Special Relativity

Fabian Hassler

2024

Prof. Dr. Fabian Hassler
Institute for Quantum Information
RWTH Aachen

Please report typos and other errors to F. Hassler.

This work is licensed under CC BY-NC 4.0. The translation has been done with the help of Mistral Small 4. In case of ambiguities, it might be helpful to consult the German version.

Contents

1 Relativity Principle **1**
- 1.1 Galilean Relativity Principle . . . 2
  - 1.1.1 Inertial frame . . . 2
  - 1.1.2 Galilean transformation . . . 3
  - 1.1.3 Covariance of Newton's equations . . . 3
- 1.2 Einstein's Relativity Principle . . . 4
  - 1.2.1 Michelson-Morley experiment . . . 4
  - 1.2.2 Einstein's postulate . . . 6

2 Lorentz Transformation **9**
- 2.1 Transformation between inertial frames . . . 9
- 2.2 Lorentz Group . . . 11
  - 2.2.1 Proper-orthochronous Lorentz Group . . . 12
- 2.3 Structure of the Poincaré Group . . . 15
- 2.4 Addition of Velocities . . . 15
- 2.5 Minkowski Diagrams . . . 18
  - 2.5.1 Relativistic Causality . . . 18
  - 2.5.2 Lorentz Boosts . . . 18

3 Relativistic Mechanics **21**
- 3.1 Four-velocity . . . 21
- 3.2 Energy-Momentum Relation . . . 22
- 3.3 Equation of Motion . . . 23
  - 3.3.1 Force-free Particle . . . 24
  - 3.3.2 Particle in a Potential . . . 24
- 3.4 Angular momentum . . . 28

4 Maxwell's Equations **31**
- 4.1 Invariants of the Field . . . 31
- 4.2 Homogeneous Maxwell Equations . . . 32
- 4.3 Action of the Electromagnetic Field . . . 32
- 4.4 Inhomogeneous Maxwell Equations . . . 34

4.5 Transformation of the Electromagnetic Fields . . . 35
4.6 Field of a Moving Charge . . . 35
4.7 Doppler Effect . . . 38
4.8 Energy-Momentum Tensor . . . 39
4.8.1 Conservation laws in the presence of charged matter . . . 41
4.8.2 Energy-momentum tensor of macroscopic bodies . . . 43
4.8.3 Ideal fluid . . . 44
4.9 Thermodynamics . . . 45
4.9.1 Blackbody radiation . . . 45
4.9.2 Ideal gas . . . 46

5 General Relativity **49**
5.1 Geodesic equation . . . 49
5.2 Equivalence principle . . . 51
5.3 Distances and time intervals . . . 52
5.4 Accelerated reference frames . . . 54
5.4.1 Rest frame of a uniformly accelerated observer . . . 55
5.4.2 Geodesics . . . 56
5.4.3 Unruh effect . . . 57
5.5 Rotating reference frame . . . 58
5.6 Weak gravitational field . . . 60
5.7 Schwarzschild metric . . . 60
5.7.1 Eddington-Finkelstein metric . . . 63
5.7.2 Non-relativistic limit . . . 63
5.7.3 Metric near the event horizon . . . 64
5.7.4 Geodetics . . . 64
5.7.5 Perihelion precession . . . 65
5.7.6 Light deflection . . . 67
5.7.7 Stability of circular orbits . . . 67

A Tensors **69**

References

J. D. Jackson
Classical Electrodynamics
(John Wiley & Sons, New York, 1998)
good, this is the standard work on electrodynamics

A. M. Steane
Relativity Made Relatively Easy
(Oxford University Press, Oxford, 2012)
modern, explains the concepts, both special and general relativity

W. Pauli
Relativity Theory
(Springer, Berlin, 2000)
classic, compact, also includes general relativity

W. Nolting
Basic Course: Theoretical Physics 4
(Springer, Berlin, 2002)
includes exercises with solutions

T. Fließbach
General Relativity
(Springer, Berlin, 2012)
contains only the special relativity needed for general relativity

L. D. Landau, E. M. Lifschitz,
Textbook of Theoretical Physics, Volume II: Classical Field Theory
Akademie Verlag Berlin, 11th edition (1976)
classic

Chapter 1

Relativity Principle

In this introductory chapter we review part of classical mechanics with special emphasis on the principle of relativity, in preparation for the *special theory of relativity* (SR). Physical processes take place in three-dimensional affine Euclidean space. Affine in this context means that there is no distinguished origin of coordinates. Only by fixing an (arbitrary) origin does the space become a vector space $\mathbb{R}^3$.

The position of an object in this space is described by the *position vector* $\boldsymbol{r}$. To describe the motion of the object we introduce the time parameter t so that $\boldsymbol{r}(t)$ represents the trajectory of the object. In classical mechanics t is a global parameter measured in units of a periodic motion (e.g. a pendulum or atomic clock). Once a unit of time has been agreed upon, times in different reference frames differ at most by a constant.

The position vector depends on the choice of origin and therefore cannot appear in physical laws. Only vectors such as the velocity vector $\boldsymbol{v}(t) = \frac{d\boldsymbol{r}(t)}{dt} = \dot{\boldsymbol{r}}$ may appear, since this is independent of a particular choice of origin.[1] Vectors are not measurable quantities. To specify a point in position space one needs, besides the choice of origin, the fixing of a unit (e.g. the prototype meter) and the choice of coordinate axes. Three mutually perpendicular rulers of unit length thus form a coordinate system. We describe this by the unit vectors $\boldsymbol{e}_i, i = 1, 2, 3$, each pointing from the start to the end of the respective ruler. Because the vectors form an orthonormal system they satisfy $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$. After choosing a basis, the vector can be uniquely represented by its Cartesian coordinates r_i, where

$$\boldsymbol{r} = \sum_{i=1}^{3} r_i \boldsymbol{e}_i \equiv r_i \boldsymbol{e}_i \,. \tag{1.1}$$

holds.[2]

[1] In general relativity the velocity vector becomes an element of the tangent space at a point $\boldsymbol{r}$.

[2] Note the Einstein summation convention: repeated indices are summed over.

We may now ask how the coordinates of two such coordinate systems are related in general. To this end we introduce a second orthonormal basis $\boldsymbol{e}'_i$ (chosen by another observer). Since the basis $\boldsymbol{e}_i$ is complete, the new basis vectors $\boldsymbol{e}'_i$ can be written as linear combinations of the old basis vectors $\boldsymbol{e}_i$, i.e. we have $\boldsymbol{e}'_i = R_{ij}\boldsymbol{e}_j$ with matrix elements $R_{ij} \in \mathbb{R}$.[3] From the requirement that both the old and the new coordinate system are orthonormal we obtain

$$\delta_{ij} = \boldsymbol{e}'_i \cdot \boldsymbol{e}'_j = R_{ik}R_{jl}\underbrace{\boldsymbol{e}_k \cdot \boldsymbol{e}_l}_{=\delta_{kl}} = R_{ik}R_{jk} = (RR^t)_{ij}, \tag{1.2}$$

i.e. the transformation is orthonormal $R \in \mathrm{O}(3)$ with $R^{-1} = R^t$. The new coordinates r'_i are obtained from

$$r_j\boldsymbol{e}_j = \boldsymbol{r} = r'_i\boldsymbol{e}'_i = r'_iR_{ij}\boldsymbol{e}_j, \tag{1.3}$$

i.e. $r_j = (R^t)_{ji}r'_i$, which is equivalent to $r'_j = R_{ji}r_i$. In the following we use the notation $\boldsymbol{r} = (r_1, r_2, r_3)$ for the coordinates of the vector $\boldsymbol{r}$ (and not for the vector itself). This allows us to write the transformation rule compactly as $\boldsymbol{r}' = R\boldsymbol{r}$.

Choosing a (possibly time-dependent) coordinate system and a time measurement defines a reference frame. A point in spacetime described by $(t, \boldsymbol{r})$ in a reference frame is called an *event*.

1.1 Galilean Relativity Principle

The relativity principle is based on the fact that there is no preferred reference frame. The motion of a physical object should depend only on its (relative) position to other objects and not on the choice of reference frame. Moreover one requires that the physical laws take the same form in certain reference frames.

1.1.1 Inertial frame

The form of the physical laws generally depends on the reference frame. In classical mechanics there exist distinguished reference frames in which the equations of motion take a particularly simple form. Such systems are called inertial frames and in them Newton's law

$$\frac{d\boldsymbol{p}}{dt} = m\frac{d}{dt}\boldsymbol{v} = \boldsymbol{F}, \tag{1.4}$$

holds, where m denotes the mass of the particle, $\boldsymbol{p}$ the momentum, and $\boldsymbol{F}$ the force acting on the particle. In particular, force-free particles move in straight lines in inertial frames according to

$$\boldsymbol{r} = \boldsymbol{v}_0t + \boldsymbol{r}_0. \tag{1.5}$$

Empirically the fixed stars constitute an inertial frame.

[3] As an example consider a reference frame S' rotated by θ about the $\boldsymbol{e}_3$-axis of S. By elementary geometry we obtain the expression $R(\theta\boldsymbol{e}_3) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

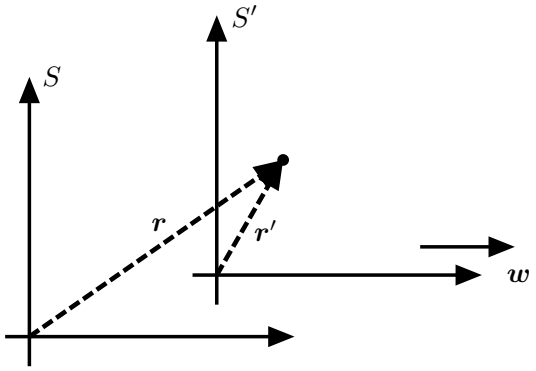

Figure 1.1: Two reference frames S and S' moving relative to each other with velocity $\boldsymbol{w}$. The coordinates $\boldsymbol{r}$ and $\boldsymbol{r}'$ of an object in the two frames are related by $\boldsymbol{r}' = \boldsymbol{r} - \boldsymbol{w}t + \boldsymbol{d}_0$ and $t' = t + t_0$ ($R = \mathbb{1}_3$).

1.1.2 Galilean transformation

We may now ask how, starting from an inertial frame S, one can find further inertial frames S'. To this end we study general transformations $\boldsymbol{r} \mapsto \boldsymbol{r}' = \boldsymbol{r}'(\boldsymbol{r}, t) = R(t)\boldsymbol{r}(t) + \boldsymbol{d}(t)$ and $t \mapsto t' = t + t_0$ with $R(t) \in \mathrm{O}(3)$. Considering the case of a force-free point mass, we require

$$0 = \frac{d^2\boldsymbol{r}'}{dt'^2} = \frac{d^2\boldsymbol{r}'}{dt^2} = \ddot{R}\boldsymbol{r} + 2\dot{R}\dot{\boldsymbol{r}} + R\ddot{\boldsymbol{r}} + \ddot{\boldsymbol{d}} = \ddot{R}\boldsymbol{r} + 2\dot{R}\dot{\boldsymbol{r}} + \ddot{\boldsymbol{d}}. \tag{1.6}$$

where we have used the general parametrization (1.5) of a force-free point mass. Since this equation must hold for all possible trajectories, we must require $\dot{R} = 0$ and $\ddot{\boldsymbol{d}} = 0$, i.e. $R(t) = R$ and $\boldsymbol{d}(t) = -\boldsymbol{w}t + \boldsymbol{d}_0$.

We have thereby determined all possible transformations mapping one inertial frame to another. These transformations

$$t' = t + t_0, \qquad \boldsymbol{r}' = R\boldsymbol{r} - \boldsymbol{w}t + \boldsymbol{d}_0, \tag{1.7}$$

are called Galilean transformations and form a 10-parameter group with parameters t_0 (1), R (3), $\boldsymbol{w}$ (3) and $\boldsymbol{d}_0$ (3 parameters). We thus see that any two inertial frames move relative to one another with constant velocity $\boldsymbol{w}$ and otherwise differ only by a shift $\boldsymbol{d}_0$ and a rotation R of the coordinate axes, see Figure 1.1.

1.1.3 Covariance of Newton's equations

An equation $G(X_{\boldsymbol{r},t}) = 0$ is called covariant or form-invariant under Galilean transformations if, when expressed in terms of the transformed quantities $X'_{\boldsymbol{r}',t'}$, it retains the same form, i.e. $G(X_{\boldsymbol{r},t}) = 0$ is equivalent to $G(X'_{\boldsymbol{r}',t'}) = 0$. We can therefore describe the Galilean relativity principle by the two equivalent statements:

- *The laws of nature have the same form in all inertial frames.*

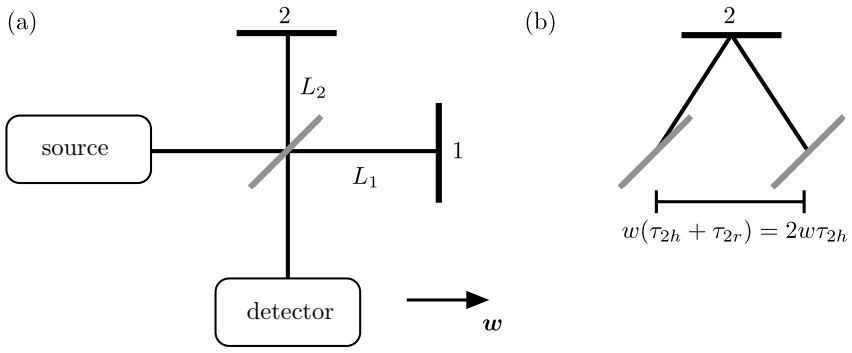

Figure 1.2: (a) Schematic setup of the Michelson interferometer to measure the Earth's velocity $\boldsymbol{w}$ relative to the stationary aether. Light emitted from the source first passes through a semi-transparent beam splitter (gray). After the two beams have traveled distances $2L_1$ and $2L_2$, their interference is recorded at the detector. In the simplest case the Earth's direction of motion is along arm 1. The propagation of light in arm 2 as seen from a stationary frame is shown in (b).

- *The laws of nature are Galilean covariant, i.e. form-invariant under Galilean transformations.*

It follows immediately that an observer cannot decide by measurements in which inertial frame he is located. In particular all inertial frames are equivalent.

Newton's equations $F = m\ddot{\boldsymbol{r}}$ are covariant provided $\boldsymbol{F}$ is a vector field, i.e. $\boldsymbol{F}$ transforms (pointwise) like a vector

$$\boldsymbol{F}'(\boldsymbol{r}') = R\boldsymbol{F}(\mathrm{r}), \tag{1.8}$$

where $\boldsymbol{r}'$ and $\boldsymbol{r}$ are related by a Galilean transformation (1.7). It is straightforward to verify that $\ddot{\boldsymbol{r}}' = R\ddot{\boldsymbol{r}}$. We can thus demonstrate the covariance of Newton's equation

$$\boldsymbol{F}'(\boldsymbol{r}') = m'\ddot{\boldsymbol{r}}' \quad \Leftrightarrow \quad R\boldsymbol{F}(\boldsymbol{r}) = m'R\ddot{\boldsymbol{r}} \quad \Leftrightarrow \quad \boldsymbol{F}(\boldsymbol{r}) = m\ddot{\boldsymbol{r}}, \tag{1.9}$$

where we had to assume that mass is invariant: $m = m'$.

1.2 Einstein's Relativity Principle

1.2.1 Michelson-Morley experiment

Maxwell's equations describe electromagnetic waves propagating at the speed of light $c = 299792458$ m/s. However, it was unclear with respect to which reference frame this speed is defined. From Galilean transformations it is easy to see that for unrotated coordinate axes ($R = \mathbb{1}_3$) velocities simply add. Consider an object moving along the path $\boldsymbol{r}'(t) = \boldsymbol{v}'t$ with velocity $\boldsymbol{v}'$ in frame S'. The same object then moves in frame S with velocity

$$\boldsymbol{v} = \dot{\boldsymbol{r}} = \boldsymbol{w} + \boldsymbol{v}'. \tag{1.10}$$

In particular, light propagates at different speeds in different reference frames.

Michelson (Nobel Prize 1907) therefore devised an experimental setup to determine in which reference frame (the aether) light waves propagate isotropically at speed c. He compared the travel times of light over distances $2L_1$, $2L_2$ in two orthogonal directions in an interferometer, see Figure 1.2(a). Since the Earth moves at $w \simeq 30$ km/s around the Sun, the travel-time differences should vary over the course of a year.

If the Earth's velocity $\boldsymbol{w}$ with respect to the aether is parallel to arm 1, we obtain the travel time $\tau_{1h} = L_1/(c-w)$ for the outward path and $\tau_{1r} = L_1/(c+w)$ for the return path. The total travel time of the light ray in arm 1 is therefore given by

$$\tau_1 = \frac{2L_1}{c}\frac{1}{1-w^2/c^2}, \tag{1.11}$$

, i.e. corrected by the factor $(1-w^2/c^2)^{-1}$ from the naive value $2L_1/c$ without taking into account the relative motion with respect to the aether.

The travel time in arm 2, which is perpendicular to the Earth's motion, is the same for the outward and return paths $\tau_{2h} = \tau_{2r}$ with $\tau_2 = 2\tau_{2h}$. With respect to the aether, the speed is simply c. However, in this reference frame the path length through the moving mirrors is longer. The total path length is given by $\ell = 2\left[L_2^2 + w^2\tau_{2h}^2\right]^{1/2}$, cf. Figure 1.2(b). Solving the equation $\tau_2 = \ell/c$ for τ_2 yields

$$\tau_2 = \frac{2L_2}{c}\frac{1}{\sqrt{1-w^2/c^2}} \tag{1.12}$$

, i.e. a reduction by another factor $(1-w^2/c^2)^{-1}$.

The interferometer measures the difference in optical path lengths

$$\delta = c(\tau_2 - \tau_1) = \frac{2}{1-w^2/c^2}\left(L_2\sqrt{1-w^2/c^2} - L_1\right) \tag{1.13}$$

with an accuracy of $\lambda \simeq 500$ nm set by the wavelength of the light. After a quarter year the apparatus is effectively rotated by 90° and we obtain a new path-length difference

$$\delta' = c(\tau_2 - \tau_1) = \frac{2}{1-w^2/c^2}\left(L_2 - L_1\sqrt{1-w^2/c^2}\right),$$

since the roles of L_1 and L_2 have meanwhile been exchanged. The difference between the path-length differences with a quarter-year time difference is given by

$$|\delta' - \delta| = \frac{2(L_2+L_1)}{1-w^2/c^2}\left(1-\sqrt{1-w^2/c^2}\right) \overset{w\ll c}{\approx} (L_1+L_2)\frac{w^2}{c^2}. \tag{1.14}$$

It is interesting to see what length of the interferometer arms is required to measure the effect of the Earth's relative motion to the aether. With $w/c \simeq 10^{-4}$ and $\delta' - \delta \simeq \lambda \simeq 10^{-7}$ m we obtain $L_1 + L_2 \simeq 10$ m.

Michelson found that the interference pattern is always the same, independent of the season and hence of the orientation of the interferometer, from which one must conclude that there is no aether (as a preferred reference frame) and that the speed of light is the same in all reference frames.

1.2.2 Einstein's postulate

The Michelson-Morley experiment led Einstein (1905) to require, in addition to the Galilean principle of relativity that the laws of physics have the same form in all inertial frames, the additional postulate

- *the speed of light has the same value c in all inertial frames.*

This principle is the foundation of the special theory of relativity. Inertial frames are defined so that free particles, like in classical mechanics, obey the law of inertia $\ddot{\boldsymbol{r}} = 0$. One sees immediately that the covariance of the physical laws can no longer be required under Galilei transformations, since these lead to the velocity-addition formula (1.10). In Chapter 2 we will investigate which transformation group is compatible with Einstein's postulate. In the following, however, we first want to become acquainted with a few simple consequences of Einstein's postulate.

As we saw in Section 1.1.2 that, under the assumption of an absolute time, the principle of relativity immediately yields the Galilei transformations as the general transformations between inertial frames, we must abandon the notion of absolute time and require in general that $t' \neq t$. The invariant value of the speed of light makes it possible to synchronize clocks 1 and 2 in the two frames S and S' and thereby to introduce, at least in part, a common time.[4]

The protocol for synchronizing two clocks in two distinct inertial frames S and S', which measure time by the same periodic process, is as follows: imagine that in frame S a light flash is emitted from clock 1 at time t_0. The beam is reflected by clock 2 (at rest in frame S') and detected by clock 1 at time $t_0 + \Delta t$. Because the speed of light is invariant, we can identify the instant of arrival of the light ray at clock 2 with the time $t_0 + \Delta t/2$ in frame S and thereby synchronize the clocks. Furthermore, the universality of the speed of light allows us to reduce length measurements to time measurements and to define the meter as the 1/299792458-th part of the distance a light ray travels in one second. The clock-synchronization problem is that the procedure is not transitive: if clocks 1 and 2 are synchronous and likewise 2 and 3, it does not follow that 1 and 3 are synchronous. This leads us to the concept of the relativity of simultaneity.

[4] As we shall see later, clocks that have been synchronized at one instant will, after having moved apart again, no longer agree.

Relativity of Simultaneity

From the absolute time in Galilean spacetime it follows that two events which occur simultaneously in the reference frame S' also occur simultaneously in S. This fact is incompatible with Einstein's postulate of the universality of the speed of light. To see this, consider two reference frames moving relative to each other with velocity $\boldsymbol{w} = (w, 0, 0)$, $w > 0$. We consider three clocks A, B, C on the x-axis that are at rest in S'. Clock B lies between clocks A and C so that the distance from A to B equals the distance from B to C. An observer in frame S' will find that a light flash emitted from B at time t'_{B} arrives simultaneously ($t'_{\mathrm{A}} = t'_{\mathrm{C}}$) at A and C.

From the viewpoint of an observer in the (moving) frame S, the light flash is emitted at time t_{B} from B and propagates with c in all directions. Since A moves toward the light flash while C moves away from it, from the perspective of S the inequality $t_{\mathrm{A}} < t_{\mathrm{C}}$ holds. Thus clocks synchronized in frame S' are not synchronized in frame S, and hence the notion of "simultaneity" depends on the reference frame. The absence of an absolute simultaneity concept is the origin of many of the initially paradoxical consequences of the relativity principle.

Time Dilation

Now consider a clock A at rest in frame S' that sends a light pulse to a mirror B at distance L_0 (orthogonal to the relative direction of motion $-\boldsymbol{w}$ of frame S). This light pulse will return to clock A after the time $\Delta t' = 2L_0/c$. In frame S the mirror B is in motion, so the light ray must traverse the path $L = \sqrt{4L_0^2 + w^2(\Delta t)^2}$, as shown in Figure 1.2(b). By Einstein's postulate the light path c is related to the time difference Δt by $L = c\,\Delta t$. Solving this equation for Δt yields the result

$$\Delta t = \frac{2L_0}{\sqrt{c^2 - w^2}} = \gamma \Delta t' \tag{1.15}$$

with the time-dilation factor

$$\gamma = \frac{1}{\sqrt{1 - w^2/c^2}} > 1. \tag{1.16}$$

From the viewpoint of S the process therefore lasts longer by the factor γ. The time measured by the clock at rest is also called the *proper time* $\Delta\tau$. It is always less than the time difference in a moving frame. If an object moves in S with variable speed $\boldsymbol{v}(t)$ from $x_{\mathrm{A}} = (t_{\mathrm{A}}, \boldsymbol{r}_{\mathrm{A}})$ to $x_{\mathrm{B}} = (t_{\mathrm{B}}, \boldsymbol{r}_{\mathrm{B}})$, one may introduce at each instant t an instantaneous rest frame S'_t with $\boldsymbol{w} = \boldsymbol{v}(t)$ and $\Delta t' = \Delta\tau = \sqrt{1 - v(t)^2/c^2}\,\Delta t$. The clock comoving with the object measures the proper time

$$\tau = \sum \Delta\tau = \int_{t_{\mathrm{A}}}^{t_{\mathrm{B}}} dt\, \sqrt{1 - v(t)^2/c^2}. \tag{1.17}$$

The proper time τ is a property of the worldline and hence independent of the inertial frame. This is immediately clear since τ is physically determined by the reading of the comoving clock.

When two observers move relative to each other, each concludes that the other's clock runs slower. This seems paradoxical at first, yet it leads to no contradiction because simultaneity is relative.

Time dilation is crucial to understanding muon decay. Cosmic rays produce muons at about 10 kilometers altitude. These muons have a relatively short lifetime of $\simeq \mu$s. In that time they would travel only a few hundred meters. Nevertheless many muons are detected at ground level. This result can be understood by noting that the muons' speed is nearly the speed of light, so $w = 0.994\,c$ implies $\gamma \simeq 10$. Thus the muons' lifetime in the Earth frame is ten times longer, and they travel several kilometers before decaying. Viewing the problem from the muons' rest frame leads us to the concept of length contraction, which we treat next.

Length Contraction

As in the preceding section, consider an arrangement at rest in frame S' consisting of a clock A and a mirror B separated by L_0. This time, however, the light path is parallel to the relative motion of the two frames (i.e., the clock assembly is rotated by 90° compared with the previous setup). The distance between A and B is determined by the light-travel time $\Delta t'$ over $L_0 = c\,\Delta t'/2$. Now consider the system in the moving frame S. Let Δt_h be the light-travel time for the outward trip (from A to B). In that time the mirror has moved by $w\,\Delta t_h$. The constancy of the speed of light requires therefore that $c\,\Delta t_h = L + w\,\Delta t_h$ with L the separation of the two mirrors in frame S. One thereby obtains for the outward-trip time $\Delta t_h = L/(c - w)$. A similar argument for the return trip yields $\Delta t_r = L/(c + w)$.

