# MATH3018: Relativity

## Lecture Notes, Spring 2021/2022

Alexander Schenkel

School of Mathematical Sciences, University of Nottingham, UK.

alexander.schenkel@nottingham.ac.uk

October 4, 2022

# Contents

# Part 1

# Preliminaries

CHAPTER 1

# Introduction

## 1.1. What is Relativity?

Relativity is one of the cornerstones of modern theoretical and mathematical physics. It unifies the concepts of space and time into a single entity called *spacetime*. Relativity comes in two variants called *special relativity* and *general relativity*. As the name suggests, special relativity may be understood as a special case of general relativity.

Special relativity develops a concept of spacetime that implements the (experimentally very well confirmed) physical postulate that the value of the speed of light is the same for all inertial observers. Observer independence of the speed of light immediately leads to novel physical effects, such as relativity of simultaneity, time dilation and length contraction. In this framework it also becomes clear that Newtonian mechanics is only an approximation (valid for particle velocities much smaller than the speed of light) of a more fundamental theory called relativistic mechanics.

General relativity is an extension of special relativity that incorporates the gravitational force as an intrinsic property of spacetime. It was developed by Einstein during the early 1900s and it is up to now our best and most elegant mathematical description of gravitation. General relativity goes far beyond Newtonian gravitation, which is recovered only for weak gravitational fields and velocities much smaller than the speed of light. It also predicts spectacular new physical effects, such as black holes, gravitational waves and gravitational time dilation.

Relativity is an indispensable tool for many areas of theoretical and mathematical physics. On the one hand, combining the principles of special relativity with quantum mechanics leads to quantum field theory, which is the theoretical foundation for elementary particle physics. In particular, the behavior of elementary particles at the Large Hadron Collider (LHC) at CERN shows strong special relativistic effects due to the fact that they propagate with almost the speed of light. On the other hand, general relativity is essential to understand and describe the structure and properties of our universe, which is studied in the field of cosmology.

## 1.2. The Module MATH3018: Relativity

The aim of this module is to provide a first introduction to both special and general relativity. We will focus mostly on the theoretical/mathematical framework underlying these theories, but we shall also highlight the relevant physical ideas and principles. Within special relativity, we shall cover the following topics:

- *Minkowski spacetime*
- *Lorentz and Poincaré transformations*

- *Relativistic mechanics*
- *Relativistic continuum mechanics*

Within general relativity, we shall cover the following topics:

- *Elementary differential geometry*
- *Physics in curved spacetimes*
- *Einstein's field equation*
- *Newtonian gravity as a limit of general relativity*
- *Examples of spacetimes, including Schwarzschild*

In particular the part on general relativity is intended to be a first elementary introduction. The University of Nottingham offers various more advanced modules whose aim is to deepen and build upon the material introduced in this module. For example, the module *MATH4015 Differential Geometry* provides a more complete introduction to the mathematical field of differential geometry and the module *MATH4016 Black Holes* offers a more detailed study of black holes.

## 1.3. About these Lecture Notes and Further Literature

Even though I try to be very careful, these lecture notes may contain typos, irregularities or inconsistencies. If you spot any of those, please inform me via email `alexander.schenkel@nottingham.ac.uk` or personally before (or after) the lectures. Of course, you are also very much welcome to visit me during my office hours.

There are many books on special and general relativity which you can find in the library or online. Obviously, it doesn't hurt and will be beneficial for you to also have a look at some of these books. It will strengthen your ability to read mathematical/physical literature and it also presents you another point of view on the material of this module. Let me however stress that, even though highly recommended, it is not strictly necessary to read additional books. I will apply the following rule:

**The exam will be about the topics discussed**
**in these lecture notes and my problem sheets.**

In case you are eager to explore further literature, you might start with the following references:

[1] James J. Callahan, *The Geometry of Spacetime: An Introduction to Special and General Relativity*, Springer (2008).

[2] Wolfgang Rindler, *Relativity: Special, General, and Cosmological*, Oxford University Press (2006).

[3] Bernard Schutz, *A First Course in General Relativity*, Cambridge University Press (2009).

[4] Sean Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, Pearson (2003).

CHAPTER 2

# Newtonian Mechanics

## 2.1. Space and Time

The aim of Newtonian mechanics is to describe time-dependent physical processes in space. A typical model for space is the 3-dimensional Cartesian space $\mathbb{R}^3$. Fixing a reference frame, i.e. a set of coordinates, we can describe a point $\mathbf{x} \in \mathbb{R}^3$ in space by three real numbers $x^1, x^2, x^3 \in \mathbb{R}$. We shall also use the vector notation

$$\mathbf{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \in \mathbb{R}^3 \ . \tag{2.1}$$

We introduce a notion of distance between two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ in space by using the *Euclidean distance formula*

$$\mathrm{dist}_E(\mathbf{x}, \mathbf{y}) := \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2} \ . \tag{2.2a}$$

It will sometimes be convenient to consider the square of the Euclidean distance, which is given by

$$\mathrm{dist}_E(\mathbf{x}, \mathbf{y})^2 = (x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2 \ . \tag{2.2b}$$

Assuming that $\mathbf{x}$ and $\mathbf{y} = \mathbf{x} + d\mathbf{x}$ are infinitesimally close points, the square of the distance between $\mathbf{x}$ and $\mathbf{y}$ is called the *Euclidean line element*. It reads as

$$ds_E^2 := \mathrm{dist}_E(\mathbf{x}, \mathbf{x} + d\mathbf{x})^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \ . \tag{2.3}$$

The line element encodes the geometry of Euclidean space.

REMARK 2.1. In mathematically precise terms, the line element (2.3) should be understood as a *quadratic form* on the space of tangent vectors $T_{\mathbf{x}}\mathbb{R}^3 = \mathbb{R}^3$ at the point $\mathbf{x} \in \mathbb{R}^3$. Let us denote this quadratic form by

$$q_E : T_{\mathbf{x}}\mathbb{R}^3 \longrightarrow \mathbb{R} \ , \quad \mathbf{v} \longmapsto (v^1)^2 + (v^2)^2 + (v^3)^2 \ , \tag{2.4}$$

where $\mathbf{v} = (v^1, v^2, v^3)$. It defines an *inner product* on $T_{\mathbf{x}}\mathbb{R}^3$ by setting

$$\begin{aligned} \langle \cdot \, , \cdot \rangle_E : T_{\mathbf{x}}\mathbb{R}^3 \times T_{\mathbf{x}}\mathbb{R}^3 &\longrightarrow \mathbb{R} \ , \\ (\mathbf{v}, \mathbf{w}) &\longmapsto \frac{1}{2}\Big(q_E(\mathbf{v} + \mathbf{w}) - q_E(\mathbf{v}) - q_E(\mathbf{w})\Big) \ . \end{aligned} \tag{2.5}$$

A brief calculation shows that $\langle \cdot \, , \cdot \rangle_E$ is the standard Euclidean inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle_E = v^1\, w^1 + v^2\, w^2 + v^3\, w^3 \ . \tag{2.6}$$

Hence, we find that the Euclidean line element (2.3) captures the same geometric information as equipping each tangent space $T_{\mathbf{x}}\mathbb{R}^3$ with the standard Euclidean inner product. The latter mathematical structure is called *Euclidean metric* on $\mathbb{R}^3$. In our introductory chapter to differential geometry we shall explain the mathematical concept of metrics in more generality and detail. △

In addition to space, Newtonian mechanics requires a concept of time. We model time by the real line $\mathbb{R}$ and denote time points by $t \in \mathbb{R}$. The distance between two points $t, u \in \mathbb{R}$ in time is given by the absolute value

$$\mathrm{dist}_{\mathrm{time}}(t,u) := |t-u| \ , \tag{2.7a}$$

and its square reads simply as

$$\mathrm{dist}_{\mathrm{time}}(t,u)^2 = (t-u)^2 \ . \tag{2.7b}$$

There is also a line element for time, namely

$$ds^2_{\mathrm{time}} := \mathrm{dist}_{\mathrm{time}}(t,t+dt)^2 = (dt)^2 \ . \tag{2.8}$$

However, we do not explicitly need this for the moment.

The simplest example of a physical process in Newtonian mechanics is a point-particle propagating through space as time passes. We model such a particle by a (smooth) map from time to space, i.e.

$$\mathbf{x} : \mathbb{R} \longrightarrow \mathbb{R}^3 \ , \quad t \longmapsto \mathbf{x}(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \\ x^3(t) \end{pmatrix} \ . \tag{2.9}$$

The interpretation is as follows: The point $\mathbf{x}(t) \in \mathbb{R}^3$ describes the position of the particle at time $t \in \mathbb{R}$. Hence, (2.9) encodes the trajectory of the particle.

Notice that we have (silently) chosen space coordinates $\mathbf{x} = (x^1, x^2, x^3)$ and a time coordinate $t$ in our description above. These coordinates fix a *reference frame* in which we measure space distances and time duration. Physically speaking, such a reference frame can be thought of as an *observer*, i.e. an imaginary person who has build up clocks and rods with respect to which they will measure distances and time. Different observers in general will choose different reference frames, i.e. different coordinates $\mathbf{x}' = (x'^1, x'^2, x'^3)$ and $t'$ for space and time. It is therefore important to investigate if the laws of Newtonian mechanics are invariant under changing the reference frame.

In Newtonian mechanics not every choice of reference frame is equally good. Newton's first law introduces a distinguished class of reference frames, called *inertial frames*, which are characterized by the following property: In an inertial frame all particles unaffected by external forces (i.e. force-free particle) are either at rest or in uniform motion. In formulas, this means that if (2.9) describes the trajectory of a force-free particle in an inertial frame, then it satisfies the differential equation

$$\frac{d^2}{dt^2}\mathbf{x}(t) = \mathbf{0} \quad \Longleftrightarrow \quad \frac{d^2}{dt^2}x^1(t) = \frac{d^2}{dt^2}x^2(t) = \frac{d^2}{dt^2}x^3(t) = 0 \ . \tag{2.10}$$

We shall see that there are different inertial frames and not just a unique one.

## 2.2. Galilean Transformations

Different reference frames are related by *coordinate transformations*: Given the space and time coordinates $\mathbf{x}$ and $t$ of a reference frame, we may choose a new set of space and time coordinates $\mathbf{x}'$ and $t'$ and thereby obtain a new reference frame. The new coordinates can be expressed in terms of the old ones, i.e. they are functions

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}, t) \quad , \qquad t' = t'(t) \; . \tag{2.11}$$

It is important to emphasize that even though the new space coordinates $\mathbf{x}'$ may depend on time $t$, the new time coordinate $t'$ *cannot* depend on $\mathbf{x}$ because time is absolute in Newtonian mechanics.

Of particular importance in Newtonian mechanics are coordinate transformations which preserve our notion of distances (cf. (2.2) and (2.7)) and the property of being an inertial frame. Such transformations may be decomposed into the following three classes:

**Transformations of space.** Our notion of distance (2.2), or equivalently the line element (2.3), is invariant under coordinate transformations of the form

$$\mathbf{x}' = R\mathbf{x} - \mathbf{b} \quad , \qquad t' = t \; , \tag{2.12}$$

where $R$ is an orthogonal $3 \times 3$-matrix (i.e. $R^{\mathrm{T}} = R^{-1}$) and $\mathbf{b} \in \mathbb{R}^3$. (Here $^{\mathrm{T}}$ denotes matrix transposition and $^{-1}$ denotes the inverse matrix.) If we apply the transformation (2.12) to a force-free particle $\mathbf{x} : \mathbb{R} \to \mathbb{R}^3$ in an inertial frame, i.e. $\frac{d^2}{dt^2}\mathbf{x}(t) = 0$, we immediately observe that

$$\frac{d^2}{dt'^2}\mathbf{x}'(t') = \frac{d^2}{dt^2}\Big(R\mathbf{x}(t) - \mathbf{b}\Big) = R\frac{d^2}{dt^2}\mathbf{x}(t) = 0 \; , \tag{2.13}$$

because $R$ and $\mathbf{b}$ are time-independent. Hence, the transformations (2.12) map inertial frames to inertial frames. Physically, we can interpret $\mathbf{x}'$ and $t'$ as the coordinates chosen by an observer whose rods are displaced (i.e. translated) by $\mathbf{b} \in \mathbb{R}^3$ and rotated by $R$ with respect to the rods of the unprimed observer.

**Transformations of time.** Our notion of time distance (2.7) is invariant under coordinate transformations of the form

$$\mathbf{x}' = \mathbf{x} \quad , \qquad t' = t - b \; , \tag{2.14}$$

where $b \in \mathbb{R}$. It is easy to see that the transformations (2.14) map inertial frames to inertial frames. Physically, we can interpret $\mathbf{x}'$ and $t'$ as the coordinates chosen by an observer whose clocks are changed by a time-shift $b$.

**Galilean boosts.** These are transformations which involve both space and time coordinates. Explicitly,

$$\mathbf{x}' = \mathbf{x} - \mathbf{v}\, t \quad , \qquad t' = t \; , \tag{2.15}$$

where $\mathbf{v} \in \mathbb{R}^3$. It is easy to check that space (2.2) and time (2.7) distances are invariant under these transformations. (Notice that a Galilean boost is nothing else but a time-dependent translation, i.e. choosing $\mathbf{b} = \mathbf{v}\, t$ in (2.12).) The transformations

(2.15) map inertial frames to inertial frames: Given a force-free particle $\mathbf{x} : \mathbb{R} \to \mathbb{R}^3$, i.e. $\frac{d^2}{dt^2}\mathbf{x}(t) = 0$, we obtain

$$\frac{d^2}{dt'^2}\mathbf{x}'(t') = \frac{d^2}{dt^2}\Big(\mathbf{x}(t) - \mathbf{v}\,t\Big) = \frac{d^2}{dt^2}\mathbf{x}(t) = 0\ , \tag{2.16}$$

because $\mathbf{v}$ is time-independent. Physically, we can interpret $\mathbf{x}'$ and $t'$ as the coordinates chosen by an observer whose clocks and rods are uniformly moving with velocity $\mathbf{v}$ with respect to the unprimed observer.

REMARK 2.2. In summary, we obtained that Newtonian mechanics (formulated in inertial frames) is invariant under coordinate transformations of the form

$$\mathbf{x}' = R\,\mathbf{x} - \mathbf{b} - \mathbf{v}\,t \quad , \qquad t' = t - b\ , \tag{2.17}$$

where $R$ is an orthogonal $3 \times 3$-matrix (i.e. $R^{\mathrm{T}} = R^{-1}$), $\mathbf{b}, \mathbf{v} \in \mathbb{R}^3$ and $b \in \mathbb{R}$. Such transformations are called *Galilean transformations* and they form a group called the *Galilean group*. △

### 2.3. Addition of Velocities

Let us fix an inertial frame with coordinates $\mathbf{x}$ and $t$. A force-free particle $\mathbf{x} : \mathbb{R} \to \mathbb{R}^3$ is characterized by the differential equation $\frac{d^2}{dt^2}\mathbf{x}(t) = 0$, hence its trajectory is of the form

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{V}\,t\ , \tag{2.18}$$

where $\mathbf{x}_0 \in \mathbb{R}^3$ is the initial position at time $t = 0$ and $\mathbf{V} \in \mathbb{R}^3$ is the velocity. Let us assume for simplicity that the velocity points along the $x^1$-direction, i.e. $\mathbf{V} = (V, 0, 0)$. Changing the inertial frame by a Galilean boost (2.15) along the $x^1$-direction, i.e. $\mathbf{v} = (v, 0, 0)$, we obtain the trajectory

$$\mathbf{x}'(t') = \mathbf{x}_0 + \mathbf{V}\,t - \mathbf{v}\,t = \mathbf{x}_0 + \begin{pmatrix} (V-v)\,t \\ 0 \\ 0 \end{pmatrix}\ , \tag{2.19}$$

where we also used that $t' = t$. The velocity of the particle determined by the primed observer is thus

$$V' = V - v\ . \tag{2.20}$$

This formula is called the *addition of velocities* formula of Newtonian mechanics.

EXAMPLE 2.3. The formula (2.20) agrees very well with our real-world experience: Consider a train passing by the train station with constant velocity $v$. Choose as unprimed inertial frame the one in which the train station is at rest and as primed inertial frame one in which the train is at rest. When an observer inside the train sees a particle moving with velocity $V'$, then an observer at the train station sees the particle moving with velocity $V = V' + v$. Figure 2.1 illustrates this fact. ▽

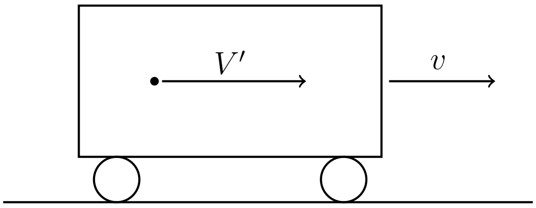


FIGURE 2.1. A train passing by the train station with constant velocity $v$. When an observer sitting inside the train sees a particle with velocity $V'$, then an observer waiting at the train station sees the same particle with velocity $V = V' + v$.

## 2.4. The Problem with the Speed of Light

The addition of velocities formula (2.20), which is a consequence of Newtonian mechanics and in particular of the form of the Galilean boost (2.15), is *incompatible* with electromagnetic phenomena such as the propagation of light. In the setup displayed in Figure 2.1, assume that we observe a propagating light ray from either the perspective of the train or the perspective of the train station. If one of the observers sees the light ray propagating with the speed of light

$$c = 299\,792\,458\,\frac{m}{s}\ , \tag{2.21}$$

then the other one would see it propagating with $c \pm v$. Hence, our addition of velocities formula (2.20) implies that the value of the speed of light depends on the inertial frame.

It is experimentally extremely well confirmed that the speed of light is a fundamental constant of physics whose value does not depend on the inertial frame. Early experiments in this direction include the famous Michelson-Morley experiment from 1887. If you would like to learn more about the details of this experiment, please search the internet (e.g. Wikipedia).

The conflict between Newtonian mechanics and the constancy of the speed of light is very dramatic: It means that Newtonian mechanics and the associated Galilean transformations are "wrong" and have to be replaced by another framework which respects the constancy of the speed of light. We shall see later that relativistic mechanics and the associated Lorentz and Poincaré transformations solve this problem. Starting from this more fundamental theory, we will obtain that Newtonian mechanics is a good approximation for situations where the velocities of particles are much smaller than the speed of light, i.e. $v \ll c$. Hence, instead of using the harsh word "wrong", it is more appropriate to say that Newtonian mechanics and Galilean transformations have a limited range of applicability (for small velocities). It is one of the skills of a good theoretical physicist to decide whether a particular problem still can be approximated by Newtonian mechanics or if one requires a relativistic model.

## 2.5. Relativity of Simultaneity

Using only the (experimentally verified) postulate that the speed of light has the same value in all inertial frames, we can arrive at the conclusion that Newton's assumption of an absolute notion of time is the cause of the insufficiency of Newtonian mechanics. We shall briefly sketch the relevant argument.

Let us consider again our example of a train passing by the train station, see Figure 2.1 for an illustration. Imagine that the observer who is at rest inside the train observes a pair of light rays that are emitted from both ends of the cabin and meet at the same time in the middle. Figure 2.2 (left) illustrates their observation. The observer will conclude that both light rays have been emitted at the same time, i.e. $t'_{\text{left}} = t'_{\text{right}}$. The observer waiting at the train station will arrive at a different conclusion, see Figure 2.2 (right) for an illustration: Because the speed of light is the same in all inertial frames, the light ray coming from the left takes more time to arrive in the middle than the one coming from the right. This is because it has to travel a longer distance. Thus, the second observer will conclude that the left light ray has been emitted *before* the right light ray, i.e. $t_{\text{left}} < t_{\text{right}}$.

In summary, we have seen that if the speed of light is the same in all inertial frames, then the concept of simultaneity necessarily has to be observer dependent. This excludes the absolute notion of time on which Newtonian mechanics is based! Special relativity is, by construction, a theory that is compatible with the constancy of the speed of light in all inertial frames. We shall develop and learn this theory during the first part of this module.

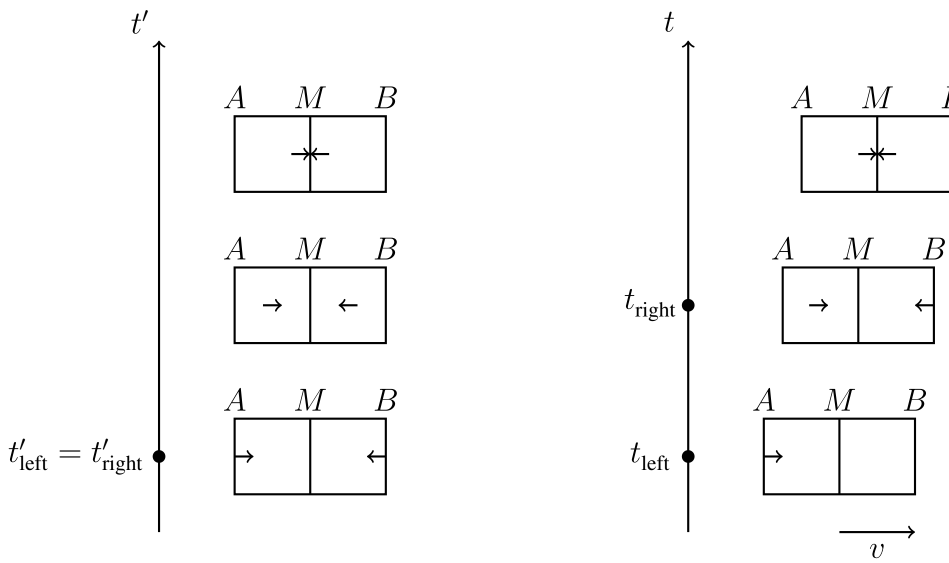


FIGURE 2.2. Light rays emitted from $A$ and $B$ meet at the same time in the middle $M$ of the train cabin. *Left:* Observer at rest inside the train. *Right:* Observer at rest at the train station.

# Part 2

# Special Relativity

CHAPTER 3

# Minkowski Spacetime

We have seen in Chapter 2.5 that the conflict between Newtonian mechanics and the constancy of the speed of light in all inertial frames is due to the fact that we treated space and time as distinct concepts. The solution to this problem is hence to introduce the new concept of *spacetime*, which unifies space and time in terms of a single structure. In contrast to our previous notions of geometry on space and time (see Chapter 2.1), there will be now a new notion of geometry on spacetime which is compatible with the constancy of the speed of light. The resulting mathematical structure is called the *Minkowski spacetime*.

## 3.1. 4-Dimensional Perspective: Events and World-Lines

The role of spacetime is to describe physical events. An *event* is characterized by a point in space $\mathbb{R}^3$ describing where it takes place, together with a point in time $\mathbb{R}$ describing when it takes place. In other words, an event is a point $x \in \mathbb{R}^4$ of a 4-dimensional Cartesian space which we shall call *spacetime*. We can express an event in terms of a set of coordinates, i.e. a reference frame for spacetime, by

$$x = \begin{pmatrix} c\,t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \in \mathbb{R}^4 \ . \tag{3.1}$$

We use the first coordinate to describe time and the other 3 coordinates to describe space. We further have multiplied time by the speed of light $c$ (which is a constant due to our postulate of constancy of the speed of light!) so that all entries of the vector have the same physical dimension [length].

