysis over Minkowski space $\mathbb{R}^4$. Its geometry is determined by the scalar product η, and the Lorentz transformations are simply the isometries of this four-dimensional pseudo-Euclidean vector space. In contrast to the tensor calculus on standard three-dimensional Euclidean space, however, we must now distinguish clearly between vectors and covectors and, more generally, between contravariant and covariant tensors. In the so-called index calculus that we shall employ in what follows, the relation between contravariant and covariant components is established by pulling indices up and down with the help of the scalar product η: numerically, this leaves the time components unchanged whereas the spatial components switch sign. On the other hand, we shall continue to follow the convention that upper and lower indices for components of three-dimensional vectors need not be distinguished and shall therefore adopt the rule that the components of three-vectors are to be identified with the corresponding components of four-vectors (upper indices) and not of four-covectors (lower indices), with one notable exception: the partial differential operator ∂ introduced above. Finally, we agree to write both four-vectors and four-covectors as rows, rather than columns. Thus, for example,

$$\begin{aligned} x^\mu &= (ct,+\boldsymbol{x})~, \\ x_\mu &= (ct,-\boldsymbol{x})~, \end{aligned} \tag{7.117}$$

$$\begin{aligned} u^\mu &= \gamma\,(c,+\boldsymbol{v})~, \\ u_\mu &= \gamma\,(c,-\boldsymbol{v})~, \end{aligned} \tag{7.118}$$

$$\begin{aligned} a^\mu &= \gamma^2\left(\frac{\gamma^2\,\boldsymbol{v}\cdot\boldsymbol{a}}{c}\,,+\boldsymbol{a}+\frac{\gamma^2\,\boldsymbol{v}\cdot\boldsymbol{a}}{c^2}\,\boldsymbol{v}\right)~, \\ a_\mu &= \gamma^2\left(\frac{\gamma^2\,\boldsymbol{v}\cdot\boldsymbol{a}}{c}\,,-\boldsymbol{a}-\frac{\gamma^2\,\boldsymbol{v}\cdot\boldsymbol{a}}{c^2}\,\boldsymbol{v}\right)~, \end{aligned} \tag{7.119}$$

but

$$\begin{aligned} \partial^\mu &= \left(\frac{1}{c}\frac{\partial}{\partial t}\,,-\boldsymbol{\nabla}\right)~, \\ \partial_\mu &= \left(\frac{1}{c}\frac{\partial}{\partial t}\,,+\boldsymbol{\nabla}\right)~. \end{aligned} \tag{7.120}$$

7.7 Relativistic Dynamics of a Point Particle

According to the postulate of relativistic covariance formulated in the previous section, we begin our approach to the formulation of the relativistic mechanics of a point particle by introducing two more four-vectors: its *four-momentum* p which is simply proportional to its four-velocity u and the *four-force* F which is simply proportional to the four-acceleration a that it suffers:

$$p^\mu = mu^\mu \ , \tag{7.121}$$

$$F^\mu = ma^\mu \ . \tag{7.122}$$

Then eq. (7.99) and eq. (7.102) imply

$$p_\mu p^\mu \equiv p^\mu p_\mu \equiv p^2 = (mc)^2 \ , \tag{7.123}$$

and

$$p_\mu F^\mu \equiv p^\mu F_\mu \equiv p \cdot F = 0 \ , \tag{7.124}$$

respectively. Here, m is a positive real constant (that is, a four-scalar) called the *rest mass* of the particle.

For the interpretation of eqns (7.121)–(7.124), we use eqns (7.118) and (7.119) to write

$$p^\mu = m\gamma\,(c, \boldsymbol{v}) \ , \tag{7.125}$$

$$F^\mu = m\gamma^2 \left(\frac{\gamma^2\, \boldsymbol{v}\cdot\boldsymbol{a}}{c} \, , \, \boldsymbol{a} + \frac{\gamma^2\, \boldsymbol{v}\cdot\boldsymbol{a}}{c^2}\, \boldsymbol{v} \right) \ . \tag{7.126}$$

Expanding in powers of $\beta = v/c$, we obtain to lowest non-trivial order

$$E = cp^0 = \frac{mc^2}{\sqrt{1 - v^2/c^2}} = mc^2 + \tfrac{1}{2}mv^2 + \ldots \ , \tag{7.127}$$

$$\boldsymbol{p} = \frac{m\boldsymbol{v}}{\sqrt{1 - v^2/c^2}} = m\boldsymbol{v} + \ldots \ , \tag{7.128}$$

$$cF^0 = \frac{\boldsymbol{v} \cdot m\boldsymbol{a}}{(1 - v^2/c^2)^2} = \boldsymbol{v} \cdot m\boldsymbol{a} + \ldots \ , \tag{7.129}$$

$$\boldsymbol{F} = \frac{m\boldsymbol{a}}{1 - v^2/c^2} + \frac{m\,(\boldsymbol{v}\cdot\boldsymbol{a})\,\boldsymbol{v}}{(1 - v^2/c^2)^2} = m\boldsymbol{a} + \ldots \ . \tag{7.130}$$

Comparing with Newtonian mechanics, we arrive directly at the desired physical interpretation of four-momentum and four-force. Their spatial components are the relativistic generalizations of the ordinary momentum $\boldsymbol{p}_N = m\boldsymbol{v}$ and of the ordinary force $\boldsymbol{F}_N = m\boldsymbol{a}$, respectively, whereas their time components (after multiplication by c) provide the relativistic generalizations of kinetic energy $E_N = \frac{1}{2}mv^2$ and of power $L_N = \boldsymbol{v}\cdot\boldsymbol{F}_N$. (The index N is to indicate that these are the well-known, non-relativistic expressions of Newtonian mechanics.) Thus eq. (7.122) is to be viewed as the *relativistic equation of motion of a point particle*; it extends and unites in a relativistically covariant fashion the Newtonian equation of motion and the resulting energy balance. From eq. (7.130) we also see that the constant m is really the rest mass of the particle, that is, the coefficient that describes its inertia in its momentary rest frame. Indeed, the classical relation $\boldsymbol{p} = m\boldsymbol{v}$ does hold, at least approximately, for small velocities. The exact relation between momentum and velocity, however, reads $\boldsymbol{p} = m\gamma\boldsymbol{v}$. This means that the inertial mass of a particle – which is the coefficient between the force acting on it and the resulting acceleration – is enhanced by a factor of γ when viewed from a moving reference frame rather than the particles's momentary rest frame. This phenomenon is known as the *relativistic enhancement of mass* and has been verified experimentally with high precision; it must in particular be taken into account in the construction of particle accelerators. From a theoretical point of view, it is however not adequate to call the factor $m\gamma$ a mass since although it is a scalar, it is not a world scalar: only the rest mass has a Lorentz invariant meaning.

As has already been mentioned and is also indicated by the notation employed in eq. (7.127), the time component of four-momentum (multiplied by a factor of c) must in relativistic mechanics be interpreted as the total energy:

$$E \;=\; cp^0 \;. \tag{7.131}$$

In particular, it contains a constant contribution, namely

$$E_0 \;=\; mc^2 \;. \tag{7.132}$$

It is proportional to the rest mass and can be called the *rest energy* of the particle. This proportionality between rest mass and rest energy is certainly the most famous prediction of special relativity. It results naturally from the necessity to reconcile the basic laws of mechanics with Einstein's principle of relativity. However, the derivation given here, although plausible, is by no means compelling, since the total energy E could differ from the time component p^0 of four-momentum by some additive constant. For example, it would be conceivable to replace eq. (7.127) by

$$E \;=\; cp^0 - mc^2 \;=\; mc^2\left(\frac{1}{\sqrt{1-v^2/c^2}} - 1\right) \;=\; \tfrac{1}{2}mv^2 \;+\; \ldots \quad , \tag{7.133}$$

a relation that in the limit of small velocities would be equally compatible with Newtonian mechnics, in which the total energy is only determined up to an additive constant anyway. In order to rule out this possibility and to fix the additive constant to the value given by eq. (7.127), it is necessary to consider the conservation laws for total energy and total spatial momentum in systems of point particles, in the absence of external forces, for example in scattering processes: these conservation laws can only be unified into a single conservation law for the total four-momentum if the total energy is defined by eq. (7.131), i.e.,

$$E \;=\; \sqrt{m^2c^4 + c^2|\boldsymbol{p}|^2} \;, \tag{7.134}$$

without any additional constant.

It must be stressed that the conservation law for four-momentum in closed systems is valid for *all* scattering processes – elastic as well as inelastic. In particular, the *binding energy* of bound states (which may for example be formed by inelastic scattering) is reflected as a *mass defect*: the mass of a bound state is less than the sum of the masses of its constituents. This fact provides the basis for energy production in stars by nuclear fusion, so that our mere existence as human beings, based on billions of years of biological evolution for which the shining of our sun has furnished the source of energy, is proof enough for the correctness of Einstein's formula (7.132).