Through addition of these two equations one obtains for the total travel time the expression $\Delta t = \Delta t_h + \Delta t_r = 2\gamma^2 L/c$. On the other hand, by virtue of the time-dilation formula $\Delta t = \gamma \Delta t'$. Solving the equation $2L_0 = c\Delta t' = 2\gamma L$ for L yields the length-contraction formula

$$L = L_0\sqrt{1 - w^2/c^2} = L_0/\gamma; \tag{1.18}$$

i.e., a moving rod appears shortened relative to its rest state by the factor γ (in the direction of motion). With length contraction we can interpret the muon problem from the muon's rest frame as a contraction of the distance between Earth's surface and the muons' production site. In this way length contraction and time dilation are mutually entailed and together lead to a consistent relativistic description.

Chapter 2

Lorentz Transformation

In the previous chapter we saw that Einstein's postulate, supported by the Michelson-Morley experiment, forces us to rethink our notions of time and space. In particular, we must abandon the idea of an absolute time. In this chapter we shall develop Einstein's postulate into a mathematical theory that allows us to formulate new laws of nature compatible with that postulate. To do so, we must first determine what the new transformations between inertial frames look like.

Just as the Galilei transformations leave Newton's equations invariant, the Lorentz transformations are the group of transformations that leave the law of inertia and Einstein's postulate invariant. In this sense, the Lorentz transformations enable one to pass from one inertial frame to any other inertial frame in special relativity. The next task—taken up in the following chapter—is to find a relativistic (i.e. Lorentz-invariant) formulation of mechanics and electrodynamics. It will turn out that the Maxwell equations, with a suitable transformation of the fields, are already Lorentz-invariant. By contrast, Newtonian mechanics requires modification, because action-at-a-distance laws (e.g. Newton's law of gravitation) are a priori non-relativistic: they refer to "the distance between two bodies at the same time," a notion incompatible with the relativity of simultaneity. Hence in relativity all action-at-a-distance laws must be replaced by fields governed by relativistic field equations such as those of Maxwell.

2.1 Transformation between inertial frames

At first sight Einstein's postulate seems to contradict the law of inertia. This impression arises only because we are deeply rooted in our classical world picture, and—as seen in the last chapter—Einstein's postulate entails the loss of an absolute notion of simultaneity and therefore introduces a time coordinate for each frame. To obtain the group of Lorentz transformations, consider two inertial frames S and S' moving relative to each other with velocity $\boldsymbol{w}$.

Consider a light flash emitted at $(t_1, \boldsymbol{r}_1)$ and arriving later at $(t_2, \boldsymbol{r}_2)$. By Einstein's

postulate both

$$(\Delta \boldsymbol{r})^2 - c^2(\Delta t)^2 = 0 \qquad \text{and} \qquad (\Delta \boldsymbol{r}')^2 - c^2(\Delta t')^2 = 0 \tag{2.1}$$

hold with $\Delta t = t_2 - t_1$ and $\Delta \boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$. The motion of a force-free particle in any inertial frame is a straight line, see (1.5). Requiring that a Lorentz transformation map inertial frames to one another immediately implies that it maps straight lines to straight lines. This is equivalent to the Lorentz transformation being a linear map from $(\Delta t, \Delta \boldsymbol{r})$ to $(\Delta t', \Delta \boldsymbol{r}')$. By linearity, (2.1) is equivalent to[1]

$$(\Delta \boldsymbol{r}')^2 - c^2(\Delta t')^2 = \kappa(\boldsymbol{w})[(\Delta \boldsymbol{r})^2 - c^2(\Delta t)^2]. \tag{2.2}$$

Now consider a further frame S'' moving with velocity $-\boldsymbol{w}$ relative to S'. We obtain

$$(\Delta \boldsymbol{r}'')^2 - c^2(\Delta t'')^2 = \kappa(-\boldsymbol{w})\kappa(\boldsymbol{w})[(\Delta \boldsymbol{r})^2 - c^2(\Delta t)^2].$$

Because S'' is at rest relative to S we have $\Delta \boldsymbol{r} = \Delta \boldsymbol{r}''$ and $\Delta t = \Delta t''$, so $\kappa(-\boldsymbol{w})\kappa(\boldsymbol{w}) = 1$. By spatial isotropy $\kappa(\boldsymbol{w})$ can depend only on the magnitude w of the relative velocity, not on its direction. Hence $\kappa(-\boldsymbol{w})\kappa(\boldsymbol{w}) = \kappa(w)^2 = 1$ and therefore $\kappa(w) = 1$.[2]

These results are best summarized by combining space and time into the *four-coordinates*

$$x = (x^\mu) = (x^0, x^1, x^2, x^3) = (ct, \boldsymbol{r}) \tag{2.3}$$

of a particle. In this way the Lorentz transformations are defined by their leaving invariant the squared interval

$$\Delta s^2 = (\Delta x') \cdot (\Delta x') = (\Delta x) \cdot (\Delta x) = (\Delta x)_\mu (\Delta x)^\mu = \eta_{\mu\nu}(\Delta x)^\mu (\Delta x)^\nu \tag{2.4}$$

[3] with the *Minkowski metric*

$$(\eta_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.5}$$

A general affine transformation between the reference frames S and S' can be written as $x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu$ with $a \in \mathbb{R}^4$ and $\Lambda \in \mathrm{GL}(4, \mathbb{R})$. The condition (2.4) leads to the restriction

$$\eta_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta (\Delta x)^\alpha (\Delta x)^\beta = \eta_{\mu\nu}(\Delta x')^\mu (\Delta x')^\nu = \eta_{\mu\nu}(\Delta x)^\mu (\Delta x)^\nu \,,$$

[1] A linear map carries a homogeneous polynomial of degree 2 to another homogeneous polynomial of degree 2. Since $(\Delta \boldsymbol{r})^2 - c^2(\Delta t)^2 = (\Delta \boldsymbol{r}')^2 - c^2(\Delta t')^2 = 0$, homogeneity at once yields (2.2).

[2] The alternative solution $\kappa(w) = -1$ is ruled out because $\kappa = 1$ at $w = 0$.

[3] Just as the rotations R leave the distance $\sqrt{(\Delta \boldsymbol{r})^2}$ invariant.

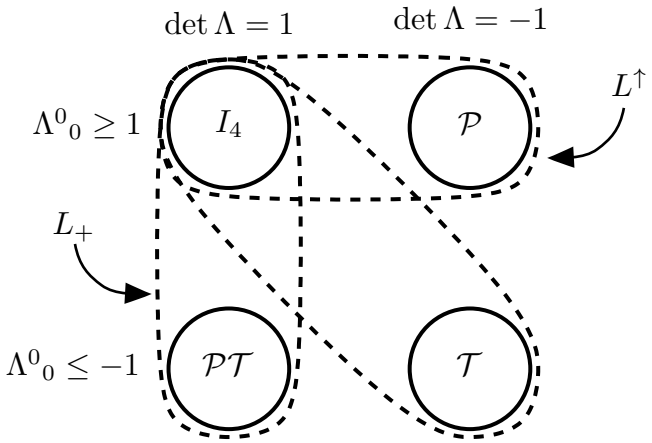

Figure 2.1: Structure of the Lorentz group: the Lorentz group L decomposes into four disjoint components characterized by $\det\Lambda = \pm 1$ and $\mathrm{sgn}(\Lambda^0{}_0)$.

which is equivalent to

$$\eta_{\mu\nu}\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta} = \eta_{\alpha\beta} \tag{2.6}$$

or in matrix notation $\Lambda^t\eta\Lambda = \eta$. Transformations which satisfy the condition (2.6) are called elements of the Lorentz group, denoted by $L = \mathrm{O}(1,3)$. The group of transformations that connect inertial systems with fixed scales is the group of inhomogeneous Lorentz transformations with $x' = \Lambda x + a$, also called the *Poincaré group*.

2.2 Lorentz Group

We consider an element $\Lambda \in L$. From the relation (2.6) a number of properties of Λ can be derived. Taking the determinant immediately yields $\det^2\Lambda = 1$. Thus L consists of two components characterized by $\det\Lambda = \pm 1$.[4] Moreover, one obtains for the (00)-component the property $(\Lambda^0{}_0)^2 - \sum_{k=1}^{3}(\Lambda^k{}_0)^2 = 1$, i.e. $(\Lambda^0{}_0)^2 \geq 1$. Hence the Lorentz group contains two disjoint components characterized by the sign of $\Lambda^0{}_0$.

[4]The inverse of Λ is obtained by multiplying on the right by $\Lambda^{-1}\eta$ as $\Lambda^{-1} = \eta\Lambda^t\eta$.

That all four cases occur is shown by the reflections

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathcal{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$
$$\mathcal{T} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathcal{PT} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{2.7}$$

with $\mathcal{P}$ the reflection at the origin of coordinates (parity) and $\mathcal{T}$ time reversal. The reflections form a subgroup of L. Further subgroups are, for example, *proper Lorentz transformations*

$$\begin{aligned} L_+ &= \{\Lambda \in L | \det \Lambda = 1\}, && \textit{orthochronous Lorentz transformations}, \\ L^\uparrow &= \{\Lambda \in L | \Lambda^0{}_0 \geq 1\}, && \textit{orthochronous proper Lorentz transformations}, \\ L_+^\uparrow &= L_+ \cap L^\uparrow. \end{aligned} \tag{2.8}$$

A Lorentz transformation maps by definition the light cone with $\Delta s^2 = 0$ onto itself. Vectors on the light cone with $v \cdot v = 0$ are also called *null vectors*. We further define vectors pointing to the exterior of the light cone as *spacelike* vectors v with $v \cdot v < 0$ and those pointing to the interior as *timelike* vectors with $v \cdot v > 0$. Timelike vectors can be further subdivided into vectors pointing to the timelike future ($v^0 > 0$) and those pointing to the timelike past ($v^0 < 0$).

That $L_+^\uparrow$ is a group is seen geometrically. A Lorentz transformation $\Lambda \in L$ maps the interior of the light cone to itself. In doing so, the two subcones (future and past) $V^\pm = \{x \mid x \cdot x > 0, \pm x^0 \geq 0\}$ are either left invariant or exchanged. The decisive factor is the sign $\mathrm{sgn}(\Lambda^0{}_0)$, since $[\Lambda(1, \mathbf{0})]^0 = \Lambda^0{}_0$. Thus $\mathrm{sgn}(\Lambda^0{}_0)$ and of course also $\det \Lambda$ are multiplicative[5] under the group operation. The orthochronous transformations therefore map the future to the future and the past to the past.

Because of the multiplicity of $\mathrm{sgn}(\Lambda^0{}_0)$ and $\det \Lambda$, every $\Lambda \in L$ can be written as the product of an element of the proper-orthochronous Lorentz group $L_+^\uparrow$ with a reflection $\{I_4, \mathcal{P}, \mathcal{T}, \mathcal{PT}\}$. We therefore restrict ourselves in what follows to the proper-orthochronous subgroup.

2.2.1 Proper-orthochronous Lorentz Group

The 4×4-matrix Λ has a total of 16 real entries. The equation (2.6) yields 10 independent equations.[6] Therefore elements $\Lambda \in L_+^\uparrow$ are determined by 6 real

[5] This follows from the general relation $\det(AB) = (\det A)(\det B)$ for arbitrary quadratic matrices A and B.

[6] Transposing (2.6) leads to the same equation system. Hence only 10 of the 16 equations are independent.

parameters.

Three of the parameters can be identified with the three-dimensional rotations. Indeed we have

$$\Lambda(R) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & R \end{array}\right) . \tag{2.9}$$

From this it follows immediately $\Lambda^t \eta \Lambda = \eta$ where $R \in \mathrm{SO}(3)$ with $R^t R = I_3$. Thus the rotations $\Lambda(R)$ form a subgroup of $L_+^\uparrow$.

Of greater interest are the remaining three-parameter transformations, for which we expect that they connect reference frames in relative motion. To this end we examine whether special solutions with the block form

$$\Lambda = \left(\begin{array}{cc|cc} a & b & 0 & 0 \\ c & d & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \tag{2.10}$$

exist, i.e. x^2 and x^3 are not transformed. Inserting into (2.6) yields the conditions

$$a^2 - c^2 = 1, \qquad ab - cd = 0, \qquad b^2 - d^2 = -1.$$

Since we are interested only in solutions from $L_+^\uparrow$, we additionally require $a > 0$ and $ad - bc = 1$.

The first equation can be satisfied by introducing the parameter χ with $a = \cosh\chi$ and $c = -\sinh\chi$. From the second equation we obtain $b = cd/a$. Substituting into $ad - bc = 1$ yields $1 = d/a$, i.e. $d = \cosh\chi$ and $b = -\sinh\chi$. Thus we obtain the special Lorentz transformations (boosts)

$$\Lambda(\chi) = \left(\begin{array}{cc|cc} \cosh\chi & -\sinh\chi & 0 & 0 \\ -\sinh\chi & \cosh\chi & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = \exp\left(\begin{array}{cc|cc} 0 & -\chi & 0 & 0 \\ -\chi & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) . \tag{2.11}$$

From the last identity it follows immediately that the boosts form a subgroup with the multiplication law

$$\Lambda(\chi_1 + \chi_2) = \Lambda(\chi_1)\Lambda(\chi_2). \tag{2.12}$$

From this one can directly conclude that $\Lambda(\chi)^{-1} = \Lambda(-\chi)$.

In the following we wish to explain the physical meaning of $\Lambda(\chi)$. To this end we write the transformation law $x = \Lambda(\chi)^{-1} x'$ of a boost in components

$$\begin{aligned} ct &= (\cosh\chi)ct' + (\sinh\chi)r_1', & r_2 &= r_2', \\ r_1 &= (\sinh\chi)ct' + (\cosh\chi)r_1', & r_3 &= r_3'. \end{aligned} \tag{2.13}$$

An object that is at rest at position $\boldsymbol{r}' = \boldsymbol{r}'(t')$ in the S' system therefore moves in the S system with the velocity[7]

$$w = \frac{dr_1}{dt} = \frac{dr_1}{dt'}\frac{dt'}{dt} = \frac{dr_1}{dt'}\left(\frac{dt}{dt'}\right)^{-1} = c\tanh\chi \tag{2.14}$$

Hence we obtain the interpretation that $\Lambda(\chi)$ transforms to a new inertial system moving with velocity w along the x axis relative to S, see Figure 1.1. By simple hyperbolic relations we obtain

$$\begin{aligned}\cosh\chi &= \frac{1}{\sqrt{1-\tanh\chi^2}} = \frac{1}{\sqrt{1-\beta^2}} = \gamma,\\ \sinh\chi &= \tanh\chi\cosh\chi = \frac{\beta}{\sqrt{1-\beta^2}} = \beta\gamma\end{aligned} \tag{2.15}$$

with $\beta = w/c$. From (2.13) this gives

$$\begin{aligned} t' &= \gamma(t - wx^1/c^2), & x'^2 &= x^2,\\ x'^1 &= \gamma(x^1 - wt), & x'^3 &= x^3. \end{aligned} \tag{2.16}$$

For $\beta \to 0$ we obtain the Galilei transformation as the limiting case

$$t' = t, \quad x'^1 = x^1 - wt, \quad x'^2 = x^2, \quad x'^3 = x^3.$$

A transformation to a system S' moving with velocity $\boldsymbol{w}$ in an arbitrary spatial direction relative to S can be obtained in principle by composing a rotation (so that $\boldsymbol{w}'$ lies along the x direction), a boost, and an inverse rotation. Much simpler, however, is the idea that one can decompose the vector $\boldsymbol{r}$ into a component $\boldsymbol{r}_\parallel$ along $\boldsymbol{w}$ and a component $\boldsymbol{r}_\perp$ orthogonal to it. Analogous to (2.16) one then obtains

$$t' = \gamma[t - (\boldsymbol{w}\cdot\boldsymbol{r}_\parallel)/c^2], \qquad \boldsymbol{r}'_\parallel = \gamma(\boldsymbol{r}_\parallel - \boldsymbol{w}t), \qquad \boldsymbol{r}'_\perp = \boldsymbol{r}_\perp. \tag{2.17}$$

Since $\boldsymbol{r}_\parallel = (\boldsymbol{r}\cdot\boldsymbol{w})\boldsymbol{w}/w^2$ and $\boldsymbol{r}_\perp = \boldsymbol{r} - \boldsymbol{r}_\parallel$ hold, one can write a general boost $x' = \Lambda(\boldsymbol{w})x$ as

$$\Lambda(\boldsymbol{w}) = \begin{pmatrix} \gamma & -\gamma\boldsymbol{w}^t/c \\ -\gamma\boldsymbol{w}/c & I_3 + (\gamma-1)\boldsymbol{w}\boldsymbol{w}^t/w^2 \end{pmatrix} \tag{2.18}$$

or in other words

$$\begin{aligned} t' &= \gamma[t - (\boldsymbol{w}\cdot\boldsymbol{r})/c^2],\\ \boldsymbol{r}' &= \boldsymbol{r} + \frac{\gamma-1}{w^2}(\boldsymbol{r}\cdot\boldsymbol{w})\boldsymbol{w} - \gamma\boldsymbol{w}t. \end{aligned}$$

Thus a boost in general depends on the three real parameters $\boldsymbol{w}$. We have $\Lambda(\boldsymbol{w}) = \Lambda(R_{\boldsymbol{w}})\Lambda(\chi)\Lambda(R^t_{\boldsymbol{w}})$ with the rotation $R_{\boldsymbol{w}}$ defined by $\boldsymbol{w} = R_{\boldsymbol{w}}(w,0,0)^t$.

The Lorentz transformations become singular for $w \to c$. The speed of light is a maximal speed. No object (and no information) can move faster than light.

[7] In the literature $\chi = \operatorname{artanh}(w/c)$ is called the *rapidity*.

2.3 Structure of the Poincaré Group

The group of affine Lorentz transformations is called the Poincaré group. From the foregoing considerations we know that a general element of the Poincaré group has the form

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}(\boldsymbol{w}, R)x^{\nu} + a^{\mu} \tag{2.19}$$

Like the Galilei group, the Poincaré group has 10 parameters: four parameters for the translations with a and the 6 parameters of the Lorentz group (3 for rotations R and 3 for boosts $\boldsymbol{w}$).

In fact one can write every Lorentz transformation Λ as a product of a rotation and a general boost with $\Lambda(\boldsymbol{w}, R) = \Lambda(R)\Lambda(\boldsymbol{w}) = \Lambda(R_2)\Lambda(\chi)\Lambda(R_1)$ with $R_1 = R^t_{\boldsymbol{w}}$ and $R_2 = RR_{\boldsymbol{w}}$. One can easily see that this decomposition is possible for every Lorentz transformation $x' = \Lambda x$. To this end consider the linear subspace $M = \{x \mid x^0 = x'^0 = 0\}$, formed by the intersection of the position vectors from S and S'. There are two possibilities: $\dim M = 2$ or $\dim M = 3$.[8]

In the case that $\dim M = 3$, both M and its orthogonal complement $M^{\perp} = \{x \mid x \cdot y = 0, \quad \forall y \in M\} = \{x \mid x^1 = x^2 = x^3 = 0\}$ are mapped to themselves. Hence Λ has directly the block form (2.9) and describes a pure rotation with $\boldsymbol{w} = 0$.

If $\dim M = 2$, we choose two orthonormal position vectors $\boldsymbol{f}_2, \boldsymbol{f}_3 \in M$. By a suitable rotation R_1 in position space $\{x^0 = 0\} \supset M$ we can arrange that $R_1\boldsymbol{f}_2 = (0, 1, 0)$ and $R_1\boldsymbol{f}_3 = (0, 0, 1)$, i.e. they coincide with the corresponding unit vectors $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ in the S system. Considering now the image vectors $\boldsymbol{f}'_2 = \Lambda\boldsymbol{f}_2$ and $\boldsymbol{f}'_3 = \Lambda\boldsymbol{f}_3$: by the definition of M these are again position vectors. Analogous to above we can, with a rotation R_2, achieve that $\boldsymbol{f}'_2 = R_2(0, 1, 0)$ and $\boldsymbol{f}'_3 = R_2(0, 0, 1)$.

Defining now $\Lambda' = \Lambda(R^t_2)\Lambda\,\Lambda(R^t_1)$, we then have by construction that the vectors $x \in N$ with $N = \{x \mid x^0 = x^1 = 0\}$ are mapped to themselves under Λ'. Moreover $N^{\perp} = \{x^2 = x^3 = 0\}$ is left invariant as a set under Λ'. [9] Therefore Λ' must have the block form (2.11) of a boost.

2.4 Addition of Velocities

It is immediately clear that Galilean velocity addition $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{v}'$ leads to contradictions with the absoluteness of the speed of light. This requires namely that with $v = c$ also $v' = c$, independently of $\boldsymbol{w}$. The relativistic formulas for velocity addition can be derived at once from the equations (2.17) for a general boost. We consider

[8] The image ΛB_0 of $B_0 = \{x \mid x^0 = 0\}$ is three-dimensional, since Λ is invertible. Intersecting this set with the three-dimensional subspace $B_0 = \{x' \mid x' = 0\}$ gives the intersection $M = (\Lambda B_0) \cap B_0$ of dimension $\dim M$. Since $\dim(\Lambda B_0) = \dim B_0 = 3$ it follows immediately that $\dim M \leq 3$. From $\dim M + \dim[(\Lambda B_0) + B_0] = \dim(\Lambda B_0) + \dim B_0 = 6$ and $\dim[(\Lambda B_0) + B_0] \leq 4$ it follows at once that $\dim M \geq 2$. Here $V + W$ denotes the sum of the vector spaces V and W.

[9] Since N remains invariant and Λ' leaves the "scalar product" $x \cdot y$ invariant, Λ' maps the orthogonal complement to the orthogonal complement.

two (axis-parallel) reference systems S and S' whose coordinates are related by (2.17) (i.e. they move with $\boldsymbol{w}$ relative to each other). Consider an object moving on the trajectory $\boldsymbol{r}'(t') = \boldsymbol{r}'_0 + \boldsymbol{v}'t'$ in S'. The same object will move on the trajectory $\boldsymbol{r}(t)$ in the S system. In four-vector notation we have the coordinates $x(t) = [ct, \boldsymbol{r}(t)]$ and $x(t') = [ct', \boldsymbol{r}'(t')]$, which are related by $x = \Lambda(\boldsymbol{w})^{-1}x' = \Lambda(-\boldsymbol{w})x'$. With the general formula for a boost, we obtain

$$\frac{dt}{dt'} = \gamma[1 + (\boldsymbol{w}\cdot\boldsymbol{v}')/c^2]$$

and

$$\frac{d\boldsymbol{r}}{dt'} = \boldsymbol{v}' + \frac{\gamma-1}{w^2}(\boldsymbol{v}'\cdot\boldsymbol{w})\boldsymbol{w} + \gamma\boldsymbol{w}.$$

The object therefore moves in the S system with the velocity

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \frac{\gamma^{-1}}{1 + (\boldsymbol{w}\cdot\boldsymbol{v}')/c^2}\left[\gamma\boldsymbol{w} + \frac{\gamma-1}{w^2}(\boldsymbol{v}'\cdot\boldsymbol{w})\boldsymbol{w} + \boldsymbol{v}'\right]. \tag{2.20}$$

Splitting the velocity $\boldsymbol{v}'$ into components parallel ($\boldsymbol{v}'_\parallel$) and perpendicular ($\boldsymbol{v}'_\perp$) to $\boldsymbol{w}$, one obtains the compact form

$$\boldsymbol{v} = \frac{\boldsymbol{w} + \boldsymbol{v}'_\parallel + \boldsymbol{v}'_\perp\sqrt{1-\beta^2}}{1 + (\boldsymbol{w}\cdot\boldsymbol{v}')/c^2}. \tag{2.21}$$

of the relativistic addition law of velocities. The inverse relation is obtained by replacing $\boldsymbol{w}$ with $-\boldsymbol{w}$.

In the limit $w, v' \ll c$ one recovers the nonrelativistic relation $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{v}'$. If $\boldsymbol{v}'$ and $\boldsymbol{w}$ are parallel ($\boldsymbol{v}'_\perp = 0$), the simplified formula $\boldsymbol{v} = (\boldsymbol{w} + \boldsymbol{v}')/[1 + wv'/c^2]$ results. In fact one can obtain this also directly from (2.12) with the hyperbolic addition formula $\tanh(\chi_1 + \chi_2) = (\tanh\chi_1 + \tanh\chi_2)/(1 + \tanh\chi_1\tanh\chi_2)$. If $\boldsymbol{v}$ is orthogonal to $\boldsymbol{w}$, the velocities add almost normally with $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{v}'\sqrt{1-\beta^2}$, where the extra factor γ^{-1} arises from the time dilation of the reference system S' relative to S.

It is instructive to derive the velocity addition also in polar coordinates. Describing the velocity $\boldsymbol{v}$ in the reference system S by the magnitude v and the angle θ via $\cos\theta = (\boldsymbol{v}\cdot\boldsymbol{w})/vw$ with the relative velocity of the reference systems and introducing analogously v', θ' in the reference system S' for $\boldsymbol{v}'$, one can rewrite the addition formula (2.21) as

$$v = \boldsymbol{v}\cdot\boldsymbol{v} = \frac{\sqrt{w^2 + v'^2 + 2wv'\cos\theta' - w^2v'^2\sin^2\theta'/c^2}}{1 + wv'\cos\theta'/c^2} \tag{2.22}$$

and

$$\tan\theta = \frac{v_\perp}{v_\parallel} = \frac{\sqrt{1-\beta^2}\,v'\sin\theta'}{w + v'\cos\theta'}. \tag{2.23}$$

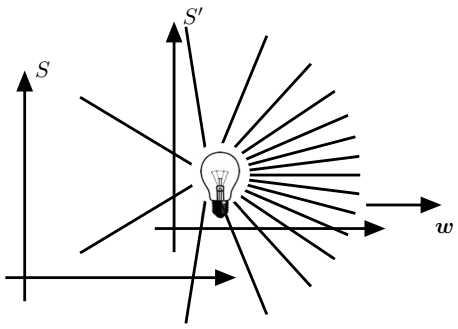

Figure 2.2: We consider a light source at rest in the S' system and emitting light uniformly in all directions. As seen in the figure, relativistic light aberration causes more light to be emitted in the forward direction when viewed from the S system. A similar effect also occurs in nonrelativistic physics if one assumes that the S' system is at rest relative to the aether. However, the relativistic and nonrelativistic formulas agree only to first order in β. Moreover the relativistic formula (2.25) for aberration is symmetric, so that an observer in the S' system observes the same effect for a light source at rest in S.