The 4-dimensional perspective alters our way of how we should think of particles propagating through space as time passes. Instead of formulating a point-particle by its trajectory in space as in (2.9), a point-particle in our 4-dimensional world is a curve in spacetime. Such curves in spacetime are called *world-lines*, see Figure 3.1 for examples.

Similarly to the situation in Newtonian mechanics, not every choice of reference frame for spacetime is equally good. We introduce a distinguished class of reference frames for spacetime by the following definition: An *inertial frame for spacetime* is a reference frame for spacetime in which the world-lines of all force-free particles are straight lines. Figure 3.1 a), b), c) gives some examples of world-lines of force-free particles in an inertial frame for spacetime. A particle under the influence of external

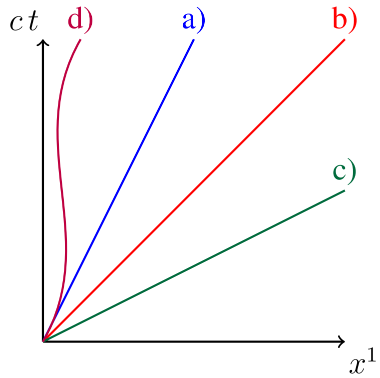


FIGURE 3.1. World-lines of particles propagating only in the $x^1$-direction, i.e. $x^2 = x^3 = 0$ are constant.

forces will have a "curved" or "bended" world-line in an inertial frame for spacetime, see Figure 3.1 d) for an example.

It will later be useful to describe world-lines also in a parametrized form by using (smooth) maps

$$x : \mathbb{R} \longrightarrow \mathbb{R}^4 \ , \quad \lambda \longmapsto x(\lambda) = \begin{pmatrix} c\,t(\lambda) \\ x^1(\lambda) \\ x^2(\lambda) \\ x^3(\lambda) \end{pmatrix} \ . \tag{3.2}$$

It is important to emphasize that the parameter $\lambda$ which parametrizes the world-line should *not* be interpreted as time; notice that the time coordinate $t(\lambda)$ is also a function of $\lambda$! The correct interpretation of $\lambda$ is just as a choice of parametrization of the world-line, whose physical content is solely given by its image in $\mathbb{R}^4$. This then brings us back to unparametrized curves as in Figure 3.1.

## 3.2. Distance Formula

Notice that the straight world-lines a), b) and c) in Figure 3.1 describe three different kinds of particles: Particle a) propagates with velocity smaller than the speed of light (i.e. $c\,t > x^1$), particle b) propagates with velocity equal to the speed of light (i.e. $c\,t = x^1$), and particle c) propagates with velocity larger than the speed of light (i.e. $c\,t < x^1$). We shall now introduce a notion of distance on spacetime which allows us to distinguish these three cases.

Let us consider as in Figure 3.1 a straight world-line. Let us further choose any two non-equal points $x, y \in \mathbb{R}^4$ lying on this world-line and consider the following quantity

$$\mathrm{dist}_M(x,y) := \sqrt{-(c\,t - c\,u)^2 + (x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2} \ , \tag{3.3a}$$

or more conveniently its square

$$\mathrm{dist}_M(x,y)^2 = -(c\,t - c\,u)^2 + (x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2 \ , \tag{3.3b}$$

where we used $x = (c\,t, x^1, x^2, x^3)$ and $y = (c\,u, y^1, y^2, y^3)$. This looks almost like (the square of) the 4-dimensional Euclidean distance formula, however with the crucial difference that the sign in front of the first term is $-1$ and not $+1$. Equation (3.3) is called the *Minkowski distance formula*. It allows us to distinguish the three different cases in Figure 3.1 a), b), c):

a) $\mathrm{dist}_M(x, y)^2 < 0$ is negative;

b) $\mathrm{dist}_M(x, y)^2 = 0$ is zero;

c) $\mathrm{dist}_M(x, y)^2 > 0$ is positive.

Hence, we can use (3.3) to detect whether a force-free particle in an inertial frame for spacetime propagates with velocity smaller, equal or larger than the speed of light.

It is convenient to introduce a bit of mathematical terminology at this point. Consider any straight world-line as in Figure 3.1. We say that the world-line is

- *time-like* if $\mathrm{dist}_M(x, y)^2 < 0$, for two points $x \neq y$ on the world-line;
- *light-like* if $\mathrm{dist}_M(x, y)^2 = 0$, for two points $x \neq y$ on the world-line; and
- *space-like* if $\mathrm{dist}_M(x, y)^2 > 0$, for two points $x \neq y$ on the world-line.

Looking back to our examples in Figure 3.1, we see that a) is time-like, b) is light-like and c) is space-like. In physics, time-like world-lines describe point-particles with a non-zero mass such as atoms, electrons, etc., while light-like world-lines describe massless point-particles such as photons (or light rays). Space-like world-lines would describe hypothetical particles called tachyons, which are not observed in nature and considered as unphysical.

REMARK 3.1. The Minkowski distance formula (3.3) is the correct relativistic generalization (and moreover a unification) of the distance formulas (2.2) and (2.7) in Newtonian mechanics. In contrast to the Newtonian distance formulas, (3.3) involves both space and time coordinates and hence it is an intrinsic quantity on spacetime and not on either space or time. We can recover (2.2) and (2.7) as special cases: Given $x = (c\,t, x^1, x^2, x^3)$ and $y = (c\,u, y^1, y^2, y^3)$ such that the time coordinates coincide, i.e. $t = u$, we obtain

$$\mathrm{dist}_M(x, y)^2 = (x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2 = \mathrm{dist}_E(\mathbf{x}, \mathbf{y})^2 \ , \qquad (3.4)$$

where $\mathbf{x} = (x^1, x^2, x^3)$ and $\mathbf{y} = (y^1, y^2, y^3)$ are the space coordinates. On the other hand, given $x = (c\,t, x^1, x^2, x^3)$ and $y = (c\,u, y^1, y^2, y^3)$ such that the space coordinates coincide, i.e. $\mathbf{x} = \mathbf{y}$, we obtain

$$\mathrm{dist}_M(x, y)^2 = -(c\,t - c\,u)^2 = -c^2\, \mathrm{dist}_{\mathrm{time}}(t, u)^2 \ , \qquad (3.5)$$

i.e. up to the prefactor $-c^2$ the non-relativistic time distance (2.7). In general, we obtain that (3.3) can be expressed as

$$\mathrm{dist}_M(x, y)^2 = -c^2\, \mathrm{dist}_{\mathrm{time}}(t, u)^2 + \mathrm{dist}_E(\mathbf{x}, \mathbf{y})^2 \ , \qquad (3.6)$$

for all points $x, y \in \mathbb{R}^4$ of spacetime. It is important to stress that neither of the two summands on the right-hand side of (3.6) will be invariant on its own under a change of inertial frame for spacetime; this is a consequence of our discussion in Chapter 2.5 where we obtained that simultaneity (i.e. the condition $t = u$) is an observer dependent concept. This fact is the reason for very interesting and unexpected relativistic phenomena like time dilation and length contraction which we will discuss later in this module. △

### 3.3. Light Cone

Our concept of time-like, light-like and space-like straight world-lines allows us to decompose spacetime $\mathbb{R}^4$ into 3 different regions, for any chosen point $y \in \mathbb{R}^4$. Let us define the subsets

$$\mathbf{I}_y := \left\{ x \in \mathbb{R}^4 \,:\, \mathrm{dist}_M(x,y)^2 < 0 \right\} , \tag{3.7a}$$

$$\mathbf{II}_y := \left\{ x \in \mathbb{R}^4 \,:\, \mathrm{dist}_M(x,y)^2 = 0 \right\} , \tag{3.7b}$$

$$\mathbf{III}_y := \left\{ x \in \mathbb{R}^4 \,:\, \mathrm{dist}_M(x,y)^2 > 0 \right\} . \tag{3.7c}$$

For a simplified 2-dimensional graphical illustration, using $y = 0$ and the restriction $x = (ct, x^1, 0, 0)$, see Figure 3.2 (left). Region $\mathbf{I}_y$ consists of all points $x \in \mathbb{R}^4$ which can be reached from $y$ via time-like world-lines. Region $\mathbf{II}_y$ consists of all points $x \in \mathbb{R}^4$ which can be reached from $y$ via light-like world-lines; this is called the *light cone*. Region $\mathbf{III}_y$ consists of all points $x \in \mathbb{R}^4$ which can be reached from $y$ via space-like world-lines.

Accepting the physical observation that all particles appearing in nature have either a time-like or a light-like world-line, we can give the following physical interpretation of Figure 3.2 (left): An observer sitting at $y$ can only receive signals/information from the past part of region $\mathbf{I}_y$ and the past part of the light cone $\mathbf{II}_y$. They can only transmit signals/information to the future part of region $\mathbf{I}_y$ and the future part of the light cone $\mathbf{II}_y$. The region $\mathbf{III}_y$ consists of all spacetime points our observer at $y$ cannot be influenced by or influence; one also says that the points in $\mathbf{III}_y$ are causally separated from $y$.

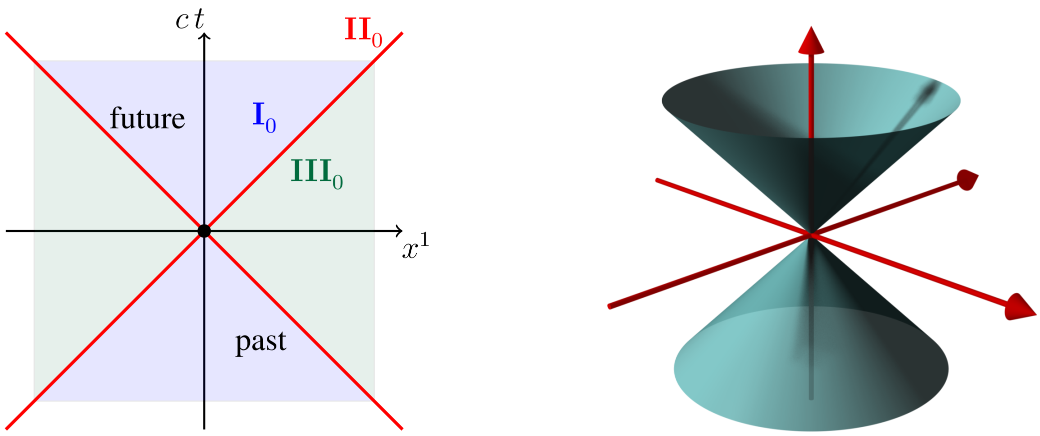


FIGURE 3.2. *Left:* 2-dimensional illustration of the regions $\mathbf{I}_y$, $\mathbf{II}_y$ and $\mathbf{III}_y$, for $y = 0$ and $x = (ct, x^1, 0, 0)$. *Right:* 3-dimensional illustration of the light cone (i.e. region $\mathbf{II}_y$).

REMARK 3.2. Our simplified 2-dimensional illustration of the light cone in Figure 3.2 (left) was done under the restriction $x = (ct, x^1, 0, 0)$. Allowing for generic

$x = (c\,t, x^1, x^2, x^3)$ (and still fixing without loss of generality $y = 0$), the light cone is determined by the equation $\mathrm{dist}_M(x, 0)^2 = 0$, or equivalently by

$$c^2\,t^2 = (x^1)^2 + (x^2)^2 + (x^3)^2 \ . \tag{3.8}$$

Geometrically, the solutions to this equation form a (3-dimensional) hyper-double cone in spacetime $\mathbb{R}^4$, which is of course not easy to visualize. To get a better feeling for how a light cone looks like, it is instructive to consider the restriction $x = (c\,t, x^1, x^2, 0)$ where we just set the third space coordinate to zero. The defining equation of the light cone (3.8) then reduces to the defining equation of the usual double cone, see Figure 3.2 (right) for a visualization. In this figure, region $\mathbf{I}_y$ is the inside part of the double cone and region $\mathbf{III}_y$ its outside part. △

### 3.4. Line Element and Metric

From the Minkowski distance formula (3.3) we can obtain the *Minkowski line element* by computing the square of the distance between two infinitesimally close points $x$ and $y = x + dx$ of spacetime $\mathbb{R}^4$. A short calculation shows that

$$ds_M^2 := \mathrm{dist}_M(x, x + dx)^2 = -c^2\,(dt)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \ . \tag{3.9}$$

The line element encodes the geometry of Minkowski spacetime.

REMARK 3.3. Similarly to Remark 2.1, it is mathematically more precise to regard the Minkowski line element (3.9) as a quadratic form on the space of tangent vectors $T_x\mathbb{R}^4$ at the point $x \in \mathbb{R}^4$ of spacetime, i.e.

$$q_M : T_x\mathbb{R}^4 \longrightarrow \mathbb{R} \ , \quad v \longmapsto -(v^0)^2 + (v^1)^2 + (v^2)^2 + (v^3)^2 \ , \tag{3.10}$$

where $v = (v^0, v^1, v^2, v^3)$. This defines an (indefinite) inner product

$$\langle \,\cdot\,,\,\cdot\,\rangle_M : T_x\mathbb{R}^4 \times T_x\mathbb{R}^4 \longrightarrow \mathbb{R} \tag{3.11a}$$

by setting

$$\langle v, w\rangle_M = \frac{1}{2}\Big(q_M(v + w) - q_M(v) - q_M(w)\Big) \ . \tag{3.11b}$$

Explicitly, the inner product is given by

$$\langle v, w\rangle_M = -v^0\,w^0 + v^1\,w^1 + v^2\,w^2 + v^3\,w^3 \tag{3.11c}$$

and it is called the *Minkowski inner product*. Equipping each tangent space $T_x\mathbb{R}^4$ with the Minkowski inner product (3.11) defines a mathematical structure called *Minkowski metric* on $\mathbb{R}^4$. This perspective will be clarified in our introduction to differential geometry later in this module. △

We can use the Minkowski line element (3.9), or equivalently the Minkowski metric (3.11), in order to generalize our definition of time-like, light-like and space-like straight world-lines to generic (i.e. not necessarily straight) world-lines. For this it is convenient to use the parametrized version (3.2) for describing world-lines. Given any world-line $x : \mathbb{R} \to \mathbb{R}^4 \ , \ \lambda \mapsto x(\lambda)$ as in (3.2), we can analyze its tangent vectors

$$\frac{dx(\lambda)}{d\lambda} \in \mathbb{R}^4 \tag{3.12}$$

at each parameter value $\lambda$. We say that the world-line $x : \mathbb{R} \to \mathbb{R}^4$ is

- *time-like* if $\langle \frac{dx(\lambda)}{d\lambda}, \frac{dx(\lambda)}{d\lambda} \rangle_M < 0$, for all $\lambda$;
- *light-like* if $\langle \frac{dx(\lambda)}{d\lambda}, \frac{dx(\lambda)}{d\lambda} \rangle_M = 0$, for all $\lambda$; and
- *space-like* if $\langle \frac{dx(\lambda)}{d\lambda}, \frac{dx(\lambda)}{d\lambda} \rangle_M > 0$, for all $\lambda$.

As a special case, we recover our previous definition which was restricted to straight world-lines. Moreover, we now can also see that the curved world-line in Figure 3.1 d) is time-like according to our more flexible definition. From now on, we shall always work with this more flexible definition of time-like, light-like and space-like.

### 3.5. Index Notation and Summation Convention

Writing out explicitly the summations in the Minkowski distance formula (3.3), the Minkowski line element (3.9) and the Minkowski inner product/metric (3.11) is typically too cumbersome in applications and calculations. The general practice in relativity is to introduce an efficient *index notation* to simplify such formulas. We shall now explain how this works.

First of all, we denote points $x \in \mathbb{R}^4$ of spacetime also by symbols like $x^\mu$, where $\mu$ is an index running over $0, 1, 2, 3$. The entries $x^i$, with $i$ an index running over $1, 2, 3$, describe the space coordinates and $x^0 = c\,t$ is up to the speed of light $c$ the time coordinate. Compare this also to (3.1). As a convention, Greek indices, e.g. $\mu, \nu, \rho, \sigma, \ldots$, always run over $0, 1, 2, 3$ and Latin indices, e.g. $i, j, k, l, \ldots$, always run over $1, 2, 3$. We next introduce the diagonal $4 \times 4$-matrix

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{3.13}$$

where it is crucial that the first diagonal element is $-1$. We denote the entries of $\eta$ by $\eta_{\mu\nu}$. (In relativity there is a difference between lower and upper indices, which will be clarified in our introduction to differential geometry later in this module.)

Using the index notation, we can write the square of the Minkowski distance formula (3.3) as

$$\mathrm{dist}_M(x, y)^2 = \sum_{\mu,\nu=0}^{3} \eta_{\mu\nu}\,(x^\mu - y^\mu)\,(x^\nu - y^\nu)\ . \tag{3.14}$$

Note that the negative sign in front of the first term in (3.3) is taken care of by the negative sign of the first diagonal entry in (3.13); that is why the negative sign in (3.13) is crucial. Still, equation (3.14) is a bit cumbersome because of writing out the summation symbols explicitly. To simplify our notations further, we use *Einstein's summation convention* which states the following: Summations over repeated upper and lower indices are suppressed from the notation. Hence, with the summation convention (3.14) simplifies further to

$$\mathrm{dist}_M(x, y)^2 = \eta_{\mu\nu}\,(x^\mu - y^\mu)\,(x^\nu - y^\nu)\ , \tag{3.15}$$

which is indeed a much more condensed and better readable expression than (3.3).

If we apply the same rules to the Minkowski line element (3.9) and the Minkowski inner product/metric (3.11), we obtain

$$ds_M^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu \tag{3.16}$$

and

$$\langle v, w \rangle_M = \eta_{\mu\nu}\, v^\mu\, w^\nu \ , \tag{3.17}$$

which are again very condensed and readable expressions.

In the following we shall use the index notation and summation convention without hesitation and further explanation whenever it is convenient.

CHAPTER 4

# Lorentz and Poincaré Transformations

## 4.1. Definition and Examples

Different reference frames for spacetime are related by *coordinate transformations*: Given the spacetime coordinates $x^\mu$ (i.e. $c\,t$, $x^1$, $x^2$ and $x^3$) of a reference frame for spacetime, we may choose a new set of spacetime coordinates $x'^\nu$ (i.e. $c\,t'$, $x'^1$, $x'^2$ and $x'^3$) and thereby obtain a new reference frame for spacetime. The new coordinates can be expressed in terms of the old ones, i.e. they are functions

$$x'^\nu = x'^\nu(x) \ , \tag{4.1}$$

or more explicitly $x'^\nu = x'^\nu(c\,t, x^1, x^2, x^3)$ for every $\nu = 0, 1, 2, 3$.

Of particular importance in special relativity are transformations which preserve the Minkowski distance formula (3.3) or equivalently the Minkowski line element $ds_M^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$, cf. (3.9). (We will see in the next section that these transformations automatically preserve the property of being an inertial frame for spacetime.) The relevant condition is thus given by

$$\eta_{\mu\nu}\, dx^\mu\, dx^\nu \overset{!}{=} \eta_{\mu\nu}\, dx'^\mu\, dx'^\nu \ . \tag{4.2}$$

Can we characterize those coordinate transformations (4.1) for which (4.2) holds true? Yes, indeed! But unfortunately this requires a bit of calculation.

First of all, let us study how the infinitesimals $dx^\mu$ behave under coordinate transformations. Inserting $x + dx$ into the right-hand side of (4.1) and using that $dx$ is infinitesimally small, i.e. Taylor expansion terminates at first order, we obtain

$$x'^\nu(x + dx) = x'^\nu(x) + \frac{\partial x'^\nu}{\partial x^\mu}\, dx^\mu \ , \tag{4.3}$$

where the symbol $\partial$ denotes partial derivatives. Comparing this with $x'^\nu + dx'^\nu$, we obtain

$$dx'^\nu = \frac{\partial x'^\nu}{\partial x^\mu}\, dx^\mu \ , \tag{4.4}$$

which is our desired transformation formula for the infinitesimals.

Inserting (4.4) twice (with appropriate index labeling) into the right-hand side of (4.2), we obtain

$$\eta_{\mu\nu}\, dx'^\mu\, dx'^\nu = \eta_{\mu\nu}\, \frac{\partial x'^\mu}{\partial x^\rho}\, \frac{\partial x'^\nu}{\partial x^\sigma}\, dx^\rho\, dx^\sigma \ . \tag{4.5}$$

Relabeling the indices (more precisely, exchanging $\mu \leftrightarrow \rho$ and $\nu \leftrightarrow \sigma$) in this formula, we find that the condition (4.2) is equivalent to

$$\eta_{\mu\nu} = \eta_{\rho\sigma}\, \frac{\partial x'^\rho}{\partial x^\mu}\, \frac{\partial x'^\sigma}{\partial x^\nu} \ . \tag{4.6}$$

In summary, condition (4.6) is necessary and sufficient for a coordinate transformation to preserve the Minkowski line element.

Next, we take the partial derivative $\frac{\partial}{\partial x^\tau}$ of both sides of (4.6). Using that $\eta_{\mu\nu}$ is constant and the Leibniz rule for partial derivatives, we find

$$0 = \eta_{\rho\sigma} \left( \frac{\partial^2 x'^\rho}{\partial x^\tau \partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} + \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial^2 x'^\sigma}{\partial x^\tau \partial x^\nu} \right) . \tag{4.7}$$

Exchanging the indices $\tau \leftrightarrow \nu$ in (4.7), we obtain the equivalent condition

$$0 = \eta_{\rho\sigma} \left( \frac{\partial^2 x'^\rho}{\partial x^\nu \partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\tau} + \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial^2 x'^\sigma}{\partial x^\nu \partial x^\tau} \right) , \tag{4.8}$$

while exchanging the indices $\tau \leftrightarrow \mu$ in (4.7), we obtain the equivalent condition

$$0 = \eta_{\rho\sigma} \left( \frac{\partial^2 x'^\rho}{\partial x^\mu \partial x^\tau} \frac{\partial x'^\sigma}{\partial x^\nu} + \frac{\partial x'^\rho}{\partial x^\tau} \frac{\partial^2 x'^\sigma}{\partial x^\mu \partial x^\nu} \right) . \tag{4.9}$$

The reason why we look at these equivalent conditions is the following: Adding or subtracting the individual sides of these equations according to the pattern

$$(4.7) \; + \; (4.8) \; - \; (4.9) \tag{4.10}$$

we obtain after using Schwarz's theorem (i.e. the order $\frac{\partial^2}{\partial x^\alpha \partial x^\beta} = \frac{\partial^2}{\partial x^\beta \partial x^\alpha}$ in which we take the 2nd partial derivatives does not matter) and symmetry of $\eta$ (i.e. $\eta_{\rho\sigma} = \eta_{\sigma\rho}$) the condition

$$0 = 2\, \eta_{\rho\sigma} \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial^2 x'^\sigma}{\partial x^\tau \partial x^\nu} . \tag{4.11}$$

Because both the matrix $\eta_{\rho\sigma}$ and the matrix $\frac{\partial x'^\rho}{\partial x^\mu}$ (called *Jacobi matrix* of the coordinate transformation) are invertible, the condition (4.11) is equivalent to

$$0 = \frac{\partial^2 x'^\sigma}{\partial x^\tau \partial x^\nu} . \tag{4.12}$$

Notice that this condition is necessary, but not sufficient, for a coordinate transformation to preserve the Minkowski line element.