A particular but extremely important situation occurs when the rest mass m vanishes. The physics of such *massless particles* is completely beyond the reach of Newtonian mechanics. Their world lines are lightlike, and according to eq. (7.92), the proper time between any two events on a lightlike curve is always zero. Proper time is therefore inadequate for parametrizing world lines of massless particles, and there is also no other physically distinguished quantity that could serve as a Lorentz invariant parameter. This implies that for massless particles, the concepts of four-velocity and of four-acceleration are ill-defined, and in fact the only concept that does remain meaningful is that of four-momentum. In particular, the concept of four-force, being the derivative of four-momentum with respect to proper time, is also ill-defined, as there is no natural Lorentz invariant parameter that could be used to replace proper time in this relation. As a result, there is no such thing as an equation of motion for massless particles. Instead, the four-momentum of a massless particle can only be changed in jumps: the world lines of massless particles are piecewise straight, beginning at the event of emission, ending at the event of absorption and possibly with discontinuities in the first derivative at events of collisions with other particles – mostly massive. The behavior of the trajectory at such events is governed by the conservation law for the total four-momentum of the particles involved in the collision.

As an application we consider Compton scattering, that is, elastic scattering between a photon (γ) and an electron (e^-).

$$
\begin{array}{rccccccc}
\text{particles:} & \gamma & + & e^- & \longrightarrow & \gamma & + & e^- \\
\text{four-momenta:} & q & + & p & = & q' & + & p'
\end{array}
$$

Obviously, $q^2 = q'^2 = 0$ and $p^2 = p'^2 = (mc)^2$, where m is the rest mass of the electron. Denoting by θ the scattering angle of the photon, as measured in the rest frame of the incoming electron, we have in this inertial system

$$
p \;=\; (mc, 0) \qquad \text{und} \qquad \boldsymbol{q} \cdot \boldsymbol{q}' \;=\; |\boldsymbol{q}|\,|\boldsymbol{q}'| \cos\theta \;=\; q^0 q'^0 \cos\theta \;,
$$

so

$$
\begin{aligned}
q - q' + p \;=\; p' \;\; &\Longrightarrow \;\; \big((q - q') + p\big)^2 \;=\; p'^2 \;\;\Longrightarrow\;\; (q - q') \cdot p \;=\; q \cdot q' \\
&\Longrightarrow \;\; mc\left(q^0 - q'^0\right) \;=\; q^0 q'^0 \,(1 - \cos\theta) \\
&\Longrightarrow \;\; mc\left(\frac{1}{q^0} - \frac{1}{q'^0}\right) \;=\; 1 - \cos\theta \;.
\end{aligned}
$$

Using the relation $E = h\nu$ between the energy E of the photon and the frequency ν of the corresponding electromagnetic wave, borrowed from quantum mechanics, we arrive at the relation between frequency loss and scattering angle that is typical for Compton scattering:

$$\frac{mc^2}{h}\left(\frac{1}{\nu'} - \frac{1}{\nu}\right) \;=\; 1 - \cos\theta \; . \tag{7.135}$$

Another nice application of the conservation law for four-momentum is the dynamics of a rocket in the context of special relativity. For the sake of simplicity, we shall once again restrict ourselves to linear motion in one space dimension and use proper time τ in the rocket as the parameter. The rocket is propelled by the emission of gas, consisting of particles of rest mass m_0, ejected with velocity $v_e = c\beta_e$, the case $m_0 = 0$, $v_e = c$, $\beta_e = 1$ (photon rocket) being explicitly included. To begin with, we define

$$\begin{aligned} M(\tau) &= \text{remaining rest mass of the rocket at proper time } \tau \; , \\ u(\tau) &= \text{four-velocity of the rocket at proper time } \tau \; , \\ p(\tau) &= \text{four-momentum of the rocket at proper time } \tau \; , \end{aligned}$$

as well as

$$\begin{aligned} q_e(\tau) \quad &= \quad \begin{array}{l}\text{four-momentum of a single particle} \\ \text{ejected by the rocket at proper time } \tau,\end{array} \\ dn(\tau) \;=\; \nu(\tau)\, d\tau \quad &= \quad \begin{array}{l}\text{number of particles ejected by the rocket} \\ \text{at proper time } \tau \text{ in proper time interval } d\tau.\end{array} \end{aligned}$$

We also write

$$\begin{aligned} u(\tau) \;=\; c\,(\gamma(\tau)\,,\, \beta(\tau)\,\gamma(\tau)) \;&=\; c\,(\cosh\theta(\tau)\,,\, \sinh\theta(\tau)) \; , \\ p(\tau) \;&=\; M(\tau)\, u(\tau) \; . \end{aligned} \tag{7.136}$$

Then the conservation law for the four-momentum of the entire system consisting of the rocket and the ejected gas reads

$$dp(\tau) \;+\; dn(\tau)\, q_e(\tau) \;=\; 0 \; ,$$

i.e.,

$$\frac{dM}{d\tau}\, u \;+\; M\,\frac{du}{d\tau} \;=\; \frac{dp}{d\tau} \;=\; -\,\nu q_e \; . \tag{7.137}$$

Together with eq. (7.99) and eq. (7.102), this equation gives after scalar multiplication with u

$$c^2\,\frac{dM}{d\tau} \;=\; u\cdot\left(\frac{dM}{d\tau}\, u \;+\; M\,\frac{du}{d\tau}\right) \;=\; -\,\nu\, u\cdot q_e \; , \tag{7.138}$$

and after scalar multiplication with itself

$$c^2\left(\frac{dM}{d\tau}\right)^2 \;+\; M^2\left(\frac{du}{d\tau}\right)^2 \;=\; \nu^2 q_e^2 \; . \tag{7.139}$$

As u is lightlike and q_e is timelike or lightlike, we have $u \cdot q_e \neq 0$ and may therefore use eq. (7.138) to eliminate the particle flux rate ν from eq. (7.139). Then

$$c^2 \left(\frac{dM}{d\tau} \right)^2 + M^2 \left(\frac{du}{d\tau} \right)^2 \;=\; \frac{q_e^2 c^4}{(u \cdot q_e)^2} \left(\frac{dM}{d\tau} \right)^2 ,$$

i.e.,

$$-\,\frac{1}{c^2} \left(\frac{du}{d\tau} \right)^2 \;=\; \left(1 - \frac{q_e^2 c^2}{(u \cdot q_e)^2} \right) \frac{1}{M^2} \left(\frac{dM}{d\tau} \right)^2 . \tag{7.140}$$

If $m_0 = 0$, we have $q_e^2 = 0$, so the first factor on the rhs of eq. (7.140) equals 1. If $m_0 > 0$, this factor can be most easily evaluated in the rocket's momentary rest frame, where

$$\begin{aligned} u \;=\; (c, 0) \;\;,\;\; q_e \;=\; m_0 c \,(\gamma_e, \beta_e \gamma_e) \;\;&\Longrightarrow\;\; u \cdot q_e \;=\; m_0 c^2 \gamma_e \;\;,\;\; q_e^2 \;=\; m_0^2 c^2 \\ &\Longrightarrow\;\; 1 - \frac{q_e^2 c^2}{(u \cdot q_e)^2} \;=\; 1 - \gamma_e^{-2} \;=\; \beta_e^2 \,. \end{aligned}$$

Thus in any case, we obtain

$$\left(\frac{d\theta}{d\tau} \right)^2 \;=\; -\,\frac{1}{c^2} \left(\frac{du}{d\tau} \right)^2 \;=\; \frac{\beta_e^2}{M^2} \left(\frac{dM}{d\tau} \right)^2 , \tag{7.141}$$

(the first of these two equalities follows from eq. (7.136)). Supposing that θ increases monotonically with τ (otherwise, we need only perform the substitution $\theta \to -\theta$), we can take the square root to arrive at

$$\frac{d\theta}{d\tau} \;=\; -\,\frac{\beta_e}{M} \frac{dM}{d\tau} \,. \tag{7.142}$$

In the case of a constant emission velocity $v_e = c\beta_e$, integration of this differential equation is elementary: the solution with the initial condition $\theta(\tau_0) = 0$ with $M(\tau_0) = M_0$ is

$$\begin{aligned} \theta(\tau) \;&=\; -\,\beta_e \ln \frac{M(\tau)}{M_0} \,, \\ \beta(\tau) \;&=\; \tanh\left(-\,\beta_e \ln \frac{M(\tau)}{M_0} \right) , \\ \gamma(\tau) \;&=\; \cosh\left(-\,\beta_e \ln \frac{M(\tau)}{M_0} \right) , \end{aligned} \tag{7.143}$$

or

$$\begin{aligned} \beta(\tau) \;&=\; \frac{1 - \left(\dfrac{M(\tau)}{M_0} \right)^{2\beta_e}}{1 + \left(\dfrac{M(\tau)}{M_0} \right)^{2\beta_e}} \,, \\ \gamma(\tau) \;&=\; \frac{1}{2} \left(\left(\frac{M(\tau)}{M_0} \right)^{\beta_e} + \left(\frac{M(\tau)}{M_0} \right)^{-\beta_e} \right) . \end{aligned} \tag{7.144}$$