From (2.22) it follows at once that with $w, v' \leq c$ also $v \leq c$ holds. Thus the speed of light plays the role of a limiting speed that cannot be exceeded.

Next consider in detail the case of a light particle with $v' = c$. We have with $v' = c$ also $v = c$ independently of $w \leq c$ and θ', which is precisely Einstein's postulate. From the formula (2.23) one obtains in this case the relativistic formula

$$\tan\theta = \frac{\sqrt{1-\beta^2}\,\sin\theta'}{\beta + \cos\theta'} \tag{2.24}$$

for light aberration. We wish to rewrite this formula slightly so that the symmetry $(\theta, \beta) \leftrightarrow (\theta', -\beta)$ becomes manifest. From (2.21) we obtain that $\operatorname{sgn}(\cos\theta) = \operatorname{sgn}(\boldsymbol{w} \cdot \boldsymbol{v}) = \operatorname{sgn}(\beta + \cos\theta')$. Thus the sign of

$$\cos\theta = \pm(1 + \tan^2\theta)^{-1/2} = \frac{\beta + \cos\theta'}{1 + \beta\cos\theta'}$$

is fixed. Taking these two equivalent equations together yields the symmetric form

$$\tan\frac{\theta}{2} = \frac{\tan\theta}{1 + 1/\cos\theta} = \sqrt{\frac{1-\beta}{1+\beta}}\,\tan\frac{\theta'}{2} \tag{2.25}$$

for light aberration, see Figure 2.2. From this aberration formula one can derive the transformation formula for a solid-angle element $d\Omega = d(\cos\theta)\,d\varphi$ of a light source. If one takes the z axis along the direction of relative motion of the reference systems, then $d\Omega/d\Omega' = d(\cos\theta)/d(\cos\theta')$. By differentiating equation $1 - \beta\cos\theta =$

$(1-\beta^2)/(1+\beta\cos\theta')$ one obtains immediately

$$d\Omega = \frac{1-\beta^2}{(1+\beta\cos\theta')^2}d\Omega'. \tag{2.26}$$

2.5 Minkowski Diagrams

Minkowski or spacetime diagrams can be used to solve problems in relativity graphically without analytic computation. They provide a simple visualization of how space and time merge into spacetime in relativity theory. In Minkowski diagrams one omits two of the three spatial dimensions and reduces everything to the representation of time and one spatial coordinate (either r_1 or r).

2.5.1 Relativistic Causality

In Newtonian mechanics causality is defined by absolute time. An event A can influence B if and only if $t_B > t_A$. In relativity theory the notion of absolute time loses its meaning. Moreover, causality becomes more restrictive because now c sets a maximal speed for any interaction. Indeed, the worldline $x = (ct, \boldsymbol{r})$ of a light pulse emitted at $t = 0$ in $\boldsymbol{r} = 0$ satisfies

$$x \cdot x = x_\mu x^\mu = \eta_{\mu\nu} x^\mu x^\nu = c^2t^2 - r^2(t) = 0. \tag{2.27}$$

It is easy to see that for objects moving with $v < c$ one has $x \cdot x > 0$. The future of the event $0 = (0, \mathbf{0})$ therefore consists of all events with $x \cdot x \geq 0$ and $x^0 \geq 0$. The past correspondingly consists of events with $x \cdot x \geq 0$ and $x^0 \leq 0$; cf. the discussion after (2.8).

Relativistic causality requires that an event A can influence another event B only if B lies in the future of A, i.e. $(x_B - x_A)^2 \geq 0$ and $x_B^0 \geq x^A$. Equivalently, an event B can be influenced by another event A only if A lies in the past of B. Because orthochronous Lorentz transformations leave both Δs^2 and the sign of Δx^0 invariant, all frames agree on the definitions of future and past and hence on relativistic causality.

Events with $x \cdot x < 0$ are spacelike separated from 0 and hence stand in no causal relation (neither future nor past). In Newtonian mechanics no such spacelike separations exist, since every pair of events lies either in the future or in the past of the other. The possible causal relations can be neatly summarized in a Minkowski diagram, see Figure 2.3(a).

Because the interval Δs^2 is Lorentz invariant, all observers (in inertial frames) agree on the causal relation between any two events.

2.5.2 Lorentz Boosts

It is instructive to see how a Lorentz boost $\Lambda(\chi)$ is represented in a Minkowski diagram. As already noted, we restrict ourselves to events with $x^2 = x^3 = 0$. Thus

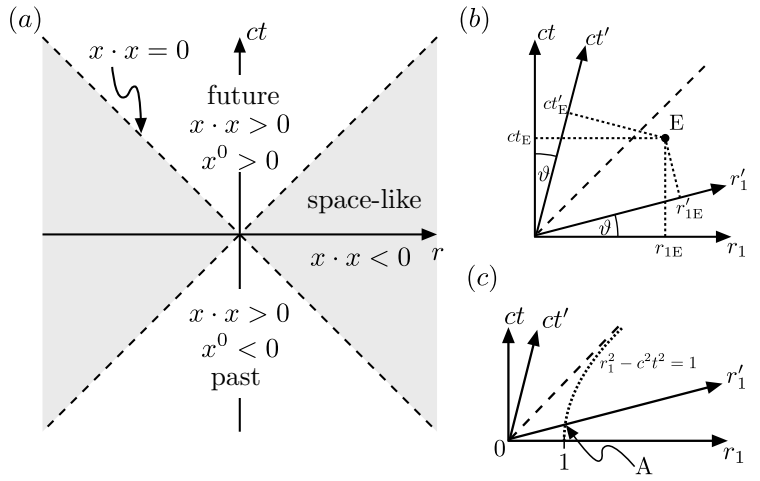

Figure 2.3: (a) Classification of all events relative to a reference event at $0 = (0, \mathbf{0})$ into future, past, and spacelike. (b) An event E is described in two coordinate systems S and S' moving relative to each other with velocity $w = \tan\vartheta\, c$ by the coordinates $(ct_E, r_{1\mathrm{E}})$ and $(ct'_E, r'_{1\mathrm{E}})$, respectively. The coordinates are related by a simple geometric construction shown in the figure. (c) Geometric construction by which, starting from a unit length in the S system, the unit length in the moving S' system can be determined. One constructs the intersection point A of the unit hyperbola $r_1^2 - c^2t^2 = 1$ with the r'_1 axis; the length $\overline{OA}$ then fixes the unit length in the S' system.

events are specified by their two-dimensional spacetime coordinates (t, r_1). Lorentz transformations are not rotations of the (ct, r_1) plane, because they leave invariant not the Euclidean distance $c^2t^2 + r_1^2$ but the Minkowski distance $c^2t^2 - r_1^2$. One can determine the orientation of the coordinate axes of the moving S' frame relative to the S frame. The ct' axis is determined by $0 = r'_1 = \gamma(r_1 - wt)$, i.e. $ct = (c/w)r_1$. Hence the ct' axis makes an angle ϑ with the ct axis, where

$$\tan\vartheta = \frac{w}{c}. \tag{2.28}$$

From the relation $|w| \leq c$ one immediately finds $|\vartheta| \leq \pi/4$ and also $\tan\vartheta = \tanh\chi$. Analogously one finds that the r'_1 axis is determined by $0 = t' = \gamma(t - wr_1/c^2)$. Thus the r'_1 axis also makes an angle ϑ with the r_1 axis. As shown in Figure 2.3(b), the axes for $w > 0$ are thereby "collapsed."

By construction a light ray always travels along the angle bisector in the positive t direction. To obtain the correct Lorentz transformation between frames, the unit lengths on the axes must change under the transformation. One can use the fact that a Lorentz transformation leaves the interval $c^2t^2 - r_1^2$ invariant. In particular, for every event at unit distance one has $r_1^2 - c^2t^2 = 1$. As seen in Figure 2.3(c), this unit hyperbola intersects the r'_1 axis at point A, which by interval invariance must lie at

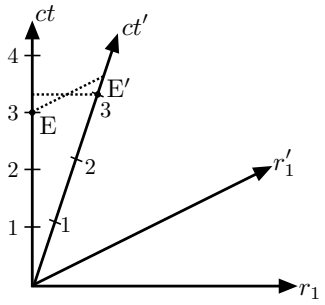

Figure 2.4: Illustration in a Minkowski diagram that mutually moving observers each find that the clock in the other frame runs slower. The events E and E' occur at time three in their respective rest frames (i.e. a clock at rest shows 3). The other observer is moving and records that the event in the other frame occurs after 3. The situation is perfectly symmetric, since each observer finds that the clock in the other frame runs slower.

$(0, 1)$. Hence the r'_1 axis is stretched relative to the r_1 axis.

One finds the stretch factor by a simple argument: every point on the r'_1 axis is given in the S system by $\tan\vartheta = ct/r_1$. Its intersection with the unit hyperbola is therefore at $(1 - \tan^2\vartheta)r_1^2 = 1$, which yields the unit length

$$\overline{0\mathrm{A}} = \sqrt{r_1^2 + c^2t^2} = \sqrt{(1 + \tan^2\vartheta)r_1^2} = \sqrt{\frac{1 + w^2/c^2}{1 - w^2/c^2}} \tag{2.29}$$

on the r'_1 axis of the S' system. The length units of the S' coordinate system must therefore be drawn stretched by the factor $\overline{0\mathrm{A}} \geq 1$, both on the r'_1 axis and (by symmetry) on the ct' axis. Only with this rescaling of the S' coordinate system does the coordinate transformation become symmetric, as required by special relativity. For example, Figure 2.4 shows that each observer finds the clock in the other frame to run slower.

Chapter 3

Relativistic Mechanics

As we saw in Chapter 1, Newtonian mechanics is not compatible with the relativity postulate. In special relativity, we must therefore render the laws of mechanics covariant under Lorentz transformations. Since four-tensors have a fixed transformation behavior, cf. Appendix A, the equations of the new mechanics must equate tensors of the same rank. As a subsidiary condition, the old mechanics should be recovered in the limit $v \ll c$.

3.1 Four-velocity

The motion of an object is specified by the trajectory $\boldsymbol{r}(t)$ with the velocity $\boldsymbol{v}(t) = \dot{\boldsymbol{r}}$. In relativistic notation, the trajectory becomes a world line

$$x(\lambda) = [x^\mu(\lambda)] = [ct, \boldsymbol{r}(t)] = [ct(\lambda), \boldsymbol{r}(\lambda)] \tag{3.1}$$

with $\boldsymbol{r}(\lambda) = \boldsymbol{r}(t(\lambda))$ and λ an arbitrary parameter that parametrizes the world line. As explained in Chapter 2, transformations between different inertial frames are effected by a Lorentz transformation Λ with $x' = \Lambda x$. Hence x^μ is a contravariant four-vector.

One would like to generalize the notion of velocity to a four-vector. To do so, $x(\tau)$ must be differentiated with respect to a Lorentz scalar that reduces to time in the limit $v \ll c$. As explained in (1.17), the proper time

$$\tau(t) = \int_0^t dt' \sqrt{1 - v(t')^2/c^2} = \int_{x(0)}^{x(t)} \sqrt{\eta_{\mu\nu} dx^\mu dx^\nu}/c \tag{3.2}$$

is precisely a Lorentz scalar. Indeed, from the transformation $x' = \Lambda x$ of a world line one obtains the invariance of proper time

$$\tau'(t') = \int_{x'(0)}^{x'(t')} \sqrt{\eta_{\mu\nu} dx'^\mu dx'^\nu} = \int_{x(0)}^{x(t)} \sqrt{\eta_{\mu\nu} dx^\mu dx^\nu} = \tau(t).$$

.

To render the parametrization in (3.1) frame-independent, it is expedient to set $\lambda = \tau$. Differentiating the world line with respect to proper time yields the covariant four-vector

$$u(\tau) = \frac{dx(\tau)}{d\tau} = \frac{dt}{d\tau}\frac{dx(t)}{dt} = \gamma(t)[c, \boldsymbol{v}(t)] \tag{3.3}$$

with $\gamma(t) = [1 - v(t)^2/c^2]^{-1/2}$, which we call the *four-velocity*. From the four-vector u^μ one obtains a Lorentz scalar by contraction with itself. In fact $u \cdot u = u_\mu u^\mu = \gamma^2(c^2 - v^2) = c^2$ holds, so that u is always normalized to c and only three parameters are independent.

3.2 Energy-Momentum Relation

With the four-velocity we can immediately generalize the momentum $\boldsymbol{p}$ to a covariant four-momentum

$$(p^\mu) = (p^0, \boldsymbol{p}) = mu = m\gamma(c, \boldsymbol{v}) \tag{3.4}$$

. In the nonrelativistic limit, when $v \ll c$ and $\gamma \to 1$, the spatial components of p^μ agree with the nonrelativistic momentum. The spatial part therefore remains conserved for a force-free particle in an inertial frame. Since p^μ is a four-vector, the entire four-momentum p^μ is conserved in every inertial frame. Analogously one may conclude that the total momentum $P = \sum_{j=1}^{N} p_j$ of N particles with four-momenta p_j is conserved in every inertial frame.

The question now arises which conserved quantity in nonrelativistic physics corresponds to the time component p^0. Expanding p^0 in v/c gives

$$p^0 = \frac{mc}{\sqrt{1 - v^2/c^2}} = mc + \frac{1}{2}mv^2/c + \dots . \tag{3.5}$$

One immediately recognizes that $p^0 c$, up to the constant mc^2, agrees with the kinetic energy E_{kin} of Newtonian mechanics. This motivates the identification $E = p^0 c$, so that

$$E = p^0 c = E_0 + E_{\text{kin}} \quad \text{with} \quad E_0 = mc^2. \tag{3.6}$$

.

The rest energy $E_0 = mc^2$ is precisely Einstein's famous formula. In contrast to nonrelativistic physics, relativistic energy contains a rest-energy contribution that depends only on the object's mass. Unlike in nonrelativistic physics, mass can be annihilated in relativity theory and converted into other forms of energy.

Consider, for illustration, the (symmetric) decay of a particle into two parts. In the rest frame of the particle one starts with the four-momentum $P^\mu = (Mc, \boldsymbol{0})$ with M the total mass of the particle. After the symmetric decay the total momentum P^μ is composed of the momenta of two particles of mass m moving with velocity $\pm\boldsymbol{v}$. The

particles have four-momenta $\gamma m(c, \pm\boldsymbol{v})$ with $\gamma = (1 - v^2/c^2)^{-1/2}$, which sum to the total momentum $P^\mu = 2\gamma m(c, \mathbf{0})$. From conservation of four-momentum we can infer that

$$2m = M\sqrt{1 - v^2/c^2} < M\,; \tag{3.7}$$

i.e., the total mass is not conserved. The rest energy of the mass defect is given by

$$(M - 2m)c^2 = 2mc^2(\gamma - 1) = 2\frac{1}{2}mv^2 + \ldots \tag{3.8}$$

and hence equals the nonrelativistic energy of the decay products (for $v \ll c$). Thus, in the decay of the particle rest energy is converted into kinetic energy of the decay products.

From the fact that the length of p^μ forms a Lorentz scalar one obtains, in addition, the relativistic energy-momentum relation[1]

$$p \cdot p = p_\mu p^\mu = (E/c)^2 - p^2 = m^2c^2, \tag{3.9}$$

which, by construction, holds in all inertial frames.

3.3 Equation of Motion

With this preparation it is now possible to formulate mechanics in a Lorentz-covariant way. The simplest route is via the Lagrangian formalism. In this formalism one assigns to each world line $x(t) = [ct, \boldsymbol{r}(t)]$ an action[2]

$$S[r(t)] = \int_{t_0}^{t_1} dt\, L(\boldsymbol{r}, \boldsymbol{v}, t) \tag{3.10}$$

with the Lagrangian L and $\boldsymbol{v} = \dot{\boldsymbol{r}}$. Hamilton's principle requires that the path of a particle be characterized by an extremum of the action with $\delta S = 0$. In this comparison, different world lines with fixed endpoints $\boldsymbol{r}(t_0) = \boldsymbol{r}_0$ and $\boldsymbol{r}(t_1) = \boldsymbol{r}_1$ are considered. The extremal principle therefore leads, via partial integration, to

$$0 = \delta S = \int_{t_0}^{t_1} dt \sum_{k=1}^{3} \Big[(\partial_{r_k} L)\delta r_k + (\partial_{v_k} L)\partial_t \delta r_k\Big] = \int_{t_0}^{t_1} dt \sum_{k=1}^{3} \left(\partial_{r_k} L - \frac{d}{dt}\partial_{v_k} L\right)\delta r_k\,.$$

Since the variation $\delta r(t)$ with $\delta r(t_0) = \delta r(t_1) = 0$ is arbitrary, Hamilton's principle yields the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial v_k} = \frac{\partial L}{\partial r_k}, \tag{3.11}$$

which play the role of the equations of motion. From the condition δS on the path it is directly evident that the equation of motion is covariant provided the action S is a Lorentz scalar.[3]

[1] Note the distinction between $p \cdot p = p_\mu p^\mu = \eta_{\mu\nu} p^\mu p^\nu$ and $p^2 = \boldsymbol{p}^2 = \sum_{k=1}^{3} p_k^3$.

[2] We use the time t in a reference frame S as curve parameter λ.

[3] In principle the actions in different inertial frames could differ by a factor.

3.3.1 Force-free Particle

In nonrelativistic mechanics a force-free particle is described by the Lagrangian $L_0 = \frac{1}{2}mv^2$. To obtain a relativistic generalization of the equations of motion for a force-free particle we need an action that is a Lorentz scalar and reproduces nonrelativistic behavior in the limit $v \ll c$. As we saw in (3.2), proper time assigns a Lorentz scalar to a world line. A natural ansatz is therefore $S_0 = -mc^2\tau$, where the prefactor E_0 is chosen so that S_0 has the dimension of an action. Indeed we obtain

$$S_0 = -mc^2\tau = -mc^2 \int dt\, \sqrt{1 - v^2/c^2} = \int dt \left(-mc^2 + \frac{1}{2}mv^2 + \dots\right), \tag{3.12}$$

so that the relativistic Lagrangian

$$L_0 = -mc^2\sqrt{1 - v^2/c^2} \tag{3.13}$$

in the limit $v \ll c$ agrees, up to the constant mc^2, with the nonrelativistic Lagrangian of a force-free particle. The relativistic equations of motion

$$\frac{d}{dt}(\gamma m\boldsymbol{v}) = \frac{d\boldsymbol{p}}{dt} = 0 \tag{3.14}$$

are obtained as the Euler-Lagrange equations of L_0.

In this form it is not immediately obvious that (3.14) is covariant. To make the covariance explicit, we note that, by the energy-momentum relation (3.9), $\dot{p}^0$ is determined together with (3.14). Indeed, differentiating (3.9) with respect to t yields

$$c\frac{dp^0}{dt} = c\frac{\boldsymbol{p}}{p^0} \cdot \frac{d\boldsymbol{p}}{dt} = \boldsymbol{v} \cdot \frac{d\boldsymbol{p}}{dt} = 0, \tag{3.15}$$

so that we arrive at the manifestly covariant equation $dp^\mu/d\tau = 0$.

3.3.2 Particle in a Potential

In Galilean physics one often considers a particle in a scalar potential $V(\boldsymbol{r}, t)$. The corresponding problem is not really fruitful in relativistic mechanics. The issue is that, starting from a scalar potential V, the potential in another reference frame automatically becomes velocity dependent $\boldsymbol{v}$. Hence it is better to start directly from a vector potential that couples to the velocity. Another point is that there is no physical force that can be described by a scalar potential, since the gravitational force is not formulated covariantly and hence only the electromagnetic force fits into the relativistic concept.

Electrodynamics comes equipped with a scalar $\varphi(\boldsymbol{r}, t)$ and a vector potential $\boldsymbol{A}(\boldsymbol{r}, t)$. The action of electromagnetic fields on a particle of charge q is described in classical mechanics by the Lagrangian

$$L = L_0 + \frac{q}{c}\boldsymbol{v} \cdot \boldsymbol{A} - q\varphi \tag{3.16}$$

Replacing the free Lagrangian $L_0 = mv^2/2$ by the relativistic generalization (3.13), one sees that the action

$$S = \int dt\, L = -\int d\tau \left(mc^2 + \frac{q}{c} u \cdot A\right) = -\int d\tau \left(mc^2 + \frac{q}{c} u^\mu A_\mu\right) \tag{3.17}$$

is already covariant provided $A = (A^\mu) = (\varphi, \boldsymbol{A})$ transforms like a four-vector.[4]

For the case that $\boldsymbol{A} = 0$ in a reference frame, one obtains a particle in a scalar potential $V(\boldsymbol{r}, t) = q\,\varphi(\boldsymbol{r}, t)$. In this special case, the Euler-Lagrange equations take the form

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{F} \tag{3.18}$$

with $\boldsymbol{p} = \gamma m \boldsymbol{v}$ and $\boldsymbol{F} = -\boldsymbol{\nabla} V$ the force acting on the particle. The equation of motion (3.18) is often written in the four-vector notation

$$\frac{dp^\mu}{d\tau} = K^\mu \tag{3.19}$$

with K^μ the *Minkowski force*. The spatial components of the Minkowski force follow directly from (3.18) as $\boldsymbol{K} = \gamma \boldsymbol{F}$, with $\gamma d\tau = dt$. The time component K^0 of the Minkowski force is determined from the additional condition

$$K^\mu u_\mu = \frac{dp^\mu}{d\tau} u_\mu = \frac{m}{2} \frac{d}{d\tau} (u^\mu u_\mu) = 0 \tag{3.20}$$

i.e. the four-acceleration $du^\mu/d\tau \propto K^\mu$ is always orthogonal to the four-velocity. One thus obtains the complete Minkowski force as[5]

$$(K^\mu) = \gamma(\boldsymbol{v} \cdot \boldsymbol{F}/c, \boldsymbol{F}). \tag{3.21}$$

As for the force-free particle, the 0-component of (3.19) is an expression of energy conservation. Indeed we obtain

$$\frac{dE}{dt} = c \frac{dp^0}{dt} = \boldsymbol{v} \cdot \boldsymbol{F}. \tag{3.22}$$

Hence $\boldsymbol{v} \cdot \boldsymbol{F}$ is the power delivered by the external force $\boldsymbol{F}$ to the particle and $E = cp^0$ is (up to the constant mc^2) the kinetic energy of the particle.

We now wish, in the general case with $\boldsymbol{A} \neq 0$ from L, to derive the Euler-Lagrange equations. For this we need the easily proven intermediate results

$$\frac{\partial L}{\partial r_k} = -q \frac{\partial \varphi}{\partial r_k} + \frac{q}{c} \sum_{l=1}^{3} v_l \frac{\partial A_l}{\partial r_k}, \qquad \frac{\partial L}{\partial v_k} = \gamma m v_k + \frac{q}{c} A_k$$

$$\text{and} \quad \frac{d}{dt} \frac{\partial L}{\partial v_k} = \frac{dp_k}{dt} + \frac{q}{c} \left(\frac{\partial A_k}{\partial t} + \sum_{l=1}^{3} \frac{\partial A_k}{\partial r_l} v_l \right).$$

[4] In Chapter 4 we will see that A^μ does indeed transform like a four-vector in electrodynamics.

[5] Written out in components, (3.20) yields $K^\mu u_\mu = \gamma c K^0 - \gamma \boldsymbol{K} \cdot \boldsymbol{v} = \gamma c K^0 - \gamma^2 \boldsymbol{F} \cdot \boldsymbol{v} = 0$, which can be solved directly for K^0.

We thereby obtain the equations of motion

$$\begin{aligned}\frac{dp_k}{dt} &= q\left(-\frac{\partial\varphi}{\partial r_k} - \frac{1}{c}\frac{\partial A_k}{\partial t}\right) + \frac{q}{c}\sum_{l=1}^{3} v_l\left(\frac{\partial A_l}{\partial r_k} - \frac{\partial A_k}{\partial r_l}\right) \\ &= qE_k + \frac{q}{c}(\boldsymbol{v}\wedge\boldsymbol{B})_k \end{aligned} \tag{3.23}$$

for a particle subject to the Lorentz force $\boldsymbol{F}_L = q\boldsymbol{E} + q\boldsymbol{v}\wedge\boldsymbol{B}/c$ in the electric field $\boldsymbol{E} = -\boldsymbol{\nabla}\varphi - \partial_t\boldsymbol{A}/c$ and magnetic field $\boldsymbol{B} = \boldsymbol{\nabla}\wedge\boldsymbol{A}$. [6]

Note that in the presence of a vector potential the canonical momentum $\boldsymbol{p}_{\text{kan}} = \partial_{\boldsymbol{v}}L = \boldsymbol{p} + q\boldsymbol{A}/c$ does not coincide with the kinematic momentum $\boldsymbol{p} = \gamma m\boldsymbol{v}$. In particular, the energy-momentum relation (3.9) refers to the kinematic and not to the canonical momentum.