Using (4.12) and (4.6), we can now find an explicit expression for all coordinate transformations preserving the Minkowski line element. From (4.12) it follows that the new spacetime coordinates $x'^\nu$ are 1st-order polynomials in the old spacetime coordinates $x^\mu$. Hence, there exist $b \in \mathbb{R}^4$ and a $4 \times 4$-matrix $\Lambda$ such that

$$x'^\nu = \Lambda^\nu{}_\mu \, x^\mu - b^\nu . \tag{4.13}$$

Notice that the Jacobi matrix of this transformation is given by

$$\frac{\partial x'^\nu}{\partial x^\mu} = \Lambda^\nu{}_\mu , \tag{4.14}$$

hence condition (4.6) reads as

$$\eta_{\mu\nu} = \eta_{\rho\sigma} \, \Lambda^\rho{}_\mu \, \Lambda^\sigma{}_\nu . \tag{4.15}$$

The coordinate transformations characterized by (4.13) and (4.15) are so important in special relativity that they deserve a name.

DEFINITION 4.1. A spacetime coordinate transformation of the form (4.13) with $\Lambda$ satisfying (4.15) is called a *Poincaré transformation*. Whenever $b = 0$, it is also called a *Lorentz transformation*.

It is sometimes convenient to express (4.13) and (4.15) in vector and matrix form, i.e. without using the index notation. The vector form of (4.13) is simply given by

$$x' = \Lambda\, x - b\ , \tag{4.16}$$

where $\Lambda\, x$ is the action of the $4 \times 4$-matrix $\Lambda$ on the vector $x$, cf. (3.1). In matrix form, the condition (4.15) reads as

$$\eta = \Lambda^{\mathrm{T}}\, \eta\, \Lambda\ , \tag{4.17}$$

where $\eta$ is the $4 \times 4$-matrix in (3.13) and $^{\mathrm{T}}$ denotes matrix transposition. (Of course, juxtaposition on the right-hand side of (4.17) means matrix multiplication.)

REMARK 4.2. In this form we observe that Lorentz transformations are quite similar to orthogonal transformations in Euclidean geometry: The orthogonality condition $R^{\mathrm{T}} = R^{-1}$ is equivalent to the condition $\mathbb{1} = R^{\mathrm{T}}\, R$, where $\mathbb{1}$ stands for the identity matrix. Inserting the identity matrix between $R^{\mathrm{T}}$ and $R$, we obtain the equivalent condition $\mathbb{1} = R^{\mathrm{T}}\, \mathbb{1}\, R$, which looks similar to (4.17). As already for the Minkowski distance formula (3.3), the (crucial!) difference between Lorentz transformations and orthogonal transformations is a negative sign (here encoded in $\eta$, cf. (3.13)). In mathematics, transformations satisfying the condition (4.17) are also called *pseudo-orthogonal transformations*. △

EXAMPLE 4.3 (Translations). The identity matrix $\Lambda = \mathbb{1}$ clearly satisfies condition (4.17). Hence, the transformation

$$x' = x - b\ , \tag{4.18}$$

for $b \in \mathbb{R}^4$, is a Poincaré transformation. Physically, the observer with coordinates $x'$ is displaced (i.e. translated) by $b$ in space and time with respect to the observer with coordinates $x$. Recall that Newtonian mechanics allows for precisely the same transformations, see Chapter 2.2. ▽

EXAMPLE 4.4 (Rotations). Let us set $b = 0$ and consider the matrix (in block form)

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}\ , \tag{4.19}$$

where $R$ is an orthogonal $3 \times 3$-matrix, i.e. $R^{\mathrm{T}}\, R = \mathbb{1}$. Notice that condition (4.17) is satisfied, which we see by the calculation (with block matrices)

$$\Lambda^{\mathrm{T}}\, \eta\, \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & R^{\mathrm{T}}\, R \end{pmatrix} = \eta\ . \tag{4.20}$$

As a consequence, the transformation

$$x' = \Lambda\, x \tag{4.21}$$

is a Lorentz transformation. Physically, the observer with coordinates $x'$ is rotated by $R$ (i.e. rotated just in space and not in time!) with respect to the observer with coordinates $x$. Recall that Newtonian mechanics allows for precisely the same transformations, see Chapter 2.2. ▽

EXAMPLE 4.5 (Lorentz boosts). Let us set $b = 0$ and consider the matrix

$$\Lambda = \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0 \\ -\sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{4.22}$$

where $\psi \in \mathbb{R}$ is a real parameter and $\sinh$ (respectively, $\cosh$) is the hyperbolic sine (respectively, cosine) function. One can easily compute by matrix multiplication

$$\Lambda^{\mathrm{T}}\,\eta\,\Lambda = \begin{pmatrix} -(\cosh\psi)^2 + (\sinh\psi)^2 & 0 & 0 & 0 \\ 0 & (\cosh\psi)^2 - (\sinh\psi)^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \eta\ , \tag{4.23}$$

where in the last step we used the well-known identity

$$(\cosh\psi)^2 - (\sinh\psi)^2 = 1\ . \tag{4.24}$$

As a consequence, the transformation

$$x' = \Lambda\, x \tag{4.25}$$

is a Lorentz transformation, which is called *Lorentz boost* along the $x^1$-direction. (There is an obvious notion of Lorentz boost along any other space direction. It is a good exercise to derive it on your own.) Notice that Lorentz boosts mix between space and time coordinates. They should be interpreted physically as transformations between observers which are uniformly moving with respect to each other. (See Chapter 4.3 below for more details on this point.) Hence, they are similar to the Galilean boosts of Chapter 2.2, however with the crucial advantage that they are compatible with constancy of the speed of light. The latter fact will be explained in the next section. ▽

## 4.2. Transformation of World-Lines

We did not confirm yet that Poincaré transformations preserve inertial frames for spacetime, i.e. that straight world-lines are mapped to straight world-lines under such transformations. Fortunately, this is an immediate consequence of our explicit expression (4.16): Because the new coordinates $x'$ are $1^{\text{st}}$-order polynomials in $x$, a straight world-line in $x$-coordinates is also straight in $x'$-coordinates.

Even more, we shall now see that Poincaré transformations preserve the property of a world-line being time-like, light-like or space-like. In particular, this implies that Poincaré transformations map light rays to light rays and hence they are compatible with the constancy of the speed of light in all inertial frames for spacetime. (This is the crucial difference to the Galilean transformations in Newtonian mechanics.) Let us provide the relevant argument: Consider a parametrized world-line $x : \mathbb{R} \to \mathbb{R}^4\,,~\lambda \mapsto x(\lambda)$ in $x$-coordinates. The corresponding parametrized world-line in $x'$-coordinates is then given by

$$x'(\lambda) = \Lambda\, x(\lambda) - b\,, \tag{4.26}$$

where $\Lambda$ and $b$ are of course independent of $\lambda$. Taking the $\lambda$-derivative as in (3.12), we obtain

$$\frac{dx'(\lambda)}{d\lambda} = \Lambda\,\frac{dx(\lambda)}{d\lambda}~. \tag{4.27}$$

Recall from the text below Eqn. (3.12) that the decision whether the world-line is time-like, light-like or space-like is done by considering the Minkowski inner product $\langle \frac{dx'(\lambda)}{d\lambda}, \frac{dx'(\lambda)}{d\lambda}\rangle_M$ of (4.27) with itself. Our claim then follows from the fact that Poincaré transformations preserve the Minkowski inner product. Explicitly,

$$\begin{aligned}\langle \frac{dx'(\lambda)}{d\lambda}, \frac{dx'(\lambda)}{d\lambda}\rangle_M &= \eta_{\mu\nu}\,\frac{dx'^{\mu}(\lambda)}{d\lambda}\,\frac{dx'^{\nu}(\lambda)}{d\lambda} = \eta_{\mu\nu}\,\Lambda^{\mu}{}_{\rho}\,\Lambda^{\nu}{}_{\sigma}\,\frac{dx^{\rho}(\lambda)}{d\lambda}\,\frac{dx^{\sigma}(\lambda)}{d\lambda}\\ &= \eta_{\rho\sigma}\,\frac{dx^{\rho}(\lambda)}{d\lambda}\,\frac{dx^{\sigma}(\lambda)}{d\lambda} = \langle \frac{dx(\lambda)}{d\lambda}, \frac{dx(\lambda)}{d\lambda}\rangle_M~,\end{aligned} \tag{4.28}$$

where the third equality follows from (4.15).

## 4.3. More on Lorentz Boosts

The most interesting Poincaré transformations are the Lorentz boosts introduced in Example 4.5, because these are genuinely different to Galilean transformations in Newtonian mechanics. We will therefore study Lorentz boosts in more detail.

For simplifying our discussion, we consider in the following only Lorentz boosts along the $x^1$-direction, i.e. spacetime coordinate transformations

$$x' = \Lambda\, x \tag{4.29}$$

with

$$\Lambda = \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0\\ -\sinh\psi & \cosh\psi & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}~, \tag{4.30}$$

where $\psi \in \mathbb{R}$ is a real parameter. Writing out (4.29) in components, we obtain

$$\begin{pmatrix} c\,t'\\ x'^1\\ x'^2\\ x'^3\end{pmatrix} = \begin{pmatrix} \cosh\psi\,c\,t - \sinh\psi\,x^1\\ -\sinh\psi\,c\,t + \cosh\psi\,x^1\\ x^2\\ x^3\end{pmatrix}~, \tag{4.31}$$

which shows explicitly how Lorentz boosts mix between space and time.

Let us now consider a particle in $x'$-coordinates that is at rest at the origin $x'^1 = x'^2 = x'^3 = 0$ of the reference frame for space. From (4.31) we obtain that in $x$-coordinates the world-line of this particle is determined by

$$0 = -\sinh\psi\, c\,t + \cosh\psi\, x^1 \ , \tag{4.32}$$

or after slight rearrangements

$$x^1 = \tanh\psi\, c\,t \ . \tag{4.33}$$

As a consequence, the unprimed observer sees the particle moving uniformly in the $x^1$-direction with velocity

$$v = \tanh\psi\, c \ . \tag{4.34}$$

(It is easy to confirm that all particles which are at rest in $x'$-coordinates will have the same velocity $v$ when observed from the unprimed observer. Checking this fact is a good exercise!) The physical interpretation of the Lorentz boost (4.31) is thus as follows: It describes the transformation from an inertial frame for spacetime to one which is moving with relative velocity $v = \tanh\psi\, c$ in the $x^1$-direction.

The following observation is crucial: Recall that $\tanh\psi$, for $\psi \in \mathbb{R}$, takes values in the open interval $(-1, 1)$, see Figure 4.1 for a visualization. As a consequence, the velocity (4.34) observed by the unprimed observer takes values in the interval $(-c, c)$, i.e.

$$v = \tanh\psi\, c \in (-c, c) \ . \tag{4.35}$$

In particular, it cannot exceed the speed of light, which is a feature we have already seen more abstractly in Chapter 4.2. Notice that this is in contrast to Galilean boosts (2.15) in Newtonian mechanics, where the relative velocity between two observers could be arbitrary (and hence leads to a conflict with constancy of the speed of light).

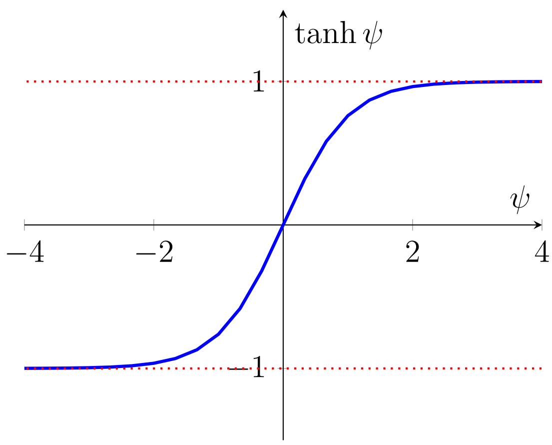


FIGURE 4.1. Plot of the hyperbolic tangent function tanh.

We finish this section by expressing the Lorentz boost (4.30) in terms of the parameter $v$ (4.35) instead of $\psi$; this will be useful for comparing it to the Galilean boost (2.15). Recall the identities

$$\cosh\psi = \frac{1}{\sqrt{1-(\tanh\psi)^2}} \quad , \qquad \sinh\psi = \tanh\psi\ \cosh\psi\ . \tag{4.36}$$

Inserting $\tanh\psi = \frac{v}{c}$, which follows from (4.35), yields

$$\cosh\psi = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \quad , \qquad \sinh\psi = \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}}\ . \tag{4.37}$$

Hence, (4.30) reads as

$$\Lambda = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{4.38}$$

More explicitly, expressing (4.31) in terms of $v$ we obtain

$$\begin{pmatrix} c\,t' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \frac{c\,t-\frac{v}{c}\,x^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{-v\,t+x^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ x^2 \\ x^3 \end{pmatrix} . \tag{4.39}$$

It is now clear that the Galilean boost (2.15) is an approximation of the Lorentz boost for $v \ll c$, i.e. $\frac{v}{c} \ll 1$ much smaller than 1: For $\frac{v}{c} \ll 1$, we find $\sqrt{1-\frac{v^2}{c^2}} \approx 1$ and $c\,t - \frac{v}{c}\,x^1 \approx c\,t$, hence (4.39) is approximately

$$\begin{pmatrix} c\,t' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \frac{c\,t-\frac{v}{c}\,x^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{-v\,t+x^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ x^2 \\ x^3 \end{pmatrix} \approx \begin{pmatrix} c\,t \\ -v\,t + x^1 \\ x^2 \\ x^3 \end{pmatrix} \qquad \text{(for } v \ll c\text{)}\ , \tag{4.40}$$

which is precisely the Galilean boost (2.15).

As a side-remark, the factor

$$\gamma := \frac{1}{\sqrt{1-\beta^2}} := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{4.41}$$

that appears in many of the expressions above is often called the *Lorentz factor* in the physics literature, where also

$$\beta := \frac{v}{c} \tag{4.42}$$

is used to denote the ratio of $v$ to the speed of light $c$.

### 4.4. Relativistic Addition of Velocities

The addition of velocities formula (2.20) of Newtonian mechanics is only an approximation (for small velocities) of a more fundamental relativistic addition of velocities formula which we shall now derive.

Assume that the observer with $x'$-coordinates observes a particle propagating along the $x'^1$-direction with velocity $v' = \tanh\psi'\, c \in (-c, c)$. The Lorentz boost

$$\Lambda' = \begin{pmatrix} \cosh\psi' & -\sinh\psi' & 0 & 0 \\ -\sinh\psi' & \cosh\psi' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.43}$$

transforms to an inertial frame for spacetime in which the particle is at rest. Our unprimed observer, who sees the primed observer moving with velocity $v = \tanh\psi\, c \in (-c, c)$ in the $x^1$-direction, has to apply the Lorentz boost

$$\Lambda'' = \Lambda'\,\Lambda \tag{4.44}$$

to transform to the rest frame of the particle. Explicitly, we obtain

$$\begin{aligned} \Lambda'' &= \begin{pmatrix} \cosh\psi' & -\sinh\psi' & 0 & 0 \\ -\sinh\psi' & \cosh\psi' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0 \\ -\sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cosh(\psi'+\psi) & -\sinh(\psi'+\psi) & 0 & 0 \\ -\sinh(\psi'+\psi) & \cosh(\psi'+\psi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \end{aligned} \tag{4.45}$$

where we used the identities

$$\cosh\psi'\ \cosh\psi + \sinh\psi'\ \sinh\psi = \cosh(\psi'+\psi)\ , \tag{4.46a}$$

$$\cosh\psi'\ \sinh\psi + \sinh\psi'\ \cosh\psi = \sinh(\psi'+\psi)\ . \tag{4.46b}$$

The velocity of the particle observed by the unprimed observer is thus

$$v'' = \tanh(\psi'+\psi)\, c = \frac{\tanh\psi'\, c + \tanh\psi\, c}{1 + \tanh\psi'\ \tanh\psi} = \frac{v'+v}{1+\frac{v'\,v}{c^2}}\ . \tag{4.47}$$

For small velocities $v' \ll c$ and $v \ll c$, we obtain the approximation

$$v'' \approx v' + v \qquad (\text{for } v' \ll c \text{ and } v \ll c)\ , \tag{4.48}$$

which is precisely the Galilean addition of velocities formula (2.20). (Here we used a slightly different notation. For a comparison between (4.48) and (2.20), set $V = v''$ and $V' = v'$ in (2.20).)

CHAPTER 5

# Relativistic Physical Effects

In this chapter we discuss two spectacular relativistic physical effects: Time dilation and length contraction. These effects are a direct consequence of the mathematical fact that, even though Poincaré transformations preserve the Minkowski distance formula (3.3) or equivalently the Minkowski line element (3.9), they *do not* preserve the individual concepts of time duration and space distance.

## 5.1. Time Dilation

Let us fix an inertial frame for spacetime with coordinates $x$ and consider a particle propagating with constant velocity $v \in (-c, c)$ in the $x^1$-direction. Using the techniques developed in Chapter 4.3, we can find a Lorentz boost to transform to a new inertial frame for spacetime with coordinates $x'$ in which our particle is at rest. The transformation explicitly reads as

$$x' = \Lambda\, x \ , \tag{5.1}$$

with

$$\Lambda = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ . \tag{5.2}$$

Let us consider any two events $A$ and $B$ lying on the world-line of our particle. Using the coordinates of the primed observer (who sees the particle at rest), we can write

$$x'_A = \begin{pmatrix} c\, t'_A \\ {x'_A}^1 \\ {x'_A}^2 \\ {x'_A}^3 \end{pmatrix} \quad , \qquad x'_B = \begin{pmatrix} c\, t'_B \\ {x'_A}^1 \\ {x'_A}^2 \\ {x'_A}^3 \end{pmatrix} \ . \tag{5.3}$$

Notice that the space coordinates of both events are the same because by hypothesis the particle is at rest in $x'$-coordinates. The primed observer will conclude that the time duration between event $A$ and $B$ is

$$\Delta t' = t'_B - t'_A \ . \tag{5.4}$$

Let us now consider the same events in the coordinates of the unprimed observer. Using the inverse Lorentz boost

$$\Lambda^{-1} = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.5}$$

and the inverse transformation formula

$$x = \Lambda^{-1}\, x' \ , \tag{5.6}$$

we obtain

$$x_A = \Lambda^{-1}\, x'_A = \begin{pmatrix} \frac{c\,t'_A + \frac{v}{c}\,{x'_A}^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{v\,t'_A + {x'_A}^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ {x'_A}^2 \\ {x'_A}^3 \end{pmatrix} \ , \qquad x_B = \Lambda^{-1}\, x'_B = \begin{pmatrix} \frac{c\,t'_B + \frac{v}{c}\,{x'_A}^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{v\,t'_B + {x'_A}^1}{\sqrt{1-\frac{v^2}{c^2}}} \\ {x'_A}^2 \\ {x'_A}^3 \end{pmatrix} \ . \tag{5.7}$$

The unprimed observer will thus conclude that the time duration between event $A$ and $B$ is

$$\Delta t = t_B - t_A = \frac{1}{c}\left(\frac{c\,t'_B + \frac{v}{c}\,{x'_A}^1}{\sqrt{1-\frac{v^2}{c^2}}} - \frac{c\,t'_A + \frac{v}{c}\,{x'_A}^1}{\sqrt{1-\frac{v^2}{c^2}}}\right) = \frac{\Delta t'}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma\,\Delta t' \ , \tag{5.8}$$

where in the last step we recognized the Lorentz factor (4.41). Equivalently, we obtain the relation

$$\Delta t' = \sqrt{1-\frac{v^2}{c^2}}\ \Delta t = \gamma^{-1}\,\Delta t \tag{5.9}$$

between the time durations $\Delta t$ and $\Delta t'$ measured by our two observers.

Let us give a physical interpretation: For $v \neq 0$, the inverse Lorentz factor $\gamma^{-1} = \sqrt{1-\frac{v^2}{c^2}} < 1$ is smaller than $1$ and it gets smaller and smaller when we increase $v$ towards $\pm c$. Together with (5.9), this implies that the primed observer (seeing the particle at rest) measures a shorter time duration between event $A$ and $B$ than the unprimed observer (seeing the particle moving with velocity $v$). In summary, our calculation above explains the popular saying that “moving clocks go slower”.

## 5.2. Length Contraction

Let us fix an inertial frame for spacetime with coordinates $x'$. Consider a resting (1-dimensional) rod of length $L'$ which is aligned along the ${x'}^1$-direction. The endpoints $A$ and $B$ of this rod are described by the world-lines (in $x'$-coordinates)

$$x'_A = \begin{pmatrix} c\,t'_A \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \qquad x'_B = \begin{pmatrix} c\,t'_B \\ L' \\ 0 \\ 0 \end{pmatrix} \quad , \tag{5.10}$$

where we assumed without loss of generality that the endpoint $A$ is located at the origin of our spatial coordinate system. (This can always be achieved by applying a translation.) Figure 5.1 provides an illustration of our present scenario.

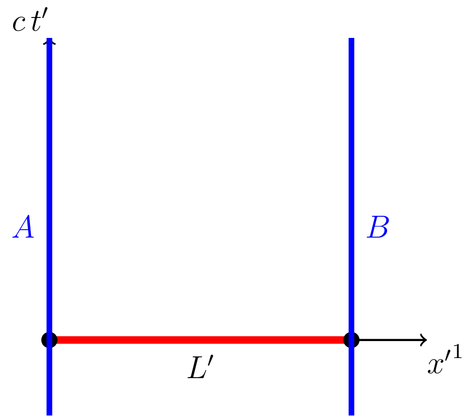


FIGURE 5.1. A 1-dimensional rod of length $L'$ resting in the $x'$-coordinate system. $A$ and $B$ describe the world-lines of its endpoints.

We can measure the length $L'$ of the rod via the following prescription: Fix an arbitrary time $t'_0$ (interpreted as the time at which we measure the length) and consider both endpoints $A$ and $B$ at this time. We obtain the events

$$x'_A|_{t'_0} = \begin{pmatrix} c\,t'_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \qquad x'_B|_{t'_0} = \begin{pmatrix} c\,t'_0 \\ L' \\ 0 \\ 0 \end{pmatrix} \quad . \tag{5.11}$$

Their distance in space is given by

$$\Delta x'^1 = {x'_B}^1|_{t'_0} - {x'_A}^1|_{t'_0} = L' \ , \tag{5.12}$$

which coincides with our original length $L'$.