In the non-relativistic limit ($\beta_e \ll 1$), this reduces to the well known equation of motion of the rocket

$$v(t) \;=\; -\,v_e \ln \frac{M(\tau)}{M_0} \;. \tag{7.145}$$

In any case, the final velocity that can be reached depends only on the emission velocity v_e and on the ratio between the rocket's initial and final mass. In order to maximize the final velocity, it thus seems desirable to take v_e as large as possible; the best choice would be a photon rocket, where $v_e = c$. However, a photon rocket has its own problems which appear most clearly when one compares the acceleration a that can be reached with the power P available in the rocket's engine. Indeed, evaluating eq. (7.142) in the rocket's momentary rest frame and using eq. (7.103) gives

$$a \;=\; -\,\frac{v_e}{M}\frac{dM}{d\tau} \;, \tag{7.146}$$

whereas the power P of the rocket's engine is given by

$$P \;=\; -\,\frac{dM}{d\tau}\,c^2 - m_0\nu c^2 \;. \tag{7.147}$$

This can be understood by going to the rocket's momentary rest frame, in which the loss of rest mass dM ($dM<0$) during the proper time interval $d\tau$ is easily seen to be made up from two contributions: the transformation of mass into energy inside the rocket's engine ($-L\,d\tau/c^2$) and the loss due to the ejection of the gas particles ($-m_0\nu d\tau$). If $m_0 > 0$, we can use eq. (7.138) to transform the second term in eq. (7.147):

$$P \;=\; -\,\frac{dM}{d\tau}\,c^2 - m_0\nu c^2 \;=\; -\,\frac{dM}{d\tau}\,c^2\left(1 - \frac{m_0c^2}{u\cdot q_e}\right) \;.$$

But in the rocket's momentary rest frame, we have

$$\begin{aligned} u \;=\; (c,0) \;\;,\;\; q_e \;=\; m_0c\,(\gamma_e, \beta_e\gamma_e) \;\;&\Longrightarrow\;\; u\cdot q_e \;=\; m_0c^2\gamma_e \\ &\Longrightarrow\;\; \frac{m_0c^2}{u\cdot q_e} \;=\; \gamma_e^{-1} \;=\; \sqrt{1-\beta_e^2} \;. \end{aligned}$$

Thus in any case, we obtain

$$P \;=\; -\,\frac{dM}{d\tau}\,c^2\,(1 - \sqrt{1-\beta_e^2}) \;. \tag{7.148}$$

Combination with eq. (7.146) yields the following relation between the specific power P/M of the rocket's engine and the acceleration a attained:

$$\frac{P}{M} \;=\; c\,\frac{1-\sqrt{1-\beta_e^2}}{\beta_e}\,a \;. \tag{7.149}$$

In the non-relativistic limit ($\beta_e \ll 1$), this reduces to

$$\frac{P}{M} \;=\; \tfrac{1}{2}v_e a \;, \tag{7.150}$$

whereas in the ultra-relativistic limit ($\beta_e = 1$), we get

$$\frac{L}{M} = ca \ . \tag{7.151}$$

Thus in order to achieve an acceleration of about 10 meters per square second, the engine of a photon rocket would have to provide a specific power of about 3 Gigawatt per kilogram – which is pure science fiction.

We conclude by computing the efficiency η of the rocket, that is, the ratio between the total kinetic energy T of the rocket and the total work W done by the rocket's engine, assuming as before that the emission velocity $v_e = c\beta_e$ is constant. If $M_0 = M(\tau_0)$ is the initial mass and $M_1 = M(\tau_1)$ is the final mass of the rocket and $x = M_1/M_0$ $(x<1)$ is their ratio, then

$$T = M_1 c^2 (\gamma_1 - 1) \ ,$$

whereas integration of eq. (7.148) gives

$$W = (M_0 - M_1)\, c^2 \left(1 - \sqrt{1 - \beta_e^2}\right) \ .$$

Using eq. (7.143) gives

$$\eta = \frac{T}{W} = \frac{x}{1-x}\,\frac{\cosh(\beta_e \ln x) - 1}{1 - \sqrt{1-\beta_e^2}} = \frac{x}{1-x}\,\frac{x^{\beta_e} + x^{-\beta_e} - 2}{2(1 - \sqrt{1-\beta_e^2})} \ . \tag{7.152}$$

In the non-relativistic limit ($\beta_e \ll 1$), this reduces to

$$\eta = \frac{x}{1-x} \ln^2 x \ , \tag{7.153}$$

with a maximum at $1/x = 4,93$, $\eta = 0,647$, whereas in the ultra-relativistic limit ($\beta_e = 1$), we get

$$\eta = \tfrac{1}{2}(1 - x) \ , \tag{7.154}$$

i.e., the efficiency of a photon rocket is always $< 50\%$.

7.8 Covariant Formulation of Electrodynamics

As has already been mentioned at the beginning of this chapter, the observed fact that the speed of light is the same in all inertial systems, raised to the status of a postulate, is the starting point of the theory of special relativity. Light being an electromagnetic phenomenon, electrodynamics should be a relativistically covariant theory. In order to transform the standard formulation of the basic laws of electrodynamics – the Lorentz force law and Maxwell's equations – into manifestly covariant form, we must first of all describe how quantities such as charge density and current density, scalar potential and vector potential or electric field and magnetic field can be combined into four-vectors or four-tensors, respectively.

The starting point for this reformulation is the fact, substantiated by extensive experimental evidence, that electric charge is an absolutely conserved quantity which, in addition, only appears in integer multiples of the so-called elementary charge and shows no velocity dependence whatsoever: this means that the electric charge of a point particle is a Lorentz invariant real constant, or four-scalar.

Now consider a charge density ρ_0 that is at rest in a given Lorentz system; thus the corresponding current density is $\boldsymbol{j}_0 = 0$. Viewed from another Lorentz system moving with velocity $\boldsymbol{v}$ with respect to the first one, we find, due to invariance of the total charge and Lorentz contraction of the volume element (cf. eq. (7.83)), the charge density $\rho = \rho_0 \gamma$ and, due to convection, the current density $\boldsymbol{j} = \rho \boldsymbol{v} = \rho_0 \gamma \boldsymbol{v}$. This suggests to combine charge density ρ and current density $\boldsymbol{j}$ into a four-vector field

$$j^\mu \;=\; (\rho c, \boldsymbol{j}) \quad , \quad j_\mu \;=\; (\rho c, -\boldsymbol{j}) \tag{7.155}$$

called the *current four-vector density*: this allows to write the conservation law for electric charge (cf. eq. (3.1)) in relativistically covariant form:

$$\partial_\mu j^\mu \;=\; 0 \; . \tag{7.156}$$

In complete analogy, the scalar potential ϕ and the vector potential $\boldsymbol{A}$ can be combined into a four-vector field

$$A^\mu \;=\; (\phi/\kappa c, \boldsymbol{A}) \quad , \quad A_\mu \;=\; (\phi/\kappa c, -\boldsymbol{A}) \tag{7.157}$$

called the *four-potential*: this allows to write the Lorentz gauge (3.46), as well as Maxwell's equations (3.54) and (3.55) for the potentials in the Lorentz gauge in relativistically covariant form

$$\partial_\mu A^\mu \;=\; 0 \; , \tag{7.158}$$

$$\Box A^\mu \;\equiv\; \partial^2 A^\mu \;=\; \kappa \mu_0 \, j^\mu \; . \tag{7.159}$$

The fields are obtained from the potentials by differentiation. Electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ can be combined into a rank 2 antisymmetric four-tensor field F, called the *field strength tensor*, according to

$$\begin{aligned} -\,F^{0i} \;&=\; F_{0i} \;=\; E_i/\kappa c \; , \\ F^{ij} \;&=\; F_{ij} \;=\; -\,\epsilon_{ijk}\, B_k \; , \end{aligned} \tag{7.160}$$

where we must remember our convention not to distinguish between upper and lower indices of three-dimensional vectors: in particular, E_i and B_i are the components of $\boldsymbol{E}$ and $\boldsymbol{B}$, respectively. In matrix notation,

$$
\begin{aligned}
F^{\mu\nu} &= \begin{pmatrix} 0 & -E_1/\kappa c & -E_2/\kappa c & -E_3/\kappa c \\ +E_1/\kappa c & 0 & -B_3 & +B_2 \\ +E_2/\kappa c & +B_3 & 0 & -B_1 \\ +E_3/\kappa c & -B_2 & +B_1 & 0 \end{pmatrix} , \\
F_{\mu\nu} &= \begin{pmatrix} 0 & +E_1/\kappa c & +E_2/\kappa c & +E_3/\kappa c \\ -E_1/\kappa c & 0 & -B_3 & +B_2 \\ -E_2/\kappa c & +B_3 & 0 & -B_1 \\ -E_3/\kappa c & -B_2 & +B_1 & 0 \end{pmatrix} .
\end{aligned}
\tag{7.161}
$$