One may also cast the equation of motion of a particle in an electromagnetic field into an explicitly covariant form. To this end one defines the (twice covariant) electromagnetic field strength tensor (Faraday tensor) as

$$F_{\mu\nu} = (dA)_{\mu\nu} \equiv \frac{\partial A_\nu}{\partial x^\mu} - \frac{dA_\mu}{\partial x^\nu}. \tag{3.24}$$

The tensor $F_{\mu\nu}$ is antisymmetric and thus has 6 independent components. By explicit calculation one finds that the components[7]

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix} \tag{3.25}$$

are directly given by the electric and magnetic fields. The equation of motion can thereby be written covariantly in the form of (3.19), where the Lorentz force becomes the manifestly covariant Minkowski force[8]

$$K^\mu = \frac{q}{c}F^{\mu\nu}u_\nu = \gamma(q\boldsymbol{v}\cdot\boldsymbol{E}/c, \boldsymbol{F}_L)^\mu \tag{3.26}$$

From the antisymmetry of F one directly infers that the constraint $K^\mu u_\mu = 0$ is satisfied, cf. Eq. (3.20).

Particle in a homogeneous electric field

The form of the equation of motion (3.19) ensures that even under a constant applied force the speed of light is never exceeded. As an example we consider a particle of

[6] We denote by $\boldsymbol{v}\wedge\boldsymbol{w}$ between three-vectors the vector product $\boldsymbol{v}\times\boldsymbol{w}$. For four-vectors it becomes the more general, exterior product $(v\wedge w)^{\mu\nu} = v^\mu w^\nu - v^\nu w^\mu$, see Chapter 3.4.

[7] In general, an antisymmetric tensor of second rank in Minkowski spacetime contains precisely one polar vector $\boldsymbol{E}$ and one axial vector $\boldsymbol{B}$.

[8] Note that the magnetic part does no work $\boldsymbol{v}\cdot\boldsymbol{F}_L = \boldsymbol{v}\cdot\boldsymbol{E}$. Furthermore one obtains $F^{\mu\nu}$ by $F_{\mu\nu}$ with the replacement $\boldsymbol{E}\mapsto -\boldsymbol{E}$.

mass m subject to a homogeneous electric field $\boldsymbol{E} = (E_0, 0, 0)$. Let $\boldsymbol{v}(0) = 0$ so that the motion proceeds only in the x-direction. We obtain the equation of motion

$$\frac{d}{dt}\frac{v_1(t)}{\sqrt{1 - v_1(t)^2/c^2}} = c\frac{d\sinh\chi(t)}{dt} = \frac{qE_0}{m} = a\,, \tag{3.27}$$

with the rapidity χ defined by $v_1 = c\tanh\chi$. The solution is therefore given by $\sinh\chi = at/c$, which corresponds to

$$v_1(t) = \frac{at}{\sqrt{1 + (at/c)^2}} = \begin{cases} at, & |t| \ll c/a, \\ c, & |t| \gg c/a \end{cases} \tag{3.28}$$

Although the velocity saturates at c, the energy

$$E = \frac{mc^2}{\sqrt{1 - v_1(t)^2/c^2}} = mc^2\sqrt{1 + (at/c)^2} \tag{3.29}$$

continues to grow without bound. The same holds, by $E^2 - p_1^2c^2 = m^2c^4$, for the relativistic momentum.

It is also instructive to compute the world line of an object undergoing constant acceleration. The trajectory of a spaceship moving with constant acceleration is given by

$$\begin{aligned} r_1(t) &= \int_0^t dt'\, v_1(t') = \frac{c^2}{a}\int d\chi\, \sinh\chi = \frac{c^2}{a}\sqrt{1 + (at/c)^2} \\ &= \text{const.} + \frac{at^2}{2} + \dots\,. \end{aligned} \tag{3.30}$$

Thus the path of the spaceship is simply a hyperbola, whereas in Newtonian mechanics the path is a parabola. An important quantity is the proper time

$$\tau(t) = \int_0^t dt'\,\sqrt{1 - v_1(t')^2/c^2} = \frac{c}{a}\int d\chi = \frac{c\,\mathrm{arsinh}(at/c)}{a}, \tag{3.31}$$

which measures the time in the local frame of the spaceship. [9]

With the hyperbolic trajectory of the constantly accelerated spaceship one can resolve the twin paradox: at the initial time $t_0 = -T/2$ the spaceship is located at the position $x_0 = r_1(-T/2)$. Let us now imagine that one of the twins boards the spaceship while the other remains at rest at the position x_0. At the time $t_1 = T/2$ the spaceship is again at the same position $x_0 = r_1(T/2)$. For the twin at rest the time T has passed in the meantime. The other twin, however, measures the proper time (3.31) inside the spaceship. For him the shorter time

$$\tau(T/2) - \tau(-T/2) = \frac{2c\,\mathrm{arsinh}(aT/2c)}{a} < T \tag{3.32}$$

[9] Parametrized by the proper time, the path of the spaceship takes the simple form $r_1(\tau) = (c^2/a)\cosh(c\tau/a)$.

has therefore passed. Hence, upon meeting again at the position x_0 the accelerated twin is younger than the one at rest. All observers will agree with this statement. Unlike the time dilation between two inertial systems, see Figure 2.4, the situation here is not symmetric, since only one of the twins was accelerated.

Particle in a homogeneous magnetic field

As a second example we now examine the case of a particle in a homogeneous magnetic field $\boldsymbol{B} = (0, 0, B_0)$. Let the initial condition $\boldsymbol{v}(0) = (v_0, 0, 0)$ be such that the motion takes place only in the xy-plane. The equation of motion has the general form

$$\dot{\boldsymbol{p}} = \frac{q}{c}\boldsymbol{v} \wedge \boldsymbol{B}. \tag{3.33}$$

Since $E/mc^2 = \gamma$ holds and a magnetic field does no work, γ is conserved. Inserting the relation $\boldsymbol{p} = m\gamma\boldsymbol{v}$ into (3.33) and writing the equation in components, we obtain

$$\dot{v}_1 = \omega v_2, \qquad \dot{v}_2 = -\omega v_1 \tag{3.34}$$

with the Larmor frequency $\omega = qB_0/mc\gamma$. The equations are solved by

$$v_1(t) = v_0 \cos(\omega t), \qquad v_2(t) = -v_0 \sin(\omega t). \tag{3.35}$$

A further integration yields the trajectory

$$r_1(t) = \rho \sin(\omega t), \qquad r_2 = \rho \cos(\omega t). \tag{3.36}$$

We recognize that the particle moves on a circular path of radius $\rho = v_0/\omega = qv_0B_0/mc\gamma$.

Because γ is constant, the proper time is simply given by $\tau = t/\gamma$. An observer moving on the circular path will therefore find that he needs the proper time $2\pi/\omega\gamma$ for one revolution. Again there is no symmetry between the two reference frames, so that all observers agree that the angular velocity $\gamma\omega$ measured in the accelerated frame is larger than that in the laboratory frame.

3.4 Angular momentum

In a closed system, besides energy and momentum conservation—which in relativity theory appears as the conservation of the four-momentum p^μ—in Newtonian mechanics the conservation of angular momentum $\boldsymbol{L} = \boldsymbol{r} \wedge \boldsymbol{p}$ also holds. The conservation of angular momentum originates from the isotropy of the physical laws.

It turns out that the relativistic generalization of angular momentum leads to the

antisymmetric angular momentum tensor[10]

$$L^{\mu\nu} = (x \wedge p)^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu = \begin{pmatrix} 0 & -cN^1 & -cN^2 & -cN^3 \\ cN^1 & 0 & L_3 & -L_2 \\ cN^2 & -L_3 & 0 & L_1 \\ cN^3 & L_2 & -L_1 & 0 \end{pmatrix}^{\mu\nu}$$
$$= \left(\begin{array}{c|c} 0 & -c\boldsymbol{N}^t \\ \hline c\boldsymbol{N} & \boldsymbol{x} \wedge \boldsymbol{p} \end{array}\right) \tag{3.37}$$

, cf. (3.25). Here we have introduced the new vector $\boldsymbol{N} = (E/c^2)\boldsymbol{r} - \boldsymbol{p}t = m\gamma(\boldsymbol{r} - \boldsymbol{v}t)$ by $\gamma = E/mc^2$, cf. (3.4), which is also called the *dynamical mass moment*.

Consider now N particles that are located in a closed system. As we have already seen in Chapter 3.2 using the example of the four-momentum, the conservation of the total angular momentum $\boldsymbol{L} = \sum_{j=1}^N \boldsymbol{L}_j$ is covariant only if the entire angular momentum tensor $L^{\mu\nu} = \sum_{j=1}^N (L_j)^{\mu\nu}$ is conserved. Hence the mass moment $\boldsymbol{N} = \sum_{j=1}^N \boldsymbol{N}_j$ must also be conserved. Since the total energy $cP^0 = \sum_{j=1}^N E_j$ is also conserved, the conservation of the dynamical mass moment can be rewritten as

$$\frac{c^2\boldsymbol{N}}{\sum_{j=1}^N E_j} = \frac{\sum_{j=1}^N (E_j\boldsymbol{r}_j - c^2\boldsymbol{p}_j t)}{\sum_{j=1}^N E_j} = \boldsymbol{R} - \boldsymbol{V}t = \text{const.} \tag{3.38}$$

Thus we see that the "mean" position vector

$$\boldsymbol{R} = \frac{\sum_{j=1}^N E_j\boldsymbol{r}_j}{\sum_{j=1}^N E_j} \tag{3.39}$$

moves uniformly with the velocity

$$\boldsymbol{V} = \frac{c^2 \sum_{j=1}^N \boldsymbol{p}_j}{\sum_{j=1}^N E_j} \tag{3.40}$$

. Equation (3.39) provides an extension of the concept of center of mass into the relativistic context. Indeed, in the limit where all velocities are small, $E_j \approx m_j c^2$ holds. Hence we recover the classical expressions

$$\boldsymbol{R} = \frac{\sum_{j=1}^N m_j\boldsymbol{r}_j}{\sum_{j=1}^N m_j} \qquad \text{and} \qquad \boldsymbol{V} = \frac{\sum_{j=1}^N \boldsymbol{p}_j}{\sum_{j=1}^N sum_j}\,. \tag{3.41}$$

However, it should be noted that $\boldsymbol{R}$ in general does not transform like the spatial components of a four-vector. Consequently the center of mass also depends on the observer and is not simply fixed relative to the particles.

[10] That $\boldsymbol{L}$ can be treated as a vector in Newtonian mechanics is due to the special status of the vector product in three dimensions. In general $L = x \wedge p$ is given by the exterior product of x and p. Hence $L^{\mu\nu}$ is an antisymmetric tensor of rank two with six independent components in four-dimensional spacetime.

Chapter 4

Covariant Formulation of Maxwell's Equations

We will see in this chapter that Maxwell's equations already satisfy Einstein's relativity principle. This is no surprise, since the principle is based on the constancy of the speed of light in all inertial systems; a fact that is already accounted for in Maxwell's equations. We will learn a manifestly covariant formulation of electromagnetism and find that Maxwell's equations do not require any correction. In fact, the special theory of relativity exists precisely to make mechanics compatible with Maxwell's theory without altering the latter.

4.1 Invariants of the Field

From the Faraday tensor $F_{\mu\nu}$ one can form invariants (scalars) that do not change under transitions to other inertial systems. These are important because they yield a scalar action whose Euler-Lagrange equations are covariant. It is clear that from $F_{\mu\nu}$ one can construct the scalar

$$\frac{1}{2}F_{\mu\nu}F^{\mu\nu} = B^2 - E^2 \tag{4.1}$$

and the pseudoscalar[1]

$$-\frac{1}{8}\varepsilon^{\mu\nu\sigma\tau}F_{\mu\nu}F_{\sigma\tau} = \boldsymbol{E}\cdot\boldsymbol{B} \tag{4.2}$$

. The pseudoscalar character of $\boldsymbol{E}\cdot\boldsymbol{B}$ is immediately clear, since it is a product of a polar ($\boldsymbol{E}$) and an axial ($\boldsymbol{B}$) vector.

By the invariance of $E^2 - B^2$ one immediately obtains that if $E > B$ holds in one reference frame, then $E' > B'$ also holds in every other inertial system. Likewise,

[1] Here we introduce the totally antisymmetric (Levi-Civita) tensor $\varepsilon^{\mu\nu\sigma\tau}$ (with the convention $\varepsilon^{0123} = 1$). The Levi-Civita tensor is an (invariant) pseudotensor, since $\varepsilon^{\alpha\beta\gamma\delta} = \Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}\Lambda^{\gamma}{}_{\sigma}\Lambda^{\delta}{}_{\tau}\varepsilon^{\mu\nu\sigma\tau} = \det(\Lambda)\varepsilon^{\alpha\beta\gamma\delta}$ holds.

from the invariance of $\boldsymbol{E} \cdot \boldsymbol{B}$ it follows that observers in all inertial systems agree on whether $\boldsymbol{E}$ and $\boldsymbol{B}$ form an obtuse or acute angle.

4.2 Homogeneous Maxwell Equations

The homogeneous Maxwell equations do not couple to charges. Their statement is that to describe the physics of the electromagnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ one may pass to the potentials φ and $\boldsymbol{A}$. Indeed, the homogeneous Maxwell equations follow directly from the existence of the potential A^μ. With $F_{\mu\nu} = \partial A_\mu/\partial x^\nu - \partial A_\nu/\partial x^\mu$ we can verify by a simple calculation[2]

$$\frac{\partial \mathcal{F}^{\mu\nu}}{\partial x^\mu} = \frac{1}{2}\varepsilon^{\mu\nu\sigma\tau}\frac{\partial F_{\sigma\tau}}{\partial x^\mu} = 0 \tag{4.3}$$

; here we have introduced the dual field-strength tensor $\mathcal{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\sigma\tau}F_{\sigma\tau}$, whose divergence vanishes according to (4.3). The dual field-strength tensor has components

$$(\mathcal{F}^{\mu\nu}) = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{pmatrix}. \tag{4.4}$$

It is obtained from $F^{\mu\nu}$ by the replacement $\boldsymbol{E} \mapsto \boldsymbol{B}$ and $\boldsymbol{B} \mapsto -\boldsymbol{E}$. Thus the condition that the divergence of $\mathcal{F}$ vanishes leads to the homogeneous Maxwell equations

$$\frac{\partial \mathcal{F}^{\mu 0}}{\partial x^\mu} \mapsto \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0, \qquad \frac{\partial \mathcal{F}^{\mu j}}{\partial x^\mu} \mapsto -\frac{1}{c}\dot{\boldsymbol{B}} - \boldsymbol{\nabla} \wedge \boldsymbol{E} = 0. \tag{4.5}$$

4.3 Action of the Electromagnetic Field

The total action of the electromagnetic field with N charged *species* can be decomposed into the parts $S = S_f + S_m + S_{mf}$. The contribution S_m is the free action of the *species* and is given, according to Chapter 3.3.1, by

$$S_m = -\sum_{j=1}^{N} m_j c^2 \int d\tau_j. \tag{4.6}$$

The interaction of the *species* with the electromagnetic field via the Lorentz force is described by S_{mf}. According to Chapter 3.3.2, this part is given by

$$S_{mf} = -\sum_{j=1}^{N} \frac{q_j}{c} \int A(x_j) \cdot u_j \, d\tau_j = -\sum_{j=1}^{N} \frac{q_j}{c} \int A(x_j) \cdot dx_j \tag{4.7}$$

[2]For example, $\varepsilon^{\mu\nu\sigma\tau}\partial^2 A_\sigma/(\partial x^\tau \partial x^\mu) = 0$ holds, since ε is antisymmetric in $\tau \leftrightarrow \mu$ while the remainder is symmetric. The same applies to the second term.

with the charge q_j of the j-th *species*, which plays the role of the coupling constant.

The last term S_f has not yet appeared. It describes the intrinsic dynamics of the fields $\boldsymbol{E}, \boldsymbol{B}$ and $F_{\mu\nu}$, respectively. In general, field equations arise as Euler-Lagrange equations of a Lagrange density $\mathcal{L}_f$. From the Lagrange density one obtains, by integration over space, the Lagrange function $L_f = \int d^3r\, \mathcal{L}_f$. After further integration over time we then obtain the action $S_f = \int dt\, L_f = \int d\Upsilon\, \mathcal{L}_f$, where we have introduced the measure $d\Upsilon = dt\, dV = dt\, d^3r$ in spacetime. By the general transformation theorem for integrals one finds that the measure is a Lorentz scalar. Indeed, a change of variables $x' = \Lambda x$ leads to $d\Upsilon' = |\det \Lambda|\, d\Upsilon$ with the functional determinant $\det \Lambda = \pm 1$. The requirement that S_f be a scalar is therefore equivalent to the Lagrange density $\mathcal{L}_f$ being a Lorentz scalar.

We know that the Maxwell equations for the electromagnetic fields must be linear in order to satisfy the superposition principle. Hence $\mathcal{L}_f$ must be quadratic in $F_{\mu\nu}$. Moreover, since the Maxwell equations are first order in time and space, $\mathcal{L}_f$ must not contain derivatives of $F_{\mu\nu}$. The only scalar that meets these requirements was introduced in (4.1). We therefore set

$$S_f = -\frac{1}{16\pi} \int d\Upsilon\, F_{\mu\nu} F^{\mu\nu} = \frac{1}{8\pi} \int d\Upsilon\, (E^2 - B^2), \tag{4.8}$$

where the prefactor is chosen so that the Euler-Lagrange equations yield the Maxwell equations in Gaussian units.

It is often useful also to express matter in terms of fields. To do so, one passes from *species* with coordinates $\boldsymbol{r}_j(t)$ to the charge density

$$\rho(\boldsymbol{r}, t) = \sum_{j=1}^{N} q_j\, \delta^{(3)}[\boldsymbol{r} - \boldsymbol{r}_j(t)], \tag{4.9}$$

, which one usually treats as a continuous function of the spatial coordinate $\boldsymbol{r}$. Although the charge q_j is a scalar, the same does not hold for the charge density because of Lorentz contraction. In fact, only $dq = \rho\, d^3r$ is invariant under coordinate transformations. Since $d\Upsilon = dt\, d^3r$ is also a scalar, ρ transforms like dt, i.e. like the spatial component of a four-vector. In fact, the expression (4.9) can be extended to the four-current density[3]

$$j^\mu = \sum_{j=1}^{N} q_j c \int dx_j^\mu\, \delta^{(4)}(x - x_j) = \sum_{j=1}^{N} q_j \frac{dx_j^\mu(t)}{dt}\, \delta^{(3)}[\boldsymbol{r} - \boldsymbol{r}_j(t)] = (c\rho, \boldsymbol{j})^\mu \tag{4.10}$$

. The spatial components form the current density vector

$$\boldsymbol{j} = \sum_{j=1}^{N} q_j \dot{\boldsymbol{r}}_j\, \delta^{(3)}(\boldsymbol{r} - \boldsymbol{r}_j)\,. \tag{4.11}$$

[3]Here $(x_j^\mu) = [t, \boldsymbol{r}_j(t)]$ denotes the trajectory of the j-th *species* and dx_j^μ the integration along this trajectory.

One has the relation

$$\frac{\partial}{\partial x^\mu}\left(\frac{dx_j^\mu}{dt}\,\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}_j(t)]\right) = \frac{\partial}{\partial t}\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}_j(t)] + \dot{\boldsymbol{r}}_j\cdot\boldsymbol{\nabla}\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}_j(t)] \tag{4.12}$$
$$= \frac{\partial}{\partial t}\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}_j(t)] - \dot{\boldsymbol{r}}_j\cdot\boldsymbol{\nabla}_{\boldsymbol{r}_j}\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}_j(t)] = 0\,.$$

From this follows the continuity equation

$$0 = \frac{\partial j^\mu}{\partial x^\mu} = \frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j} \tag{4.13}$$

in manifestly covariant form. With the four-current density the action can be written as

$$S = -\sum_{j=1}^{N}\int mc^2 d\tau_j - \frac{1}{c}\int d\Upsilon\, A_\mu j^\mu - \frac{1}{16\pi}\int d\Upsilon F_{\mu\nu}F^{\mu\nu}. \tag{4.14}$$

4.4 Inhomogeneous Maxwell Equations

Applying Hamilton's principle to (4.14) yields two kinds of equations. First, one may keep the fields fixed and vary the trajectories. This leads, for each *species*, to the relativistic equation of motion (3.23). On the other hand, one may keep the trajectories (and hence j^μ) invariant and vary the fields. This yields equations that describe the influence of matter on the fields. The equations for the fields are obtained as Euler-Lagrange equations

$$\frac{\partial}{\partial x^\nu}\frac{\partial\mathcal{L}}{\partial(\partial A_\mu/\partial x^\nu)} = \frac{\partial\mathcal{L}}{\partial A_\mu}, \tag{4.15}$$

for the potentials A_μ. Note here that the Lagrange density $\mathcal{L} = \mathcal{L}_{mf} + \mathcal{L}_f = -A_\mu j^\mu/c - F_{\mu\nu}F^{\mu\nu}/16\pi$ contains no matter contribution S_m to the action, since the latter does not depend on A_μ.

First we evaluate the partial derivatives of the Lagrange density with respect to the potentials. We obtain (since $F_{\mu\nu}$ depends only on $\partial A_\mu/\partial x^\nu$)

$$\frac{\partial\mathcal{L}}{\partial A_\mu} = -\frac{1}{c}j^\mu, \qquad \frac{\partial\mathcal{L}}{\partial(\partial A_\mu/\partial x^\nu)} = -\frac{1}{8\pi}F^{\sigma\tau}\frac{\partial F_{\sigma\tau}}{\partial(\partial A_\mu/\partial x^\nu)} = \frac{1}{4\pi}F^{\mu\nu}. \tag{4.16}$$

Thus the Euler-Lagrange equations become

$$\frac{\partial}{\partial x^\nu}F^{\mu\nu} = -\frac{4\pi}{c}j^\mu. \tag{4.17}$$

Setting $\mu = 0$ in (4.17) yields

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho\,. \tag{4.18}$$

With $\mu = j$ one obtains in addition the three equations

$$\frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} - \boldsymbol{\nabla} \wedge \boldsymbol{B} = -\frac{4\pi}{c}\boldsymbol{j}\,. \tag{4.19}$$

Together, (4.18) and (4.19) form the inhomogeneous Maxwell equations. The covariance of these equations follows directly from their equivalence with (4.17).

Taking the four-divergence of (4.17) yields

$$\frac{\partial^2 F^{\mu\nu}}{\partial x^\mu \partial x^\nu} = -\frac{4\pi}{c}\frac{\partial j^\mu}{\partial x^\nu}. \tag{4.20}$$

The left-hand side vanishes identically because it is a contraction of a symmetric tensor $\partial^2/\partial x^\mu \partial x^\nu$ with an antisymmetric tensor $F^{\mu\nu}$. Hence we obtain the continuity equation $\partial j^\mu/\partial x^\mu = 0$, see (4.13).

4.5 Transformation of the Electromagnetic Fields

Since the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are components of the antisymmetric field-strength tensor $F_{\mu\nu}$, we can derive their transformation behavior from the general transformation law $F'_{\mu\nu} = \Lambda_\mu{}^\sigma \Lambda_\nu{}^\tau F_{\sigma\tau}$. Under a special Lorentz transformation we obtain from $F' = \Lambda(-\chi) F \Lambda^t(-\chi)$ after a short calculation[4]

$$\begin{aligned} E'_1 &= E_1, & B'_1 &= B_1, \\ E'_2 &= \gamma(E_2 - \beta B_3), & B'_2 &= \gamma(B_2 + \beta E_3), \\ E'_3 &= \gamma(E_3 + \beta B_2), & B'_3 &= \gamma(B_3 - \beta E_2). \end{aligned} \tag{4.21}$$

As in (2.17) this result can be generalized to an arbitrary Lorentz boost, and we obtain

$$\begin{aligned} \boldsymbol{E}'_\parallel &= \boldsymbol{E}_\parallel, & \boldsymbol{E}'_\perp &= \gamma\left(\boldsymbol{E}_\perp + \frac{1}{c}\boldsymbol{w} \wedge \boldsymbol{B}_\perp\right), \\ \boldsymbol{B}'_\parallel &= \boldsymbol{B}_\parallel, & \boldsymbol{B}'_\perp &= \gamma\left(\boldsymbol{B}_\perp - \frac{1}{c}\boldsymbol{w} \wedge \boldsymbol{E}_\perp\right). \end{aligned} \tag{4.22}$$

4.6 Field of a Moving Charge

In covariant form, the inhomogeneous Maxwell equations read (4.17). To solve the homogeneous equations, we may introduce the potential A^μ, which brings the equation into the form

$$\Box A^\mu - \frac{\partial^2 A^\nu}{\partial x_\mu \partial x^\nu} = \frac{4\pi}{c} j^\mu \tag{4.23}$$

[4] Note that $\Lambda_\mu{}^\nu(\chi) = \Lambda^\mu{}_\nu(-\chi)$.

with the d'Alembert operator $\Box = \partial^2/c^2\partial t^2 - \nabla^2$. Since the potentials are determined only up to a gradient[5] one may impose in addition the condition $\partial A^\nu/\partial x_\nu = 0$, known as the Lorenz gauge. Thus the vector potential satisfies the wave equation

$$\Box A^\mu = \frac{4\pi}{c} j^\mu. \tag{4.24}$$

One can determine the general solution of (4.24) with the help of the Green's function $D(x)$, which solves the inhomogeneous equation

$$\Box D(x) = \delta^{(4)}(x) \tag{4.25}$$

. Using known methods from Fourier analysis, one can show that

$$D_r(x) = -\frac{1}{(2\pi)^4} \int d^4k \, \frac{e^{-ik_\mu x^\mu}}{k^\mu k_\mu + i0^+ \operatorname{sgn}(k_0)} \tag{4.26}$$

solves the above equation for $D(x)$, where the imaginary term in the denominator, as we will show below, ensures that $D_r(x)$ vanishes for $x^0 < 0$, thereby making D_r the *retarded* Green's function.