Let us now consider the same problem from the perspective of an observer who sees the rod moving with constant velocity $v \in (-c, c)$ along the $x^1$-direction. The coordinates of this observer are denoted by $x$ and they are obtained as in the previous section via the inverse Lorentz boost (5.5) and the transformation formula

$x = \Lambda^{-1}\, x'$. The world-lines of the two endpoints (5.10) read in $x$-coordinates as

$$x_A = \Lambda^{-1}\, x'_A = \begin{pmatrix} \frac{c\, t'_A}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{v\, t'_A}{\sqrt{1-\frac{v^2}{c^2}}} \\ 0 \\ 0 \end{pmatrix} \quad , \qquad x_B = \Lambda^{-1}\, x'_B = \begin{pmatrix} \frac{c\, t'_B + \frac{v}{c} L'}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{v\, t'_B + L'}{\sqrt{1-\frac{v^2}{c^2}}} \\ 0 \\ 0 \end{pmatrix} \quad . \tag{5.13}$$

To measure the length of the rod in $x$-coordinates, we use the same prescription as above: Fix an arbitrary time $t_0$, which we interpret as the time at which we measure the length. Considering the events $x_A|_{t_0}$ and $x_B|_{t_0}$ fixes the time coordinates $t'_A$ and $t'_B$ in (5.13) via the equations

$$c\, t_0 = \frac{c\, t'_A}{\sqrt{1 - \frac{v^2}{c^2}}} \quad , \qquad c\, t_0 = \frac{c\, t'_B + \frac{v}{c} L'}{\sqrt{1 - \frac{v^2}{c^2}}} \ . \tag{5.14}$$

The solution of these equations is

$$t'_A = \sqrt{1 - \frac{v^2}{c^2}}\; t_0 \quad , \qquad t'_B = \sqrt{1 - \frac{v^2}{c^2}}\; t_0 - \frac{v}{c^2}\, L' \ . \tag{5.15}$$

Inserting this into (5.13), we obtain after a short calculation

$$x_A|_{t_0} = \begin{pmatrix} c\, t_0 \\ v\, t_0 \\ 0 \\ 0 \end{pmatrix} \quad , \qquad x_B|_{t_0} = \begin{pmatrix} c\, t_0 \\ v\, t_0 + \sqrt{1 - \frac{v^2}{c^2}}\; L' \\ 0 \\ 0 \end{pmatrix} \quad . \tag{5.16}$$

The distance in space between $A$ and $B$ is thus given by

$$\Delta x^1 = x^1_B|_{t_0} - x^1_A|_{t_0} = \sqrt{1 - \frac{v^2}{c^2}}\; L' = \gamma^{-1}\, L' \ , \tag{5.17}$$

where in the last step we recognized the inverse of the Lorentz factor (4.41). Because $\Delta x^1$ is the length of the rod measured by the unprimed observer, we also shall use the notation $L = \Delta x^1$. Hence, we obtain the relation

$$L' = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\; L = \gamma\, L \tag{5.18}$$

between the lengths $L$ and $L'$ measured by our two observers.

Let us give an interpretation: For $v \neq 0$, the Lorentz factor $\gamma = 1/\sqrt{1 - \frac{v^2}{c^2}} > 1$ is greater than $1$ and it gets bigger and bigger when we increase $v$ towards $\pm c$. Together with (5.18), this implies that the primed observer (seeing the rod at rest) measures a longer rod than the unprimed observer (seeing the rod moving with velocity $v$). In summary, our calculation above explains the popular saying that “moving rods are shorter”.

CHAPTER 6

# Relativistic Mechanics

In this chapter we develop a relativistic generalization of Newtonian mechanics. We introduce $4$-dimensional generalizations of velocities, momenta, accelerations and forces, which are called $4$-velocities, $4$-momenta, $4$-accelerations and $4$-forces. This will lead us to one of the most famous equation in physics: $E = m\,c^2$.

## 6.1. Proper Time

Recall that a point-particle is described by its (not necessarily straight) world-line in Minkowski spacetime

$$x : \mathbb{R} \longrightarrow \mathbb{R}^4 \ , \quad \lambda \longmapsto x(\lambda) = \begin{pmatrix} c\,t(\lambda) \\ x^1(\lambda) \\ x^2(\lambda) \\ x^3(\lambda) \end{pmatrix} \ , \tag{6.1}$$

where $\lambda$ is an arbitrary parametrization. Unless otherwise stated, we assume in this chapter that the world-line is time-like, i.e.

$$\langle \frac{dx(\lambda)}{d\lambda}, \frac{dx(\lambda)}{d\lambda} \rangle_M = \eta_{\mu\nu}\, \frac{dx^\mu(\lambda)}{d\lambda}\, \frac{dx^\nu(\lambda)}{d\lambda} < 0 \ , \tag{6.2}$$

for all $\lambda \in \mathbb{R}$.

For a time-like world-line, we may always choose coordinate time $t$ (of any inertial frame for spacetime) as parameter. (Can you prove this?) In this parametrization we have

$$x(t) = \begin{pmatrix} c\,t \\ x^1(t) \\ x^2(t) \\ x^3(t) \end{pmatrix} = \begin{pmatrix} c\,t \\ \mathbf{x}(t) \end{pmatrix} \ , \tag{6.3}$$

where we introduced the space 3-vector $\mathbf{x}(t) = (x^1(t), x^2(t), x^3(t))$ for later convenience. The derivative with respect to $t$ then reads as

$$\frac{dx(t)}{dt} = \begin{pmatrix} c \\ \frac{d\mathbf{x}(t)}{dt} \end{pmatrix} = \begin{pmatrix} c \\ \mathbf{v}(t) \end{pmatrix} \ , \tag{6.4}$$

where $\mathbf{v}(t) = (v^1(t), v^2(t), v^3(t))$ is the velocity (in general time-dependent) measured by the observer with $x$-coordinates. Condition (6.2) then simply says that

$$\mathbf{v}(t)^2 < c^2 \,, \tag{6.5}$$

for all $t$, where by $\mathbf{v}(t)^2 = v^1(t)^2 + v^2(t)^2 + v^3(t)^2$ we denote the square of the Euclidean norm of the velocity $\mathbf{v}(t)$. In words, the norm of the velocity $\mathbf{v}(t)$ is always smaller than the speed of light for any time-like world-line.

The problem with our description above is that it is coordinate (i.e. observer) dependent. The parametrization of the world-line by coordinate time $t$, and as a consequence also the velocity $\mathbf{v}(t)$, *do not* transform well under Poincaré transformations; this is similar to the fact that neither time nor space are individually observer independent concepts as they transform into each other by Poincaré transformations. The solution to this problem will be to introduce an intrinsic Poincaré-invariant parametrization of the world-line by its proper time.

*Proper time* $\tau$ is by definition the time experienced by the particle itself. More precisely, it is the time which a clock that is attached to the particle would show, i.e. a clock that is moving with the same velocity as the particle. Due to the relativistic effect of time dilation, the proper time $\tau$ will in general be different to the coordinate time $t$. Using our time dilation formula (5.9), we obtain that an infinitesimal change in proper time $d\tau$ is related to an infinitesimal change in coordinate time $dt$ by the formula

$$d\tau = \sqrt{1 - \frac{\mathbf{v}(t)^2}{c^2}}\, dt \,. \tag{6.6}$$

Notice that, in contrast to (5.9), the velocity in this formula may depend on time $t$ because the world-line of our particle is not necessarily straight. Equation (6.6) may also be understood as a differential equation

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{\mathbf{v}(t)^2}{c^2}} \,, \tag{6.7}$$

whose solution $\tau = \tau(t)$ is the transformation formula from coordinate time $t$ to proper time $\tau$. The solution of this differential equation is given by integration

$$\tau(t) = \int_{t_0}^{t} \sqrt{1 - \frac{\mathbf{v}(t')^2}{c^2}}\, dt' \,, \tag{6.8}$$

where $t_0 \in \mathbb{R}$ is an arbitrary constant which is used to fix the origin of proper time. (Notice that $\tau(t_0) = 0$.) The following argument shows that proper time, defined by (6.6), is indeed a Poincaré-invariant concept: Taking the square of this expression, we compute

$$\begin{aligned} (d\tau)^2 &= \left(1 - \frac{\mathbf{v}(t)^2}{c^2}\right) (dt)^2 = \left(1 - \frac{1}{c^2}\left(\frac{d\mathbf{x}(t)}{dt}\right)^2\right) (dt)^2 \\ &= (dt)^2 - \frac{1}{c^2}\, (d\mathbf{x}(t))^2 = -\frac{1}{c^2}\, ds_M^2 \,, \end{aligned} \tag{6.9}$$

or equivalently

$$-c^2\, (d\tau)^2 = ds_M^2 \,. \tag{6.10}$$

In words, this calculation shows that (up to the Poincaré-invariant factor $-c^2$), the square of proper time between two (infinitesimally close) points of the world-line is given by the Minkowski line element, which is a Poincaré-invariant quantity. As a consequence, proper time is a Poincaré-invariant quantity, i.e. $\tau' = \tau$.

To conclude this subsection, let us notice that (6.10) provides an equally good, but simpler definition of proper time than our original definition in (6.6). Hence, you can decide for yourself which of these definitions you prefer.

## 6.2. Kinematics

In the following we use proper time $\tau$ to parametrize our world-line, i.e. we consider the map

$$x : \mathbb{R} \longrightarrow \mathbb{R}^4 \ , \quad \tau \longmapsto x(\tau) = \begin{pmatrix} c\, t(\tau) \\ \mathbf{x}(\tau) \end{pmatrix} \ . \tag{6.11}$$

In contrast to our previous description (6.3) using coordinate time $t$, the parametrization by proper time (6.11) transforms well under Poincaré transformations: Given a Poincaré transformation $x' = \Lambda\, x - b$, the world-line (6.11) transforms as

$$x'(\tau') = \Lambda\, x(\tau) - b \ , \tag{6.12}$$

because, according to our discussion above, $\tau' = \tau$ is a Poincaré-invariant quantity.

Let us now consider the $\tau$-derivative of the world-line (6.11). We define

$$u(\tau) := \frac{dx(\tau)}{d\tau} \tag{6.13}$$

and call $u(\tau)$ the $4$-*velocity* at proper time $\tau$ of the particle. Notice that the $4$-velocity transforms well under Poincaré transformations $x' = \Lambda\, x - b$,

$$u'(\tau') = \frac{dx'(\tau')}{d\tau'} = \frac{d}{d\tau}\Big(\Lambda\, x(\tau) - b\Big) = \Lambda\, \frac{dx(\tau)}{d\tau} = \Lambda\, u(\tau) \ , \tag{6.14}$$

because $b$ and $\Lambda$ are independent of $\tau$, and $\tau' = \tau$ is Poincaré-invariant. Taking the second $\tau$-derivative of the world-line (6.11) we define

$$a(\tau) := \frac{d^2x(\tau)}{d\tau^2} = \frac{du(\tau)}{d\tau} \tag{6.15}$$

and call $a(\tau)$ the $4$-*acceleration* at proper time $\tau$ of the particle. Again, it is easy to see that the $4$-acceleration transforms as

$$a'(\tau') = \Lambda\, a(\tau) \tag{6.16}$$

under Poincaré transformations $x' = \Lambda\, x - b$.

It is sometimes convenient to express the $4$-velocity $u$ in terms of the usual velocity $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ seen by our observer with $x$-coordinates. For this we decompose (6.13) into a time and space component

$$u = \begin{pmatrix} u^0 \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} c\, \frac{dt}{d\tau} \\ \frac{d\mathbf{x}}{d\tau} \end{pmatrix} \ . \tag{6.17}$$

(We suppress here and in the following the arguments $\tau$ in $u(\tau)$ and $x(\tau)$ to simplify notations.) Using (6.6) we find

$$u^0 = c\,\frac{dt}{d\tau} = \frac{c}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}} \tag{6.18a}$$

and

$$\mathbf{u} = \frac{d\mathbf{x}}{d\tau} = \frac{d\mathbf{x}}{dt}\,\frac{dt}{d\tau} = \frac{\mathbf{v}}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\;. \tag{6.18b}$$

In particular, notice that the space component $\mathbf{u}$ of the 4-velocity $u$ is *not* the velocity $\mathbf{v}$ measured by our observer with $x$-coordinates, but scaled by the Lorentz factor $1/\sqrt{1-\frac{\mathbf{v}^2}{c^2}}$. From (6.18) we immediately obtain

$$\langle u,u\rangle_M = \eta_{\mu\nu}\,u^\mu\,u^\nu = -(u^0)^2 + \mathbf{u}^2 = \frac{1}{1-\frac{\mathbf{v}^2}{c^2}}\Big(-c^2+\mathbf{v}^2\Big) = -c^2 \tag{6.19}$$

for the Minkowski inner product of $u$ with itself. Hence, the 4-velocity of any time-like world-line is normalized according to

$$\langle u,u\rangle_M = -c^2\;. \tag{6.20}$$

In particular, we rediscover that $u$ is time-like, which we of course already knew before.

We conclude this section by analyzing the Minkowski inner product $\langle a,u\rangle_M$ between the 4-velocity and the 4-acceleration. Taking the $\tau$-derivative of the normalization (6.20) yields

$$0 = \frac{d}{d\tau}\langle u,u\rangle_M = \eta_{\mu\nu}\,\frac{d}{d\tau}\Big(u^\mu\,u^\nu\Big) = \eta_{\mu\nu}\,\Big(\frac{du^\mu}{d\tau}\,u^\nu + u^\mu\,\frac{du^\nu}{d\tau}\Big) = 2\eta_{\mu\nu}a^\mu\,u^\nu\;, \tag{6.21}$$

where we used that $\eta$ is symmetric, i.e. $\eta_{\mu\nu} = \eta_{\nu\mu}$. It follows that

$$\langle a,u\rangle_M = 0\;, \tag{6.22}$$

i.e. the 4-acceleration $a$ is "orthogonal" (in the sense of the Minkowski inner product) to the 4-velocity $u$. In particular, one can prove that $a$ is space-like because $u$ is time-like and non-zero. (Try to prove this!)

### 6.3. Dynamics

The obvious relativistic generalization of Newton's second law is given by

$$m_0\,a = F\;, \tag{6.23a}$$

or in our index notation by

$$m_0\,a^\mu = F^\mu\;. \tag{6.23b}$$

It is important to stress that (6.23) is an equation involving the 4-acceleration $a$ and equating it with a quantity $F$ called 4-*force*. The parameter $m_0$ is called the *rest mass* of the particle. Notice that for (6.23) to transform well under Poincaré transformations $x' = \Lambda\,x - b$, the 4-force $F$ has to transform as

$$F' = \Lambda\,F \tag{6.24}$$

and the rest mass has to be Poincaré-invariant. (This follows from the transformation behavior of the 4-acceleration (6.16).) Moreover, because of (6.22) any 4-force has to satisfy

$$\langle F, u \rangle_M = 0 \,, \tag{6.25}$$

i.e. it has to be "orthogonal" (in the Minkowski sense) to the 4-velocity $u$.

It is fully justified if you wonder at this point about the physical interpretation of the 4-force. Where does it come from? How does it look like? We are unfortunately facing here the same issues as in Newtonian mechanics: A force is introduced as a concept which influences the movement of particles, however neither Newtonian mechanics nor our relativistic mechanics provides an explanation where the force comes from. In practice, physicists study data coming from experiments and develop a model for a force (or 4-force) which describes these particular observations. We will study a physically relevant example of a 4-force, the so-called Lorentz force, at the end of this chapter. It describes the force experienced by charged particles in electric and magnetic fields.

The following observation is very important: Setting $F = 0$ in (6.23), we obtain the equation

$$a = \frac{du}{d\tau} = 0 \,. \tag{6.26}$$

As a consequence, the 4-velocity $u = \mathrm{const}$ is constant. Using further (6.18), this implies that also the "ordinary" velocity $\mathbf{v}$ measured by the observer with $x$-coordinates is constant. We thus obtain that the solution $x(\tau)$ to (6.23) for $F = 0$ describes a straight world-line. In other words, force-free particles move according to our equation (6.23) on straight world-lines (as they should do!).

Again from formal analogy to Newtonian mechanics, we define

$$p := m_0\, u \tag{6.27}$$

and call $p$ the 4-*momentum* of the particle. The relativistic equation of motion (6.23) can then also be written as

$$\frac{dp}{d\tau} = F \,, \tag{6.28a}$$

or in our index notation as

$$\frac{dp^\mu}{d\tau} = F^\mu \,. \tag{6.28b}$$

This form of the relativistic equation of motion is suitable for comparison to Newtonian mechanics. Let us decompose $p$ and $F$ into time and space components

$$p = \begin{pmatrix} p^0 \\ \mathbf{p} \end{pmatrix} \quad , \qquad F = \begin{pmatrix} F^0 \\ \mathbf{F} \end{pmatrix} \,, \tag{6.29}$$

and consider the space component of (6.28), i.e.

$$\frac{d\mathbf{p}}{d\tau} = \mathbf{F} \,. \tag{6.30}$$

Using further the relation between proper time $\tau$ and coordinate time $t$ given in (6.6), we obtain for the coordinate time derivative

$$\frac{d\mathbf{p}}{dt} = \frac{d\tau}{dt}\frac{d\mathbf{p}}{d\tau} = \sqrt{1-\frac{\mathbf{v}^2}{c^2}}\,\mathbf{F}\,. \tag{6.31}$$

Introducing the "effective force"

$$\mathbf{K} := \sqrt{1-\frac{\mathbf{v}^2}{c^2}}\,\mathbf{F}\,, \tag{6.32}$$

we obtain the equation

$$\frac{d\mathbf{p}}{dt} = \mathbf{K}\,, \tag{6.33}$$

which looks precisely like the Newtonian equation of motion written in terms of the momentum of the particle.

A few physical remarks are in order: Firstly, the magnitude of the effective force $\mathbf{K}$ (6.32) is velocity dependent. It goes to zero when the velocity approaches the speed of light, i.e. fast particles experience a weaker effective force. Secondly, expressing $\mathbf{p}$ in terms of the "ordinary" velocity $\mathbf{v}$ measured by the observer with $x$-coordinates by using (6.18), we obtain

$$\mathbf{p} = m_0\,\mathbf{u} = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\,\mathbf{v} = m\,\mathbf{v}\,, \tag{6.34}$$

where in the last equality we introduced the "effective mass"

$$m = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\,. \tag{6.35}$$

The latter quantity goes to $\infty$ when the velocity approaches the speed of light, i.e. fast particles have a larger effective mass. Both of these features physically prevent us from accelerating massive particles to the speed of light (or faster than the speed of light): When the particle gets faster, it experiences a weaker effective force and a larger effective mass, which makes it harder and harder to accelerate it further.

As a final remark, let us study the limit for slow particles $\mathbf{v}^2 \ll c^2$. In this case we have

$$m = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}} \approx m_0 \tag{6.36}$$

and

$$\mathbf{K} = \sqrt{1-\frac{\mathbf{v}^2}{c^2}}\,\mathbf{F} \approx \mathbf{F}\,. \tag{6.37}$$

Using also (6.34), we find that our equation of motion (6.33) can be approximated in this limit by

$$m_0\,\frac{d\mathbf{v}}{dt} \approx \mathbf{F}\,, \tag{6.38}$$

which is Newton's second law. Hence, relativistic mechanics may be approximated by Newtonian mechanics if the velocities are much smaller than the speed of light.

## 6.4. Energy

We now will study the time component of our relativistic equation of motion (6.28). It reads as

$$\frac{dp^0}{d\tau} = F^0 \, . \tag{6.39}$$

Using (6.25), i.e.

$$-F^0 \, u^0 + \mathbf{F} \cdot \mathbf{u} = 0 \, , \tag{6.40}$$

where $\mathbf{F} \cdot \mathbf{u} = F^1 \, u^1 + F^2 \, u^2 + F^3 \, u^3$ is the Euclidean inner product, and (6.18), we can express

$$F^0 = \frac{\mathbf{F} \cdot \mathbf{u}}{u^0} = \frac{\mathbf{F} \cdot \mathbf{v}}{c} = \frac{\mathbf{K} \cdot \mathbf{v}}{c \sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} \tag{6.41}$$

in terms of the ordinary velocity $\mathbf{v}$ and the effective force $\mathbf{K}$ (6.32). As a consequence, the 4-force $F$ is completely specified by the effective force $\mathbf{K}$ and the equation

$$F = \frac{1}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} \begin{pmatrix} \frac{\mathbf{K} \cdot \mathbf{v}}{c} \\ \mathbf{K} \end{pmatrix} \, . \tag{6.42}$$

Using further (6.6) to express the $\tau$-derivative in terms of a $t$-derivative, we obtain that (6.39) is equivalent to

$$\frac{d(c \, p^0)}{dt} = \mathbf{K} \cdot \mathbf{v} \, . \tag{6.43}$$

The quantity on the right-hand side of (6.43), i.e. the Euclidean inner product between force and velocity, is called *power* in Newtonian physics. Because power is by definition the rate of consuming energy per time, i.e. $\frac{dE}{dt} = \mathrm{power}$, it is justified to call

$$E := c \, p^0 = c \, m_0 \, u^0 = \frac{m_0 \, c^2}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} = m \, c^2 \tag{6.44}$$

the *energy* of the particle. (In the last equality we used the definition (6.36) of the effective mass $m$ of the particle.) Then (6.43) is given by

$$\frac{dE}{dt} = \mathbf{K} \cdot \mathbf{v} \, , \tag{6.45}$$

i.e. it is just the balance of energy equation from physics. With this notational convention, the 4-momentum takes the form

$$p = \begin{pmatrix} \frac{E}{c} \\ \mathbf{p} \end{pmatrix} \, , \tag{6.46}$$

i.e. its time component is (up to a factor $1/c$) the energy of the particle. This is why the 4-momentum $p$ is also often called the *energy-momentum* 4-*vector* in the literature.

Let us now study the energy (6.44) in the case where the velocity is much smaller than $c$. Using first-order Taylor expansion in $\frac{\mathbf{v}^2}{c^2}$, we obtain

$$E = m_0\, c^2\, \frac{1}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}} \approx m_0\, c^2 \left(1 + \frac{\mathbf{v}^2}{2\, c^2}\right) = m_0\, c^2 + \frac{1}{2} m_0\, \mathbf{v}^2\ . \tag{6.47}$$

The second term is precisely the (non-relativistic) kinetic energy of the particle from Newtonian physics. The first term is more interesting: It says that a massive particle (with rest mass $m_0$) carries an energy, even when it is at rest (i.e. $\mathbf{v} = 0$). The quantity

$$E_0 = m_0\, c^2 \tag{6.48}$$

is called the *rest energy* of the particle.

Equations (6.44) and (6.48) establish a relationship between mass and energy. This *mass-energy equivalence* is a direct consequence of special relativity. It is physically extremely important to understand, for example, nuclear fission reactions. In nuclear fission, a heavy (unstable) particle decays into lighter particles and thereby releases a huge amount of energy in the form of gamma-radiation. This energy is used for example in nuclear power plants. Notice that the origin of this energy is the rest energy of our heavy particle.

## 6.5. Energy-Momentum Relation

In Newtonian mechanics, we have the famous relationship $E_{\mathrm{kin}} = \frac{\mathbf{p}^2}{2m}$ between the (kinetic) energy and the momentum of the particle. This relation is not anymore valid in special relativity, but it has to be replaced by a more fundamental energy-momentum relation which we shall now derive.