This allows to rewrite the definition of the fields in terms of the potentials, given by eqns (3.40) and (3.42), in relativistically covariant form:

$$F_{\mu\nu} \;=\; \partial_\mu A_\nu - \partial_\nu A_\mu \; . \tag{7.162}$$

Moreover, the homogeneous Maxwell equations (3.5-b) and (3.5-c) assume the relativistically covariant form

$$\partial_\kappa F_{\mu\nu} \,+\, \partial_\mu F_{\nu\kappa} \,+\, \partial_\nu F_{\kappa\mu} \;=\; 0 \; , \tag{7.163}$$

whereas the inhomogeneous Maxwell equations (3.5-a) and (3.5-d) assume the relativistically covariant form

$$\partial_\mu F^{\mu\nu} \;=\; \kappa\mu_0 \, j^\nu \; . \tag{7.164}$$

The transformation law of electromagnetic fields under a boost $B(\boldsymbol{v}) = \Lambda(\boldsymbol{n}, \theta)$ can be read off from the fact that F is a four-tensor. A short calculation using eqns (7.47)–(7.51) gives:

$$
\begin{aligned}
\boldsymbol{E}'_\parallel &= \boldsymbol{E}_\parallel \quad , \quad \boldsymbol{E}'_\perp \;=\; \gamma\big(\boldsymbol{E}_\perp \,+\, \kappa c\beta \, \boldsymbol{n} \times \boldsymbol{B}\big) \; , \\
\boldsymbol{B}'_\parallel &= \boldsymbol{B}_\parallel \quad , \quad \boldsymbol{B}'_\perp \;=\; \gamma\Big(\boldsymbol{B}_\perp \,-\, \frac{\beta}{\kappa c} \boldsymbol{n} \times \boldsymbol{E}\Big) \; .
\end{aligned}
\tag{7.165}
$$

The notation in terms of differential forms is also much simpler and more transparent in its relativistically covariant form, that is, in four-dimensional Minkowski space-time, than it is in ordinary three-dimensional Euclidean space, with time as an additional parameter. In order to see this, we introduce the canonical basis $\{e_0, e_1, e_2, e_3\}$ of Minkowski space defined in eq. (7.9) and the corresponding dual basis $\{e^0, e^1, e^2, e^3\}$ and define two one-forms j, A and one two-form F on Minkowski space, as follows:

$$j \;=\; j_\mu \, e^\mu \quad , \quad A \;=\; A_\mu \, e^\mu \; , \tag{7.166}$$

$$F \;=\; \tfrac{1}{2} \, F_{\mu\nu} \, e^\mu \wedge e^\nu \; . \tag{7.167}$$

The important equations can then be written down using the exterior derivative d and the star operator $*$ for differential forms on Minkowski space; the latter is explicitly given by

$$
\begin{array}{c}
*\,(1) \;=\; e^0 \wedge e^1 \wedge e^2 \wedge e^3 \;, \\
*\left(e^0\right) \;=\; e^1 \wedge e^2 \wedge e^3 \quad , \quad *\left(e^i\right) \;=\; \tfrac{1}{2}\,\epsilon_{ijk}\, e^0 \wedge e^j \wedge e^k \;, \\
*\left(e^0 \wedge e^i\right) \;=\; -\,\tfrac{1}{2}\,\epsilon_{ijk}\, e^j \wedge e^k \quad , \quad *\left(e^j \wedge e^k\right) \;=\; +\,\tfrac{1}{2}\,\epsilon_{jkl}\, e^0 \wedge e^l \;, \\
*\left(e^1 \wedge e^2 \wedge e^3\right) \;=\; e^0 \quad , \quad *\left(e^0 \wedge e^j \wedge e^k\right) \;=\; \epsilon_{jkl}\, e^l \;, \\
*\left(e^0 \wedge e^1 \wedge e^2 \wedge e^3\right) \;=\; -\,1 \;,
\end{array}
\tag{7.168}
$$

with the convention

$$
\epsilon^{0123} \;=\; -\,1 \quad , \quad \epsilon_{0123} \;=\; +\,1 \;. \tag{7.169}
$$

Then the conservation law (7.156) for electric charge and the Lorentz gauge (7.158) assume the form

$$
d * j \;=\; 0 \tag{7.170}
$$

and

$$
d * A \;=\; 0 \tag{7.171}
$$

respectively, whereas the definition of the fields in terms of the potentials becomes

$$
F \;=\; dA \;. \tag{7.172}
$$

Finally, the homogeneous and inhomogeneous Maxwell equations read

$$
dF \;=\; 0 \tag{7.173}
$$

and

$$
-\,*\,d * F \;=\; \kappa\mu_0\, j \tag{7.174}
$$

respectively.

Apart from Maxwell's equations, we have as a fundamental law of electrodynamics the Lorentz force law, whose relativistically covariant form has not yet been given. It states that the four-force F^μ exerted by the electromagnetic field strength tensor $F^{\mu\nu}$ on a point particle with charge q, moving along its world line $\tau \mapsto x^\mu(\tau)$ with four-velocity u^μ (here, the parameter τ is proper time), is the *Lorentz four-force*

$$
F^\mu \;=\; \kappa q u_\nu F^{\mu\nu} \;. \tag{7.175}
$$

When dealing not with a charged point particle but with a general charge and current distribution given by a current four-vector density j^μ, the Lorentz four-force must be replaced by the *Lorentz four-force density*

$$
f^\mu \;=\; \kappa j_\nu F^{\mu\nu} \;. \tag{7.176}
$$

The spatial components of the expressions (7.175) and (7.176) are the relativistic generalizations of the force laws (3.2) and (3.3), respectively, whereas their time components (after multiplication by c) provide the relativistic generalizations of the expressions $q\boldsymbol{v}\cdot E$ and $\boldsymbol{j}\cdot\boldsymbol{E}$ for the power and the power density, respectively, referring to the work done by the field on the matter distribution. Moreover, the two equations (7.175) and (7.176) can in relativistically covariant manner be identified as two variants of one and the same force law if one considers that a point particle with charge q, moving along its world line $\tau \mapsto x^\mu(\tau)$ with four-velocity u^μ (here, the parameter τ is proper time), produces the current four-vector density

$$j^\mu(x) \;=\; qc \int d\tau \; u^\mu(\tau)\; \delta(x - x(\tau)) \;, \tag{7.177}$$

whereas the four-force F^μ to which it is subjected along its world line corresponds to a four-force density

$$f^\mu(x) \;=\; c \int d\tau \; F^\mu(\tau)\; \delta(x - x(\tau)) \;. \tag{7.178}$$

Both of these are, as expected, concentrated on the particle's world line in a δ-function like manner.

7.9 The Energy-Momentum Tensor of the Electromagnetic Field

In Chapter 3, we have discussed separately the energy balance and the momentum balance in the presence of electromagnetic fields. These two can now be unified in a remarkable way. To this end, we use the Maxwell equations (7.164) and (7.163) to rewrite the Lorentz four-force density f^μ given by eq. (7.176) as follows:

$$\begin{aligned}
f^\mu \;&=\; \kappa\, F^{\mu\kappa}\, j_\kappa \;=\; \frac{1}{\mu_0}\, F^{\mu\kappa}\, \partial^\nu F_{\nu\kappa} \\
&=\; \frac{1}{\mu_0}\, \big(\partial^\nu (F^{\mu\kappa} F_{\nu\kappa}) \;-\; (\partial^\nu F^{\mu\kappa})\, F_{\nu\kappa} \big) \\
&=\; \frac{1}{\mu_0}\, \big(\partial^\nu (F^{\mu\kappa} F_{\nu\kappa}) \;-\; \tfrac{1}{2} (\partial^\nu F^{\mu\kappa})\, F_{\nu\kappa} \;-\; \tfrac{1}{2} (\partial^\kappa F^{\mu\nu})\, F_{\kappa\nu} \big) \\
&=\; \frac{1}{\mu_0}\, \big(\partial^\nu (F^{\mu\kappa} F_{\nu\kappa}) \;-\; \tfrac{1}{2} (\partial^\nu F^{\mu\kappa} + \partial^\kappa F^{\nu\mu})\, F_{\nu\kappa} \big) \\
&=\; \frac{1}{\mu_0}\, \big(\partial^\nu (F^{\mu\kappa} F_{\nu\kappa}) \;+\; \tfrac{1}{2} (\partial^\mu F^{\kappa\nu})\, F_{\nu\kappa} \big) \\
&=\; \frac{1}{\mu_0}\, \big(\partial^\nu (F^{\mu\kappa} F_{\nu\kappa}) \;-\; \tfrac{1}{4} \partial^\mu (F^{\nu\kappa} F_{\nu\kappa}) \big) \;.
\end{aligned}$$