For the integration over k_0, we require the poles of the integrand in (4.26). These are located at $k_0^\pm = \pm|\boldsymbol{k}| - i0^+$. If $x^0 < 0$, the integration path must be closed in the upper complex half-plane due to the factor $e^{-ik_0x^0}$ in the denominator. Since the integrand has no poles in this half-plane, the integral vanishes. If $x^0 > 0$, the path must be closed in the lower half-plane. In this case, both poles $k_0^\pm$ contribute, and we obtain[6]

$$\begin{aligned} D_r(x) &= \frac{\Theta(t)}{(2\pi)^3} \int d^3k \, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \frac{\sin(c|\boldsymbol{k}|t)}{|\boldsymbol{k}|} \\ &= \frac{\Theta(t)}{(2\pi)^2} \int_0^\infty dk \int_{-1}^1 d(\cos\theta) \, e^{ikr\cos\theta} \frac{\sin(ckt)}{k} \\ &= \frac{\Theta(t)}{2\pi^2 r} \int_0^\infty dk \, \sin(kr)\sin(ckt) = \frac{\Theta(t)}{4\pi r}\delta(ct - r) \end{aligned} \tag{4.27}$$

, where $\Theta(t)$ is the Heaviside step function. In the form $D_r(x) = \Theta(x^0)\delta(x \cdot x)/2\pi$, it is easy to see that the function $D_r(x)$ is indeed a Lorentz scalar under orthochronous Lorentz transformations.[7]

The general solution of the Maxwell equation (4.24) is therefore given by

$$A^\mu(x) = A_0^\mu(x) + \frac{4\pi}{c} \int d^4x' \, D_r(x - x') j^\mu(x') \tag{4.28}$$

[5] The fact that both A and A' lead to the same field-strength tensor $F = F'$ as long as $A_\mu = A'_\mu + \partial\chi/\partial x^\mu$ with an arbitrary function $\chi(x)$ is called the *gauge freedom* of the Maxwell equations.

[6] Here θ denotes the angle between $\boldsymbol{k}$ and $\boldsymbol{x}$ and $k = |\boldsymbol{k}|$.

[7] Using the general formula $\delta[f(x)] = \sum_j \delta(x - x_j)/|f'(x_j)|$, where x_j are the (simple) zeros of $f(x)$, one immediately obtains $\Theta(t)\delta(ct - r)/r = \Theta(t)[\delta(ct - r) + \delta(ct + r)]/r = 2\Theta(t)\delta[(ct)^2 - r^2]$.

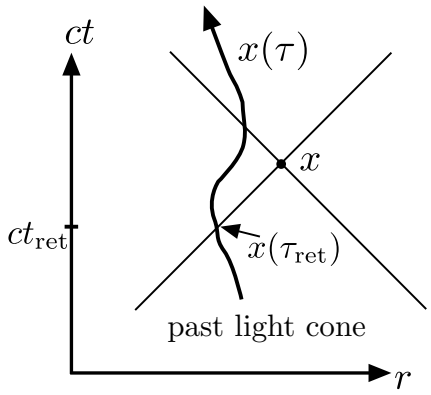

Figure 4.1: Only the intersection point $x^\mu(\tau_{\rm ret})$ of the world-line $x^\mu(\tau)$ with the past light cone emanating from the observation point x contributes to the potential at x from the charge moving on $x^\mu(\tau)$.

with $A_0(x)$ a solution of the homogeneous wave equation. Since we assume that the source is localized in space and time, in the limit $x^0 \to -\infty$ the second term vanishes and $A_0(x)$ corresponds to the incoming solution at the initial time.

Consider a point charge q moving along the trajectory $\boldsymbol{r}(t)$ (in the system S). The charge and current densities are then given by

$$\rho(\boldsymbol{r},t) = q\,\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}(t)], \qquad \boldsymbol{j}(\boldsymbol{r},t) = q\,\boldsymbol{v}(t)\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}(t)] \tag{4.29}$$

with $\boldsymbol{v}(t) = \dot{\boldsymbol{r}}(t)$ the velocity in S. These densities can be written in covariant form by parametrizing the world-line of the particle by the proper time τ. A straightforward calculation shows that the four-current density

$$j^\mu(x) = qc\int d\tau\, u^\mu(\tau)\delta^{(4)}[x-x(\tau)] \tag{4.30}$$

is the covariant formulation of (4.29). Substituting (4.30) into (4.28) (with $A_0^\mu(x)\equiv 0$) and integrating over x' yields

$$A^\mu(x) = 2q\int d\tau\, u^\mu(\tau)\Theta[x^0-x^0(\tau)]\delta[(x-x(\tau))\cdot(x-x(\tau))]. \tag{4.31}$$

The remaining integration over τ contributes only when $\tau_{\rm ret}$ is a null vector with $x-x(\tau_{\rm ret})$ and $x^0 > x^0(\tau_{\rm ret})$; i.e. the contribution comes from the past light cone emanating from the observation point x. Because of the condition $|\dot{\boldsymbol{r}}| < c$ there is only one such intersection point, see Figure 4.1. One thus obtains the four-potential

$$A^\mu(x) = \left[\frac{q\,u^\mu(\tau)}{u_\nu(\tau)(x^\nu - x^\nu(\tau))}\right]_{\rm ret}, \tag{4.32}$$

where "ret" refers to the fact that the particle position and velocity must be evaluated at the retarded time $\tau = \tau_{\rm ret}$. The potential (4.32) is called the Liénard-Wiechert potential.

For applications it is often useful to split the potential into space and time components. With $R\,\boldsymbol{n} = \boldsymbol{r} - \boldsymbol{r}(t_{\rm ret})$, $R = c(t - t_{\rm ret})$ we obtain

$$u_\mu[x^\mu - x^\mu(\tau_{\rm ret})] = \gamma c^2(t - t_{\rm ret}) - \gamma\,\boldsymbol{v}\cdot[\boldsymbol{r} - \boldsymbol{r}(t_{\rm ret})] = \gamma c R(1 - \boldsymbol{v}\cdot\boldsymbol{n}/c)\,;$$

here R denotes the distance and $\boldsymbol{n}$ the direction from $\boldsymbol{r} - \boldsymbol{r}(t_{\rm ret})$. The potentials are therefore given by

$$\varphi(\boldsymbol{r},t) = \left[\frac{q}{(1 - \boldsymbol{v}\cdot\boldsymbol{n}/c)R}\right]_{\rm ret}, \qquad \boldsymbol{A}(\boldsymbol{r},t) = \left[\frac{q\boldsymbol{v}/c}{(1 - \boldsymbol{v}\cdot\boldsymbol{n}/c)R}\right]_{\rm ret}. \tag{4.33}$$

4.7 Doppler Effect

In the derivation of (4.26) we introduced the four-wave vector $k^\mu = (\omega/c, \boldsymbol{k})$ with $k_\mu x^\mu = \omega t - \boldsymbol{k}\cdot\boldsymbol{r}$. From the electrodynamics lecture it is known that the electromagnetic field of a monochromatic light wave is given by

$$\boldsymbol{E}(x) = \boldsymbol{\mathcal{E}}\cos(k_\mu x^\mu) \quad \text{and} \quad \boldsymbol{B}(x) = \hat{\boldsymbol{k}}\wedge\boldsymbol{E}(x) \tag{4.34}$$

where $\boldsymbol{\mathcal{E}} \perp \boldsymbol{k}$ fixes the (linear) polarization and $\hat{\boldsymbol{k}} = \boldsymbol{k}/k$. Since the maxima and minima of the electromagnetic fields are measurable, the phase $k_\mu x^\mu$ must be a Lorentz scalar and hence k_μ is a covariant four-vector. Using (2.17) one therefore obtains the transformation rules

$$\omega' = \gamma(\omega - \boldsymbol{w}\cdot\boldsymbol{k}), \qquad \boldsymbol{k}'_\parallel = \gamma(\boldsymbol{k}_\parallel - \boldsymbol{w}\,\omega/c^2), \qquad \boldsymbol{k}'_\perp = \boldsymbol{k}_\perp. \tag{4.35}$$

For light waves the relativistic dispersion relation $k_0 = \omega/c = |\boldsymbol{k}|$ holds, so that $k_\parallel$ and $k_\perp$ are determined by ω and the angle θ between $\boldsymbol{k}$ and $\boldsymbol{w}$ with $k_\perp = c\,\omega\sin\theta$ and $k_\parallel = c\,\omega\cos\theta$. From (4.35) one finds that these quantities transform as

$$\omega' = \gamma\omega(1 - \beta\cos\theta), \qquad \tan\theta' = \frac{k'_\perp}{k'_\parallel} = \frac{\sin\theta}{\gamma(\cos\theta - \beta)} \tag{4.36}$$

the first formula describes the relativistic Doppler shift, while the second formula represents the light aberration, cf. (2.24).

We further wish to determine the transformation rule of the amplitude $\mathcal{E}$. For this, we will use that $\hat{\boldsymbol{k}}, \hat{\boldsymbol{\mathcal{E}}}, \hat{\boldsymbol{\mathcal{B}}}$ and $\boldsymbol{\mathcal{B}} = \hat{\boldsymbol{k}}\wedge\boldsymbol{\mathcal{E}}$ form an orthonormal basis. To simplify the calculation, we choose the coordinate system such that $\boldsymbol{w} = (c\beta, 0, 0)^t$. The fields $\boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{B}}$ then transform according to (4.21) and we obtain

$$\begin{aligned}
\mathcal{E}'^2 &= \mathcal{E}_1^2 + \gamma^2(\mathcal{E}_2 - \beta\mathcal{B}_3)^2 + \gamma^2(\mathcal{E}_3 + \beta\mathcal{B}_2)^2 \\
&= \gamma^2\mathcal{E}^2[1 - 2\beta(\hat{\mathcal{E}}_2\hat{\mathcal{B}}_3 - \hat{\mathcal{E}}_3\hat{\mathcal{B}}_2) + \beta^2(1 - \hat{\mathcal{B}}_1^2 - \hat{\mathcal{E}}_1^2)] \\
&= \gamma^2\mathcal{E}^2(1 - 2\beta\hat{k}_1 + \beta^2\hat{k}_1^2) = \gamma^2\mathcal{E}^2(1 - \beta\hat{k}_1)^2;
\end{aligned}$$

here we have used the orthonormality in the form $(\hat{\boldsymbol{\mathcal{E}}} \wedge \hat{\boldsymbol{\mathcal{B}}})_1 = \hat{\mathcal{E}}_2\hat{\mathcal{B}}_3 - \hat{\mathcal{E}}_3\hat{\mathcal{B}}_2 = \hat{k}_1$ and $\hat{k}_1^2 + \hat{\mathcal{E}}_1^2 + \hat{\mathcal{B}}_1^2 = 1$. With $\hat{k}_1 = \cos\theta$ we can write the transformation of the wave amplitude in general as

$$\mathcal{E}' = \frac{1-\beta\cos\theta}{\sqrt{1-\beta^2}}\mathcal{E} \quad \Leftrightarrow \quad \mathcal{E} = \frac{1+\beta\cos\theta'}{\sqrt{1-\beta^2}}\mathcal{E}'. \tag{4.37}$$

4.8 Energy-Momentum Tensor

Analogous to the transition from the Lagrangian to the Hamiltonian in particle mechanics, one can pass from the Lagrangian density to the Hamiltonian density. Let us first consider the case of free fields with $S_{mf} = 0$. Then $\mathcal{L}$ does not depend explicitly on x_μ and we expect that there is a conservation law associated with the energy density

$$T^\mu{}_\nu = \frac{\partial\mathcal{L}_f}{\partial(\partial A_\sigma/\partial x^\mu)}\frac{\partial A_\sigma}{\partial x^\nu} - \delta^\mu{}_\nu\mathcal{L}_f = -\frac{1}{4\pi}F^{\mu\sigma}\frac{\partial A_\sigma}{\partial x^\nu} - \delta^\mu{}_\nu\mathcal{L}_f \tag{4.38}$$

In fact, after a short calculation we obtain that

$$\begin{aligned}
\frac{\partial T^\mu{}_\nu}{\partial x^\mu} &= -\frac{1}{4\pi}\frac{\partial}{\partial x^\mu}\left(F^{\mu\sigma}\frac{\partial A_\sigma}{\partial x^\nu}\right) - \frac{\partial\mathcal{L}_f}{\partial x^\nu} \\
&= -\frac{1}{4\pi}F^{\mu\sigma}\frac{\partial^2 A_\sigma}{\partial x^\mu x^\nu} + \frac{1}{16\pi}\frac{\partial F_{\mu\sigma}F^{\mu\sigma}}{\partial x^\nu} \\
&= -\frac{1}{4\pi}F^{\mu\sigma}\frac{\partial^2 A_\sigma}{\partial x^\mu x^\nu} + \frac{1}{8\pi}F^{\mu\sigma}\frac{\partial F_{\mu\sigma}}{\partial x^\nu} \\
&= -\frac{1}{4\pi}F^{\mu\sigma}\frac{\partial^2 A_\sigma}{\partial x^\mu x^\nu} + \frac{1}{8\pi}F^{\mu\sigma}\left(\frac{\partial^2 A_\sigma}{\partial x^\mu\partial x^\nu} - \frac{\partial^2 A^\mu}{\partial x^\sigma\partial x^\nu}\right) = 0,
\end{aligned} \tag{4.39}$$

where in the second step we have used the equation of motion $\partial F^{\mu\sigma}/\partial x^\mu = 0$. Integrating the equation (4.39) over space at fixed time, one finds that

$$0 = \int d^3r\,\frac{\partial T^{\mu\nu}}{\partial x^\mu} = \frac{1}{c}\frac{\partial}{\partial t}\int d^3r\,T^{0\nu} + \int d^3r\,\frac{\partial}{\partial r^j}T^{j\nu} = \frac{1}{c}\frac{\partial}{\partial t}\int d^3r\,T^{0\nu}, \tag{4.40}$$

where we have assumed that the fields and hence $T^{\mu\nu}$ fall off sufficiently rapidly so that the integration over the divergence gives no contribution. Thus we obtain the conservation $\partial P^\mu/\partial t = 0$ of the four-vector

$$P^\mu = \frac{1}{c}\int d^3r\,T^{0\mu}, \tag{4.41}$$

which can be identified with the four-momentum vector of the electromagnetic field.

The tensor $T^{\mu\nu}$ is not uniquely determined by the property $\partial T^{\mu\nu}/\partial x^\mu = 0$. Indeed, we can add to $T^{\mu\nu}$ an arbitrary four-divergence $\partial\psi^{\mu\nu\sigma}/\partial x^\sigma$ of a tensor ψ with

$\psi^{\mu\nu\sigma} = -\psi^{\sigma\nu\mu}$ without changing the four-momentum P^μ. We can make the energy-momentum tensor unique by imposing the additional requirement that the four-dimensional angular momentum tensor $L^{\mu\nu}$ is given by the usual form

$$L^{\mu\nu} = \int (x^\mu dP^\nu - x^\nu dP^\mu) = \frac{1}{c}\int d^3r\,(x^\mu T^{0\nu} - x^\nu T^{0\mu}) \tag{4.42}$$

Just as the vanishing of the four-divergence $\partial T^{\mu\nu}/\partial x^\mu$ represents the local formulation of four-momentum conservation, the vanishing of the four-divergence $\partial(x^\mu T^{\sigma\nu} - x^\nu T^{\sigma\mu})/\partial x^\sigma$ yields the angular momentum conservation $\partial L^{\mu\nu}/\partial t = 0$. The local form of angular momentum conservation therefore requires that

$$0 = \frac{\partial(x^\mu T^{\sigma\nu} - x^\nu T^{\sigma\mu})}{\partial x^\sigma} = T^{\mu\nu} - T^{\nu\mu} \tag{4.43}$$

i.e. $T^{\mu\nu}$ is a symmetric tensor.

To symmetrize the tensor in (4.38) we add the term

$$\frac{F^{\mu\sigma}}{4\pi}\frac{\partial A_\nu}{\partial x^\sigma} = \frac{1}{4\pi}\frac{\partial(F^{\mu\sigma}A_\nu)}{\partial x^\sigma}$$

This yields the correctly symmetrized expression for the energy-momentum tensor of the electromagnetic field

$$T^{\mu\nu} = \frac{1}{4\pi}\Big(-F^{\mu\sigma}F^{\nu}{}_{\sigma} + \frac{1}{4}\eta^{\mu\nu}F_{\sigma\tau}F^{\sigma\tau}\Big). \tag{4.44}$$

As a side effect, the symmetrized form of $T^{\mu\nu}$ now depends only on the field strength tensor and no longer directly on the potentials. Moreover, it is traceless, with $T^{\mu}{}_{\mu} = 0$.

One can express the energy-momentum tensor directly in terms of the fields $\boldsymbol{E}$ and $\boldsymbol{B}$. Inserting (3.25) gives

$$(T^{\mu\nu}) = \left(\begin{array}{c|c} u & \boldsymbol{S}^t/c \\ \hline \boldsymbol{S}/c & -\sigma \end{array}\right) \tag{4.45}$$

with the energy density

$$u = \frac{1}{8\pi}(E^2 + B^2), \tag{4.46}$$

the Poynting vector

$$\boldsymbol{S} = \frac{c}{4\pi}\boldsymbol{E}\wedge\boldsymbol{B} \tag{4.47}$$

and the Maxwell stress tensor

$$\sigma_{kl} = \frac{1}{4\pi}\left[E_kE_l + B_kB_l - \frac{1}{2}\delta_{kl}(E^2+B^2)\right]. \tag{4.48}$$

The local conservation law $\partial T^{\mu\nu}/\partial x^\mu$ contains the Poynting theorem (with $\nu = 0$)

$$0 = \frac{\partial T^{\mu 0}}{\partial x^\mu} = \frac{1}{c}\left(\frac{\partial u}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{S}\right) \tag{4.49}$$

and momentum conservation (with $\nu = j$)

$$0 = \frac{\partial T^{\mu j}}{\partial x^\mu} = \frac{1}{c^2}\frac{\partial S_j}{\partial t} - \frac{\partial \sigma_{jk}}{\partial r_k} \tag{4.50}$$

with the momentum flux density σ_{jk}.

By construction the angular-momentum density is likewise conserved in the form $L^{\mu\nu\sigma} = T^{\mu\nu}x^\sigma - T^{\mu\sigma}x^\nu$,

$$\frac{\partial L^{\mu\nu\sigma}}{\partial x^\mu} = \frac{\partial (T^{\mu\nu}x^\sigma - T^{\mu\sigma}x^\nu)}{\partial x^\mu} = 0. \tag{4.51}$$

From the continuity equation (4.43) one obtains local conserved quantities corresponding to the tensors $L^{0\mu\nu}$. As discussed in Chapter 3.4, this tensor consists both of the angular-momentum density $L_k \propto \varepsilon_{klm}L^{0lm}$ and of the moment of mass $N^k \propto L^{00k}$, which is connected with conservation of center-of-mass motion.

4.8.1 Conservation laws in the presence of charged matter

So far we have considered the charge-free field. If now charged particles are present, the energy-momentum tensor acquires a contribution from the charged particles, which we now wish to determine. In the presence of charge, the divergence of the energy-momentum tensor becomes

$$\frac{\partial T^\mu{}_\nu}{\partial x^\mu} = \frac{1}{4\pi}\left(-F^{\mu\sigma}\frac{\partial F_{\nu\sigma}}{\partial x^\mu} - F_{\nu\sigma}\frac{\partial F^{\mu\sigma}}{\partial x^\mu} + \frac{1}{2}F^{\sigma\tau}\frac{\partial F_{\sigma\tau}}{\partial x^\nu}\right). \tag{4.52}$$

Into this equation we substitute the Maxwell equations (4.3) and (4.17) in the form

$$\frac{\partial F_{\sigma\tau}}{\partial x^\nu} = -\frac{\partial F_{\tau\nu}}{\partial x^\sigma} - \frac{\partial F_{\nu\sigma}}{\partial x^\tau}, \quad \frac{\partial F^{\mu\sigma}}{\partial x^\mu} = \frac{4\pi}{c}j^\sigma$$

and obtain

$$\frac{\partial T^\mu{}_\nu}{\partial x^\mu} = \frac{1}{4\pi}\left(-\frac{1}{2}F^{\sigma\tau}\frac{\partial F_{\tau\nu}}{\partial x^\sigma} - \frac{1}{2}F^{\sigma\tau}\frac{\partial F_{\nu\sigma}}{\partial x^\tau} - F^{\mu\sigma}\frac{\partial F_{\nu\sigma}}{\partial x^\mu} - \frac{4\pi}{c}F_{\nu\sigma}j^\sigma\right).$$

By renaming summation indices we can easily see that the first three terms cancel each other. We thus arrive at the result

$$\frac{\partial T^{\mu\nu}}{\partial x^\mu} = -\frac{1}{c}F^{\nu\sigma}j_\sigma. \tag{4.53}$$

The right-hand side is interpreted as the Lorentz four-force density (four-momentum transfer per 4D volume element).

The time component of the equation (4.53) leads to the energy conservation

$$\frac{\partial u}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{S} = -\boldsymbol{j}\cdot\boldsymbol{E} \tag{4.54}$$

and the spatial component leads to the momentum conservation

$$\frac{1}{c^2}\frac{\partial S_j}{\partial t} - \frac{\partial \sigma_{jk}}{\partial r_k} = -\left[\rho E + \frac{1}{c}(\boldsymbol{j} \wedge \boldsymbol{B})\right]_j \tag{4.55}$$

of the field in the presence of matter, described by $j^\mu = (\rho c, \boldsymbol{j})$. The source term $\boldsymbol{j} \cdot \boldsymbol{E}$ in (4.54) precisely describes the power of the field on charged particles via the Lorentz force $\boldsymbol{v} \cdot \boldsymbol{F}_L = q\boldsymbol{v} \cdot \boldsymbol{E}$, see (3.22). The source of momentum conservation (4.55) is, according to Newton's second law, simply the Lorentz force $\boldsymbol{F}_L = q\boldsymbol{E} + q\boldsymbol{v} \wedge \boldsymbol{B}/c$.

The expression (4.41) shows that one can (locally) assign to the fields an energy density $u = T^{00}$ and a momentum density $\boldsymbol{S}_i/c^2 = T^{0i}/c$. These satisfy the local conservation laws (4.54) and (4.55). Here $\boldsymbol{S}$ denotes the energy- and $-\sigma$ the momentum-flux density. The source terms describe the energy and momentum transfer between matter $\leftrightarrow$ and fields.

A further way to express the local energy-momentum conservation is obtained by assigning to the matter an energy-momentum tensor

$$T_m^{\mu\nu}(x) = \sum_{j=1}^{N} \int d\tau_j \, m_j u_j^\mu u_j^\nu \, \delta^{(4)}(x - x_j) = \sum_{j=1}^{N} m_j u_j^\nu(t) \frac{dx_j^\mu(t)}{dt} \, \delta^{(3)}[\boldsymbol{r} - \boldsymbol{r}_j(t)] \tag{4.56}$$

, cf. (4.10). The divergence of the energy-momentum tensor of matter is given by

$$\frac{\partial T_m^{\mu\nu}}{\partial x^\mu} = \sum_{j=1}^{N} m_j c \frac{du_j^\nu}{dt} \, \delta^{(3)}[\boldsymbol{r} - \boldsymbol{r}_j(t)] \tag{4.57}$$

where we have used (4.12). Particles with charge q_j satisfy the equation of motion

$$m_j \frac{du_j^\nu}{d\tau} = \frac{q_j}{c} F^{\nu\sigma} u_\sigma \tag{4.58}$$

see (3.26), or equivalently

$$m_j \frac{du_j^\nu}{dt} = \frac{q_j}{c} F^{\nu\sigma} \frac{dx_{j\sigma}}{dt} \, . \tag{4.59}$$

Hence we obtain the result

$$\frac{\partial T_m^{\mu\nu}}{\partial x^\mu} = \frac{1}{c} F^{\nu\sigma} j_\sigma \tag{4.60}$$

for the divergence of the energy-momentum tensor of matter. A comparison with (4.53) shows that the energy conservation of the total system

$$\frac{\partial T_{\text{tot}}^{\mu\nu}}{\partial x^\mu} = \frac{\partial (T^{\mu\nu} + T_m^{\mu\nu})}{\partial x^\mu} = 0 \tag{4.61}$$

holds.

4.8.2 Energy-momentum tensor of macroscopic bodies

Analogously to (4.9) we wish to treat an assembly of mass points as a continuous system with the mass density

$$\mu(\boldsymbol{r},t)=\sum_{j=1}^{N} m_j\delta^{(3)}[\boldsymbol{r}-\boldsymbol{r}_j(t)]\,. \tag{4.62}$$

We first describe the energy-momentum tensor $T_r^{\mu\nu}$ in the local rest frame. The energy density $T_r^{00}=\mu c^2$ then has only the contribution from the rest energy and the momentum density T_r^{0i} vanishes, $T_r^{0i}=T_r^{i0}=0$.

We wish to describe an isotropic system, so that in the local rest frame the force is the same in all directions. We consider a surface element dS with normal $\boldsymbol{n}$. The momentum flux through this surface element is just the force

$$dF_k=\sigma_{kl}n_l\,dS, \tag{4.63}$$

acting on the surface element. The isotropic force is described by the pressure p with

$$dF_k=-p\,n_k\,dS. \tag{4.64}$$

Equating the two expressions yields the stress tensor $\sigma_{kl}=-p\,\delta_{kl}$ and hence the energy-momentum tensor

$$(T_r^{\mu\nu})=\begin{pmatrix}\mu c^2 & 0 & 0 & 0\\ 0 & p & 0 & 0\\ 0 & 0 & p & 0\\ 0 & 0 & 0 & p\end{pmatrix} \tag{4.65}$$

in the local rest frame.