Recalling the normalization $\langle u, u\rangle_M = -c^2$ of the 4-velocity (6.20) and the definition $p = m_0\, u$ of the 4-momentum, we immediately obtain the equation

$$\langle p, p\rangle_M = m_0^2\, \langle u, u\rangle_M = -m_0^2\, c^2\ . \tag{6.49}$$

Using further that

$$\langle p, p\rangle_M = \eta_{\mu\nu}\, p^\mu\, p^\nu = -(p^0)^2 + \mathbf{p}^2 = -\frac{E^2}{c^2} + \mathbf{p}^2\ , \tag{6.50}$$

we can express the energy $E$ as a function of $\mathbf{p}$ via

$$E^2 = m_0^2\, c^4 + \mathbf{p}^2\, c^2\ . \tag{6.51}$$

This is the relativistic *energy-momentum relation*.

Let us consider the positive square root of (6.51), i.e.

$$E = \sqrt{m_0^2\, c^4 + \mathbf{p}^2\, c^2} = \sqrt{E_0^2 + \mathbf{p}^2\, c^2}\ , \tag{6.52}$$

where in the last equality we used the definition (6.48) of the rest energy $E_0$. For small momenta $\mathbf{p}^2\, c^2 \ll E_0^2$, we can Taylor expand (6.52) to first-order in $\mathbf{p}^2\, c^2/E_0^2$ and obtain

$$E = E_0 \sqrt{1 + \frac{\mathbf{p}^2\, c^2}{E_0^2}} \approx E_0 + \frac{\mathbf{p}^2}{2\, m_0} \qquad (\text{for } \mathbf{p}^2\, c^2/E_0^2 \ll 1)\ , \tag{6.53}$$

which is (up to the rest energy contribution) the Newtonian energy-momentum relation. Another interesting limit of (6.52) is for $E_0^2/\mathbf{p}^2\,c^2 \ll 1$, i.e. the rest energy is negligible to the $\mathbf{p}^2\,c^2$ contribution. We obtain by zeroth-order Taylor expansion in $E_0^2/\mathbf{p}^2\,c^2$

$$E = c\,|\mathbf{p}|\,\sqrt{\frac{E_0^2}{\mathbf{p}^2\,c^2} + 1} \approx c\,|\mathbf{p}| \qquad (\text{for } E_0^2/\mathbf{p}^2\,c^2 \ll 1)\ , \tag{6.54}$$

where $|\mathbf{p}| = \sqrt{\mathbf{p}^2}$ is the Euclidean norm of $\mathbf{p}$. Physically, this shows that for very high momentum $\mathbf{p}$ (compared to the rest energy), the energy-momentum relation becomes linear, while for very low momentum (compared to the rest energy) it is quadratic.

In your module on electromagnetism, you will learn that electromagnetic waves (i.e. light) have a linear energy-momentum relation *for all values of* $\mathbf{p}$. Comparison with (6.52) then leads us to the conclusion that light particles (i.e. photons) have zero rest mass $m_0 = 0$, which is indeed an experimentally confirmed fact.

## 6.6. Example: Charged Particle in a Constant Electric Field

The prime example of a relativistic $4$-force is the *Lorentz force*, which acts on charged particles propagating through an electric and magnetic field. It is completely characterized by the effective force

$$\mathbf{K} = q\,\Big(\mathbf{E} + \mathbf{v}\times\mathbf{B}\Big)\ , \tag{6.55}$$

where $q$ is the electric charge of the particle, $\mathbf{E}$ is the electric field and $\mathbf{B}$ is the magnetic field. (The symbol $\times$ denotes the cross-product of $3$-vectors.) In the following, we consider only the case $\mathbf{B} = \mathbf{0}$ of vanishing magnetic field. Using (6.42), we find that the corresponding $4$-force is

$$F = \frac{1}{\sqrt{1-\frac{\mathbf{v}^2}{c^2}}}\begin{pmatrix}\frac{q}{c}\,\mathbf{E}\cdot\mathbf{v}\\ q\,\mathbf{E}\end{pmatrix} = \frac{q}{c}\begin{pmatrix}\mathbf{E}\cdot\mathbf{u}\\ \mathbf{E}\,u^0\end{pmatrix}\ , \tag{6.56}$$

where in the last equality we used (6.18).

Let us now solve the relativistic equation of motion

$$m_0\,\frac{du}{d\tau} = F \tag{6.57}$$

under the initial conditions (at $\tau = 0$) given by

$$u(0) = \begin{pmatrix}c\\0\\0\\0\end{pmatrix}\ ,\qquad x(0) = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}\ . \tag{6.58}$$

(Physically, these initial conditions say that the particle is initially located at $x = 0$ and that it has zero initial velocity $\mathbf{v} = 0$.) To simplify our calculations, let us assume that the electric field is constant and that it points towards the $x^1$-direction, i.e. $\mathbf{E} = (E^1, 0, 0)$. Writing out all four components of (6.57) and dividing by $m_0$ yields

$$\frac{d}{d\tau}\begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix} = \underbrace{\frac{q\,E^1}{m_0\,c}}_{=\frac{\alpha}{c}} \begin{pmatrix} u^1 \\ u^0 \\ 0 \\ 0 \end{pmatrix} . \tag{6.59}$$

Using our initial conditions, we immediately find that $u^2 = u^3 = 0$ and $x^2 = x^3 = 0$, hence we can focus on the system of differential equations

$$\frac{du^0}{d\tau} = \frac{\alpha}{c}\,u^1 \quad , \qquad \frac{du^1}{d\tau} = \frac{\alpha}{c}\,u^0 \tag{6.60}$$

given by the first two components. Using our initial conditions, the solution is

$$u^0 = c\,\cosh\left(\frac{\alpha}{c}\,\tau\right) \quad , \qquad u^1 = c\,\sinh\left(\frac{\alpha}{c}\,\tau\right) . \tag{6.61}$$

To obtain the world-line $x(\tau)$, we have to integrate the equations

$$\frac{d(c\,t)}{d\tau} = u^0 = c\,\cosh\left(\frac{\alpha}{c}\,\tau\right) \quad , \qquad \frac{dx^1}{d\tau} = u^1 = c\,\sinh\left(\frac{\alpha}{c}\,\tau\right) . \tag{6.62}$$

Using again our initial conditions, we obtain

$$c\,t = \frac{c^2}{\alpha}\,\sinh\left(\frac{\alpha}{c}\,\tau\right) \quad , \qquad x^1 = \frac{c^2}{\alpha}\left(\cosh\left(\frac{\alpha}{c}\,\tau\right) - 1\right) . \tag{6.63}$$

The world-line describing the solution $x(\tau)$ is thus a hyperbola in the $c\,t - x^1$-plane. See Figure 6.1 for a visualization.

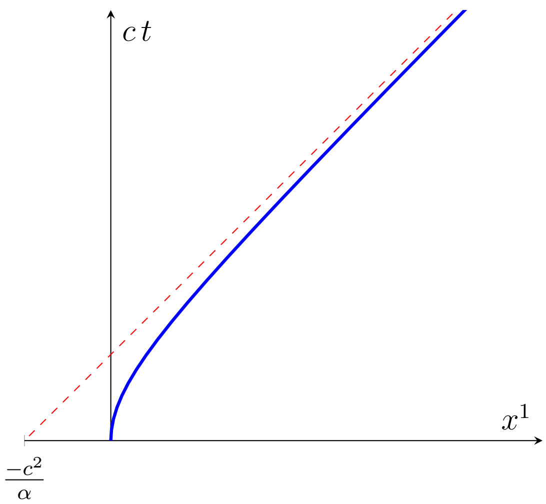


FIGURE 6.1. World-line of a charged particle under the influence of the Lorentz force corresponding to a constant electric field $\mathbf{E} = (E^1, 0, 0)$. The dashed red line is light-like.

Let us conclude this chapter by discussing the physics of our solution (6.63). For small $\tau$, we can expand (6.63) as

$$c\,t \approx c\,\tau \quad , \qquad x^1 \approx \frac{c^2}{\alpha}\left(1 + \frac{1}{2}\frac{\alpha^2}{c^2}\tau^2 - 1\right) = \frac{\alpha}{2}\tau^2\,. \tag{6.64}$$

Inserting the first equation into the second one, we obtain

$$x^1(t) \approx \frac{\alpha}{2}t^2\,, \tag{6.65}$$

which is the trajectory of a particle in Newtonian mechanics that experiences a constant acceleration $\alpha = \frac{q\,E^1}{m_0}$. The small time behavior of our world-line can thus be approximated very well by Newtonian mechanics. For large $\tau$ the situation is different: While a constant acceleration in Newtonian mechanics would accelerate our particle to an arbitrary high velocity, the velocity of the world-line (6.63) is bounded by the speed of light $c$. To see this fact, let us rearrange the first equation in (6.63) as

$$\frac{\alpha}{c}\tau = \sinh^{-1}\left(\frac{\alpha}{c}t\right)\,. \tag{6.66}$$

Inserting this into the second equation in (6.63) yields

$$x^1 = \frac{c^2}{\alpha}\left(\cosh\sinh^{-1}\left(\frac{\alpha}{c}t\right) - 1\right) = \frac{c^2}{\alpha}\left(\sqrt{1 + \frac{\alpha^2}{c^2}t^2} - 1\right)\,. \tag{6.67}$$

We now can determine the velocity (in the $x^1$-direction) of our particle

$$v = \frac{dx^1}{dt} = \frac{\alpha\,t}{\sqrt{1 + \frac{\alpha^2}{c^2}t^2}}\,. \tag{6.68}$$

Figure 6.2 visualizes the velocity as a function of time $t$. Notice in particular that $v < c$ for all times and that $v \to c$ for $t \to \infty$.

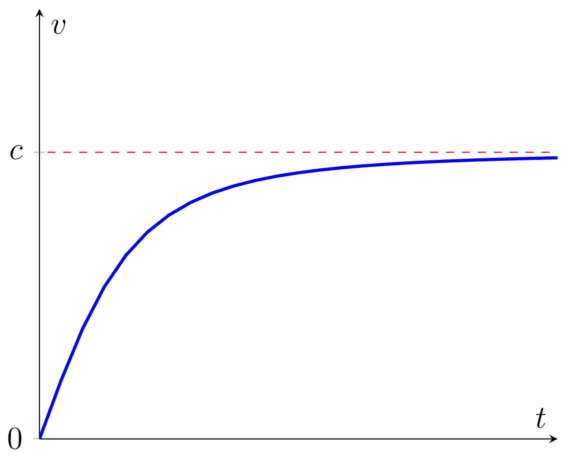


FIGURE 6.2. Time-dependence of the velocity of a charged particle in a constant electric field.

CHAPTER 7

# Relativistic Continuum Mechanics

In this chapter we introduce the concept of energy-momentum tensors for relativistic point-particles and also for continuous distributions of relativistic matter, such as dust or fluids. The energy-momentum tensor will play a crucial role in general relativity, where it is the source of the gravitational field.

## 7.1. Energy-Momentum Tensor of Point-Particles

Let us consider a point-particle with rest mass $m_0 > 0$ which we describe by the world-line (parametrized by proper time $\tau$)

$$x_P : \mathbb{R} \longrightarrow \mathbb{R}^4 \ , \quad \tau \longmapsto x_P(\tau) = \begin{pmatrix} c\, t_P(\tau) \\ \mathbf{x}_P(\tau) \end{pmatrix} \ . \tag{7.1}$$

We denote the world-line by $x_P$, i.e. with a subscript '$P$' (meaning particle), because the symbol $x$ will play a different role in what follows. As in the previous chapter, we denote the $4$-velocity of our particle by

$$u_P = \frac{dx_P}{d\tau} = \begin{pmatrix} u_P^0 \\ \mathbf{u}_P \end{pmatrix} \ . \tag{7.2}$$

We also assume that our particle satisfies the relativistic equation of motion without an external $4$-force, i.e.

$$m_0 \, \frac{du_P}{d\tau} = 0 \ . \tag{7.3}$$

We have learned in Chapter 6 that a point-particle carries energy and momentum. We now would like to think of this energy and momentum as being localized in Minkowski spacetime at those points/events $x \in \mathbb{R}^4$ which lie in the image of the particle's world-line. This requires us to introduce a suitable notion of energy and momentum *density*, which will be realized by the concept of *energy-momentum tensors* (also called *stress-energy tensors* in the literature).

In full generality, an energy-momentum tensor is an assignment

$$T : \mathbb{R}^4 \longrightarrow \mathrm{Mat}_4 \ , \quad x \longmapsto T(x) \tag{7.4}$$

of a $4 \times 4$-matrix to each point $x \in \mathbb{R}^4$ of Minkowski spacetime that satisfies various conditions which we shall discuss in more detail below. We will often write this $4 \times 4$-matrix in our index notation as

$$T^{\mu\nu}(x) \ . \tag{7.5}$$

The conditions to be satisfied by an energy-momentum tensor are as follows:

(i) $T$ is symmetric, i.e.

$$T^{\mu\nu}(x) = T^{\nu\mu}(x) \; ; \tag{7.6}$$

(ii) $T$ transforms under a Poincaré transformation $x' = \Lambda\, x - b$ as

$$T'^{\mu\nu}(x') = \Lambda^{\mu}{}_{\rho}\, \Lambda^{\nu}{}_{\sigma}\, T^{\rho\sigma}(x) \; ; \tag{7.7}$$

(iii) $T$ is conserved, i.e. the $4$-divergence vanishes

$$\frac{\partial}{\partial x^{\mu}} T^{\mu\nu}(x) = 0 \; . \tag{7.8}$$

We propose the following energy-momentum tensor for our point-particle

$$T^{\mu\nu}_{P}(x) = \int_{\mathbb{R}} m_0\, u^{\mu}_{P}\, u^{\nu}_{P}\, \delta^{(4)}(x - x_P(\tau))\, c\, d\tau \; , \tag{7.9}$$

where $\delta^{(4)}$ is the $4$-dimensional Dirac-delta function. Recall that it is defined by the integration property

$$\int_{\mathbb{R}^4} \delta^{(4)}(x - y)\, f(x)\, d^4 x = f(y) \; , \tag{7.10}$$

for all (smooth) functions $f : \mathbb{R}^4 \to \mathbb{R}\, ,\; x \mapsto f(x)$ on Minkowski spacetime. Condition (i) is clearly satisfied. Condition (ii) requires a brief calculation

$$\begin{aligned} T'^{\mu\nu}_{P}(x') &= \int_{\mathbb{R}} m_0\, u'^{\mu}_{P}\, u'^{\nu}_{P}\, \delta^{(4)}(x' - x'_P(\tau))\, c\, d\tau \\ &= \Lambda^{\mu}{}_{\rho}\, \Lambda^{\nu}{}_{\sigma} \int_{\mathbb{R}} m_0\, u^{\rho}_{P}\, u^{\sigma}_{P}\, \delta^{(4)}\Big(\Lambda(x - x_P(\tau))\Big)\, c\, d\tau \\ &= \Lambda^{\mu}{}_{\rho}\, \Lambda^{\nu}{}_{\sigma}\, T^{\rho\sigma}_{P}(x) \; , \end{aligned} \tag{7.11}$$

where we used the transformation property (6.14) of $4$-velocities, the standard property $\delta^{(4)}(\Lambda(x - x_P(\tau))) = \frac{1}{|\det \Lambda|}\, \delta^{(4)}(x - x_P(\tau)) = \delta^{(4)}(x - x_P(\tau))$ of the Dirac-delta function, and the fact that proper time is a Poincaré-invariant quantity. Verifying condition (iii) requires a slightly longer and more complicated calculation. We will first do this calculation and then give explanations afterwards:

$$\begin{aligned} \frac{\partial}{\partial x^{\mu}} T^{\mu\nu}_{P}(x) &= \int_{\mathbb{R}} m_0\, u^{\mu}_{P}\, u^{\nu}_{P}\, \frac{\partial}{\partial x^{\mu}} \delta^{(4)}(x - x_P(\tau))\, c\, d\tau \\ &= -\int_{\mathbb{R}} m_0\, u^{\mu}_{P}\, u^{\nu}_{P}\, \frac{\partial}{\partial x^{\mu}_{P}} \delta^{(4)}(x - x_P(\tau))\, c\, d\tau \\ &= -\int_{\mathbb{R}} m_0\, u^{\nu}_{P}\, \frac{d}{d\tau} \delta^{(4)}(x - x_P(\tau))\, c\, d\tau \\ &= \int_{\mathbb{R}} m_0\, \frac{du^{\nu}_{P}}{d\tau}\, \delta^{(4)}(x - x_P(\tau))\, c\, d\tau \\ &= 0 \; . \end{aligned} \tag{7.12}$$

In the first step we just pulled the derivative under the integral. The second step is a consequence of the chain rule for differentiation. In step three we realized (again via the chain rule for differentiation) that

$$u^{\mu}_{P}\, \frac{\partial}{\partial x^{\mu}_{P}} = \frac{dx^{\mu}_{P}}{d\tau}\, \frac{\partial}{\partial x^{\mu}_{P}} = \frac{d}{d\tau} \; . \tag{7.13}$$

The fourth step is just integration by parts. In step five we used the relativistic equation of motion (7.3) for our force-free particle. It is an important observation to notice that condition (iii) for an energy-momentum tensor is linked to the equation of motion for the matter it describes.

Let us now give a physical interpretation of the quantity $T_P^{\mu\nu}(x)$ defined in (7.9). First of all, the physical dimension of $T_P^{\mu\nu}(x)$ is $\mathrm{kg\ m^2/s^2\ m^{-3}}$, i.e. energy per 3-volume. Second, because of the Dirac-delta function, $T_P^{\mu\nu}(x) = 0$ if $x \in \mathbb{R}^4$ does not lie on the world-line of the particle; the energy-momentum tensor is thus localized on the particle's world-line. To study in more detail the individual components of $T_P^{\mu\nu}(x)$, and hence to provide a physical interpretation, we shall rewrite (7.9) in the following equivalent form

$$T_P^{\mu\nu}(x) = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}_P^2}{c^2}}}\, v_P^\mu\, v_P^\nu\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))\,, \tag{7.14}$$

where $\mathbf{v}_P = \frac{d\mathbf{x}_P}{dt}$ is the ordinary velocity and

$$v_P = \sqrt{1-\frac{\mathbf{v}_P^2}{c^2}}\, u_P = \begin{pmatrix} c \\ \mathbf{v}_P \end{pmatrix}\,. \tag{7.15}$$

(Equation (7.14) can be obtained from (7.9) by computing the integral over proper time. It is a good exercise to do this calculation.)

This form of writing the energy-momentum tensor is particularly useful for physically analyzing its components: For the $00$-component we find

$$T_P^{00}(x) = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}_P^2}{c^2}}}\, c^2\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t)) = E_P\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))\,, \tag{7.16}$$

where in the second equality we used our previous definition of energy (6.44). Thus, we obtain that $T_P^{00}(x)$ describes the energy density. For the $0i$-component, with $i = 1, 2, 3$, we find

$$T_P^{0i}(x) = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}_P^2}{c^2}}}\, c\, v_P^i\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t)) = c\, p_P^i\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))\,, \tag{7.17}$$

where $p_P^i$ are the space components of the $4$-momentum $p_P = m_0\, u_P$ of the particle. Thus, we obtain that $T_P^{0i}(x)$ describes (up to a factor $c$) the $i$-component of the momentum density. Another way of interpreting $T_P^{0i}(x)$ is obtained by writing

$$T_P^{0i}(x) = E_P\, \frac{v_P^i}{c}\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))\,. \tag{7.18}$$

Hence, $T_P^{0i}(x)$ also describes the flow of energy density in the $i$-direction. This perspective is also useful to interpret the components $T_P^{ij}(x)$, with $i, j = 1, 2, 3$: We obtain

$$T_P^{ij}(x) = \frac{m_0}{\sqrt{1-\frac{\mathbf{v}_P^2}{c^2}}}\, v_P^i\, v_P^j\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t)) = p_P^i\, v_P^j\, \delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))\,, \tag{7.19}$$

i.e. $T_P^{ij}(x)$ describes the flow of the $i$-component of the momentum density in the $j$-direction. Alternatively, writing

$$T_P^{ij}(x) = v_P^i \, p_P^j \, \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t)) \ , \tag{7.20}$$

we find that $T_P^{ij}(x)$ also describes the flow of the $j$-component of the momentum density in the $i$-direction.

We conclude this section by studying a system of $N \in \mathbb{N}$ non-interacting point-particles. Assume that particle $a = 1, \dots, N$ has rest mass $m_{0a} > 0$ and that it is described by a world-line

$$x_a : \mathbb{R} \longrightarrow \mathbb{R}^4 \ , \quad \tau \longmapsto x_a(\tau) \tag{7.21}$$

satisfying the force-free relativistic equation of motion

$$m_{0a} \, \frac{du_a}{d\tau} = 0 \ . \tag{7.22}$$

The total energy-momentum tensor for our system of $N$ particles is given by the sum

$$T^{\mu\nu}_{\text{tot}}(x) = \sum_{a=1}^{N} T^{\mu\nu}_a(x) = \sum_{a=1}^{N} \int_{\mathbb{R}} m_{0a} \, u_a^\mu \, u_a^\nu \, \delta^{(4)}(x - x_a(\tau)) \, c \, d\tau \tag{7.23}$$

of the individual energy-momentum tensors. Our conditions (i), (ii) and (iii) for an energy-momentum tensor are clearly fulfilled.

## 7.2. Energy-Momentum Tensor of Relativistic Dust and Fluids

In practice, it is often necessary to consider forms of matter which consists of extremely many particles. For example, 1 liter of water consists of about $10^{25}$ particles ($H_2O$-molecules) and a handful of dust consists of a huge number of dust particles. Describing such kinds of matter by keeping track of the world-line of each individual particle and using (7.23) as energy-momentum tensor is clearly very inconvenient, complicated and inefficient.

A more clever and practical way to describe matter consisting of a huge number of particles is in terms of a *continuum approximation*: We consider a function

$$\rho : \mathbb{R}^4 \longrightarrow \mathbb{R} \ , \quad x \longmapsto \rho(x) \tag{7.24}$$

on the Minkowski spacetime taking values in the real numbers. The function $\rho$ is physically interpreted as *mass density*, i.e. the distribution of mass/matter throughout Minkowski spacetime. More precisely, the value $\rho(x)$ describes the mass localized in an infinitesimally small volume at the spacetime point $x$. The physical dimension of $\rho$ has to be $\mathrm{kg\,m^{-3}}$. To describe the dynamics of our distribution of mass, we further introduce a $4$-vector-valued function

$$U : \mathbb{R}^4 \longrightarrow \mathbb{R}^4 \ , \quad x \longmapsto U(x) = \begin{pmatrix} U^0(x) \\ \mathbf{U}(x) \end{pmatrix} \tag{7.25}$$

on the Minkowski spacetime. We will often write this vector in our index notation as

$$U^\mu(x) \ . \tag{7.26}$$

The physical interpretation of $U(x)$ is as the $4$-velocity of our distribution of mass at the spacetime point $x$; it describes how the mass density $\rho(x)$ at $x$ is flowing/moving.