Thus we have

$$f^\mu \;=\; -\; \partial_\nu T^{\mu\nu} \;, \tag{7.179}$$

with a rank 2 four-tensor field T, called the *energy-momentum tensor of the electromagnetic field*, defined by

$$T^{\mu\nu} \;=\; -\; \frac{1}{\mu_0}\, \big(\eta_{\kappa\lambda} F^{\mu\kappa} F^{\nu\lambda} \;-\; \tfrac{1}{4} \eta^{\mu\nu} F^{\kappa\lambda} F_{\kappa\lambda} \big) \;. \tag{7.180}$$

It is symmetric and traceless:

$$T^{\mu\nu} \;=\; T^{\nu\mu} \quad , \quad T^{\mu}_{\ \mu} \;=\; 0 \; . \tag{7.181}$$

Its components have the following meaning:

- Its purely time component is the energy density ρ^E of the electromagnetic field; cf. eq. (3.58):
$$T^{00} \;=\; T_{00} \;=\; \rho^E \; .$$
- Its mixed components are (apart from factors of c or $1/c$) the components of the Poynting vector $\boldsymbol{S}$ that describes both the energy flux density and the momentum density of the electromagnetic field; cf. eqns (3.59), (3.60) and (3.85):
$$T^{0i} \;=\; T^{i0} \;=\; -\,T_{0i} \;=\; -\,T_{i0} \;=\; S_i/c \;=\; j_i^E/c \;=\; \rho_i^P\,c \; .$$
- Its purely spatial components are (apart from a sign) the components of Maxwell's stress tensor that describes the momentum flux density of the electromagnetic field; cf. eqns (3.86) and (3.87):
$$T^{ik} \;=\; T^{ki} \;=\; T_{ik} \;=\; T_{ki} \;=\; -\,T_{ik}^{\mathrm{Max}} \;=\; j_{ik}^P \; .$$

In particular, the relativistically covariant formulation of electrodynamics reveals the equality (up to a factor of c^2) between the energy flux density and the momentum density – a fact already mentioned, though without further explanation, in Chapter 3 – to be part of the statement that the energy-momentum tensor is symmetric. The remaining part of the same statement is that Maxwell's stress tensor is symmetric – another fact already observed and commented upon in Chapter 3. Moreover, the statement that the energy-momentum tensor is also traceless means that the energy density is equal to the negative of the trace of Maxwell's stress tensor, that is, the energy density of the electromagnetic field describes precisely the volume expanding part of the forces that it exerts.

7.10 Radiation by a Moving Point Charge: Liénard-Wiechert Potentials

In this section we want to determine the electromagnetic potentials and fields generated by an arbitrary moving point particle with charge q: these can be given in closed form. From Chapter 6.2 we know that the solution of the field equation (7.159) for the potentials in the Lorentz gauge and with retarded boundary conditions can be obtained by convolution of the source with the retarded Green function of the wave operator, that is,[2]

$$A^\mu(x) \;=\; \frac{\kappa\mu_0}{4\pi} \int d^4x' \; G_{\mathrm{ret}}(x-x')\; j^\mu(x') \; , \tag{7.182}$$

[2] The normalization of the Green functions of the wave operator used here, which is natural in the context of a relativistically covariant theory, deviates from that used in Chapter 6 by a factor of c.

where

$$\Box\, G_{\rm ret}(x - x') \;=\; 4\pi\, \delta(x - x') \tag{7.183}$$

and explicitly

$$G_{\rm ret}(x - x') \;=\; 2\, \theta(x^0 - x'^0)\, \delta((x - x')^2)\ . \tag{7.184}$$

Substituting eq. (7.177) and using eq. (3.27), we get

$$A^\mu(x) \;=\; \frac{q}{2\pi\kappa c\epsilon_0} \int d\tau\, d^4x'\ \theta(x^0 - x'^0)\, \delta((x - x')^2)\, u^\mu(\tau)\, \delta(x' - x(\tau))$$

and can perform the integration over x':

$$A^\mu(x) \;=\; \frac{q}{2\pi\kappa c\epsilon_0} \int d\tau\ \theta(x^0 - x^0(\tau))\, \delta((x - x(\tau))^2)\, u^\mu(\tau)\ . \tag{7.185}$$

The remaining integration over τ can be performed as well, observing that the argument of the δ function, viewed as a function of τ, has precisely two simple zeroes, namely at the points $\tau_{\rm ret}$ and $\tau_{\rm adv}$ given by[3]

$$x^0 - x^0(\tau_{\rm ret}) \;=\; |\boldsymbol{x} - \boldsymbol{x}(\tau_{\rm ret})| \quad , \quad x^0(\tau_{\rm adv}) - x^0 \;=\; |\boldsymbol{x}(\tau_{\rm adv}) - \boldsymbol{x}|\ . \tag{7.186}$$

Obviously, $\tau_{\rm ret}$ and $\tau_{\rm adv}$ are precisely the parameter values that correspond to the intersection points between the particle's world line and the backward and forward light cone attached to the point x, respectively; the four-velocity and the four-acceleration at these intersection points will be denoted by $u_{\rm ret}$ and $a_{\rm ret}$ and by $u_{\rm adv}$ and $a_{\rm adv}$, respectively. Moreover, we also introduce the four-scalars

$$\rho_{\rm ret} \;=\; \frac{1}{c}\, u_{\rm ret} \cdot (x - x_{\rm ret}) \quad , \quad \rho_{\rm adv} \;=\; \frac{1}{c}\, u_{\rm adv} \cdot (x_{\rm adv} - x)\ , \tag{7.187}$$

and for reasons to become clear soon, the four-vectors

$$n_{\rm ret} \;=\; \frac{x - x_{\rm ret}}{\rho_{\rm ret}} - \frac{u_{\rm ret}}{c} \quad , \quad n_{\rm adv} \;=\; \frac{x_{\rm adv} - x}{\rho_{\rm adv}} - \frac{u_{\rm adv}}{c}\ . \tag{7.188}$$

(Observe that $x - x_{\rm ret}$ and $x_{\rm adv} - x$ are positive lightlike four-vectors whereas $u_{\rm ret}$ and $u_{\rm adv}$ are positive timelike four-vectors; therefore, $\rho_{\rm ret} > 0$ and $\rho_{\rm adv} > 0$.) In the momentary rest frame of the point particle, we have

$$\begin{array}{lllllll} u^\mu_{\rm ret} = (c, 0) & , & x^\mu - x^\mu_{\rm ret} = (\rho_{\rm ret}, \rho_{\rm ret}\boldsymbol{n}) & , & n^\mu_{\rm ret} = (0, \boldsymbol{n})\ . \\ u^\mu_{\rm adv} = (c, 0) & , & x^\mu_{\rm adv} - x^\mu = (\rho_{\rm adv}, \rho_{\rm adv}\boldsymbol{n}) & , & n^\mu_{\rm adv} = (0, \boldsymbol{n})\ . \end{array} \tag{7.189}$$

(Cf. Fig. 7.7.) Then

$$\begin{aligned} \frac{d}{d\tau}(x - x(\tau))^2\Big|_{\tau=\tau_{\rm ret}} &= -2c\,\rho_{\rm ret} \neq 0 \\ \frac{d}{d\tau}(x - x(\tau))^2\Big|_{\tau=\tau_{\rm adv}} &= +2c\,\rho_{\rm adv} \neq 0 \end{aligned} \tag{7.190}$$

and thus

[3] The indices "ret" and "adv" stand for "retarded" and "advanced", respectively.

Fig. 7.7: Calculating the potentials generated by a moving point charge (one space dimension is suppressed)

$$\delta\left((x - x(\tau))^2\right) \;=\; \frac{\delta(\tau - \tau_{\rm ret})}{2c\,\rho_{\rm ret}} + \frac{\delta(\tau - \tau_{\rm adv})}{2c\,\rho_{\rm adv}} \;. \tag{7.191}$$

Due to the θ-function, only the retarded contribution remains, and we arrive at the so-called *Liénard-Wiechert potentials*

$$A^\mu(x) \;=\; \frac{\kappa\mu_0\, q}{4\pi}\, \frac{u^\mu_{\rm ret}}{\rho_{\rm ret}} \;. \tag{7.192}$$