To find the form of the energy-momentum tensor in a general frame of reference, one must bring it into a covariant form using the four-velocity $u^\mu=\gamma(c,\boldsymbol{v})$. The tensor $T^{\mu\nu}$ must be chosen so that in the rest frame with $u^\mu=(c,\boldsymbol{0})$ it assumes the above form. It is easy to see that

$$T^{\mu\nu}=\left(\mu+\frac{p}{c^2}\right)u^\mu u^\nu-p\,\eta^{\mu\nu} \tag{4.66}$$

satisfies this requirement.[8] This expression represents the energy-momentum tensor of a macroscopic body. From (4.66) one obtains directly

$$u=\frac{\mu c^2+pv^2/c^2}{1-v^2/c^2},\quad \boldsymbol{S}=\frac{(\mu c^2+p)\boldsymbol{v}}{1-v^2/c^2},\quad \sigma_{kl}=-\frac{(\mu+p/c^2)v_kv_l}{1-v^2/c^2}-p\delta_{kl}.$$

From the expression (4.66) it follows at once

$$T^\mu{}_\mu=\mu c^2-3p. \tag{4.67}$$

[8] One may also obtain (4.66) by Lorentz-transforming the expression in the local rest frame.

By the stability condition $T^{\mu}{}_{\mu} \geq 0$ one obtains the inequality[9]

$$p \leq \frac{\mu c^2}{3}, \tag{4.68}$$

where the equality sign is never attained for massive particles.

As already noted for (4.44), one finds that for the energy-momentum tensor of the radiation field (without matter) $T^{\mu}{}_{\mu} = 0$ holds. Thus for the isotropic radiation field the radiation pressure (with $\sigma_{kl} = -p\delta_{kl}$)

$$p = \frac{u}{3}. \tag{4.69}$$

Equation (4.69) is the equation of state of the electromagnetic field. From it, using the first and second law of thermodynamics without additional assumptions, one can derive the Stefan-Boltzmann law. By the first law

$$dU = TdS - pdV\,. \tag{4.70}$$

Hence we obtain

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial S}{\partial V}\right)_T - p \tag{4.71}$$

by differentiating with respect to V at fixed T. The second law yields the Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V$$

for the free energy. From (4.71) it follows with $U = uV$ and $p = u/3$ immediately

$$u(T) = \frac{T}{3}\frac{\partial u(T)}{\partial T} - \frac{u(T)}{3}\,. \tag{4.72}$$

Equivalently one finds

$$\frac{\partial \log u(T)}{\partial \log T} = \frac{T}{u(T)}\frac{\partial u(T)}{\partial T} = 4. \tag{4.73}$$

The last equation, upon integration, yields the Stefan-Boltzmann law $u(T) \propto T^4$.

4.8.3 Ideal fluid

We consider an assembly of mass points with density μ and (isotropic) pressure p, which move locally with velocity $\boldsymbol{v}(\boldsymbol{r}, t)$. The energy-momentum tensor $T^{\mu\nu}$ has the form

$$T^{\mu\nu} = \left(\mu + \frac{p}{c^2}\right) u^{\mu}u^{\nu} - p\eta^{\mu\nu}$$

[9] This condition follows from (4.56) with $T^{\mu}{}_{\mu} = \sum_j \int d\tau_j\, m_j c^2 \delta^{(4)}(x - x_j) \geq 0$.

with the local four-velocity $u^\mu(\boldsymbol{r},t) = [c, \boldsymbol{v}(\boldsymbol{r},t)]/\sqrt{1-v^2/c^2}$, see (4.66). If the mass points represent an ideal fluid without viscosities, the local energy-momentum conservation holds[10]

$$0 = \frac{\partial T^{\mu\nu}}{\partial x^\mu} = \left(\mu + \frac{p}{c^2}\right)\left(u^\nu \frac{\partial u^\mu}{\partial x^\mu} + u^\mu \frac{\partial u^\nu}{\partial x^\mu}\right) + u^\nu u^\mu \frac{\partial}{\partial x^\mu}\left(\mu + \frac{p}{c^2}\right) - \frac{\partial p}{\partial x_\nu}.$$

Contracting with u_ν we obtain the equation

$$\begin{aligned} 0 &= \left(\mu + \frac{p}{c^2}\right)\left(c^2 \frac{\partial u^\mu}{\partial x^\mu} + u^\mu \overbrace{u_\nu \frac{\partial u^\nu}{\partial x^\mu}}^{=0}\right) + c^2 u^\mu \frac{\partial}{\partial x^\mu}\left(\mu + \frac{p}{c^2}\right) - u^\mu \frac{\partial p}{\partial x^\mu} \\ &= c^2 \frac{\partial(\mu u^\mu)}{\partial x^\mu} + p \frac{\partial u^\mu}{\partial x^\mu}. \end{aligned} \tag{4.74}$$

In the non-relativistic limit $u^\mu = (c, \boldsymbol{v})$ and $p \ll \mu c^2$. Hence (4.74) simplifies to the continuity equation

$$0 = \frac{\partial(\mu u^\mu)}{\partial x^\mu} = \frac{\partial \mu}{\partial t} + \boldsymbol{\nabla} \cdot (\mu \boldsymbol{v}).$$

Subtracting from $\partial T^{\mu\nu}/\partial x^\mu$ the equation (4.74) multiplied by u^ν/c^2, we obtain the additional conditions

$$0 = \left(\mu + \frac{p}{c^2}\right) u^\mu \frac{\partial u^\nu}{\partial x^\mu} - \frac{\partial p}{\partial x_\nu} + \frac{u^\mu u^\nu}{c^2} \frac{\partial p}{\partial x^\mu}. \tag{4.75}$$

In the non-relativistic limit the spatial components with $\nu = j$ reduce to the Euler equations

$$0 = \mu u^\mu \frac{\partial v_j}{\partial x^\mu} + \frac{\partial p}{\partial r_j} \quad \Leftrightarrow \quad \mu \frac{\partial \boldsymbol{v}}{\partial t} + \mu(\boldsymbol{v} \cdot \boldsymbol{\nabla})\boldsymbol{v} = -\boldsymbol{\nabla} p,$$

which relate the convective time derivative of the velocity field to the pressure gradient.

The energy tensor of the ideal fluid plays a central role in cosmology, because the constituents of the universe, due to their weak interactions, behave to good approximation like ideal fluids.

4.9 Thermodynamics

4.9.1 Blackbody radiation

The principle of relativity allows us to derive from Planck's radiation law

$$K'_{\omega'} = \frac{\hbar \omega'^3}{4\pi^3 c^2 (e^{\hbar\omega'/k_B T} - 1)} \tag{4.76}$$

[10] We assume that the mass points are uncharged and hence there is no coupling to the electromagnetic field.

for the intensity $dI' = K'_{\omega'} d\omega' d\Omega'$ of blackbody radiation at temperature T and angular frequency ω' in the rest frame, the corresponding quantity in the laboratory frame moving relative to it. The intensity is determined by the magnitude of the Poynting vector. For a light wave (4.34) we have the intensity $dI \propto |\mathcal{E} \wedge \mathcal{B}| \propto \mathcal{E}^2$. From (4.37) we obtain $dI'/dI = \gamma^2(1 - \beta\cos\theta)^2$ and with the formulas (4.36) for the Doppler shift of the angular frequency and the transformation formula for a solid angle element

$$d\Omega' = \frac{1-\beta^2}{(1-\beta\cos\theta)^2} d\Omega, \tag{4.77}$$

see (2.26), we find

$$K'_{\omega'} = K_\omega \gamma^3 (1 - \beta\cos\theta)^3 . \tag{4.78}$$

Planck's radiation law therefore takes in the laboratory frame the form

$$K_\omega = \frac{K'_{\omega'}}{\gamma^3(1-\beta\cos\theta)^3} = \frac{\hbar\omega^3}{4\pi^3 c^2 [e^{(\hbar\omega/k_B T)\gamma(1-\beta\cos\theta)} - 1]}. \tag{4.79}$$

A comparison with (4.76) shows that most factors cancel and only the Doppler shift $\omega' = \gamma\omega(1 - \beta\cos\theta)$ remains.

After integrating (4.79) over all frequencies we obtain the directional dependence of the intensity of blackbody radiation[11]

$$I(\theta) = \int_0^\infty d\omega\, K_\omega = \frac{\sigma}{\pi} \frac{T^4}{\gamma^4(1-\beta\cos\theta)^4} \tag{4.80}$$

in the laboratory frame with the Stefan-Boltzmann constant $\sigma = \pi^2 k_B^4/60\hbar^3 c^2$. The intensity is maximal in the forward direction with $\theta \approx 0$, see Figure 2.2.

The total energy present in the cavity of volume V is obtained from the relation $|\boldsymbol{S}| = cu$, which holds for electromagnetic radiation. Hence

$$U = V \int d\Omega \frac{I(\theta)}{c} = \frac{4\gamma^2(1+\frac{1}{3}\beta^2)\sigma}{c} V T^4 = \gamma(1+\tfrac{1}{3}\beta^2)U', \tag{4.81}$$

where we have used that $V' = \gamma V$. Equation (4.81) can also be obtained directly from the transformation $u = T^{00} = \Lambda^0{}_\mu(-\chi)\Lambda^0{}_\nu(-\chi)T'^{\mu\nu} = \gamma^2(u' - \beta^2\sigma'_{11}) = \gamma^2(u' + \frac{1}{3}\beta^2 u')$, using that $\sigma_{11} = -p = -\frac{1}{3}u$ holds, cf. (4.69).

4.9.2 Ideal gas

We expect relativistic effects to become noticeable in the treatment of ideal gases of particles of mass m in statistical mechanics once the typical speed v_{typ}, given by $m v_{\text{typ}}^2 = k_B T$, becomes comparable to the speed of light. We therefore introduce the dimensionless parameter

$$\alpha = \frac{c^2}{v_{\text{typ}}^2} = \frac{mc^2}{k_B T} \tag{4.82}$$

[11] We use the result $\int_0^\infty d\omega\, \omega^3/(e^{\alpha\omega} - 1) = \pi^4/15\alpha^4$.

In the non-relativistic limit $\alpha \gg 1$ and one achieves $\alpha = 1$ at temperature $T_c = mc^2/k_B$. Relativistic effects in the treatment of ideal gases are academic, since T_c for hydrogen already at 10^{13} K. Nevertheless we wish to determine the thermodynamic quantities for a relativistic ideal gas.

The energy of a particle is given by the energy-momentum relation

$$E(p) = \sqrt{m^2c^4 + p^2c^2}. \tag{4.83}$$

The free energy has the form $F = -Nk_BT \log Z_1$, where (with $p = mc \sinh \chi$)

$$\begin{aligned} Z_1 &= V \int d^3p\, e^{-E(p)/k_BT} = 4\pi V \int_0^\infty dp\, p^2 e^{-E(p)/k_BT} \\ &= 4\pi V m^3 c^3 \int_0^\infty d\chi\, \cosh\chi \sinh^2\chi\, e^{-\alpha\cosh\chi} \\ &= 4\pi V m^3 c^3 \frac{K_2(\alpha)}{\alpha} \end{aligned} \tag{4.84}$$

is the single-particle partition function; here we have introduced the modified Bessel function of the second kind $K_2(x) = \int_0^\infty dt\, \cosh(2t) \exp(-x \cosh t)$. The free energy is therefore, up to a constant, given by

$$F(V,T,N) = -Nk_BT[\log V + \log K_2(\alpha) - \log \alpha]. \tag{4.85}$$

The remaining thermodynamic quantities follow from the usual relations

$$p = -\frac{\partial F}{\partial V}, \qquad U = F - T\frac{\partial F}{\partial T} = -T\frac{\partial (F/T)}{\partial \log T}. \tag{4.86}$$

The first equation yields the equation of state

$$p = \frac{Nk_BT}{V}, \tag{4.87}$$

which agrees with that of the non-relativistic ideal gas. However, the equation for the internal energy is modified and we now have (using that $d \log T = -d \log \alpha$)

$$U = Nk_BT\left[1 - \alpha \frac{d \log K_2(\alpha)}{d\alpha}\right]. \tag{4.88}$$

For $\alpha \gg 1$ we have the asymptotic form $K_2(\alpha) \simeq \sqrt{\pi/2\alpha}\, e^{-\alpha}$ and hence $d \log K_2(\alpha)/d\alpha \simeq -1 - 1/(2\alpha)$ and

$$U \simeq Nk_BT(\alpha + \tfrac{3}{2}) = Nmc^2 + \tfrac{3}{2}Nk_BT, \tag{4.89}$$

which, apart from the rest energy, agrees with the familiar result. By expanding $K_2(\alpha)$ to next order one obtains the first relativistic correction to the internal energy

$$\Delta U = \frac{15k_BT}{8mc^2} Nk_BT. \tag{4.90}$$

Chapter 5

General Relativity

We have so far written the equation of motion of a particle only in inertial frames. From classical mechanics we know that in non-inertial reference frames additional fictitious forces appear. The *general theory of relativity* (GR) extends the special theory of relativity by also admitting non-inertial reference frames. As we will see later, the gravitational force is then reinterpreted simply as a fictitious force in a non-inertial frame.

5.1 Geodesic equation

The description of fictitious forces in relativity is not so straightforward, because accelerated reference frames, as we have already seen in the exercises with the example of Thomas precession, immediately lead to curved spaces. In these spaces the Minkowski arc length ds, which is an invariant, is described by the (differential) expression (metric)[1]

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \tag{5.1}$$

cf. (2.4). This expression is physically invariant, since ds/c gives the proper time of an observer moving on a worldline from x^μ to $x^\mu + dx^\mu$. Here the coordinate-dependent symmetric matrix $g_{\mu\nu}(x)$ takes over the role of the Minkowski metric $\eta_{\mu\nu}$ in an inertial system. In any other reference frame $x'^\mu \equiv x'^\mu(x)$ one therefore has unchanged

$$ds^2 = g'_{\mu\nu}(x') dx'^\mu dx'^\nu \,. \tag{5.2}$$

Thus the metric tensor transforms as a rank-2 covariant tensor

$$g_{\mu'\nu'}(x) = g'_{\sigma\tau}(x') \frac{\partial x'^\sigma}{\partial x^\mu} \frac{\partial x'^\tau}{\partial x^\nu} \,. \tag{5.3}$$

[1] It is conventional to write the squared distance $ds^2 = (ds)^2$ in order to minimize the number of parentheses. If needed, we therefore use parentheses in the expression $d(s^2) = 2sds$.

Geodesics $x(\tau)$ ("shortest" connections) are characterized by the condition

$$\delta \int_{(1)}^{(2)} ds = \delta \int_{t^{(1)}}^{t^{(2)}} dt \, \sqrt{g_{\mu\nu}(x)\dot{x}^\mu \dot{x}^\nu} \tag{5.4}$$

where the endpoints $x^{(1)}$ and $x^{(2)}$ are held fixed, cf. (3.13). As already seen in Chapter 3.3.1, force-free particles move on worldlines that correspond to an extremum of the proper time $\int d\tau = \int ds/c$. Correspondingly, force-free particles in general relativity follow geodesics. Gravity here acts as a fictitious force that modifies the explicit form of the geodesics. It is thus "hidden" in the metric.

To obtain an expression for the geodesic equation we must vary the action (5.4). It must be noted that the variation is to be carried out not at fixed arc length but at fixed endpoints. One therefore has to introduce a parametrization whose limits $t^{(1)}$ and $t^{(2)}$ are fixed, with $x(t^{(1)}) = x^{(1)}$ and $x(t^{(2)}) = x^{(2)}$, with the goal of varying the right-hand side of (5.4). The variation is rather cumbersome because of the square root. In practice one therefore uses the following trick. One introduces the variational problem

$$0 = \delta \int_{(1)}^{(2)} dt \, [\lambda^{-1} g_{\mu\nu}(x)\dot{x}^\mu \dot{x}^\nu + \lambda] \tag{5.5}$$

with an auxiliary function $\lambda(t)$. Varying $\lambda(t)$ leads to the condition that $\lambda^{-2} g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 1$, or $g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = \lambda^2$. At the extremum (i.e., on the geodesic) (5.5) is therefore of the form $0 = \delta \int dt \, \lambda$. Inserting $\lambda = (g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu)^{1/2} = ds/dt$ shows that the variational problem (5.5) is equivalent to (5.4).

If we now parametrize the worldline by $c\, d\tau' = \lambda(t)\, dt$ we obtain

$$\begin{aligned} 0 &= \delta \int_{(1)}^{(2)} d\tau' \left[\frac{g_{\mu\nu}(x)}{c} \frac{dx^\mu}{d\tau'} \frac{dx^\nu}{d\tau'} + c\, d\tau' \right] \\ &= \frac{1}{c} \delta \int_{(1)}^{(2)} d\tau' \, g_{\mu\nu}(x) \frac{dx^\mu}{d\tau'} \frac{dx^\nu}{d\tau'} + \underbrace{\delta \int_{(1)}^{(2)} c\, d\tau'}_{=0} . \end{aligned} \tag{5.6}$$

At the extremum $d\tau' = \lambda\, dt/c = ds/c$ corresponds precisely to the proper time $d\tau$ on the geodesic.[2] Thus we have shown that (5.4) is equivalent to

$$0 = \delta \int d\tau \, g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = \delta \int d\tau \, g_{\alpha\beta}(x) \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} , \tag{5.7}$$

provided the worldline on the geodesic is parametrized by the proper time.

The Euler-Lagrange equations for the variational principle (5.7) are the geodesic equation. From $(d/d\tau)(\partial_{dx^\nu/d\tau} L) - \partial_{x^\nu} L = 0$, using the Lagrangian from (5.7), one

[2] Away from the geodesic τ' must of course differ from the proper time, since we wish to hold $\tau'^{(1)}$, $\tau'^{(2)}$ fixed even though the proper time changes.

obtains the relations

$$0 = 2\frac{d}{d\tau}\left(g_{\nu\beta}\frac{dx^\beta}{d\tau}\right) - \frac{\partial g_{\alpha\beta}}{\partial x^\nu}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 2g_{\nu\beta}\frac{d^2x^\beta}{d\tau^2} + \left(2\frac{\partial g_{\nu\beta}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu}\right)\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau}\,. \quad (5.8)$$

Multiplying these equations by the matrix inverse to $g_{\mu\nu}$, $g^{\mu\nu}$, yields the equations of motion (geodesic equation)

$$\frac{d^2x^\mu(\tau)}{d\tau^2} = -\Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha(\tau)}{d\tau}\frac{dx^\beta(\tau)}{d\tau}, \quad (5.9)$$

with τ the proper time and the Christoffel symbols[3] (symmetrized over $\alpha \leftrightarrow \beta$)

$$\Gamma^\mu{}_{\alpha\beta} = \frac{1}{2}g^{\mu\nu}\left(\frac{\partial g_{\alpha\nu}}{\partial x^\beta} + \frac{\partial g_{\beta\nu}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu}\right)\,. \quad (5.10)$$

5.2 Equivalence principle

Already Newton had noticed that the inertial mass m_t, which determines the dynamics of a body in an inertial system via $m_t\ddot{\boldsymbol{r}} = \boldsymbol{F}$, equals the gravitational mass m_s of the body that appears in the law of gravitation $\boldsymbol{F} = m_s\boldsymbol{g}$. This equality can be tested, for example, by experiments with simple pendulums of different composition, since their period is proportional to $\sqrt{m_t/m_s}$.

Assuming the equivalence of gravitational and inertial mass, one can measure no gravity in a freely falling frame in a homogeneous and stationary gravitational field. For example, for N mass points with pairwise interactions $\boldsymbol{F}(\boldsymbol{r}_j - \boldsymbol{r}_k)$ in the nonrelativistic limit the equations of motion are

$$m_j\ddot{\boldsymbol{r}}_j = m_j\boldsymbol{g} + \sum_k \boldsymbol{F}(\boldsymbol{r}_j - \boldsymbol{r}_k), \quad (5.11)$$

where the vectors $\boldsymbol{r}_j$ are measured relative to the Earth, which produces the gravitational force $m_j\boldsymbol{g}$. If we now transform to the co-falling system

$$\boldsymbol{r}' = \boldsymbol{r} - \frac{1}{2}\boldsymbol{g}t^2, \qquad t' = t, \quad (5.12)$$

the equations of motion become

$$m_j\ddot{\boldsymbol{r}}'_j = \sum_k \boldsymbol{F}(\boldsymbol{r}'_j - \boldsymbol{r}'_k). \quad (5.13)$$

An experimenter in the freely falling frame finds the same laws of mechanics as an experimenter in an inertial frame without gravitational force.

The *equivalence principle* (Einstein 1907) generalizes this situation to all laws of nature:

[3] The Christoffel symbols are not tensors, since they do not have the correct transformation properties under coordinate transformations.

> Given an event in an arbitrary gravitational field, one can always find a reference frame in which, within a sufficiently small region, the laws of nature take on the same form as in an unaccelerated Cartesian coordinate system in the absence of gravitational forces.

Such a reference frame is called a *local inertial frame*.[4] In a local inertial frame the laws of special relativity hold. The gravitational force is merely a fictitious force that can be eliminated by choosing an appropriate local coordinate system. Thus the equality of gravitational and inertial mass becomes an equivalence. However, spacetime is in general curved. The equation of motion of a force-free particle[5] is described by the geodesic equation (5.9), where the right-hand side takes over the role of the gravitational force (divided by the mass of the particle).

As an example we consider the above transformation (5.12). Since the primed coordinate system is a local inertial system, the Minkowski expression $ds^2 = c^2dt'^2 - d\boldsymbol{r}'^2$ holds for the proper time. Because the proper time is invariant, in the unprimed system one finds

$$ds^2 = c^2dt'^2 - d\boldsymbol{r}'^2 = c^2dt^2 - d\boldsymbol{r}^2 - g^2t^2dt^2 + 2t\,\boldsymbol{g}\cdot d\boldsymbol{r}\,dt = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{5.14}$$

with the metric

$$g_{\mu\nu}(x) = \left(\begin{array}{c|c} 1 - g^2t^2/c^2 & \boldsymbol{g}^t\,t/c \\ \hline \boldsymbol{g}t/c & -I_3 \end{array}\right). \tag{5.15}$$

The only nonvanishing elements of the Christoffel symbol are then given by

$$\Gamma^j{}_{00} = -\frac{g_j}{c^2}, \tag{5.16}$$

so that the geodesic equations take the form

$$\frac{d^2t}{d\tau^2} = 0, \qquad \frac{d^2\boldsymbol{r}}{d\tau^2} = \boldsymbol{g}\left(\frac{dt}{d\tau}\right)^2$$

The first equation is solved by $t = \alpha\tau + \beta$ and the second is then equivalent to Newton's law $m\ddot{\boldsymbol{r}} = \boldsymbol{F}_G$ with the gravitational force $\boldsymbol{F}_G = m\boldsymbol{g}$, cf. (5.11), without interaction.

5.3 Distances and time intervals

In general relativity the coordinates x^μ may be chosen arbitrarily. The connection between different reference frames is provided by the invariant proper time $d\tau = ds/c$ with the arc length (5.1). One therefore asks how the general coordinates relate to the actual distances and time intervals on which all observers (coordinate systems) can agree.

[4] In a local inertial frame one has $\Gamma^\mu{}_{\alpha\beta} = 0$.

[5] In general relativity a particle acted on only by gravitational force and not by Lorentz force is a force-free particle, since the former is merely a fictitious force.

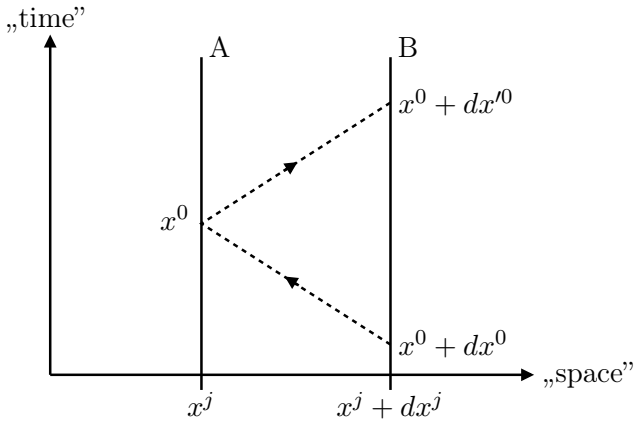

Figure 5.1: Synchronization of two neighboring coordinate systems at spatial points A and B by exchange of light signals. A light pulse is emitted at location B at time $x^0 + dx^0$, hits the mirror at location A at time x^0, before returning to location B at time $x^0 + dx'^0$.

As in special relativity, the proper time τ at a spacetime point is a covariant quantity, namely the time that a clock at rest in the local rest frame would measure. For such a clock one has $dx^1 = dx^2 = dx^3 = 0$ and the Minkowski distance ds is then simply $c\,d\tau$. We thus obtain

$$ds^2 = c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu = g_{00}(dx^0)^2$$

and hence

$$d\tau = \frac{1}{c}\sqrt{g_{00}}\,dx^0, \tag{5.17}$$

from which

$$\tau = \frac{1}{c}\int dx^0\,\sqrt{g_{00}} \tag{5.18}$$

follows as the relation between the proper time and the coordinates x^μ.

Next we wish to determine an expression for the spatial distance dl. It is not possible simply to set $dx^0 = 0$ in the expression (5.1), since in GR the proper time at different spatial points is related to the coordinate dx^0 in different ways.