Inspired by our energy-momentum tensor for point-particles (7.9) and particle systems (7.23), we propose the following energy-momentum tensor for our continuous distribution of matter

$$T^{\mu\nu}_{\text{dust}}(x) = \rho(x)\, U^\mu(x)\, U^\nu(x)\,. \tag{7.27}$$

This form of matter is called *dust* and one should think of it as a continuous approximation of a cloud of extremely many non-interacting (dust) particles.

Let us now study the conditions (i), (ii) and (iii) that any energy-momentum tensor has to satisfy. Condition (i) is clearly satisfied, $T^{\mu\nu}_{\text{dust}}(x) = T^{\nu\mu}_{\text{dust}}(x)$. Condition (ii) is satisfied provided that we introduce the following transformation rules for $\rho$ and $U$ under Poincaré transformations $x' = \Lambda\, x - b$:

$$\rho'(x') = \rho(x) \quad , \qquad U'^\mu(x') = \Lambda^\mu{}_\nu\, U^\nu(x)\,. \tag{7.28}$$

These transformation rules are very well-motivated because the mass density $\rho$ is a scalar quantity and $U$ should transform similarly to the 4-velocity of point-particles (6.14). Condition (iii) is *not* automatically satisfied. Recalling that for our example of point-particles condition (iii) was equivalent to the relativistic equation of motion, we *define* the equation of motion for dust to be

$$\frac{\partial}{\partial x^\mu} T^{\mu\nu}_{\text{dust}}(x) = 0\,. \tag{7.29}$$

REMARK 7.1. To get a better understanding of this equation of motion, let us study its Newtonian limit for small velocities. We can then write

$$U(x) = \frac{1}{\sqrt{1-\frac{\mathbf{v}(x)^2}{c^2}}} \begin{pmatrix} c \\ \mathbf{v}(x)\end{pmatrix} \approx \begin{pmatrix} c \\ \mathbf{v}(x)\end{pmatrix}\,, \tag{7.30}$$

where $\mathbf{v}(x)$ is the velocity of the dust at $x$ that is measured by our observer with $x$-coordinates. In this approximation, we obtain for the 0-component of (7.29)

$$0 = \frac{\partial}{\partial x^\mu} T^{\mu 0}_{\text{dust}}(x) \approx \frac{\partial}{\partial x^\mu}\Big(\rho\, U^\mu\, c\Big) = U^\mu \frac{\partial \rho}{\partial x^\mu}\, c + \rho\, \frac{\partial U^\mu}{\partial x^\mu}\, c\,. \tag{7.31}$$

Writing out the (implicit) summations and using again that $U^0 \approx c$ is constant, we find that this equation is equivalent to

$$0 = \frac{\partial \rho}{\partial t} + \mathbf{v}\cdot\nabla\rho + \rho\,\nabla\cdot\mathbf{v} = \frac{d\rho}{dt} + \rho\,\nabla\cdot\mathbf{v}\,, \tag{7.32}$$

which is precisely the continuity equation of Newtonian continuum mechanics. (Here $\nabla = (\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3})$ is the gradient you know from 3-dimensional vector calculus and $\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla$ is the total time derivative.) For the $i$-component of (7.29), with $i = 1, 2, 3$, we obtain in our Newtonian approximation

$$0 = \frac{\partial}{\partial x^\mu} T^{\mu i}_{\text{dust}}(x) \approx \underbrace{\left(\frac{\partial \rho}{\partial t} + \mathbf{v}\cdot\nabla\rho + \rho\,\nabla\cdot\mathbf{v}\right)}_{=0} v^i + \rho\, U^\mu \frac{\partial v^i}{\partial x^\mu}\,. \tag{7.33}$$

Notice that the first term vanishes by using Eqn. (7.32). As a consequence, we obtain the equation

$$0 = \rho \, \frac{\partial \mathbf{v}}{\partial t} + \rho \, (\mathbf{v} \cdot \nabla) \, \mathbf{v} = \rho \, \frac{d\mathbf{v}}{dt} \ . \tag{7.34}$$

This equation is called the (pressureless) Euler equation of Newtonian continuum mechanics. In summary, we have seen in this remark that our equation of motion for dust (7.29) is a relativistic generalization of the continuity equation (7.32) and the (pressureless) Euler equation (7.34) of Newtonian continuum mechanics. △

To conclude this chapter, let us briefly mention that the energy-momentum tensor of fluids is rather similar to the one of dust (7.27). The only difference between a fluid and dust is that the former causes *pressure*. We describe the pressure density by a function

$$P : \mathbb{R}^4 \longrightarrow \mathbb{R} \ , \quad x \longmapsto P(x) \tag{7.35}$$

on the Minkowski spacetime taking values in the real numbers; $P(x)$ is the pressure localized in an infinitesimally small $3$-volume at $x$. The energy-momentum tensor for a fluid is then obtained by adding suitable pressure terms to (7.27). Explicitly,

$$T^{\mu\nu}_{\text{fluid}} = \Big(\rho + \frac{P}{c^2}\Big) \, U^\mu \, U^\nu + P \, \eta^{\mu\nu} \ , \tag{7.36}$$

where $\eta$ is the Minkowski metric (3.13). We will not study this energy-momentum tensor in more detail and just mention that in the Newtonian limit (for small velocities) the relativistic equation of motion for fluids $\frac{\partial}{\partial x^\mu} T^{\mu\nu}_{\text{fluid}} = 0$ is approximated by the usual equations of Newtonian hydrodynamics, i.e. an appropriate continuity equation and the Euler equation.

# Part 3

# General Relativity

CHAPTER 8

# The Physical Ideas Underlying General Relativity

## 8.1. Incompatibility of Newtonian Gravity and Special Relativity

Newton's law of gravitation describes how a mass density $\rho(t, \mathbf{x})$ generates a gravitational potential $\Phi(t, \mathbf{x})$. The relevant equation is

$$\Delta\Phi = 4\pi\, G\, \rho\,, \tag{8.1}$$

where

$$G = 6.67408 \times 10^{-11}\ \mathrm{m}^3\, \mathrm{kg}^{-1}\, \mathrm{s}^{-2} \tag{8.2}$$

is Newton's constant and

$$\Delta = \nabla^2 = \sum_{i=1}^{3} \frac{\partial^2}{\partial x^{i\,2}} \tag{8.3}$$

is the Laplace operator. (Notice that $\Delta$ involves only space derivatives, but no time derivatives.) It is obvious that Eqn. (8.1) is *not* form-invariant under Poincaré transformations: Performing for example a Lorentz boost, the Laplace operator in (8.1) would transform to an operator involving time derivatives because Lorentz boosts mix between space and time. As a consequence, Newtonian gravitation is incompatible with special relativity because its fundamental law (8.1) is not form-invariant under changing the inertial frame for spacetime.

An alternative way to see the incompatibility between Newtonian gravitation and special relativity is to look at the solution of (8.1), i.e.

$$\Phi(t, \mathbf{x}) = -G \int_{\mathbb{R}^3} \frac{\rho(t, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|}\, d^3\mathbf{y}\ . \tag{8.4}$$

We see that the value of $\Phi$ at $\mathbf{x}$ will respond instantaneously to a change in $\rho$ at any other point $\mathbf{y}$. As we argued in Chapter 2.5, simultaneity of events in special relativity is a concept that is observer dependent. This again leads to the conclusion that the laws of Newtonian gravitation are incompatible with special relativity.

General relativity, which we will develop in the remaining part of the module, will be the cure to this incompatibility.

## 8.2. Equivalence Principle

The physics of a non-relativistic point-particle in a gravitational potential $\Phi(t, \mathbf{x})$ is described by Newton's second law

$$m_0\, \frac{d^2}{dt^2}\mathbf{x}(t) = m_G\, \mathbf{g}(t, \mathbf{x}(t))\ , \tag{8.5}$$

where the gravitational field g is defined by

$$\mathbf{g} = -\nabla\Phi \ . \tag{8.6}$$

Notice that we distinguished between the *inertial mass* $m_0$ and the *gravitational mass* $m_G$, which in principle could be different! An experimentally verified fact is that inertial mass and gravitational mass are the same, i.e.

$$m_0 = m_G \ , \tag{8.7}$$

for *all* different forms of matter/particles. Informally, one can say that "all kinds of matter/particles fall equally fast in a gravitational field".

Because of the equivalence between inertial and gravitational mass, the corresponding mass terms drop out of Newton's second law and we obtain the equation

$$\frac{d^2}{dt^2}\mathbf{x}(t) = \mathbf{g}(t, \mathbf{x}(t)) \ , \tag{8.8}$$

which describes the dynamics of any particle in a gravitational field. Consider now for simplicity a constant gravitational field $\mathbf{g} = \mathrm{const}$. Defining new coordinates $t'$ and $\mathbf{x}'$ by the coordinate transformation formulas

$$t' = t \quad , \qquad \mathbf{x}' = \mathbf{x} - \frac{1}{2}\mathbf{g}\, t^2 \ , \tag{8.9}$$

we observe that the trajectory of our particle in $x'$-coordinates satisfies the equation

$$\frac{d^2}{dt'^2}\mathbf{x}'(t') = \frac{d^2}{dt^2}\Big(\mathbf{x}(t) - \frac{1}{2}\mathbf{g}\, t^2\Big) = \frac{d^2}{dt^2}\mathbf{x}(t) - \mathbf{g} = \mathbf{g} - \mathbf{g} = \mathbf{0} \ . \tag{8.10}$$

In words, transforming to a *non-inertial* (uniformly accelerated) reference frame, we were able to transform away our constant gravitational field g.

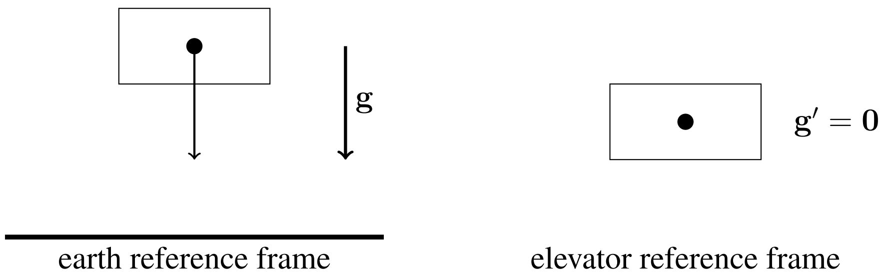


FIGURE 8.1. An elevator containing a particle. Both are falling freely in the gravitational field of earth. *Left:* An observer sitting on earth sees that there is a force acting on the particle, the gravitational force. *Right:* An observer sitting in the elevator sees no force acting on the particle.

Physically, you can understand this fact by thinking of the following scenario, see Figure 8.1: Think of a freely-falling elevator in the gravitational field of earth. Any particle inside the elevator will experience (due to equivalence of inertial and gravitational mass) the same acceleration as the elevator itself, so it is also freely falling. From the perspective of an observer living on earth, the particle experiences a force, the gravitational force. However, from the perspective of an observer sitting inside the elevator, the same particle appears to be force-free. As a side-remark, notice that this is the reason why astronauts in the international space station do not experience gravitational forces! They and the space station itself are freely falling in the gravitational field of earth.

Given a non-constant gravitational field $\mathbf{g}(t, \mathbf{x})$, we may choose an arbitrary event $t_0$ and $\mathbf{x}_0$ and think of $\mathbf{g}$ as approximately constant in a sufficiently small region around this point. Using a similar coordinate transformation formula as above, we can find a reference frame $t'$ and $\mathbf{x}'$ in which the gravitational field is approximately transformed away in our small region. We call such reference frames *local inertial frames* (at $t_0$ and $\mathbf{x}_0$), because in these frames all trajectories of particles experiencing only the gravitational force will be approximately straight lines near the point $t_0$ and $\mathbf{x}_0$. The possibility that we can find local inertial frames is called the *equivalence principle*.

## 8.3. Geometrization of Gravitation

We have seen in the previous section that Newtonian gravitational fields can approximately be transformed away locally in small regions by coordinate transformations to non-inertial reference frames. We should think of these transformations as "absorbing" the local effects of gravitation into our spacetime geometry. Einstein's deep idea was to "capture" or "absorb" *all* effects of gravitation in the geometry of spacetime.

The good thing about Einstein's idea to geometrize gravitation is that it can be applied to special relativity. In special relativity, we focused on the Minkowski spacetime $\mathbb{R}^4$ with line element $ds_M^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$. Force-free particles move on straight world-lines in Minkowski spacetime. In the presence of gravitation, the geometry of spacetime changes in the sense that the Minkowski line element is generalized to

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu \ , \tag{8.11}$$

where $g_{\mu\nu}(x)$ is the so-called metric tensor; it contains all information about the gravitational field. Particles under the influence of gravitation will be described by the "straightest world-lines" in our spacetime geometry given by (8.11), which are so-called geodesics. Hence, gravitation should not be regarded as an external force, but as a consequence of the geometry of spacetime.

The equivalence principle then demands that, for any spacetime point $x_0 \in \mathbb{R}^4$, we can find new spacetime coordinates $x'$, such that we have approximately

$$ds^2 \approx \eta_{\mu\nu}\, dx'^\mu\, dx'^\nu \tag{8.12}$$

in a small region around $x_0$. In such local inertial frames for spacetime, we thus have approximately transformed away the effects of gravitation and our spacetime looks locally like Minkowski spacetime. We shall give a precise statement of this fact later in this module.

## 8.4. Energy-Momentum Tensor as Source Term

It is natural to ask what is the "correct" equation of motion for the metric field $g_{\mu\nu}$? Recalling Newton's law of gravitation (8.1), we see that gravitation should be sourced by a mass density. We have seen that in special relativity mass is equivalent to energy ($E = m\,c^2$) and that energy mixes under Lorentz transformations with momentum $\mathbf{p}$. A natural candidate for a source term for the metric field $g_{\mu\nu}$ is thus given by an energy-momentum tensor $T^{\mu\nu}$ of some matter model. The relativistic generalization of (8.1) should then be of the form

$$G^{\mu\nu} = \kappa\,T^{\mu\nu}\,, \tag{8.13}$$

where $G^{\mu\nu}$ is a suitable tensor constructed out of the metric $g_{\mu\nu}$ and $\kappa$ is some numerical constant. In this module, we will learn how to construct $G^{\mu\nu}$ and how to determine $\kappa$ such that 1.) (8.13) is form-invariant under general coordinate transformations and 2.) in the non-relativistic limit (8.13) is approximately Newton's law of gravitation (8.1).

CHAPTER 9

# Elementary Differential Geometry

In this chapter we give an elementary introduction to those aspects of differential geometry that are required for formalizing Einstein's theory of general relativity. Throughout the rest of this module, we shall make the simplifying assumption that there exists a globally defined coordinate system on spacetime, i.e. that every event $x$ can be characterized uniquely by four numbers $x^\mu$, i.e. $x^0$, $x^1$, $x^2$ and $x^3$. There exist examples where one needs more than one coordinate system (that are only locally defined) to describe spacetime, which leads to the mathematical concept of manifolds. All our differential geometric techniques which we introduce in this chapter can be generalized to manifolds; this is the subject of the more advanced module *MATH4015 Differential Geometry*. Hence, if you like the elementary ideas of differential geometry presented in this chapter, you should consider taking the module on differential geometry in your future studies.

REMARK 9.1. In this chapter we restrict ourselves to $4$-dimensional spacetimes because these are the ones required in general relativity. However, it is important to emphasize that all differential geometric concepts and techniques which we develop make obviously sense in any spacetime dimension $m \geq 2$; we just have to change the range of our indices $\mu, \nu, \rho, \dots$ from $0, 1, 2, 3$ to $0, 1, \dots, m-1$. We will see examples of lower dimensional spacetimes in the problem sheets. △

## 9.1. General Coordinates and Coordinate Transformations

Similarly to special relativity, we describe physical events in terms of points of a $4$-dimensional spacetime $M$. Spacetime in general relativity is not necessarily all of $\mathbb{R}^4$, but it also could just be an open subset

$$M \subseteq \mathbb{R}^4 . \tag{9.1}$$

(More generally, spacetime could even be a $4$-dimensional manifold, but as mentioned above we will not consider this more general class of spacetimes in this module.) A spacetime point $x \in M$ can be described by a choice of four coordinates $x^\mu$, with $\mu = 0, 1, 2, 3$.

The crucial difference between special relativity and general relativity is that there is no concept of inertial frames for the latter. This means that all observers have to be treated on the same footing. Mathematically, this amounts to the statement that any choice of coordinates $x^\mu$ is equally good in general relativity and that its laws have to be formulated in such a way that they are form-invariant under general coordinate transformations.

A *general coordinate transformation* from a set of coordinates $x^\mu$ to another set of coordinates $x'^\mu$ is a (smooth and) invertible map

$$x^\mu \longmapsto x'^\mu = x'^\mu(x) \ , \tag{9.2}$$

or more explicitly $x'^\mu(x) = x'^\mu(x^0, x^1, x^2, x^3)$. Here invertible means that there exists another (smooth) map

$$x'^\mu \longmapsto x^\mu = x^\mu(x') \ , \tag{9.3}$$

such that

$$x^\mu(x'(x)) = x^\mu \quad , \qquad x'^\mu(x(x')) = x'^\mu \ . \tag{9.4}$$

Associated to any general coordinate transformation is its *Jacobi matrix*

$$J^\mu{}_\nu(x) = \frac{\partial x'^\mu}{\partial x^\nu} \ , \tag{9.5}$$

which is of course a function depending on $x$. Because any general coordinate transformation is invertible, the Jacobi matrix is invertible (at every point $x \in M$) and its inverse is given by

$$J^{-1\mu}{}_\nu(x) = \frac{\partial x^\mu}{\partial x'^\nu} \ . \tag{9.6}$$

In fact, this statement follows from the chain rule for partial differentiation

$$J^{-1\mu}{}_\nu(x)\ J^\nu{}_\rho(x) = \frac{\partial x^\mu}{\partial x'^\nu}\,\frac{\partial x'^\nu}{\partial x^\rho} = \frac{\partial x^\mu}{\partial x^\rho} = \delta^\mu_\rho \ , \tag{9.7}$$

where

$$\delta^\mu_\rho = \begin{cases} 1 & , \ \text{if } \mu = \rho \ , \\ 0 & , \ \text{else} \ , \end{cases} \tag{9.8}$$

is the *Kronecker-delta*. Similarly, one finds that $J^\mu{}_\nu(x)\ J^{-1\nu}{}_\rho(x) = \delta^\mu_\rho$.

EXAMPLE 9.2 (Cylindrical coordinates). Let $M = \mathbb{R}^4$ and consider cylindrical coordinates $x^0 = ct$, $x^1 = r$, $x^2 = \varphi$ and $x^3 = z$ around the spatial $z$-axis. The coordinate transformation to Cartesian coordinates is given by

$$\begin{aligned} x'^0 &= x^0 = ct \ , \\ x'^1 &= r\ \cos\varphi \ , \\ x'^2 &= r\ \sin\varphi \ , \\ x'^3 &= x^3 = z \ . \end{aligned} \tag{9.9}$$

The associated Jacobi matrix reads as

$$J^\mu{}_\nu(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & -r\ \sin\varphi & 0 \\ 0 & \sin\varphi & r\ \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ . \tag{9.10}$$

▽

EXAMPLE 9.3 (Spherical coordinates). Let $M = \mathbb{R}^4$ and consider spherical coordinates $x^0 = ct$, $x^1 = r$, $x^2 = \theta$ and $x^3 = \varphi$. The coordinate transformation to Cartesian coordinates is given by

$$\begin{aligned} x'^0 &= x^0 = c\,t\;, \\ x'^1 &= r\,\cos\varphi\,\sin\theta\;, \\ x'^2 &= r\,\sin\varphi\,\sin\theta\;, \\ x'^3 &= r\,\cos\theta\;. \end{aligned} \tag{9.11}$$

The associated Jacobi matrix reads as

$$J^\mu{}_\nu(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi\,\sin\theta & r\,\cos\varphi\,\cos\theta & -r\,\sin\varphi\,\sin\theta \\ 0 & \sin\varphi\,\sin\theta & r\,\sin\varphi\,\cos\theta & r\,\cos\varphi\,\sin\theta \\ 0 & \cos\theta & -r\,\sin\theta & 0 \end{pmatrix}\;. \tag{9.12}$$

▽

## 9.2. Tangent and Cotangent Spaces

Let us consider parametrized world-lines

$$x^\mu : \mathbb{R} \longrightarrow M\;,\quad \lambda \longmapsto x^\mu(\lambda) \tag{9.13}$$

in our spacetime $M$. Let us fix any point $x_0 \in M$ and assume that our world-line satisfies

$$x^\mu(0) = x_0^\mu\;, \tag{9.14}$$

i.e. it passes through our point at parameter $\lambda = 0$. We take the $\lambda$-derivative at $\lambda = 0$ and obtain a 4-vector

$$\frac{dx^\mu(\lambda)}{d\lambda}\Big|_{\lambda=0} \in T_{x_0}M = \mathbb{R}^4 \tag{9.15}$$

that describes the *tangent vector* of our world-line at $x_0$. The vector space $T_{x_0}M$ we introduced in (9.15) is called the *tangent space* at $x_0$: Its elements are by definition the tangent vectors of all possible world-lines passing through $x_0$.

REMARK 9.4. The mathematically inclined reader may appreciate the following more abstract, but precise, definition of $T_{x_0}M$. Let us denote by $C^\infty_{x_0}(\mathbb{R}, M)$ the set of all smooth curves (world-lines) $x^\mu : \mathbb{R} \to M$ that satisfy $x^\mu(0) = x_0^\mu$. We introduce the following equivalence relation on this set

$$x^\mu \sim \widetilde{x}^\mu \quad :\Longleftrightarrow \quad \frac{dx^\mu(\lambda)}{d\lambda}\Big|_{\lambda=0} = \frac{d\widetilde{x}^\mu(\lambda)}{d\lambda}\Big|_{\lambda=0}\;, \tag{9.16}$$

i.e. two curves passing through $x_0 \in M$ are equivalent if and only if their tangents at $x_0$ coincide. The tangent space at $x_0 \in M$ is then defined as the quotient

$$T_{x_0}M := C^\infty_{x_0}(\mathbb{R}, M)/\sim\;. \tag{9.17}$$

It is a good exercise to explore how one can introduce a natural vector space structure on (9.17). △

Given any tangent vector $t^\mu \in T_{x_0}M$, which by definition may be represented by a world-line $x^\mu(\lambda)$ satisfying

$$t^\mu = \frac{dx^\mu(\lambda)}{d\lambda}\Big|_{\lambda=0} , \tag{9.18}$$

it is interesting to ask how $t^\mu$ transforms under general coordinate transformations $x^\mu \mapsto x'^\mu(x)$. Using the chain rule for differentiation, we compute

$$t'^\mu = \frac{dx'^\mu(\lambda)}{d\lambda}\Big|_{\lambda=0} = \frac{\partial x'^\mu}{\partial x^\nu}\,\frac{dx^\nu(\lambda)}{d\lambda}\Big|_{\lambda=0} = \frac{\partial x'^\mu}{\partial x^\nu}(x_0)\, t^\nu \, . \tag{9.19}$$

Hence, tangent vectors $t^\mu \in T_{x_0}M$ transform under general coordinate transformations according to

$$t'^\mu = J^\mu{}_\nu(x_0)\, t^\nu \, , \tag{9.20}$$

where the Jacobi matrix $J^\mu{}_\nu(x_0)$ is evaluated at our base-point $x_0$.