To determine the corresponding field strengths, it must be observed that $\tau_{\rm ret}$ depends implicitly on x, and therefore the same holds for $x_{\rm ret}$, $u_{\rm ret}$, $a_{\rm ret}$, $\rho_{\rm ret}$ and $n_{\rm ret}$. Thus we first differentiate the equation that determines $\tau_{\rm ret}$ with respect to x

$$\begin{aligned}
0 \;&=\; \partial^\mu\left((x - x_{\rm ret})^2\right) \;=\; 2\eta_{\kappa\lambda}\,(x^\kappa - x^\kappa_{\rm ret})\,\partial^\mu(x^\lambda - x^\lambda_{\rm ret}) \\
&=\; 2(x^\mu - x^\mu_{\rm ret}) \,-\, 2\eta_{\kappa\lambda}\,(x^\kappa - x^\kappa_{\rm ret})\,\frac{dx^\lambda}{d\tau}\bigg|_{\tau=\tau_{\rm ret}} \partial^\mu\tau_{\rm ret} \\
&=\; 2(x^\mu - x^\mu_{\rm ret}) \,-\, 2(x - x_{\rm ret})\cdot u_{\rm ret}\;\partial^\mu\tau_{\rm ret} \;,
\end{aligned}$$

obtaining

$$\partial^\mu\tau_{\rm ret} \;=\; \frac{x^\mu - x^\mu_{\rm ret}}{c\rho_{\rm ret}} \;. \tag{7.193}$$

This implies

$$\begin{aligned}
\partial^\mu x^\nu_{\rm ret} \;&=\; \frac{dx^\nu}{d\tau}\bigg|_{\tau=\tau_{\rm ret}} \partial^\mu\tau_{\rm ret} \;=\; u^\nu_{\rm ret}\,\frac{x^\mu - x^\mu_{\rm ret}}{c\rho_{\rm ret}} \;, \\
\partial^\mu u^\nu_{\rm ret} \;&=\; \frac{du^\nu}{d\tau}\bigg|_{\tau=\tau_{\rm ret}} \partial^\mu\tau_{\rm ret} \;=\; a^\nu_{\rm ret}\,\frac{x^\mu - x^\mu_{\rm ret}}{c\rho_{\rm ret}} \;,
\end{aligned}$$

or

$$\partial^\mu x^\nu_{\rm ret} \;=\; (n^\mu_{\rm ret} + u^\mu_{\rm ret}/c)\, u^\nu_{\rm ret}/c \;, \tag{7.194}$$

$$\partial^\mu u^\nu_{\rm ret} \;=\; (n^\mu_{\rm ret} + u^\mu_{\rm ret}/c)\, a^\nu_{\rm ret}/c \;, \tag{7.195}$$

as well as

$$\begin{aligned} \partial^\mu \rho_{\rm ret} \;&=\; \frac{1}{c}\eta_{\kappa\lambda}\left((\partial^\mu u^\kappa_{\rm ret})(x^\lambda - x^\lambda_{\rm ret}) \;+\; u^\kappa_{\rm ret}(\eta^{\mu\lambda} - \partial^\mu x^\lambda_{\rm ret})\right) \\ &=\; \frac{1}{c}\, u^\mu_{\rm ret} \;+\; \frac{1}{c^2}\eta_{\kappa\lambda}\left(a^\kappa_{\rm ret}(x^\lambda - x^\lambda_{\rm ret}) \;-\; u^\kappa_{\rm ret}u^\lambda_{\rm ret}\right)(n^\mu_{\rm ret} + u^\mu_{\rm ret}/c) \;\;, \end{aligned}$$

i.e.,

$$\partial^\mu \rho_{\rm ret} \;=\; -\; n^\mu_{\rm ret} \;+\; \frac{1}{c^2}\,\rho_{\rm ret}\,(a_{\rm ret}\cdot n_{\rm ret})\,(n^\mu_{\rm ret} + u^\mu_{\rm ret}/c) \;\;. \tag{7.196}$$

Hence we get

$$\begin{aligned} \partial^\mu A^\nu(x) \;&=\; \frac{\kappa\mu_0\, q}{4\pi}\;\frac{\rho_{\rm ret}(\partial^\mu u^\nu_{\rm ret}) - (\partial^\mu \rho_{\rm ret})u^\nu_{\rm ret}}{\rho^2_{\rm ret}} \\ &=\; \frac{\kappa\mu_0\, q}{4\pi}\left(\frac{(n^\mu_{\rm ret} + u^\mu_{\rm ret}/c)\, a^\nu_{\rm ret}}{c\rho_{\rm ret}} \;+\; \frac{n^\mu_{\rm ret}u^\nu_{\rm ret}}{\rho^2_{\rm ret}}\right. \\ &\qquad\left. -\;\frac{(a_{\rm ret}\cdot n_{\rm ret})\,(n^\mu_{\rm ret} + u^\mu_{\rm ret}/c)\, u^\nu_{\rm ret}}{c^2\rho_{\rm ret}}\right) \;\;, \end{aligned}$$

so finally

$$\begin{aligned} F^{\mu\nu}(x) \;&=\; \frac{\kappa\mu_0\, q}{4\pi\rho^2_{\rm ret}}\,(n^\mu_{\rm ret}u^\nu_{\rm ret} - n^\nu_{\rm ret}u^\mu_{\rm ret}) \\ &\quad+\; \frac{\kappa\mu_0\, q}{4\pi c\rho_{\rm ret}}\left(\frac{1}{c}\,(u^\mu_{\rm ret}a^\nu_{\rm ret} - u^\nu_{\rm ret}a^\mu_{\rm ret}) \;+\; (n^\mu_{\rm ret}a^\nu_{\rm ret} - n^\nu_{\rm ret}a^\mu_{\rm ret})\right. \\ &\qquad\left. -\;\frac{1}{c}\,(a_{\rm ret}\cdot n_{\rm ret})\,(n^\mu_{\rm ret}u^\nu_{\rm ret} - n^\nu_{\rm ret}u^\mu_{\rm ret})\right) \;\;. \end{aligned} \tag{7.197}$$

We also want to write down the non-relativistic form of these formulae. To this end, we set

$$\boldsymbol{r} \;=\; \boldsymbol{x} - \boldsymbol{x}_{\rm ret} \quad,\quad r \;=\; |\boldsymbol{r}| \quad,\quad \boldsymbol{n} \;=\; \boldsymbol{r}/r \tag{7.198}$$

and obtain

$$u^\mu_{\rm ret} \;=\; \gamma_{\rm ret}\, c\,(1, \boldsymbol{\beta}_{\rm ret}) \;\;,$$

$$a^\mu_{\rm ret} \;=\; \gamma^2_{\rm ret}\left(\gamma^2_{\rm ret}\,\boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{a}_{\rm ret} \;,\, \boldsymbol{a}_{\rm ret} \;+\; \gamma^2_{\rm ret}\,(\boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{a}_{\rm ret})\,\boldsymbol{\beta}_{\rm ret}\right) \;\;,$$

$$\rho_{\rm ret} \;=\; \gamma_{\rm ret}(r - \boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{r}) \;=\; \gamma_{\rm ret}\, r\,(1 - \boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{n}) \;\;,$$

$$n^\mu_{\rm ret} \;=\; \left(\frac{1}{\gamma_{\rm ret}\,(1 - \boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{n})} \;-\; \gamma_{\rm ret} \;,\; \frac{\boldsymbol{n}}{\gamma_{\rm ret}\,(1 - \boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{n})} \;-\; \gamma_{\rm ret}\boldsymbol{\beta}_{\rm ret}\right) \;\;,$$

and after a short calculation

$$a_{\rm ret}\cdot n_{\rm ret} \;=\; -\;\gamma_{\rm ret}\,\frac{\boldsymbol{a}_{\rm ret}\cdot\boldsymbol{n}}{1 - \boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{n}} \;+\; \gamma^3_{\rm ret}\,(\boldsymbol{\beta}_{\rm ret}\cdot\boldsymbol{a}_{\rm ret}) \;\;.$$

This implies for the scalar potential

$$\phi(x) \;=\; \frac{q}{4\pi\epsilon_0}\,\frac{1}{r}\,\frac{1}{1-\boldsymbol{\beta}_{\mathrm{ret}}\cdot\boldsymbol{n}} \tag{7.199}$$

and for the vector potential

$$\boldsymbol{A}(x) \;=\; \frac{\kappa\mu_0\, q}{4\pi}\,\frac{1}{r}\,\frac{c\boldsymbol{\beta}_{\mathrm{ret}}}{1-\boldsymbol{\beta}_{\mathrm{ret}}\cdot\boldsymbol{n}}\;, \tag{7.200}$$

whereas the electric field $\boldsymbol{E}(x)$ and the magnetic field $\boldsymbol{B}(x)$ can after some calculation be shown to be given by the following expressions:

$$\begin{aligned} \boldsymbol{E}(x) \;&=\; \frac{q}{4\pi\epsilon_0}\,\frac{1}{r^2}\,\frac{\boldsymbol{n}-\boldsymbol{\beta}_{\mathrm{ret}}}{\gamma_{\mathrm{ret}}^2\,(1-\boldsymbol{\beta}_{\mathrm{ret}}\cdot\boldsymbol{n})^3} \\ &\quad+\frac{q}{4\pi\epsilon_0}\,\frac{1}{r}\,\frac{\boldsymbol{n}\times((\boldsymbol{n}-\boldsymbol{\beta}_{\mathrm{ret}})\times\boldsymbol{a}_{\mathrm{ret}})}{c^2\,(1-\boldsymbol{\beta}_{\mathrm{ret}}\cdot\boldsymbol{n})^3}\;, \end{aligned} \tag{7.201}$$

$$\boldsymbol{B}(x) \;=\; \frac{1}{\kappa c}\,\boldsymbol{n}\times\boldsymbol{E}(x)\;. \tag{7.202}$$

Thus the field strengths decompose into *velocity fields* that do not depend on the acceleration and fall off at infinity like $1/r^2$ and *acceleration fields* or *radiation fields* that depend linearly on the acceleration and fall off at infinity like $1/r$.