To define the spatial distance we use a similar idea as in the derivation of length contraction in special relativity. We consider two infinitesimally neighboring spatial points A and B with coordinates x^j and $x^j + dx^j$. We imagine that at location B a light signal is emitted at time $x^0 + dx^0$, reflected at A at time x^0, and detected later at time $x^0 + dx'^0$ at location B (with fixed spatial coordinates $dx^j = 0$), see Figure 5.1. The time $d\tau$ (measured at location A) required for this is just twice the distance $2dl/c$ between the two locations.

The Minkowski distance of two events is given by

$$ds^2 = g_{jk} dx^j dx^k + 2g_{0j} dx^0 dx^j + g_{00}(dx^0)^2. \tag{5.19}$$

Events connected by a light signal satisfy $ds^2 = 0$. Solving $ds^2 = 0$ for dx^0 yields the two solutions

$$dx^0 = \frac{1}{g_{00}}\left[-g_{0j}dx^j - \sqrt{(g_{0j}g_{0k} - g_{jk}g_{00})dx^j dx^k}\right]$$
$$dx'^0 = \frac{1}{g_{00}}\left[-g_{0j}dx^j + \sqrt{(g_{0j}g_{0k} - g_{jk}g_{00})dx^j dx^k}\right],$$

which correspond exactly to the emission and reception of the light signal at location B. The proper-time difference $d\tau$ is then just $dx'^0 - dx^0$ times the time-dilation factor $\sqrt{g_{00}}/c$. The distance dl therefore comes out as

$$dl^2 = (cd\tau/2)^2 = \gamma_{jk}dx^j dx^k \tag{5.20}$$

with the spatial metric

$$\gamma_{jk} = -g_{jk} + \frac{g_{0j}g_{0k}}{g_{00}}. \tag{5.21}$$

It is important to remember that γ_{jk} in general depends on x^0, so that the spatial metric changes with time. It therefore makes no sense to speak of distances greater than infinitesimal ones, since distances depend on the worldline on which they are measured.

As in special relativity, light signals can also be used to synchronize clocks at different spatial points and thereby to define the notion of "simultaneity." To this end we again consider the exchange of light signals between spacetime points A and B. We define the midpoint $x^0 + (dx'^0 + dx^0)/2c$ between the time of emission $(x^0 + dx^0)$ and reception $(x^0 + dx'^0)$ of the light signal at location B as simultaneous with time x^0 at location A. The time difference Δt (measured on the time scale x^0/c) between the time at location B and the time at location A is then given by

$$\Delta t = \frac{1}{2c}\left(dx'^0 + dx^0\right) = -\frac{g_{0j}dx^j}{cg_{00}}. \tag{5.22}$$

Knowledge of the time difference allows one to synchronize the clocks at the infinitesimally neighboring positions A and B relative to the observer A. One can continue this synchronization successively, but it is not possible to synchronize clocks on a closed curve, since in general $\oint \Delta t \neq 0$. Since in general relativity clocks at different spatial points in the same reference frame run at different rates (g_{00} in general depends on x^j), the synchronization does not persist. The impossibility of synchronizing clocks is, however, a property of the reference frame, since one can always choose a system with $g_{0j} = 0$.

5.4 Accelerated reference frames

In this chapter we study two simple examples of accelerated reference frames in flat Minkowski spacetime. These examples can also be treated within the framework

of special relativity (which we have partly done in the exercises), by describing everything from an inertial frame and then transforming to the local rest frame in each case. One example was the expression (1.17) for the proper time of a generally accelerated observer.

5.4.1 Rest frame of a uniformly accelerated observer

We consider a reference frame undergoing uniform acceleration a along the x direction. An example is the rest frame of a charged particle in a homogeneous electric field $\boldsymbol{E} = (am/q, 0, 0)$. Since the y and z coordinates are unaffected by the acceleration, we restrict ourselves to the 2×2 metric in the (ct, x) subspace. The worldline was determined in (3.30); it is given by a hyperbola of the form

$$x = h \cosh \chi = \frac{c^2}{a} \sqrt{1 + (at/c)^2} \tag{5.23}$$

with vertex $h = c^2/a$ and $\sinh \chi = at/c$, i.e., $ct = h \sinh \chi$ and $dx/dt = c \tanh \chi$.

We now wish to introduce new coordinates. We change from x and t to $h \geq 0$ and χ. An observer at rest with respect to h therefore undergoes uniform acceleration in the inertial system. The coordinate χ describes the rapidity (which grows linearly) and h the (inverse) acceleration. We obtain

$$dx = \cosh \chi \, dh + h \sinh \chi \, d\chi, \qquad c \, dt = \sinh \chi \, dh + h \cosh \chi \, d\chi$$

and hence

$$ds^2 = c^2 dt^2 - dx^2 = h^2 \, d\chi^2 - dh^2. \tag{5.24}$$

with the (diagonal) metric tensor

$$(g_{\mu\nu}) = \begin{pmatrix} h^2 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.25}$$

which is also called the *Rindler metric*. The Rindler metric refers to the Rindler coordinates $(x^\mu) = (\chi, h)$. Here $x^0 = \chi$ is a timelike coordinate (since $ds^2 > 0$ for $d\chi \neq 0$ at $dh = 0$) and $x^1 = h$ a spacelike coordinate (since $ds^2 < 0$ for $dh \neq 0$ at $d\chi = 0$).

The proper time of a uniformly accelerated observer (h fixed) is given by

$$\tau = \frac{1}{c} \int dx^0 \sqrt{g_{00}} = \frac{1}{c} \int d\chi \, h = \frac{c\chi}{a} = \frac{c \operatorname{arsinh}(at/c)}{a}, \tag{5.26}$$

cf. (3.31).

The Rindler coordinates represent only events with $0 \leq x \leq ct$, see Figure 5.2. Indeed all events with $x^2 = c^2 t^2$ lie at $h = 0$, where the Rindler metric is no longer invertible. Thus $h = 0$ represents the edge of the Rindler map. A further peculiarity

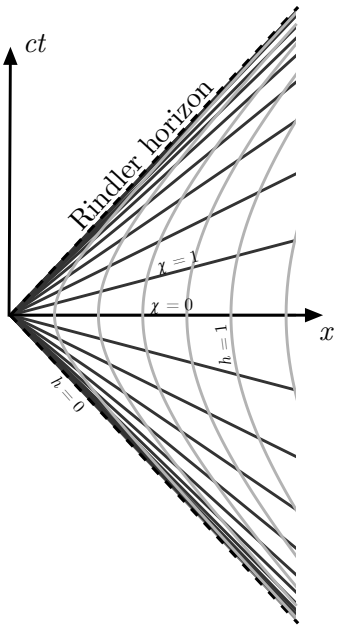

Figure 5.2: The Rindler coordinates are shown in the (t, x) plane, which describes a flat Minkowski spacetime. The χ axes (dark) run mainly horizontally and the h axes (light) vertically. Both become collinear at the Rindler horizon with $x^2 = c^2t^2$ (dashed line).

of the line $h = 0$ arises because no observer moving with constant finite acceleration ($h =$ const.) can receive light from the region with $ct \geq x$.[6] Nor can any of the observers send light into the region with $ct \leq -x$. From the definition of h one sees moreover that at the horizon the acceleration a diverges. In this sense the line $h = 0$ represents an (event) horizon, the so-called *Rindler horizon*. As we will see later, this horizon corresponds to the *event horizon* of a *black hole*, since the Rindler metric is precisely the Schwarzschild metric of an observer at fixed distance from the center (in accordance with the equivalence of acceleration and gravitation).

5.4.2 Geodesics

The geodesics of the Rindler coordinates are of course simply straight lines in the t, x coordinates. It is instructive to solve the geodesic equation (5.9) directly in this simple case in order to "check" this fact. Most Christoffel symbols vanish. The only nonvanishing elements are $\Gamma^0{}_{01} = \Gamma^0{}_{10} = h^{-1}$ and $\Gamma^1{}_{00} = h$. Written out, the geodesic equations (5.9) are therefore

$$\frac{d^2\chi}{d\tau^2} + \frac{2}{h}\frac{d\chi}{d\tau}\frac{dh}{d\tau} = 0, \qquad \frac{d^2h}{d\tau^2} + h\frac{d\chi}{d\tau}\frac{d\chi}{d\tau} = 0. \tag{5.27}$$

[6] Communication with the region $ct > x$ is not, however, in principle impossible. For example, an observer can stop and then cross the Rindler horizon, after which he can receive signals from the region $ct > x$.

The first equation is equivalent to

$$h^{-2}\frac{d}{d\tau}\Big(h^2\frac{d\chi}{d\tau}\Big) = 0$$

with the general solution $d\chi/d\tau = c\alpha/h^2$, $\alpha \in \mathbb{R}$. Moreover we know that for a massive particle

$$c^2 = \frac{ds^2}{d\tau^2} = h^2\left(\frac{d\chi}{d\tau}\right)^2 - \left(\frac{dh}{d\tau}\right)^2$$

holds, from which $dh/d\tau = \pm c\sqrt{\alpha^2/h^2 - 1}$ follows (and hence $h \leq \alpha$). Together with the equation for $d\chi/d\tau$ we obtain

$$\frac{dh}{d\chi} = \pm h\sqrt{1 - h^2/\alpha^2}.$$

It is helpful to replace the variable h by β with

$$\pm\beta = \operatorname{artanh}(\sqrt{1 - h^2/\alpha^2}) = \operatorname{arcosh}(\alpha/h).$$

Thus $d\log h/d\chi = \tanh\beta$. One shows directly that

$$\frac{d\beta}{d\chi} = \pm\frac{d\operatorname{arcosh}(\alpha/h)}{d\log h}\tanh\beta = \mp\frac{1}{\sqrt{1 - h^2/\alpha^2}}\tanh\beta = -1$$

and hence $\beta = -\chi + \beta_0$, $\beta_0 \in \mathbb{R}$. For the velocity one therefore has

$$\begin{aligned}\frac{v}{c} = \frac{1}{c}\frac{dx}{dt} &= \left[\frac{d(h\sinh\chi)}{d\chi}\right]^{-1}\frac{d(h\cosh\chi)}{d\chi} = \frac{\tanh\chi + (d\log h/d\chi)}{1 + (d\log h/d\chi)\tanh\chi} \\ &= \tanh(\beta + \chi) = \tanh\beta_0,\end{aligned}$$

i.e., it is constant along the geodesics. At time $t = 0$ the geodesic is at location $x_0 = h$ with

$$\cosh\beta = \cosh\beta_0 = \alpha/|x_0|,$$

i.e., $x_0 = \pm\sqrt{1 - v^2/c^2}\,\alpha$. The integration constants β_0 and α therefore determine the velocity v and the initial position x_0. The geodesics are thus straight lines

$$x = x_0 + vt$$

in the original coordinates.

5.4.3 Unruh effect

The Unruh effect arises because the quantum fluctuations of the electromagnetic field appear to a uniformly accelerated observer as blackbody radiation. The velocity of an observer undergoing uniform acceleration (h fixed) increases steadily. Therefore the frequency of a source in the inertial system is increasingly Doppler shifted. As

an example we consider a plane wave proportional to $\psi = e^{i(x+ct)\omega/c}$ propagating opposite to the direction of acceleration. In Rindler coordinates the wave has the form

$$\psi = \exp\left(ihe^{\chi}\omega/c\right), \tag{5.28}$$

where the factor he^{χ} corresponds precisely to the blueshift. For the accelerated observer (h fixed) the wave has the spectrum[7]

$$S(\Omega) \propto \tilde{S}(\Omega) = \left| \int_{-\infty}^{\infty} d\tau \, e^{i\Omega\tau} \psi \right|^2 \tag{5.29}$$

with the proper time $\tau = h\chi/c = c\chi/a$.

The integral $\tilde{S}$ can be evaluated, and we obtain

$$\tilde{S}(\Omega) = \frac{2\pi h}{\Omega c} \frac{1}{e^{2\pi h\Omega/c} - 1}, \tag{5.30}$$

independent of the frequency ω in the rest frame. The spectral distribution looks like blackbody radiation at a temperature

$$T_U = \frac{\hbar c}{2\pi k_B h} = \frac{\hbar a}{2\pi k_B c}. \tag{5.31}$$

5.5 Rotating reference frame

As a second example, let us consider the case of a uniformly rotating reference frame with angular velocity Ω. Since no acceleration occurs along the rotation axis, we restrict the discussion to the two spatial directions orthogonal to the rotation axis. We denote the coordinates in the unaccelerated inertial system by t', x', y' with the metric

$$ds^2 = c^2 dt'^2 - dx'^2 - dy'^2. \tag{5.32}$$

The (polar) coordinates in the rotating system are denoted by r, φ, t. They are related via $x' = r\cos(\varphi + \Omega t), y' = r\sin(\varphi + \Omega t), t' = t$ to the coordinates in the inertial system. Substituting this into (5.32), we obtain

$$ds^2 = (c^2 - \Omega^2 r^2)\, dt^2 - dr^2 - r^2\, d\varphi^2 - 2\Omega r^2\, dt\, d\varphi, \tag{5.33}$$

which corresponds to the metric

$$(g_{\mu\nu}) = \begin{pmatrix} c^2 - \Omega^2 r^2 & 0 & -\Omega r^2 \\ 0 & -1 & 0 \\ -\Omega r^2 & 0 & -r^2 \end{pmatrix} \tag{5.34}$$

[7] Note that $\omega \mapsto \omega + i0^+$ and $\Omega \mapsto \Omega - i0^+$ are for convergence. These shifts correspond to the description of emission (ω) and absorption (Ω).

in the coordinates $(x^\mu) = (t, r, \varphi)$; this metric is called the *Born metric*. The rotating coordinate system only makes sense for $r \leq c/\Omega$. Beyond that g_{00} becomes negative because the rotational velocity exceeds the speed of light.

In a rotating reference frame, clocks can no longer be synchronized unambiguously at all locations. Computing the time difference along a closed curve parameterized by $r(\varphi)$, we obtain from (5.22)

$$\Delta t = -\frac{1}{c}\oint \frac{g_{0j}}{g_{00}} dx^j = \frac{1}{c^2}\oint \frac{\Omega r^2 d\varphi}{1-\Omega^2 r^2/c^2}. \tag{5.35}$$

For "small" rotational velocities $\Omega r \ll c$ we may expand the denominator. We obtain

$$\Delta t = \frac{\Omega}{c^2}\int r^2 d\varphi = \frac{2\Omega}{c^2} S \tag{5.36}$$

with S, the signed area enclosed by the closed curve.

This result directly explains the Sagnac effect, discovered in 1913 by Georges Sagnac. Consider a light ray propagating (e.g. in a waveguide) along the curve $r(\varphi)$. The propagation speed is c when clocks along the path of the light ray are synchronized and proper time is used for time measurements. We are interested in the travel-time difference of two light rays propagating with or against the rotation direction to leading order in $\Omega r/c$. The difference between the world time t and the proper time τ is given by $1/\sqrt{g_{00}} = 1 + \mathcal{O}(\Omega r/c)^2$ and is therefore subdominant. For the light ray propagating in the direction of rotation, we obtain from (5.36) the travel time

$$\Delta t_{\circlearrowleft} = \frac{L}{c} + \frac{2\Omega}{c^2}|S|$$

with L the length of the waveguide; the second term arises here because of the synchronization of the clocks. For the light ray propagating against the rotation direction we obtain, by contrast, the travel time

$$\Delta t_{\circlearrowright} = \frac{L}{c} - \frac{2\Omega}{c^2}|S|.$$

Hence the travel-time difference is $\Delta t = 4\Omega|S|/c^2$, corresponding to a phase shift $\Delta\varphi = 2\pi c\Delta t/\lambda$. If the two light rays are allowed to interfere in a Sagnac interferometer, one therefore obtains an interference pattern with periodicity $\Delta\varphi = 2\pi\mathbb{Z}$. Sagnac interferometers are used to measure rotations absolutely.

Next we determine the spatial metric γ_{jk}. Using (5.21) we obtain

$$dl^2 = dr^2 + \frac{r^2}{1-\Omega^2 r^2/c^2} d\varphi^2. \tag{5.37}$$

Computing the circumference at fixed radius r, we obtain

$$U = \int_0^{2\pi} d\varphi \frac{r}{\sqrt{1-\Omega^2 r^2/c^2}} = \frac{2\pi r}{\sqrt{1-\Omega^2 r^2/c^2}} > 2\pi r.$$

An observer in the accelerated reference frame will interpret this result as a hyperbolic curvature of space.

5.6 Weak gravitational field

In a weak gravitational field $\phi(\boldsymbol{r}) \ll c^2$ and at low velocities the geodesic equations (5.9) should reduce to the Newtonian equations $\ddot{\boldsymbol{r}} = -\boldsymbol{\nabla}\phi(\boldsymbol{r})$. The nonrelativistic action for a particle in a gravitational field has the form

$$S_{\rm nr} = \int dt \left(mc^2 + \tfrac{1}{2}mv^2 - m\phi\right) = -mc \int dt \left(c - \frac{v^2}{2c} + \frac{\phi}{c}\right),$$

where, compared with the conventional expression, we have added the constant $\int dt\, mc^2$ (rest energy). Comparing with the covariant expression $-mc\int ds$, see (3.12) and (5.4), yields the arc length

$$ds = \left(c - \frac{v^2}{2c} + \frac{\phi}{c}\right) dt. \tag{5.38}$$

Squaring this expression, keeping only the leading terms in v/c and ϕ/c^2, gives the result

$$ds^2 = (c^2 + 2\phi)dt^2 - d\boldsymbol{r}^2, \tag{5.39}$$

where we have used $d\boldsymbol{r} = \boldsymbol{v}\, dt$. The component g_{00} therefore takes the value

$$g_{00} = 1 + \frac{2\phi}{c^2}; \tag{5.40}$$

in addition $g_{0j} = 0$ and $g_{jk} = -\delta_{jk}$. The spatial metric is therefore simply $\gamma_{jk} = \delta_{jk}$, i.e. space is Euclidean.

When computing the Christoffel symbols, note that in the approximation discussed, because of the extra factor c, the terms $dx^0/d\tau = cdt/d\tau$ in the geodesic equations dominate over the terms $dx^j/d\tau$. The relevant Christoffel symbol is therefore $\Gamma^j{}_{00} = \partial\phi/c^2\partial x^j$, so that the geodesic equation reduces to Newton's equation with $\boldsymbol{g} = -\boldsymbol{\nabla}\phi$, cf. Eq. (5.16).

Using the expression (5.17) we obtain the gravitational redshift

$$d\tau = \sqrt{1 + 2\phi/c^2}dt \approx (1 + \phi/c^2)dt, \tag{5.41}$$

i.e. the proper time of an observer in the gravitational field with $\phi < 0$ passes more slowly.

5.7 Schwarzschild metric

We have seen that matter in a general reference frame moves on a geodesic with $ds > 0$. Light moves on *null geodesics* with $ds = 0$. By the equivalence principle, the motion of matter in an arbitrary gravitational field is therefore determined. However, one needs the metric in a space with bodies that mutually influence one another via gravitational forces. This is precisely what the Einstein equations accomplish; they

are the field equations for the metric tensor $g_{\mu\nu}(x)$. In these, the energy-momentum distribution of matter and radiation acts as the source of the field. Thus the metric on the one hand and matter and radiation on the other influence each other—unlike in special relativity. We will not be able to study the Einstein equations in this introduction, since we lack the necessary prerequisites from differential geometry.

Schwarzschild found the spherically symmetric, exact solution (*Schwarzschild metric*)

$$ds^2 = \left(1 - \frac{r_s}{r}\right) c^2 dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) \tag{5.42}$$

of the Einstein field equations in vacuum with a point source at the origin, where r_s is a parameter whose physical meaning we will learn later. The chart parameters are defined on $r > r_s$, $0 \le \theta < \pi$ and $0 \le \varphi < 2\pi$. For $r \to \infty$ the metric becomes asymptotically Minkowskian (in polar coordinates)

$$ds^2 = c^2 dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\varphi^2), \tag{5.43}$$

so that t denotes the time and $\boldsymbol{r}$ the spatial coordinates of an (infinitely distant) observer in his inertial frame. At the Schwarzschild radius $r = r_s$ the chart becomes singular and the opening of the light cones with $ds = 0$ tends to zero. This manifests itself in the fact that the redshift diverges, since the proper time is given by

$$d\tau = \sqrt{1 - r_s/r}\, dt, \tag{5.44}$$

so that for $r \to r_s$ the time for the external observer appears to pass ever more slowly. However, only the chart, not the metric itself, becomes singular. This can be seen from the fact that $\det g = -r^4 \sin^2\theta$ does not vanish at $r = r_s$. In fact one sees that at the Schwarzschild radius the signature of the coordinates t and r is exchanged and thus radius and time exchange their roles.

One can visualize the spatial metric

$$dl^2 = \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) \tag{5.45}$$

of the Schwarzschild solution as follows. One reduces the variables by restricting to the equatorial plane with $\theta = \pi/2$. The curvature is obtained by embedding in Euclidean space (with cylindrical coordinates)

$$dl^2 = dr^2 + r^2 d\varphi^2 + dw^2 \tag{5.46}$$

with the additional dimension w. Restricting the space to the *Flamm paraboloid* with

$$w^2 = 4r_s(r - r_s)\,, \tag{5.47}$$

see Figure 5.3, one obtains $2w dw = 4\sqrt{r_s(r - r_s)}\, dw = 4r_s\, dr$ and hence the metric

$$dl^2 = \frac{r_s}{r - r_s} dr^2 + dr^2 + r^2 d\varphi^2 = \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2 d\varphi^2\,, \tag{5.48}$$

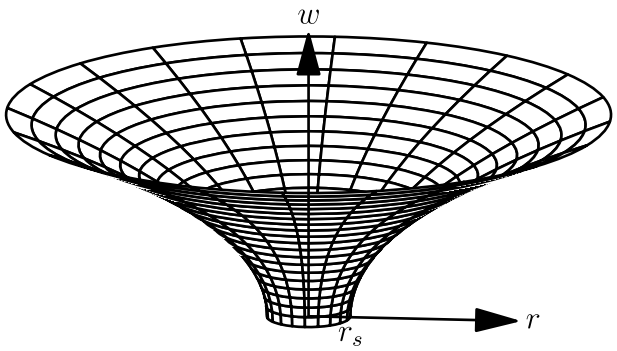

Figure 5.3: Flamm's paraboloid: spatial metric of the Schwarzschild solution ($dt = 0$). The horizontal plane corresponds to the radius r and the angle φ (at the equator with $\theta = \pi/2$). The vertical direction (w) is an additional dimension. The distances on the paraboloid correspond to the spacelike distances of the Schwarzschild solution. The paraboloid exhibits at every point a saddle-shaped (hyperbolic) curvature.

which is precisely the spatial metric of the Schwarzschild solution at the equator.

Consider, as an example, light moving radially toward the Schwarzschild radius. Since light moves on a null geodesic, its motion is determined by $ds = d\theta = d\varphi = 0$. The null geodesic equation therefore takes the simple form[8]

$$\left(1 - \frac{r_s}{r}\right) c^2 dt^2 = \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 \tag{5.49}$$

and hence

$$\frac{cdt}{dr} = \pm(1 - r_s/r)^{-1}, \tag{5.50}$$

where the sign + (−) corresponds to the outgoing (ingoing) radial null geodesic. Integration yields the result

$$c(t - t_0) = \pm(r + r_s \log|r - r_s|). \tag{5.51}$$

Thus to an infinitely distant observer it appears as though the light requires infinite time to reach the Schwarzschild radius. This is a manifestation of the infinite redshift.

The radial distance l of an observer at position r from the Schwarzschild radius is given by

$$\begin{aligned} l &= \int_{r_s}^{r} dl = \int_{r_s}^{r} \frac{dr'}{\sqrt{1 - r_s/r'}} \\ &= \sqrt{r(r - r_s)} + r_s \log\left(\sqrt{r/r_s - 1} + \sqrt{r/r_s}\right) > r - r_s. \end{aligned} \tag{5.52}$$

This distance is always greater than $r - r_s$, but it is not infinite. Hence the Schwarzschild radius lies at a well-defined spatial distance from an external observer, see Figure 5.3.

[8] This form also follows from the general geodesic equation (5.66) with $V(r) = 0$, because of $L = l = 0$.

5.7.1 Eddington-Finkelstein metric

To remove the singularity of the Schwarzschild chart we introduce a new time coordinate v so that the ingoing null geodetics become straight lines with $v =$ const .. This is achieved by the following coordinate transformation,

$$cv = ct + r + r_s \log|r - r_s|, \tag{5.53}$$

so that the outgoing radial null geodetic equation becomes $v = t_0$, cf. (5.51). Differentiation yields

$$c\,dv = c\,dt + (1 - r_s/r)^{-1} dr$$

and hence one obtains the Eddington-Finkelstein metric

$$ds^2 = c^2(1 - r_s/r)dv^2 - 2cdvdr - r^2(d\theta^2 + \sin^2\theta\, d\varphi^2). \tag{5.54}$$

One sees that the line element is regular at $r = r_s$. Thus the Eddington-Finkelstein metric analytically continues the Schwarzschild metric. Considering now null geodetics in the new coordinates, by construction one has $v = t_0$ for the ingoing geodetics. Ingoing light rays therefore simply pass through the Schwarzschild radius.