The dual notion of a tangent vector is called a *cotangent vector*. We define the cotangent space at $x_0 \in M$ as the dual vector space

$$T^*_{x_0}M := (T_{x_0}M)^* = \mathbb{R}^{4*} \simeq \mathbb{R}^4 \, . \tag{9.21}$$

Elements in $T^*_{x_0}M$ are denoted by symbols like $\omega_\mu \in T^*_{x_0}M$. As we shall clarify below, it is crucial to denote cotangent vectors $\omega_\mu$ by a lower index and not by an upper index $t^\mu$ as for tangent vectors. Being an element of the dual vector space means that each $\omega_\mu \in T^*_{x_0}M$ defines a linear map

$$\omega_\mu : T_{x_0}M \longrightarrow \mathbb{R} \, , \quad t^\mu \longmapsto \omega_\mu\, t^\mu \, , \tag{9.22}$$

where (according to Einstein's summation convention) the index $\mu$ is summed over $0, 1, 2, 3$.

Cotangent vectors transform differently under general coordinate transformations than tangent vectors; this is the reason why we have to distinguish between lower and upper indices on $\omega_\mu$ and respectively $t^\mu$. The transformation property of cotangent vectors can be derived from (9.22). Because real numbers do not transform under coordinate transformations, we have the equality

$$\omega'_\mu\, t'^\mu = \omega_\mu\, t^\mu \, , \tag{9.23}$$

for all cotangent vectors $\omega_\mu \in T^*_{x_0}M$ and all tangent vectors $t^\mu \in T_{x_0}M$. Inserting our transformation formula (9.20) for tangent vectors into this equation, we obtain

$$\omega'_\mu\, t'^\mu = \omega'_\mu\, J^\mu{}_\nu(x_0)\, t^\nu = \omega_\nu\, t^\nu \, , \tag{9.24}$$

from which it follows that

$$\omega_\nu = \omega'_\mu\, J^\mu{}_\nu(x_0) \, , \tag{9.25}$$

or equivalently

$$\omega'_\mu = {J^{-1}}^\nu{}_\mu(x_0)\, \omega_\nu \tag{9.26}$$

if we invert the Jacobi matrix.

In summary, the crucial point of this section is the following: For any spacetime point $x_0 \in M$, there exists a notion of tangent and cotangent vectors. These are elements $t^\mu \in T_{x_0}M$ and $\omega_\mu \in T^*_{x_0}M$ of 4-dimensional vector spaces. We have to use a different index notation to distinguish between tangent and cotangent vectors.

(Explicitly: Tangent vectors have upper indices and cotangent vectors have lower indices.) This is because they have different transformation rules under general coordinate transformations

$$t'^{\mu} = J^{\mu}{}_{\nu}(x_0)\, t^{\nu} \quad , \qquad \omega'_{\mu} = J^{-1\nu}{}_{\mu}(x_0)\, \omega_{\nu} \ , \tag{9.27}$$

which are determined by either the Jacobi or by the inverse Jacobi matrix at $x_0$.

## 9.3. Tensor Fields

The concept of *fields* on a spacetime $M$ is fundamental in general relativity. (Warning: Do *not* mix up fields on spacetime with the notion of field you learned in your algebra modules!) Loosely speaking, a field $A$ on $M$ is an assignment

$$x \longmapsto A(x) = \text{“some physical quantity at } x\text{”} \tag{9.28}$$

of a physical quantity $A(x)$ to each spacetime point $x \in M$. There are different kinds of fields, depending on which kinds of physical quantities they assign.

The simplest kind of field is a *scalar field* $\Phi$: It is an assignment

$$x \longmapsto \Phi(x) \in \mathbb{R} \tag{9.29}$$

of a real number $\Phi(x)$ to each spacetime point $x$. In other words, a scalar field is a real-valued function $\Phi : M \to \mathbb{R}$ on $M$. Under general coordinate transformations $x^{\mu} \mapsto x'^{\mu}(x)$, scalar fields transform as

$$\Phi'(x') = \Phi(x) \ , \tag{9.30}$$

for each $x \in M$.

The next-to-simplest field is a *vector field*: It is an assignment

$$x \longmapsto A^{\mu}(x) \in T_x M \tag{9.31}$$

of a tangent vector $A^{\mu}(x)$ at $x$ to each spacetime point $x$. Using our transformation rule for tangent vectors (9.20), vector fields transform under general coordinate transformations as

$$A'^{\mu}(x') = J^{\mu}{}_{\nu}(x)\ A^{\nu}(x) \ , \tag{9.32}$$

i.e. with the Jacobi matrix $J^{\mu}{}_{\nu}(x)$ at the point $x$ corresponding to $A^{\nu}(x)$.

The dual notion of a vector field is called a *covector field*: It is an assignment

$$x \longmapsto B_{\mu}(x) \in T^*_x M \tag{9.33}$$

of a cotangent vector $B_{\mu}(x)$ at $x$ to each spacetime point $x$. Using our transformation rule for cotangent vectors (9.26), covector fields transform under general coordinate transformations as

$$B'_{\mu}(x') = J^{-1\nu}{}_{\mu}(x)\ B_{\nu}(x) \ , \tag{9.34}$$

i.e. with the inverse Jacobi matrix $J^{-1\nu}{}_{\mu}(x)$ at the point $x$ corresponding to $B_{\nu}(x)$.

An important feature of fields is that they can be combined to produce new fields (possibly of a different kind). Let us for example consider a vector field $A^{\mu}$ and a covector field $B_{\mu}$. We can define a scalar field $\Phi$ by setting

$$\Phi(x) := B_{\mu}(x)\, A^{\mu}(x) \ , \tag{9.35}$$

for each spacetime point $x$. (Recall our summation convention! The index $\mu$ is summed over $0, 1, 2, 3$.) In fact, it is easy to see that $\Phi$ has the correct transformation law under general coordinate transformations

$$\begin{aligned}\Phi'(x') &= B'_\mu(x')\, A'^\mu(x') = {J^{-1}}^\nu{}_\mu(x)\, B_\nu(x)\, J^\mu{}_\rho(x)\, A^\rho(x) \\ &= B_\nu(x)\, \underbrace{{J^{-1}}^\nu{}_\mu(x)\, J^\mu{}_\rho(x)}_{=\delta^\nu_\rho}\, A^\rho(x) = B_\mu(x)\, A^\mu(x)\ ,\end{aligned} \tag{9.36}$$

where we used that ${J^{-1}}^\nu{}_\mu(x)$ is the inverse of $J^\mu{}_\rho(x)$. The procedure (9.35) is also called *index contraction*, because it takes the sum over one upper and one lower index to produce a quantity without indices.

Another way to produce new fields out of old ones is to take their product (more precisely, tensor product). Given for example two vector fields $A^\mu$ and $\widetilde{A}^\mu$, we can consider the quantity

$$C^{\mu\nu}(x) := A^\mu(x)\, \widetilde{A}^\nu(x)\ , \tag{9.37}$$

which has two independent upper indices. Under a general coordinate transformation, the field $C^{\mu\nu}$ transforms as

$$\begin{aligned}C'^{\mu\nu}(x') &= A'^\mu(x')\, \widetilde{A}'^\nu(x') = J^\mu{}_\rho(x)\, A^\rho(x)\, J^\nu{}_\sigma(x)\, \widetilde{A}^\sigma(x) \\ &= J^\mu{}_\rho(x)\, J^\nu{}_\sigma(x)\, C^{\rho\sigma}(x)\ ,\end{aligned} \tag{9.38}$$

i.e. with one Jacobi matrix for each index. Similarly, given a vector field $A^\mu$ and a covector field $B_\nu$ we can consider the quantity

$$C_\mu^\nu(x) := B_\mu(x)\, A^\nu(x)\ , \tag{9.39}$$

which has two independent indices, one upper and one lower. (Notice that there is no sum here: The indices $\mu$ and $\nu$ are different!) Under a general coordinate transformation, the field $C_\mu^\nu$ transforms as

$$\begin{aligned}C'^\nu_\mu(x') &= B'_\mu(x')\, A'^\nu(x') = {J^{-1}}^\rho{}_\mu(x)\, B_\rho(x)\, J^\nu{}_\sigma(x)\, A^\sigma(x) \\ &= {J^{-1}}^\rho{}_\mu(x)\, J^\nu{}_\sigma(x)\, C_\rho^\sigma(x)\ ,\end{aligned} \tag{9.40}$$

i.e. with one inverse Jacobi matrix for the lower index and one Jacobi matrix for the upper index. Notice that our index contraction in (9.35) can be obtained by setting $\mu$ and $\nu$ equal in $C_\mu^\nu$ (and summing over it according to our summation convention), i.e.

$$C_\mu^\mu(x) = B_\mu(x)\, A^\mu(x) = \Phi(x) \tag{9.41}$$

is a scalar field.

We take these examples as motivation to define the concept of $(p, q)$*-tensor fields*, where $p$ and $q$ are non-negative integers. A $(p, q)$-tensor field is an assignment

$$x \longmapsto A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(x) \in T_xM^{\otimes p} \otimes T_x^*M^{\otimes q} \tag{9.42}$$

of a tensor product of $p$ tangent vectors and $q$ cotangent vectors at $x$ to each spacetime point $x$. It thus has $p$ upper indices and $q$ lower indices. Using our transformation formulas for tangent vectors (9.20) and cotangent vectors (9.26), $(p, q)$-tensor fields transform under general coordinate transformations as

$$A'^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(x') = J^{\mu_1}{}_{\rho_1}(x) \,\cdots\, J^{\mu_p}{}_{\rho_p}(x) \; J^{-1\sigma_1}{}_{\nu_1}(x) \,\cdots\, J^{-1\sigma_q}{}_{\nu_q}(x) \; A^{\rho_1\cdots\rho_p}_{\sigma_1\cdots\sigma_q}(x) \;, \tag{9.43}$$

i.e. with one Jacobi matrix for each upper index and one inverse Jacobi matrix for each lower index. Notice that

- $(0, 0)$-tensor fields are scalar fields;
- $(1, 0)$-tensor fields are vector fields; and
- $(0, 1)$-tensor fields are covector fields.

Moreover, (9.37) is a $(2, 0)$-tensor field and (9.39) is a $(1, 1)$-tensor field.

Index contractions and products make perfectly sense for $(p, q)$-tensor fields: Let $A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}$ be a $(p, q)$-tensor field. For any $i \in \{1, 2, \dots, p\}$ and any $j \in \{1, 2, \dots, q\}$, we consider

$$A^{\mu_1\cdots\mu_{i-1}\rho\mu_{i+1}\cdots\mu_p}_{\nu_1\cdots\nu_{j-1}\rho\nu_{j+1}\cdots\nu_q}(x) \;, \tag{9.44}$$

i.e. we contract the upper index $\mu_i$ with the lower index $\nu_j$. (Again, summation over $\rho$ is understood!) It is a good exercise to check that (9.44) defines a $(p-1, q-1)$-tensor field, i.e. it has the right transformation law under general coordinate transformations. Concerning products, given any $(p, q)$-tensor field $A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}$ and any $(l, m)$-tensor field $B^{\rho_1\cdots\rho_l}_{\sigma_1\cdots\sigma_m}$, then

$$A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(x) \; B^{\rho_1\cdots\rho_l}_{\sigma_1\cdots\sigma_m}(x) \tag{9.45}$$

defines a $(p + l, q + m)$-tensor field. Again, it is a good exercise to check explicitly the transformation law of (9.45).

## 9.4. Line Elements, Metric Tensors and Light Cones

Einstein's idea of geometrization of gravitation amounts to replacing the line element $ds_M^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$ of Minkowski spacetime by a more general line element

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu \;. \tag{9.46}$$

Notice that the coefficients $g_{\mu\nu}(x)$ are symmetric, i.e. $g_{\mu\nu}(x) = g_{\nu\mu}(x)$, and that they are allowed to depend on the spacetime point $x$. Physically, $g_{\mu\nu}(x)$ encodes the gravitational field in terms of the geometry of spacetime.

Let us recall from (4.4) the transformation rule for the infinitesimal coordinate changes $dx^\mu$ under general coordinate transformations; explicitly,

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}\, dx^\nu = J^\mu{}_\nu(x)\, dx^\nu \;. \tag{9.47}$$

This allows us to derive a transformation formula for $g_{\mu\nu}$ from the requirement that the line element does not change under general coordinate transformations, i.e.

$$ds'^2 = g'_{\mu\nu}(x')\, dx'^\mu\, dx'^\nu \stackrel{!}{=} ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu \;. \tag{9.48}$$

Let us do the relevant calculation

$$g'_{\mu\nu}(x')\, dx'^\mu\, dx'^\nu = g'_{\mu\nu}(x)\; J^\mu{}_\rho(x)\; J^\nu{}_\sigma(x)\; dx^\rho\, dx^\sigma \stackrel{!}{=} g_{\mu\nu}(x)\, dx^\mu\, dx^\nu \;, \tag{9.49}$$

which results in the transformation formula

$$g'_{\mu\nu}(x') = J^{-1\rho}{}_{\mu}(x)\; J^{-1\sigma}{}_{\nu}(x)\; g_{\rho\sigma}(x)\;. \tag{9.50}$$

Hence, we find that $g_{\mu\nu}$ is a symmetric $(0,2)$-tensor field, according to our theory developed in the previous section.

Not every symmetric $(0,2)$-tensor field $g_{\mu\nu}$ will describe a reasonable spacetime geometry, but there are further conditions on $g_{\mu\nu}$ that have to be satisfied. The equivalence principle demands that, for each spacetime point $x_0 \in M$, we must be able to find a choice of coordinates such that $g_{\mu\nu}(x_0)$ (at this point and in these coordinates) is equal to the Minkowski metric $\eta_{\mu\nu}$. This is true provided that $g_{\mu\nu}$ satisfies the following

ASSUMPTION 9.5. For all spacetime points $x_0 \in M$, the symmetric real-valued matrix $g_{\mu\nu}(x_0)$ has $3$ positive eigenvalues and $1$ negative eigenvalue.

REMARK 9.6. The $3$ positive eigenvalues correspond to space directions and the negative eigenvalue corresponds to a time direction. This is in full analogy to the Minkowski metric, which has one negative and three positive eigenvalues. △

Let us now show that, under Assumption 9.5, we can find for each point $x_0 \in M$ a choice of coordinates in which $g_{\mu\nu}(x_0)$ is the Minkowski metric. Recall from linear algebra that every symmetric matrix $g_{\mu\nu}(x_0)$ can be diagonalized by an orthogonal matrix $T^{\mu}{}_{\nu}$, i.e.

$$T^{\mu}{}_{\rho}\, T^{\nu}{}_{\sigma}\, g_{\mu\nu}(x_0) = \begin{pmatrix} \lambda_- & 0 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_2 & 0\\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}\;. \tag{9.51}$$

Due to our assumption, $\lambda_- < 0$ is negative and $\lambda_1, \lambda_2, \lambda_3 > 0$ are positive. We hence can introduce the diagonal matrix

$$S^{\mu}{}_{\nu} = \begin{pmatrix} \sqrt{-\lambda_-}^{-1} & 0 & 0 & 0\\ 0 & \sqrt{\lambda_1}^{-1} & 0 & 0\\ 0 & 0 & \sqrt{\lambda_2}^{-1} & 0\\ 0 & 0 & 0 & \sqrt{\lambda_3}^{-1} \end{pmatrix}\;, \tag{9.52}$$

and find that

$$(T\,S)^{\mu}{}_{\alpha}\,(T\,S)^{\nu}{}_{\beta}\, g_{\mu\nu}(x_0) = S^{\rho}{}_{\alpha}\, S^{\sigma}{}_{\beta}\, T^{\mu}{}_{\rho}\, T^{\nu}{}_{\sigma}\, g_{\mu\nu}(x_0) = \eta_{\alpha\beta}\;. \tag{9.53}$$

Consider now the general coordinate transformation

$$x'^{\mu} = (T\,S)^{-1\,\mu}{}_{\nu}\, x^{\nu}\;, \tag{9.54}$$

whose Jacobi matrix is

$$J^{\mu}{}_{\nu}(x) = (T\,S)^{-1\,\mu}{}_{\nu}\;. \tag{9.55}$$

Then $g_{\mu\nu}$ reads in $x'$-coordinates as

$$g'_{\alpha\beta}(x') = {J^{-1}}^{\mu}{}_{\alpha}(x)\, {J^{-1}}^{\nu}{}_{\beta}(x)\, g_{\mu\nu}(x) = (T\,S)^{\mu}{}_{\alpha}\, (T\,S)^{\nu}{}_{\beta}\, g_{\mu\nu}(x)\,. \tag{9.56}$$

It follows from (9.53) that this is precisely the Minkowski metric at the point $x_0 \in M$, which completes our proof.

A symmetric $(0,2)$-tensor field $g_{\mu\nu}$ satisfying Assumption 9.5 is the fundamental concept in general relativity. Hence, it deserves a name.

DEFINITION 9.7. A *Lorentzian metric* on a spacetime $M$ is a symmetric $(0,2)$-tensor field $g_{\mu\nu}$ that satisfies Assumption 9.5.

A Lorentzian metric $g_{\mu\nu}$ on a spacetime $M$ allows us to decompose the tangent space $T_{x_0}M$ at any point $x_0 \in M$ into time-like, light-like and space-like tangent vectors. This is in complete analogy to the case of the Minkowski spacetime, see the text below Eqn. (3.12). We say that a non-zero tangent vector $0 \neq t^\mu \in T_{x_0}M$ is

- *time-like* if $g_{\mu\nu}(x_0)\, t^\mu\, t^\nu < 0$;
- *light-like* if $g_{\mu\nu}(x_0)\, t^\mu\, t^\nu = 0$; and
- *space-like* if $g_{\mu\nu}(x_0)\, t^\mu\, t^\nu > 0$.

As a straightforward generalization of this, we say that a parametrized world-line $x^\mu : \mathbb{R} \to M\,,\ \lambda \mapsto x^\mu(\lambda)$ in $M$ is

- *time-like* if $g_{\mu\nu}(x(\lambda))\, \frac{dx^\mu(\lambda)}{d\lambda}\, \frac{dx^\nu(\lambda)}{d\lambda} < 0$, for all $\lambda$;
- *light-like* if $g_{\mu\nu}(x(\lambda))\, \frac{dx^\mu(\lambda)}{d\lambda}\, \frac{dx^\nu(\lambda)}{d\lambda} = 0$, for all $\lambda$; and
- *space-like* if $g_{\mu\nu}(x(\lambda))\, \frac{dx^\mu(\lambda)}{d\lambda}\, \frac{dx^\nu(\lambda)}{d\lambda} > 0$, for all $\lambda$.

Again similar to the Minkowski spacetime, the light-like tangent vectors in $T_{x_0}M$ define the light cone. As the Lorentzian metric $g_{\mu\nu}(x)$ is allowed to depend on the spacetime point $x$, the form of the light cone of a general Lorentzian metric also depends on $x$.

EXAMPLE 9.8 (Friedmann-Robertson-Walker spacetime). Let $M = \mathbb{R}^4$ be all of $\mathbb{R}^4$ and choose Cartesian coordinates $x^\mu$. (We write $x^0 = c\,t$ and interpret $t$ as a time coordinate.) Let us consider the Lorentzian metric on $M$ which is defined by the line element

$$ds^2 = -c^2\,(dt)^2 + a(t)^2\,\Big((dx^1)^2 + (dx^2)^2 + (dx^3)^2\Big)\,, \tag{9.57}$$

where $a(t) > 0$ is a positive function of only the time coordinate. The associated metric tensor reads in our choice of coordinates as

$$g_{\mu\nu}(x) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a(t)^2 & 0 & 0 \\ 0 & 0 & a(t)^2 & 0 \\ 0 & 0 & 0 & a(t)^2 \end{pmatrix}\,. \tag{9.58}$$

This spacetime is called the *Friedmann-Robertson-Walker spacetime*. It is used in cosmology to describe the time evolution of our universe. The function $a(t)$ is called the *scale factor* of the universe, because it scales the space distances measured with respect to our coordinate frame. When $a(t)$ grows in time, the universe is expanding. Notice that for constant scale factor $a(t) = 1$, the Friedmann-Robertson-Walker spacetime becomes the Minkowski spacetime. $\triangledown$

### 9.5. Raising and Lowering Indices

Let $M \subseteq \mathbb{R}^4$ be a spacetime with Lorentzian metric $g_{\mu\nu}$. Because of Assumption 9.5, the metric tensor is point-wise invertible. We denote its inverse by $g^{\mu\nu}$, call it the *inverse metric* and note that it is defined by the equations

$$g^{\mu\nu}(x)\, g_{\nu\rho}(x) = \delta^\mu_\rho = g_{\rho\nu}(x)\, g^{\nu\mu}(x)\ . \tag{9.59}$$

As a consequence of (9.59), the inverse metric is a $(2, 0)$-tensor field, i.e. it transforms according to

$$g'^{\mu\nu}(x') = J^\mu{}_\rho(x)\, J^\nu{}_\sigma(x)\, g^{\rho\sigma}(x) \tag{9.60}$$

under general coordinate transformations.

Using the Lorentzian metric $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$, we can raise and lower indices of $(p, q)$-tensor fields. For example, given a vector field $A^\mu$ we define a covector field $A_\mu$ by setting

$$A_\mu(x) = g_{\mu\nu}(x)\, A^\nu(x)\ . \tag{9.61}$$

Given a covector field $B_\mu$ we define a vector field $B^\mu$ by setting

$$B^\mu(x) = g^{\mu\nu}(x)\, B_\nu(x)\ . \tag{9.62}$$

These tensor fields have indeed the correct transformation behavior under general coordinate transformations, because we have seen before that index contractions (9.44) and products (9.45) of tensor fields produce again tensor fields (of a different type). The definitions above are consistent: Lowering and then raising the index of a vector field $A^\mu$ gives

$$A^\mu(x) = g^{\mu\nu}(x)\, A_\nu(x) = \underbrace{g^{\mu\nu}(x)\, g_{\nu\rho}(x)}_{=\delta^\mu_\rho}\, A^\rho(x) = A^\mu(x)\ , \tag{9.63}$$

while raising and then lowering the index of a covector field $B_\mu$ gives

$$B_\mu(x) = g_{\mu\nu}(x)\, B^\nu(x) = \underbrace{g_{\mu\nu}(x)\, g^{\nu\rho}(x)}_{=\delta^\rho_\mu}\, B_\rho(x) = B_\mu(x)\ . \tag{9.64}$$

The same constructions of course apply to generic $(p, q)$-tensor fields: Given for example a $(2, 1)$-tensor field $C^{\mu\nu}_\rho$, we can raise the index $\rho$ to obtain a $(3, 0)$-tensor field

$$C^{\mu\nu\sigma}(x) = g^{\sigma\rho}(x)\, C^{\mu\nu}_\rho(x)\ , \tag{9.65}$$

or we can lower the indices $\mu$ and $\nu$ to obtain a $(0, 3)$-tensor field

$$C_{\alpha\beta\rho}(x) = g_{\alpha\mu}(x)\, g_{\beta\nu}(x)\, C^{\mu\nu}_\rho(x)\ . \tag{9.66}$$

Of course, we can also just lower $\mu$ or $\nu$ to obtain a $(1, 2)$-tensor field.