Further insight into the properties of the Liénard-Wiechert potentials and the corresponding field strengths can be gained by investigating the Poynting vector that results from eqns (7.201) and (7.202). We do not want to go into this in more detail and just present the result for the total power P_{Rad} radiated through a closed surface located at infinity (and at rest with respect to the chosen inertial system):

$$P_{\mathrm{Rad}} \;=\; \frac{q^2}{6\pi\epsilon_0 c^3}\,\frac{\boldsymbol{a}_{\mathrm{ret}}^2-(\boldsymbol{\beta}_{\mathrm{ret}}\times\boldsymbol{a}_{\mathrm{ret}})^2}{(1-\beta_{\mathrm{ret}}^2)^3} \;=\; -\,\frac{q^2\,a_{\mathrm{ret}}^2}{6\pi\epsilon_0 c^3}\;. \tag{7.203}$$

This is the correct relativistic generalization of the non-relativistic Larmor formula (6.70): all one has to do is to replace in that formula the square of the ordinary acceleration by the negative square of the four-acceleration. In particular, the result is a four-scalar and thus independent of the choice of reference frame.

For a more thorough treatment of radiation by moving point charges we refer to [Jackson, Chap. 14].

7.11 Relativistic Hydrodynamics

To conclude this chapter, we want to show briefly that hydrodynamics – for practical applications the most important classical field theory besides electrodynamics – can also be naturally extended to the relativistic domain. Such an extension is of concrete interest in astrophysics, where fluxes of matter at high velocities do occur in the vicinity of compact objects such as neutron stars or black holes.

As in the case of electrodynamics, the relativistic formulation of hydrodynamics requires transforming the basic dynamical variables of the non-relativistic theory into four-scalars or into components of four-vectors or four-tensors. To this end, we begin by replacing, as in the case of a single point particle, the ordinary non-relativistic velocity field $\boldsymbol{v}$ by the *four-velocity field* u, that is, the timelike four-vector field with components

$$u^\mu \;=\; \gamma\,(c, \boldsymbol{v}) \quad , \quad u_\mu \;=\; \gamma\,(c, -\boldsymbol{v}) \;, \tag{7.204}$$

where

$$\gamma \;=\; \frac{1}{\sqrt{1 - \boldsymbol{v}^2/c^2}} \;. \tag{7.205}$$

Thus u satisfies the normalization condition

$$\eta_{\mu\nu} u^\mu u^\nu \;=\; c^2 \;, \tag{7.206}$$

which after differentiation implies the transversality of the four-velocity gradient:

$$u_\nu \, \partial_\mu u^\nu \;=\; 0 \;. \tag{7.207}$$

Similarly, the procedure for formulating continuity equations for extensive quantities can be carried over to the relativistic theory with no essential modifications. Every extensive quantity a comes with an associated *current four-vector density* $j^{\mu\,a}$ and an associated *source density* q^a, the first being a timelike or lightlike four-vector field with components

$$j^{\mu\,a} \;=\; (\rho^a c, \boldsymbol{j}^a) \quad , \quad j^a_\mu \;=\; (\rho^a c, -\boldsymbol{j}^a) \tag{7.208}$$

and the second being a four-scalar field; together, they satisfy the relativistically covariant continuity equation

$$\partial_\mu j^{\mu\,a} \;=\; q^a \;. \tag{7.209}$$

The full tensorial nature of $j^{\mu\,a}$ and of q^a, however, can only be inferred from that of a itself: if a is a four-scalar, such as electric charge, $j^{\mu\,a}$ will be a four-vector field and q^a will be a four-scalar field, if a is a four-vector, such as four-momentum (energy + momentum), $j^{\mu\,a}$ will be a rank 2 four-tensor field and q^a will be a four-vector field, etc.. For conserved quantities such as these and in closed systems, of course, we have $q^a = 0$.

Another important aspect which can be carried over to the relativistic theory without difficulties is the splitting into convective and conductive parts which however must here be performed for the entire current four-vector density:

$$j^{\mu\,a} \;=\; j^{\mu\,a}_{\text{conv}} + j^{\mu\,a}_{\text{cond}} \;. \tag{7.210}$$

As in the non-relativistic theory, the convective part is due to transport along with the fluid whereas the conductive part describes transport of the quantity a in the absence of fluid flow. This means that in the momentary rest frame of the fluid, the spatial components of the convective part must vanish, implying that in an arbitrary inertial frame, the convective part will be given by projection along the four-velocity field u and the conductive part by projection orthogonal to the four-velocity field u:

$$j^{\mu\,a}_{\mathrm{conv}} \;=\; \frac{1}{c^2}\,u^\mu u^\nu\,j^a_\nu\;. \tag{7.211}$$

$$j^{\mu\,a}_{\mathrm{cond}} \;=\; \Big(\eta^{\mu\nu} - \frac{1}{c^2}\,u^\mu u^\nu\Big)\,j^a_\nu\;. \tag{7.212}$$

If u itself is not a four-scalar but rather a four-vector or four-tensor, the decomposition into components parallel and orthogonal to u should also be performed with respect to the other indices.

These considerations can in particular be used to derive the energy-momentum tensor of relativistic hydrodynamics, starting out from the expressions (2.14) for the momentum density and, in particular, (2.15) with (2.21) for the momentum flux density of non-relativistic hydrodynamics. In these, the transition to the relativistically covariant form is easily performed and leads to the postulate that the *energy-momentum tensor of hydrodynamics* is composed of two parts, namely a convective part,

$$T^{\mu\nu}_{\mathrm{conv}} \;=\; \rho\,u^\mu u^\nu\;, \tag{7.213}$$

and a conductive part,

$$T^{\mu\nu}_{\mathrm{cond}} \;=\; p\,\Big(\frac{1}{c^2}\,u^\mu u^\nu - \eta^{\mu\nu}\Big) \,+\, \sigma'^{\,\mu\nu}\;, \tag{7.214}$$

so that

$$T^{\mu\nu} \;=\; \Big(\rho + \frac{p}{c^2}\Big)\,u^\mu u^\nu \,-\, p\eta^{\mu\nu} \,+\, \sigma'^{\,\mu\nu}\;, \tag{7.215}$$

where ρ denotes the *mass density* and p denotes the scalar *pressure*, as measured in the momentary rest frame of the fluid, that is, from a Lorentz system in which the fluid (at the given point) is at rest. σ' is the *friction tensor*, whose conductive nature is expressed through the orthogonality relation

$$u_\mu\,\sigma'^{\,\mu\nu} \;=\; 0\;. \tag{7.216}$$

Finally, just as in non-relativistic hydrodynamics, it is required that (for normally flowing fluids without "‘internal"’ angular momenta or torques) the friction tensor must be symmetric:

$$\sigma'^{\,\mu\nu} \;=\; \sigma'^{\,\nu\mu}\;. \tag{7.217}$$

The same must then hold for the full energy-momentum tensor:

$$T^{\mu\nu} \;=\; T^{\nu\mu}\;. \tag{7.218}$$

In the momentary rest frame of the fluid, the spatial components of the four-velocity field u vanish and the energy-momentum tensor assumes the simple form

$$T^{\mu\nu}_{\mathrm{RF}} \;=\; \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \sigma' & \\ 0 & & & \end{pmatrix}. \tag{7.219}$$

For comparison with the non-relativistic expressions one must however use a different Lorentz system in which the absolute value v of the velocity $\boldsymbol{v}$ of the fluid (at the given point) does not vanish but is small as compared to the velocity of light c, so that one may perform an expansion in powers of $\beta = v/c$ and truncate after the first non-trivial term. It is easy to see that this reproduces the known expressions for the non-relativistic momentum density and momentum flux density, whereas the energy denstiy $T^{00} = T_{00} = \rho^E$ receives an additional contribution from the pressure. Indeed, neglecting the friction term, we get

$$\rho^E \; = \; \rho c^2 \gamma^2 \, + \, p \, (\gamma^2 - 1) \; ,$$

and therefore after expansion up to second order

$$\rho^E \; = \; \rho c^2 \, + \Big(\rho + \frac{p}{c^2} \Big) \, v^2 \; .$$

For normal fluids and under usual conditions, however, one has $p \ll \rho c^2$, so that this contribution can be neglected.