The outgoing radial null geodetic is determined by $ds = d\theta = d\varphi = 0$. The solution with $dv = 0$ gives the ingoing null geodetic. Hence for an outgoing light particle we find the geodetic equation

$$\frac{cdv}{dr} = \frac{2}{1 - r_s/r} \tag{5.55}$$

with the solution [cf. (5.50)]

$$c(v - v_0) = 2(r + r_s \log|r - r_s|), \tag{5.56}$$

or, written out in the original coordinates,

$$c(t - t_0) = r + r_s \log|r - r_s|. \tag{5.57}$$

Thus outgoing light particles cannot cross the Schwarzschild radius. The surface $r = r_s$ acts like a semipermeable membrane that allows light particles to pass only from outside to inside. Therefore the Schwarzschild radius is also called the *event horizon*, since it forms the boundary of all events that can in principle be observed by an external observer.

5.7.2 Non-relativistic limit

For small velocities $dr/dt \ll c$ and large distances $r \gg r_s$ (i.e. in the non-relativistic limit) the Schwarzschild metric becomes

$$ds^2 = \left(1 - \frac{r_s}{r}\right) c^2 dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\varphi^2). \tag{5.58}$$

A comparison with (5.39) shows that the Schwarzschild metric describes a gravitational potential of the form $\phi = -c^2 r_s/2r$. For an object of mass M classical physics gives $\phi = -GM/r$ with the gravitational constant $G = 6.7 \times 10^{-11}\,\mathrm{m^3/kg\,s^2}$. One thus obtains the formula

$$r_s = \frac{2GM}{c^2} \tag{5.59}$$

for the Schwarzschild radius, expressed in terms of the mass of the central object. For the Sun with $M_\odot \approx 2 \times 10^{30}$ kg one has, for example, $r_s \approx 3$ km.

5.7.3 Metric near the event horizon

Near the event horizon we introduce the new radial coordinate h with $r = r_s + h^2/4r_s$. Expanding the Schwarzschild metric for $h \ll r_s$ leads to the metric

$$ds^2 = \frac{h^2}{4r_s^2} c^2 dt^2 - dh^2 - r_s^2 (d\theta^2 + \sin^2\theta\, d\varphi^2). \tag{5.60}$$

Hence the spacetime near the event horizon is given by the Rindler metric (5.24). Here the time t of an infinitely distant observer plays the role of the rapidity with $\chi = ct/2r_s$ and the distance from the event horizon h corresponds to the acceleration via $a = c^2/h$. Thus one can understand the environment of a black hole with Figure 5.2. The Rindler horizon with $h = 0$ plays the role of the event horizon where the chart becomes singular. The event horizon lies for the distant observer at $t = \pm\infty$ (corresponding in the figure to $\chi = \pm\infty$), since objects moving toward it become infinitely redshifted. Because the Schwarzschild metric near the event horizon reduces to the Rindler metric, the Unruh effect also occurs there. In this context the radiation emitted from the event horizon by quantum fluctuations is called *Hawking radiation*. To compute the Hawking temperature T_H it is important to note that the radiation emitted near the event horizon is redshifted before it reaches the distant observer (with proper time t). We thus obtain the result

$$T_H = \frac{\sqrt{g_{00}}}{c} T_U = \frac{\hbar c}{4\pi k_B r_s} = \frac{\hbar c^3}{8\pi G M k_B}, \tag{5.61}$$

where we used $\sqrt{g_{00}}/c = h/2r_s$ and (5.31). The Hawking temperature of the Sun is therefore $T_H \approx 6 \times 10^{-8}$ K.

5.7.4 Geodetics

The Lagrangian for the variational principle of geodetics (5.7) for the Schwarzschild metric reads

$$L = \left(1 - \frac{r_s}{r}\right) c^2 \frac{dt^2}{d\tau^2} - \left(1 - \frac{r_s}{r}\right)^{-1} \frac{dr^2}{d\tau^2} - r^2 \left(\frac{d\theta^2}{d\tau^2} + \sin^2\theta \frac{d\varphi^2}{d\tau^2}\right). \tag{5.62}$$

The Euler-Lagrange equation for θ

$$-\frac{d}{d\tau}\left(r^2 \frac{d\theta}{d\tau}\right) + \left(r \frac{d\varphi}{d\tau}\right)^2 \sin\theta\, \cos\theta = 0 \tag{5.63}$$

is identically satisfied for a trajectory in the equatorial plane with $\theta = \pi/2$, which we may assume without loss of generality.[9] Hence the Lagrangian becomes

$$L = \left(1 - \frac{r_s}{r}\right) c^2 \frac{dt^2}{d\tau^2} - \left(1 - \frac{r_s}{r}\right)^{-1} \frac{dr^2}{d\tau^2} - r^2 \frac{d\varphi^2}{d\tau^2} \,. \tag{5.64}$$

Since the variables φ and t are cyclic, the conservation laws $r^2 d\varphi/d\tau = \ell$ (angular momentum) and $c(1 - r_s/r)dt/d\tau = \mathcal{E}$ ("energy") hold. Because the Lagrangian is given by $L = (ds/d\tau)^2$ it is also conserved. For massive objects one has $L = c^2$, while for light $L = 0$.

Inserting the cyclic variables into (5.64) gives

$$L = \left(1 - \frac{r_s}{r}\right)^{-1} \left(\mathcal{E}^2 - \frac{dr^2}{d\tau^2}\right) - \frac{\ell^2}{r^2}. \tag{5.65}$$

Hence we obtain

$$\frac{dr^2}{d\tau^2} + V(r) = \mathcal{E}^2 \tag{5.66}$$

with the "potential"

$$V(r) = \left(1 - \frac{r_s}{r}\right) \left(L + \frac{\ell^2}{r^2}\right). \tag{5.67}$$

It is convenient to introduce the variable $u = 1/r$. From $du/d\tau = -u^2 dr/d\tau$ and $d\varphi/d\tau = \ell u^2$ one finds for $u' = du/d\varphi$ the result

$$u'^2 = \left(\frac{du/d\tau}{d\varphi/d\tau}\right)^2 = \frac{\mathcal{E}^2 - V}{\ell^2} = \frac{\mathcal{E}^2}{\ell^2} - (1 - r_s u)\left(\frac{L}{\ell^2} + u^2\right) . \tag{5.68}$$

Differentiating the equation with respect to φ and dividing by $2u'$ yields the result

$$u'' + u - \frac{r_s L}{2\,\ell^2} = \frac{3 r_s u^2}{2}. \tag{5.69}$$

The equation can be solved explicitly with elliptic functions. In the following we consider a few special cases.

5.7.5 Perihelion precession

Consider the motion of a massive object with $L = c^2$; then (5.69) becomes

$$u'' + u - c^2 \frac{m}{\ell^2} = 3mu^2 \tag{5.70}$$

with $m = r_s/2 = GM/c^2$. In Newtonian physics the motion of a body around a central body (Kepler problem) satisfies

$$\dot{r}^2 - \frac{2GM}{r} + \frac{\ell^2}{r^2} = 2E, \qquad \dot{\varphi} = \ell/r^2. \tag{5.71}$$

[9] Because the problem is spherically symmetric one can always choose a coordinate system so that the trajectory lies in this plane.

Hence for $u = 1/r$ we obtain

$$u'' + u - \frac{GM}{\ell^2} = 0. \tag{5.72}$$

Comparing (5.70) and (5.69) shows that (apart from $\mathcal{E}^2 - 1 = 2E$) the term proportional to u^2 represents the relativistic correction. This term corresponds to the term $\propto r^{-3}$ in (5.67).

The general solution of (5.72) has the form

$$u_0 = \frac{1}{d}[1 + \varepsilon \cos(\varphi - \varphi_0)] \tag{5.73}$$

with $d = \ell^2/GM$. For $0 < \varepsilon < 1$ the orbit is an ellipse. The perihelion lies at $\varphi = \varphi_0 + 2\pi\mathbb{Z}$. By a rotation of the coordinate system we can arrange that $\varphi_0 = 0$.

Next we determine relativistic corrections to the Kepler orbits. To this end we set $u = u_0 + u_1$ and obtain (to first order in m)

$$u_1'' + u_1 = \frac{3m}{d^2}(1 + 2\varepsilon \cos\varphi + \varepsilon^2 \cos^2\varphi). \tag{5.74}$$

With the initial conditions $u_1 = u_1' = 0$ at the perihelion ($\varphi = 0$) the three equations ($A_1 = 3m/d^2, A_2 = 6m\varepsilon/d^2, A_3 = 3m\varepsilon^2/d^2$)

$$u_1'' + u_1 = \begin{cases} A_1 \\ A_2 \cos\varphi \\ A_3 \cos^2\varphi \end{cases} \tag{5.75}$$

have the solutions

$$u_1 = \begin{cases} A_1(1 - \cos\varphi) \\ \frac{1}{2}A_2\varphi \sin\varphi \\ A_3[\frac{1}{2} - \frac{1}{6}\cos(2\varphi) - \frac{1}{3}\cos\varphi]. \end{cases} \tag{5.76}$$

Only the second term contributes to

$$u'(2\pi) = A_2\pi = \frac{6\pi m\varepsilon}{d^2}. \tag{5.77}$$

Because $u''(2\pi) = -\varepsilon/d$ (to 0th order) the perihelion shift (the shift of the zero of u') is given by

$$\Delta\varphi = -\frac{u'(2\pi)}{u''(2\pi)} = \frac{6\pi m}{d} = \frac{6\pi GM}{c^2 a(1 - \varepsilon^2)}, \tag{5.78}$$

where a is the semi-major axis of the elliptical orbit. For Mercury one finds $\Delta\varphi \approx 43''$ per century ($''$ = arcseconds), which is confirmed today to within 1%. Note that other perturbations are about a factor ten larger.

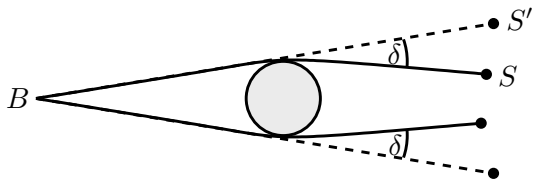

Figure 5.4: Sketch of the deflection of light emitted by a star at position S and detected at B. The light is deflected by the mass of the Sun (grey circle) and the star therefore appears at position S', corresponding to an angular shift of δ.

5.7.6 Light deflection

Next we consider lightlike geodetics with $L = 0$. Then (5.69) becomes

$$u'' + u = 3mu^2. \tag{5.79}$$

The equation $u'' + u = 0$ describes precisely light rays with $u_0 = b^{-1} \sin\varphi$, where the perihelion is chosen at $\varphi = \pi/2$ and b is the impact parameter. The zeros $\varphi = 0, \pi$ of u_0 correspond to the directions of the asymptotes for $r \to \infty$.

As above we treat the influence of general relativity by perturbation theory. To first order in m we must solve the equation

$$u_1'' + u_1 = \frac{3m}{b^2} \sin^2\varphi. \tag{5.80}$$

We obtain with $u_1 = u_1' = 0$ at the perihelion the solution in the form

$$\begin{aligned} u &= \frac{1}{b}\sin\varphi + \frac{3m}{b^2}\left[\tfrac{1}{2} + \tfrac{1}{6}\cos(2\varphi) - \tfrac{1}{3}\sin\varphi\right] \\ &= \frac{\varphi}{b} + \frac{3m}{b^2}(\tfrac{1}{2} + \tfrac{1}{6}) + O(m^2), \end{aligned} \tag{5.81}$$

i.e. the zero of u is shifted to $\varphi_\infty = -2m/b$. The total deflection is

$$\delta = 2|\varphi_\infty| = \frac{4GM}{c^2 b} \approx \frac{1.75''}{b/R_\odot}, \tag{5.82}$$

see Figure 5.4. This effect can be observed during a total solar eclipse. The experimental agreement is about 1%.

5.7.7 Stability of circular orbits

In this section we investigate for which radii $r \geq r_s$ stable circular motion is possible. For an object on a circular orbit the radial force

$$F_r = -\frac{\partial V}{\partial r} = -\frac{1}{r^4}\left(c^2 r_s r^2 - 2\ell^2 r + 3\ell^2 r_s\right) \tag{5.83}$$

must vanish. This is the case when the two attractive forces, Newtonian gravity (first term) and the relativistic term (third term), cancel the centrifugal force (second term). The angular momentum is then given by

$$\ell^2 = \frac{c^2 r^2}{2r/r_s - 3} \tag{5.84}$$

For a given angular momentum the quadratic equation (5.83) has the two solutions

$$r_> = \frac{\ell^2}{c^2 r_s}\left(1 + \sqrt{1 - \frac{3c^2 r_s^2}{\ell^2}}\right), \qquad r_< = \frac{\ell^2}{c^2 r_s}\left(1 - \sqrt{1 - \frac{3c^2 r_s^2}{\ell^2}}\right) = \frac{3\ell^2}{c^2 r_>}.$$

For small radii (smaller than $r_<$) the attractive relativistic term dominates in (5.83). For radii between $r_<$ and $r_>$ the total force is repulsive, while for large radii (larger than $r_>$) it becomes attractive again. Hence the orbit at $r_<$ is unstable against small perturbations: if the radius decreases slightly the particle is accelerated toward the center.

The orbit at $r_>$ is stable. It exists, however, only for $r \geq 3r_s$ ($\ell^2 \leq 3r_s^2 c^2$). At $r = 3r_s$ we have the maximum angular momentum $|\ell| = r^2\,|d\varphi/d\tau| = \sqrt{3}cr_s$.

The unstable orbit at $r_<$ occurs for $\frac{3}{2}r_s \leq r \leq 3r_s$, where the lower bound corresponds to an angular momentum $|\ell| \to \infty$. For radii $r < \frac{3}{2}r_s$ there are no circular orbits (not even unstable ones).

Next we wish to determine the speed $v = r\,d\varphi/d\tau_s$ of a body orbiting a black hole at radius r, measured by the proper time τ_s of a stationary observer at the same radius. We know that $\ell/r = rd\varphi/d\tau$. Hence we need only convert the time to the proper time τ_s of a stationary observer. We have $d\tau/d\tau_s = \gamma^{-1}$, with $\gamma = (1 - v^2/c^2)^{-1/2}$ the time dilation of the moving body relative to the stationary observer. Thus we obtain the equation $v = \gamma^{-1}\ell/r$, whose solution is

$$v = \frac{c}{\sqrt{2(r/r_s - 1)}}, \tag{5.85}$$

, where we have used (5.84).

At the critical radius $r = 3r_s$ for stable circular motion the speed is therefore $v = \frac{1}{2}c$, and at the radius $r = \frac{3}{2}r_s$ below which no circular motion exists we have $v = c$. Circular motion is therefore stable for $0 < v < \frac{1}{2}c$ and unstable at $\frac{1}{2}c < v < c$.

Since light always propagates at the speed of light, the circular null geodesics exist only at the radius $r = \frac{3}{2}r_s$ (this result can be derived directly as a solution of the null geodesic equation). Light particles are therefore bound to the radius $r = \frac{3}{2}r_s$, and this radius is therefore also called the *photon sphere*.

Appendix A

Tensors

We consider a real (finite-dimensional) vector space V. Then r-fold covariant and s-fold contravariant tensors are multilinear maps

$$\mathrm{T} : \underbrace{V \times \cdots \times V}_{r \text{ times}} \times \underbrace{V^* \times \cdots \times V^*}_{s \text{ times}} \mapsto \mathbb{R} \tag{A.1}$$

with V^* the dual space to V, i.e. V^* is the vector space of linear maps $V \mapsto \mathbb{R}$. Note that the space of (r, s)-tensors $\mathcal{T}^s_r(V)$ again forms a vector space. For example

$$\mathcal{T}^0_0 = \mathbb{R}, \qquad \mathcal{T}^0_1 = V^*, \qquad \mathcal{T}^1_0 = V^{**} = V. \tag{A.2}$$

In particular, an inner product $g(\boldsymbol{v}, \boldsymbol{v}) \in \mathcal{T}^0_2$, i.e. twice covariant.

Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ be a basis of V and $\boldsymbol{e}^{*1}, \ldots, \boldsymbol{e}^{*n}$ the corresponding dual basis of V^* with $\boldsymbol{e}^{*i}(\boldsymbol{e}_j) = \delta^i{}_j$. Using a basis we can express the tensor T through its components

$$T_{i_1 \ldots i_r}{}^{j_1 \cdots j_s} = \mathrm{T}(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_r}, \boldsymbol{e}^{*j_1}, \ldots, \boldsymbol{e}^{*j_s}) \tag{A.3}$$

via

$$\mathrm{T} = T_{i_1 \ldots i_r}{}^{j_1 \cdots j_s} \boldsymbol{e}^{*i_1} \otimes \cdots \boldsymbol{e}^{*i_r} \otimes \boldsymbol{e}_{j_1} \otimes \cdots \otimes \boldsymbol{e}_{j_s}, \tag{A.4}$$

where we have used the Einstein summation convention. Here we have introduced the basis

$$\boldsymbol{e}^{*i_1} \otimes \cdots \boldsymbol{e}^{*i_r} \otimes \boldsymbol{e}_{j_1} \otimes \cdots \otimes \boldsymbol{e}_{j_s}(\boldsymbol{e}_{l_1}, \ldots, \boldsymbol{e}_{l_r}, \boldsymbol{e}^{*k_1}, \ldots, \boldsymbol{e}^{*k_s}) = \delta^{i_1}{}_{l_1} \cdots \delta^{i_r}{}_{l_r} \delta^{j_1}{}_{k_1} \cdots \delta^{j_s}{}_{k_s}$$

of the multilinear map, see (A.1).

On the tensors the following operations can be defined:

Addition and multiplication by a scalar: With $\mathrm{S}, \mathrm{T} \in \mathcal{T}^s_r$ and $a, b \in \mathbb{R}$ also $a\mathrm{S} + b\mathrm{T} \in \mathcal{T}^s_r$ is given by

$$(a\mathrm{S} + b\mathrm{T})_{i_1 \ldots i_r}{}^{j_1 \cdots j_s} = aS_{i_1 \ldots i_r}{}^{j_1 \cdots j_s} + bT_{i_1 \ldots i_r}{}^{j_1 \cdots j_s};$$

i.e. $\mathcal{T}^s_r$ is a vector space.

Tensor product: With $\mathrm{S} \in \mathcal{T}^{s_1}_{r_1}$ and $\mathrm{T} \in \mathcal{T}^{s_1}_{r_2}$ the tensor product $\mathrm{S} \otimes \mathrm{T} \in \mathcal{T}^{s}_{r}$ is defined by

$$(\mathrm{S} \otimes \mathrm{T})_{i_1 \dots i_r}{}^{j_1 \dots j_s} = S_{i_1 \dots i_{r_1}}{}^{j_1 \dots j_{s_1}} T_{i_{r_1+1} \dots i_r}{}^{j_{r_1+1} \dots j_s},$$

where $r = r_1 + r_2$ and $s = s_1 + s_2$.

Contraction: From a tensor $\mathrm{T} \in \mathcal{T}^{s}_{r}$ one obtains by contraction a tensor $\mathrm{S} = K^l_k \mathrm{T} \in \mathcal{T}^{s-1}_{r-1}$. With $1 \le k \le r$ and $1 \le l \le s$ we define

$$S_{i_1 \dots i_{r-1}}{}^{j_1 \dots j_{s-1}} = T_{i_1 \dots i_{k-1} i i_{k+1} i_r}{}^{j_1 \dots j_{l-1} i j_{l+1} \dots j_s}.$$

Note: for $\mathrm{T} \in \mathcal{T}^1_1$ this corresponds K^1_1 exactly to taking the trace, $K^1_1 \mathrm{T} = \operatorname{tr} \mathrm{T}$.

Inner Product

In physical applications we usually have a distinguished inner product $g(\boldsymbol{v}, \boldsymbol{w}) \in \mathcal{T}^0_2$. In components this is expressed by the symmetric matrix $g_{ij} = g_{ij}$. This allows one to raise or lower indices. In particular, the inner product in components is given by

$$g(\boldsymbol{v}, \boldsymbol{w}) = g_{ij} v^i w^j = v_j w^j = v^i w_j \tag{A.5}$$

with the operation of lowering $v_j = g_{ji} v^i$ and $w_j = g_{ji} w^i$. By inverting this relation we obtain $v^i = g^{ij} v_j$ with the inverse $g^{ij} = (g^{-1})_{ij}$, so that $g^{ij} g_{jk} = \delta^i{}_k$.

Euclidean Space

The Euclidean space $\mathbb{R}^3$ with Cartesian basis is defined by the scalar product $g_{ij} = \delta_{ij}$. In this space raising and lowering indices is trivial and therefore the indices are usually written only downstairs. In general, we are interested in how tensors transform under a change of basis. As already seen in Chap. 1, the components of vectors $\boldsymbol{r} \in \mathcal{T}^1_0$ transform as $\boldsymbol{r}' = R\boldsymbol{r}$. Because of the trivial scalar product there is no difference between co- and contravariant transformation and we obtain that a general tensor transforms as

$$T'_{i_1 \dots i_n} = R_{i_1 j_1} \cdots R_{i_n j_n} T_{j_1 \dots j_n} \tag{A.6}$$

.

Minkowski Space

The Minkowski space is defined by the Lorentz scalar product $g_{ij} = \eta_{ij}$ with

$$(\eta_{\mu\nu}) = (\eta^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{A.7}$$

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

Consider now a basis transformation Λ, under which the components of a four-vector x transform as $x' = \Lambda x$ [1] We write this, taking into account the correct superscript position, as $x'^{\nu} = \Lambda^{\nu}{}_{\mu} x^{\mu}$. By lowering the indices we analogously obtain the transformation rule for a covariant vector as $x'_{\nu} = \Lambda_{\nu}{}^{\mu} x_{\mu}$ with

$$\Lambda_{\nu}{}^{\mu} = \eta_{\nu\sigma} \Lambda^{\sigma}{}_{\tau} \eta^{\tau\mu} = \left(\Lambda^{-1}\right)^{\mu}{}_{\nu}. \tag{A.8}$$

A general tensor therefore transforms under a basis transformation as

$$T'_{i_1 \dots i_r}{}^{j_1 \dots j_s} = \Lambda_{i_1}{}^{\mu_1} \cdots \Lambda_{i_r}{}^{\mu_r} \Lambda^{j_1}{}_{\nu_1} \cdots \Lambda^{j_s}{}_{\nu_s} T_{\mu_1 \dots \mu_r}{}^{\nu_1 \dots \nu_s}. \tag{A.9}$$

Tensors that transform in this way are also called four- or Lorentz tensors.

Tensor Fields

Note that, due to the fixed transformation rule (A.6) or (A.9), equations between tensors of the same rank are automatically covariant. This nice property can be generalized (pointwise) to tensor fields $T_{i_1 \dots i_r}{}^{j_1 \dots j_s}(t, \boldsymbol{r})$. Tensor fields additionally allow, besides the operations mentioned above, the definition of derivative operations.

In Euclidean space the gradient $\mathrm{S} = \partial \mathrm{T}$ of $T_{i_1 \dots i_n}(\boldsymbol{r})$ is automatically a tensor of rank $n+1$ defined by

$$S_{k i_1 \dots i_n} = (\partial \mathrm{T})_{k i_1 \dots i_n} = \frac{\partial}{\partial r_k} T_{i_1 \dots i_n}(\boldsymbol{r}). \tag{A.10}$$

With the chain rule

$$\frac{\partial}{\partial r'_k} = \frac{\partial r_l}{\partial r'_k} \frac{\partial}{\partial r_l} = \frac{\partial [(R^{-1})_{lm} r'_m + \text{const}]}{\partial r'_k} \frac{\partial}{\partial r_l} = (R^{-1})_{lk} \frac{\partial}{\partial r_l} = R_{kl} \frac{\partial}{\partial r_l}$$

it is easy to see that S indeed transforms as a tensor under coordinate changes. In fact we then obtain

$$\frac{\partial}{\partial r'_k} T'_{i_1 \dots i_n}(\boldsymbol{r}') = R_{kl} R_{i_1 j_1} \cdots R_{i_n j_n} \frac{\partial}{\partial r_k} T_{i_1 \dots i_n}(\boldsymbol{r}).$$

In Minkowski space one can show that the gradient $\mathrm{S} = \partial \mathrm{T}$ of a (r, s)-tensor is a $(r+1, s)$-tensor. Analogous to the Euclidean case one obtains with the chain rule

$$\frac{\partial}{\partial x'^{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} \frac{\partial}{\partial x^{\nu}} = \frac{\partial [(\Lambda^{-1})^{\nu}{}_{\sigma} x'^{\sigma} + \text{const}]}{\partial x'^{\mu}} \frac{\partial}{\partial x^{\nu}} = (\Lambda^{-1})^{\nu}{}_{\mu} \frac{\partial}{\partial x^{\nu}} = \Lambda_{\mu}{}^{\nu} \frac{\partial}{\partial x^{\nu}}.$$

and therefore S indeed transforms as a $(r+1, s)$-tensor with the transformation law

$$\begin{aligned} S'_{\mu i_1 \dots i_r}{}^{j_1 \dots j_s}(x') &= \partial_{x'^{\mu}} T'_{i_1 \dots i_r}{}^{j_1 \dots j_s}(x') \\ &= \Lambda_{\mu}{}^{\nu} \Lambda_{i_1}{}^{\mu_1} \cdots \Lambda_{i_r}{}^{\mu_r} \Lambda^{j_1}{}_{\nu_1} \cdots \Lambda^{j_s}{}_{\nu_s} \frac{\partial}{\partial x^{\nu}} T_{\mu_1 \dots \mu_r}{}^{\nu_1 \dots \nu_s}(x) \\ &= \Lambda_{\mu}{}^{\nu} \Lambda_{i_1}{}^{\mu_1} \cdots \Lambda_{i_r}{}^{\mu_r} \Lambda^{j_1}{}_{\nu_1} \cdots \Lambda^{j_s}{}_{\nu_s} S_{\nu \mu_1 \dots \mu_r}{}^{\nu_1 \dots \nu_s}(x). \end{aligned}$$

[1] We denote the four-vector, as in the main text, not in bold face.