External forces acting on the fluid will be represented by a *four-force density* f^μ, so that the energy-momentum balance can be written in the standard relativistic form:[4]

$$f^\mu \; = \; \partial_\nu T^{\mu\nu} \; . \tag{7.220}$$

After insertion of the explicit form (7.215) of the energy-momentum tensor, this relation assumes the following form

$$\partial_\mu \Big(\Big(\rho + \frac{p}{c^2} \Big) \, u^\mu u^\nu \Big) \, - \, \partial^\nu p \, + \, \partial_\mu {\sigma'}^{\mu\nu} \; = \; f^\nu \; . \tag{7.221}$$

Making use of eq. (7.206) and eq. (7.207), we can decompose this equation into its components along u and transversal to u. The longitudinal component is simply obtained by contraction with u, with the result

$$\partial_\mu \big((\rho c^2 + p) \, u^\mu \big) \; - \; u^\nu \, \partial_\nu p \, + \, u_\nu \, \partial_\mu {\sigma'}^{\mu\nu} \; = \; u_\nu f^\nu \; , \tag{7.222}$$

or

$$\partial_\mu \left(\rho c^2 u^\mu \right) + \, p \, \partial_\mu u^\mu \, + \, u_\nu \, \partial_\mu {\sigma'}^{\mu\nu} \; = \; u_\nu f^\nu \; . \tag{7.223}$$

The transversal components are obtained by inserting eq. (7.222) back into eq. (7.221) and rearranging terms:

$$\begin{aligned} \Big(\rho + \frac{p}{c^2} \Big) \, u^\mu \, \partial_\mu u^\nu \; &- \; \Big(\eta^{\mu\nu} - \frac{1}{c^2} \, u^\mu u^\nu \Big) \, \partial_\mu p \; + \; \Big(\eta^{\mu\nu} - \frac{1}{c^2} \, u^\mu u^\nu \Big) \, \eta_{\lambda\mu} \, \partial_\kappa {\sigma'}^{\kappa\lambda} \\ &= \; \Big(\eta^{\mu\nu} - \frac{1}{c^2} \, u^\mu u^\nu \Big) \, f_\mu \; . \end{aligned} \tag{7.224}$$

This is the equation of motion of relativistic hydrodynamics.

Further information can only be gained by making hypotheses about the specific form of the friction tensor. The two simplest choices, whose non-relativistic version has already been discussed in Chapter 2, are the following:

[4] The difference in sign as compared to eq. (7.179) stems from the fact that we are now considering forces acting on the fluid and not forces exerted by the fluid.

- *Ideal Fluid* or *Perfect Fluid*: No friction.

$$\sigma'^{\mu\nu} = 0 \,. \tag{7.225}$$

In this case, the equation of motion (7.224) is the correct relativistic generalization of the Euler equation (2.24).

- *Newtonian Fluid*: Friction tensor proportional to the transverse part of the four-velocity gradient.
Projecting the symmetrized four-velocity gradient

$$\partial^\mu u^\nu + \partial^\nu u^\mu$$

to its component orthogonal to the four-velocity field u and using the constraint (7.207) gives the expression

$$(\partial^\perp u)^{\mu\nu} = \partial^\mu u^\nu + \partial^\nu u^\mu - \frac{1}{c^2}\, u^\kappa\, \partial_\kappa (u^\mu u^\nu) \,, \tag{7.226}$$

which can be further split into a tracefree part and a part proportional to the transverse unit tensor:

$$(\partial_0^\perp u)^{\mu\nu} = (\partial^\perp u)^{\mu\nu} - \tfrac{2}{3}\left(\eta^{\mu\nu} - \frac{1}{c^2}\, u^\mu u^\nu\right) \partial_\kappa u^\kappa \,. \tag{7.227}$$

$$(\partial_1^\perp u)_{\mathrm{tr}}^{\mu\nu} = \tfrac{2}{3}\left(\eta^{\mu\nu} - \frac{1}{c^2}\, u^\mu u^\nu\right) \partial_\kappa u^\kappa \,. \tag{7.228}$$

Then

$$\sigma'^{\mu\nu} = \eta\, (\partial_0^\perp u)^{\mu\nu} + \tfrac{3}{2}\, \zeta\, (\partial_1^\perp u)^{\mu\nu} \,, \tag{7.229}$$

where the coefficients η and ζ are, as in the non-relativistic theory, the *viscosity* and the *volume viscosity* of the fluid, respectively.
In this case, the equation of motion (7.224) is the correct relativistic generalization of the Navier-Stokes equation (2.42).

To conclude, we want to add a few comments on the different role that is played, in particular, by the concept of mass when it comes to comparing non-relativistic and relativistic physics. In non-relativistic physics, mass is a separate conserved quantity, independent from other conserved quantities such as energy or momentum, and moreover it can only be transported by convection and not by conduction. But in the transition to relativistic physics, it completely loses its particular status. Indeed, it may at first sight seem plausible to formulate a relativistic version of the non-relativistic conservation law for mass, replacing the ordinary mass density by the rest mass density and requiring this to be a four-scalar field. However, this procedure fails due to the fact that rest mass is not an extensive quantity. (For example, the rest mass of a helium kernel is less than the sum of the rest masses of two protons and two neutrons.) Therefore, a conservation law (or even a continuity equation) for rest mass cannot exist. And indeed, the naive relativistic generalization of this conservation law, namely the equation $\partial_\mu(\rho u^\mu) = 0$, is simply wrong. Rather, eq. (7.223) shows that even in the absence of external forces and of friction, the rest mass current density ρu^μ is not conserved: the work done by the pressure can change the energy inside a given volume. The explanation is, of course, that in relativistic physics, mass is equivalent to energy which, mathematically, is not a four-scalar but rather the time component of a four-vector and which, physically, is a sum of various contributions that are subject to exchange: one of them is the rest energy of the particles that make up the fluid, but – in contrast to the non-relativistic theory – this is not separately conserved.

In this situation, it seems reasonable to replace mass density by particle number density since this will be conserved even in the relativistic theory, at least as long as particle number changing processes of creation, annihilation or transformation may be disregarded. Under this assumption, the particle number density n, as measured in the momentary rest frame of the fluid, will be conserved in the sense that the particle number current four-vector density $n^\mu = n u^\mu$ satisfies

$$\partial_\mu \left(n u^\mu \right) \;=\; 0 \; . \tag{7.230}$$

Expanding the first term in eq. (7.222) by n and inserting eq. (7.230) gives

$$u^\mu \, \partial_\mu \Big(\frac{\rho c^2 + p}{n} \Big) \;-\; \frac{1}{n} \, u^\nu \, \partial_\nu p \;+\; \frac{1}{n} \, u_\nu \, \partial_\mu \sigma'^{\,\mu\nu} \;=\; \frac{1}{n} \, u_\nu f^\nu \; , \tag{7.231}$$

or after transforming the friction term using eq. (7.216) and eq. (7.226)

$$u^\mu \, \partial_\mu \Big(\frac{\rho c^2 + p}{n} \Big) \;-\; \frac{1}{n} \, u^\nu \, \partial_\nu p \;=\; \frac{1}{n} \, u_\nu f^\nu \;+\; \frac{1}{2n} \, (\partial^\perp u)_{\mu\nu} \, \sigma'^{\,\mu\nu} \; . \tag{7.232}$$

In the language of thermodynamics, this equation is a statement about the variation of the enthalpy per particle along the flow lines. Taking into account the identity

$$dH \;=\; T \, dS \;+\; V \, dp$$

(which holds for constant particle number), it means that the variation of the entropy per particle along the flow lines is given by

$$T \, u^\mu \partial_\mu \Big(\frac{S}{N} \Big) \;=\; \frac{1}{n} \, u_\nu f^\nu \;+\; \frac{1}{2n} \, (\partial^\perp u)_{\mu\nu} \, \sigma'^{\,\mu\nu} \; . \tag{7.233}$$

In particular, for Newtonian fluids,

$$\begin{aligned} T \, u^\mu \partial_\mu \Big(\frac{S}{N} \Big) \;&=\; \frac{1}{n} \, u_\nu f^\nu \\ &\quad + \frac{1}{2n} \Big(\eta \, (\partial_0^\perp u)_{\mu\nu} (\partial_0^\perp u)^{\mu\nu} \;+\; \tfrac{3}{2} \, \zeta \, (\partial_1^\perp u)_{\mu\nu} (\partial_1^\perp u)^{\mu\nu} \Big) \; . \end{aligned} \tag{7.234}$$

Observe that the second term, being a sum of squares, cannot be negative (provided, of course, that the viscosity coefficients η and ζ are non-negative). Thus if the four-force density is orthogonal to the four-velocity field,[5] this shows that friction generates entropy: the entropy along the flow lines is constant for ideal fluids and monotonically increasing for Newtonian fluids – in perfect agreement with the second fundamental theorem of thermodynamics.

[5] This happens, e.g., for electromagnetic forces, where f^μ is proportional to $u_\nu F^{\mu\nu}$, so that due to the antisymmetry of the electromagnetic field tensor, $u_\mu f^\mu$ must vanish.