Raising and lowering indices can be used to define point-wise (indefinite) inner products for vector fields and covector fields: Given two vector fields $A^\mu$ and $\widetilde{A}^\nu$, we define

$$\langle A, \widetilde{A}\rangle(x) := g_{\mu\nu}(x)\, A^\mu(x)\, \widetilde{A}^\nu(x) = A_\mu(x)\, \widetilde{A}^\mu(x) = A^\mu(x)\, \widetilde{A}_\mu(x)\ . \tag{9.67}$$

Similarly, given two covector fields $B_\mu$ and $\widetilde{B}_\nu$, we define

$$\langle B, \widetilde{B}\rangle(x) := g^{\mu\nu}(x)\, B_\mu(x)\, \widetilde{B}_\nu(x) = B^\mu(x)\, \widetilde{B}_\mu(x) = B_\mu(x)\, \widetilde{B}^\mu(x)\ . \tag{9.68}$$

Notice that we also used here that $g_{\mu\nu}$ and $g^{\mu\nu}$ are symmetric.

## 9.6. Covariant Derivatives

To formulate dynamical laws involving $(p, q)$-tensor fields, it is unavoidable to take their partial derivatives. For example, in the special case of the Minkowski spacetime, we have seen that the equation of motion for an energy-momentum tensor $T^{\mu\nu}$ is given by the conservation law $\frac{\partial}{\partial x^\mu} T^{\mu\nu}(x) = 0$ involving a 4-divergence.

Let $M$ be any spacetime with Lorentzian metric $g_{\mu\nu}$. For a scalar (i.e. $(0,0)$-tensor) field $\Phi$, we observe that the partial derivative

$$\partial_\mu \Phi(x) := \frac{\partial}{\partial x^\mu} \Phi(x) \tag{9.69}$$

is a $(0,1)$-tensor field. (In the following we shall often use the short notation $\partial_\mu = \frac{\partial}{\partial x^\mu}$.) In fact, given a general coordinate transformation $x^\mu \mapsto x'^\mu(x)$, the expression in (9.69) transforms as

$$\partial'_\mu \Phi'(x') = \frac{\partial}{\partial x'^\mu}\, \Phi'(x') = \frac{\partial x^\nu}{\partial x'^\mu}\, \frac{\partial}{\partial x^\nu} \Phi(x) = {J^{-1}}^\nu{}_\mu(x)\, \partial_\nu \Phi(x)\ , \tag{9.70}$$

which is the correct way for a $(0,1)$-tensor field. For later use, we note from this calculation that the partial differential operator $\partial_\mu$ transforms as

$$\partial'_\mu = {J^{-1}}^\nu{}_\mu(x)\, \partial_\nu \tag{9.71}$$

under general coordinate transformations.

Unfortunately, this result does *not* generalize to generic $(p,q)$-tensor fields. To see where the problem comes from, let us consider a vector (i.e. $(1,0)$-tensor) field $A^\mu$ and form its partial derivative

$$\partial_\nu A^\mu(x)\ . \tag{9.72}$$

The expression in (9.72) does *not* transform like a $(1,1)$-tensor field, but rather like

$$\begin{aligned} \partial'_\nu A'^\mu(x') &= {J^{-1}}^\rho{}_\nu(x)\, \partial_\rho\Big(J^\mu{}_\sigma(x)\, A^\sigma(x)\Big) \\ &= {J^{-1}}^\rho{}_\nu(x)\, J^\mu{}_\sigma(x)\, \partial_\rho A^\sigma(x) + {J^{-1}}^\rho{}_\nu(x)\, \Big(\partial_\rho J^\mu{}_\sigma(x)\Big)\, A^\sigma(x)\ , \end{aligned} \tag{9.73}$$

where we have used the Leibniz rule for partial derivatives. While the first term is the desired transformation rule of $(1,1)$-tensor field, the second term is problematic and it spoils the transformation rule of the expression (9.72).

The problem with the bad transformation property of partial derivatives of vector fields can be solved by introducing the concept of *covariant derivatives*. By definition, a covariant derivative of vector fields $A^\mu$ is given by

$$\nabla_\nu A^\mu(x) := \partial_\nu A^\mu(x) + \Gamma^\mu_{\nu\rho}(x)\, A^\rho(x) \,, \tag{9.74}$$

where the components $\Gamma^\mu_{\nu\rho}(x)$ are called *Christoffel symbols*. The Christoffel symbols are *not* $(1,2)$-tensor fields, but we demand that they transform under general coordinate transformations as

$$\begin{aligned} \Gamma'^\mu_{\ \nu\rho}(x') = J^\mu{}_\sigma(x)\; {J^{-1}}^\alpha{}_\nu(x)\; {J^{-1}}^\beta{}_\rho(x)\, \Gamma^\sigma_{\alpha\beta}(x) \\ - {J^{-1}}^\alpha{}_\nu(x)\; {J^{-1}}^\beta{}_\rho(x) \left(\partial_\alpha J^\mu{}_\beta(x)\right) . \end{aligned} \tag{9.75}$$

The motivation for this transformation formula for the Christoffel symbols is that it ensures that (9.74) transforms as a $(1,1)$-tensor field. Let us confirm this statement by a straightforward, however slightly lengthy, calculation. (We will suppress here and also in the following the arguments $x$ and $x'$ for the benefit of a shorter notation.) Explicitly,

$$\begin{aligned} \nabla'_\nu A'^\mu &= \partial'_\nu A'^\mu + \Gamma'^\mu_{\ \nu\rho}\, A'^\rho \\ &\overset{(9.73)}{=} {J^{-1}}^\rho{}_\nu\, J^\mu{}_\sigma\, \partial_\rho A^\sigma + {J^{-1}}^\rho{}_\nu \left(\partial_\rho J^\mu{}_\sigma\right) A^\sigma + J^\rho{}_\sigma\, \Gamma'^\mu_{\ \nu\rho}\, A^\sigma \\ &= {J^{-1}}^\rho{}_\nu\, J^\mu{}_\sigma\, \nabla_\rho A^\sigma - \left({J^{-1}}^\rho{}_\nu\, J^\mu{}_\beta\, \Gamma^\beta_{\rho\sigma} - {J^{-1}}^\rho{}_\nu \left(\partial_\rho J^\mu{}_\sigma\right) - J^\rho{}_\sigma\, \Gamma'^\mu_{\ \nu\rho}\right) A^\sigma \\ &\overset{(9.75)}{=} {J^{-1}}^\rho{}_\nu\, J^\mu{}_\sigma\, \nabla_\rho A^\sigma \,. \end{aligned} \tag{9.76}$$

Hence, by construction, the covariant derivative (9.74) of a vector field is a $(1,1)$-tensor field.

Before showing how to determine the Christoffel symbols from the Lorentzian metric $g_{\mu\nu}$, we study how covariant derivatives act on arbitrary $(p,q)$-tensor fields. As the covariant derivative on scalar (i.e. $(0,0)$-tensor) fields $\Phi$ we can take the partial derivative, because we have seen above that

$$\nabla_\mu \Phi := \partial_\mu \Phi \tag{9.77}$$

already transforms as a $(0,1)$-tensor. We next define the covariant derivative for covector fields $B_\mu$ implicitly by the requirement

$$\nabla_\mu\left(B_\nu\, A^\nu\right) = \left(\nabla_\mu B_\nu\right) A^\nu + B_\nu \left(\nabla_\mu A^\nu\right) , \tag{9.78}$$

for all vector fields $A^\nu$ and covector fields $B_\nu$, which resembles a Leibniz rule for covariant differentiation. (Notice that the left-hand side involves just a scalar field, hence the covariant derivative agrees with the partial derivative.) Bringing the second term on the right-hand side to the left-hand side, we obtain

$$\begin{aligned} \left(\nabla_\mu B_\nu\right) A^\nu &= \partial_\mu\left(B_\nu\, A^\nu\right) - B_\nu \left(\nabla_\mu A^\nu\right) \\ &= \left(\partial_\mu B_\nu\right) A^\nu + B_\nu \left(\partial_\mu A^\nu\right) - B_\nu \left(\partial_\mu A^\nu\right) - B_\nu\, \Gamma^\nu_{\mu\rho}\, A^\rho \\ &= \left(\partial_\mu B_\nu - \Gamma^\rho_{\mu\nu}\, B_\rho\right) A^\nu \,, \end{aligned} \tag{9.79}$$

where in the second step we used the Leibniz rule for the partial derivative and inserted (9.74). In the third step we relabeled the indices. Because Eqn. (9.79) holds true for all vector fields $A^\nu$, we can read off the definition of covariant derivative for covector fields

$$\nabla_\mu B_\nu := \partial_\mu B_\nu - \Gamma^\rho_{\mu\nu}\, B_\rho \ . \tag{9.80}$$

Notice the different sign in front the Christoffel symbols in the covariant derivative for vector fields (9.74) and respectively for covector fields (9.80). The covariant derivative on arbitrary $(p,q)$-tensor fields $A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}$ can again be obtained from a Leibniz rule requirement. We will not discuss its derivation in detail, but just mention that the final result is given by the formula

$$\begin{aligned}\nabla_\mu A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q} = \partial_\mu A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q} + \Gamma^{\mu_1}_{\mu\rho}\, A^{\rho\cdots\mu_p}_{\nu_1\cdots\nu_q} + \cdots + \Gamma^{\mu_p}_{\mu\rho}\, A^{\mu_1\cdots\rho}_{\nu_1\cdots\nu_q}\\ - \Gamma^{\rho}_{\mu\nu_1}\, A^{\mu_1\cdots\mu_p}_{\rho\cdots\nu_q} - \cdots - \Gamma^{\rho}_{\mu\nu_q}\, A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\rho}\ ,\end{aligned} \tag{9.81}$$

where every upper index is treated like a vector field (9.74) and every lower index like a covector field (9.80). It is then straightforward to verify that the covariant derivative of $(p,q)$-tensor fields satisfies the Leibniz rule

$$\nabla_\mu\Big(A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}\, B^{\rho_1\cdots\rho_l}_{\sigma_1\cdots\sigma_m}\Big) = \Big(\nabla_\mu A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}\Big)\, B^{\rho_1\cdots\rho_l}_{\sigma_1\cdots\sigma_m} + A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}\,\Big(\nabla_\mu B^{\rho_1\cdots\rho_l}_{\sigma_1\cdots\sigma_m}\Big)\ , \tag{9.82}$$

for all $(p,q)$-tensor fields $A^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}$ and all $(l,m)$-tensor fields $B^{\rho_1\cdots\rho_l}_{\sigma_1\cdots\sigma_m}$.

EXAMPLE 9.9 (Covariant derivative of the metric tensor). An important special case is the covariant derivative of the Lorentzian metric $g_{\mu\nu}$. From our general formula (9.81), it follows that

$$\nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\sigma_{\mu\nu}\, g_{\sigma\rho} - \Gamma^\sigma_{\mu\rho}\, g_{\nu\sigma}\ , \tag{9.83}$$

where the minus signs are due to the fact that the metric has lower indices. ▽

We now come to the important point of proving that, under the assumptions stated below, the Christoffel symbols are uniquely determined by the Lorentzian metric $g_{\mu\nu}$ on the spacetime $M$.

ASSUMPTION 9.10. On a spacetime $M$ with Lorentzian metric $g_{\mu\nu}$, we can demand that the covariant derivative satisfies the following conditions.

(i) *Metric compatibility:* The covariant derivative of the metric tensor is zero, i.e. $\nabla_\mu g_{\nu\rho} = 0$.

(ii) *Symmetry:* The Christoffel symbols are symmetric, i.e. $\Gamma^\mu_{\nu\rho} = \Gamma^\mu_{\rho\nu}$.

A covariant derivative satisfying (i) and (ii) is called *Levi-Civita connection* in the mathematical literature.

REMARK 9.11. Condition (ii) has a geometric meaning. Using the transformation rule (9.75) of the Christoffel symbols under general coordinate transformations, one finds that their antisymmetric combination

$$\mathrm{Tor}^\mu_{\nu\rho} := \Gamma^\mu_{\nu\rho} - \Gamma^\mu_{\rho\nu} \tag{9.84}$$

is a $(1,2)$-tensor field. This tensor field is called the *torsion tensor* and our condition (ii) simply demands the covariant derivative to be torsion-free, i.e. $\mathrm{Tor}^\mu_{\nu\rho} = 0$. △

PROPOSITION 9.12 (Existence and uniqueness of the Levi-Civita connection). *Let $M$ be a spacetime with Lorentzian metric $g_{\mu\nu}$. There exists a unique covariant derivative satisfying the two conditions (i) and (ii) stated in Assumption 9.10. It is defined by the Christoffel symbols*

$$\Gamma^{\rho}_{\mu\nu} = \frac{1}{2}\, g^{\rho\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right) . \tag{9.85}$$

PROOF. Let us first assume that there exists a covariant derivative satisfying the conditions in Assumption 9.10 and show that it is necessarily unique. From condition (i), i.e. $\nabla_\mu g_{\nu\rho} = 0$, it follows that also the combination

$$\nabla_\mu g_{\nu\rho} + \nabla_\nu g_{\rho\mu} - \nabla_\rho g_{\mu\nu} = 0 \tag{9.86}$$

vanishes. Inserting three times (9.83), we find

$$\begin{aligned}
0 &= \partial_\mu g_{\nu\rho} - \Gamma^\sigma_{\mu\nu}\, g_{\sigma\rho} - \Gamma^\sigma_{\mu\rho}\, g_{\nu\sigma} \\
&\quad + \partial_\nu g_{\rho\mu} - \Gamma^\sigma_{\nu\rho}\, g_{\sigma\mu} - \Gamma^\sigma_{\nu\mu}\, g_{\rho\sigma} \\
&\quad - \partial_\rho g_{\mu\nu} + \Gamma^\sigma_{\rho\mu}\, g_{\sigma\nu} + \Gamma^\sigma_{\rho\nu}\, g_{\mu\sigma} \\
&= \partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu} - \left(\Gamma^\sigma_{\mu\nu} + \Gamma^\sigma_{\nu\mu}\right) g_{\sigma\rho} \\
&\quad - \left(\Gamma^\sigma_{\mu\rho} - \Gamma^\sigma_{\rho\mu}\right) g_{\nu\sigma} - \left(\Gamma^\sigma_{\nu\rho} - \Gamma^\sigma_{\rho\nu}\right) g_{\sigma\mu} \\
&= \partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu} - 2\, \Gamma^\sigma_{\mu\nu}\, g_{\sigma\rho} .
\end{aligned} \tag{9.87}$$

In the second equality we used that the metric tensor is symmetric, i.e. $g_{\mu\nu} = g_{\nu\mu}$, and in the third equality we used condition (ii) saying that the Christoffel symbols are symmetric. Multiplying this formula with the inverse of the metric $g_{\sigma\rho}$ we obtain our desired formula (9.85). This shows that a covariant derivative satisfying the conditions in Assumption 9.10 is uniquely specified by the formula (9.85). Moreover, because (9.85) defines a covariant derivative satisfying the desired conditions in Assumption 9.10, the Levi-Civita connection always exists. □

REMARK 9.13. It is a good, however slightly lengthy, exercise to confirm by an explicit calculation that (9.85) transforms under general coordinate transformations as required in Eqn. (9.75). △

In the following, we always use the covariant derivative defined by (9.85).

EXAMPLE 9.14. In any inertial frame for the Minkowski spacetime, i.e. $M = \mathbb{R}^4$ with Lorentzian metric $\eta_{\mu\nu}$, the Christoffel symbols (9.85) vanish because $\eta_{\mu\nu}$ is constant. In particular, in any inertial frame for the Minkowski spacetime all covariant derivatives coincide with partial derivatives. Due to our transformation rule (9.75), this is *not* the case for non-inertial frames for the Minkowski spacetime, where the Christoffel symbols may be non-trivial. We shall later see examples of this. ▽

Let us conclude this section by showing that our covariant derivative defined by (9.85) is compatible with raising and lowering the indices. First of all, we notice that from $\nabla_\mu g_{\nu\rho} = 0$ it follows that

$$\nabla_\mu g^{\nu\rho} = 0 \tag{9.88}$$

for the inverse metric. In fact, taking the covariant derivative of the defining equation $g^{\mu\nu}\, g_{\nu\rho} = \delta^\mu_\rho$, we obtain by using the Leibniz rule (9.82)

$$\begin{aligned}\nabla_\sigma\Big(g^{\mu\nu}\, g_{\nu\rho}\Big) &= \Big(\nabla_\sigma g^{\mu\nu}\Big)\, g_{\nu\rho} + g^{\mu\nu}\,\Big(\nabla_\sigma g_{\nu\rho}\Big) = \Big(\nabla_\sigma g^{\mu\nu}\Big)\, g_{\nu\rho} \\ &\overset{!}{=} \nabla_\sigma \delta^\mu_\rho = \partial_\sigma \delta^\mu_\rho + \Gamma^\mu_{\sigma\nu}\,\delta^\nu_\rho - \Gamma^\nu_{\sigma\rho}\,\delta^\mu_\nu = \Gamma^\mu_{\sigma\rho} - \Gamma^\mu_{\sigma\rho} = 0\,. \end{aligned} \tag{9.89}$$

Then (9.88) follows by multiplying this formula with the inverse of the metric $g_{\nu\rho}$. Let us consider a $(1,0)$-tensor field $A^\mu$ and lower its index via $A_\mu = g_{\mu\nu}\, A^\nu$ to obtain a $(0,1)$-tensor field. We obtain for the covariant derivative

$$\nabla_\rho A_\mu = \nabla_\rho\Big(g_{\mu\nu}\, A^\nu\Big) = \underbrace{\Big(\nabla_\rho\, g_{\mu\nu}\Big)}_{=0}\, A^\nu + g_{\mu\nu}\,\Big(\nabla_\rho A^\nu\Big) = g_{\mu\nu}\,\Big(\nabla_\rho A^\nu\Big)\,, \tag{9.90}$$

i.e. it does not matter if we first lower the index and then take the covariant derivative or, the other way around, first take the covariant derivative and then lower the index. Similarly, let us consider a $(0,1)$-tensor field $B_\mu$ and raise its index via $B^\mu = g^{\mu\nu}\, B_\nu$ to obtain a $(1,0)$-tensor field. We obtain for the covariant derivative

$$\nabla_\rho B^\mu = \nabla_\rho\Big(g^{\mu\nu}\, B_\nu\Big) = \underbrace{\Big(\nabla_\rho\, g^{\mu\nu}\Big)}_{=0}\, B_\nu + g^{\mu\nu}\Big(\nabla_\rho B_\nu\Big) = g^{\mu\nu}\Big(\nabla_\rho B_\nu\Big)\,, \tag{9.91}$$

i.e. it does not matter if we first raise the index and then take the covariant derivative or, the other way around, first take the covariant derivative and then raise the index.

## 9.7. Riemann Curvature Tensor

Recall that by Schwarz's theorem the order in which we take the 2nd partial derivatives does not matter, i.e.

$$\partial_\mu \partial_\nu \Phi = \partial_\nu \partial_\mu \Phi\,, \tag{9.92}$$

for all scalar fields $\Phi$. This statement is no longer true for covariant derivatives: The 2nd covariant derivative of, say, a covector field $B_\rho$ is given by

$$\begin{aligned}\nabla_\mu\nabla_\nu B_\rho &= \partial_\mu\Big(\nabla_\nu B_\rho\Big) - \Gamma^\sigma_{\mu\nu}\,\nabla_\sigma B_\rho - \Gamma^\sigma_{\mu\rho}\,\nabla_\nu B_\sigma \\ &= \partial_\mu\Big(\partial_\nu B_\rho - \Gamma^\beta_{\nu\rho}\, B_\beta\Big) - \Gamma^\sigma_{\mu\nu}\,\Big(\partial_\sigma B_\rho - \Gamma^\beta_{\sigma\rho} B_\beta\Big) - \Gamma^\sigma_{\mu\rho}\,\Big(\partial_\nu B_\sigma - \Gamma^\beta_{\nu\sigma} B_\beta\Big) \\ &= \partial_\mu\partial_\nu B_\rho - \Big(\partial_\mu \Gamma^\beta_{\nu\rho}\Big)\, B_\beta - \Gamma^\beta_{\nu\rho}\,\partial_\mu B_\beta - \Gamma^\sigma_{\mu\nu}\,\partial_\sigma B_\rho + \Gamma^\sigma_{\mu\nu}\,\Gamma^\beta_{\sigma\rho} B_\beta \\ &\quad - \Gamma^\sigma_{\mu\rho}\,\partial_\nu B_\sigma + \Gamma^\sigma_{\mu\rho}\,\Gamma^\beta_{\nu\sigma} B_\beta\,. \end{aligned} \tag{9.93}$$

Using again Schwarz's theorem for partial derivatives and also symmetry of the Christoffel symbols, $\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu}$, we find that the difference between $\nabla_\mu\nabla_\nu B_\rho$ and $\nabla_\nu\nabla_\mu B_\rho$ is

$$\nabla_\mu\nabla_\nu B_\rho - \nabla_\nu\nabla_\mu B_\rho = \Big(\partial_\nu \Gamma^\beta_{\mu\rho} - \partial_\mu \Gamma^\beta_{\nu\rho} + \Gamma^\sigma_{\mu\rho}\,\Gamma^\beta_{\nu\sigma} - \Gamma^\sigma_{\nu\rho}\,\Gamma^\beta_{\mu\sigma}\Big)\, B_\beta\,. \tag{9.94}$$

Notice that there is no derivative term on the covector field $B_\beta$ in (9.94). Because the left-hand side of (9.94) is a $(0,3)$-tensor field and the right-hand side involves the $(0,1)$-tensor field $B_\beta$ without derivatives, the quantity $R^\beta_{\rho\nu\mu}$ defined by

$$\nabla_\mu\nabla_\nu B_\rho - \nabla_\nu\nabla_\mu B_\rho =: R^\beta_{\rho\nu\mu}\, B_\beta\,, \tag{9.95a}$$

or explicitly by

$$R^{\beta}_{\rho\nu\mu} := \partial_\nu \Gamma^{\beta}_{\mu\rho} - \partial_\mu \Gamma^{\beta}_{\nu\rho} + \Gamma^{\sigma}_{\mu\rho}\, \Gamma^{\beta}_{\nu\sigma} - \Gamma^{\sigma}_{\nu\rho}\, \Gamma^{\beta}_{\mu\sigma} \;, \tag{9.95b}$$

is a $(1,3)$-tensor field. This tensor field is called the *Riemann curvature tensor*. Recalling our explicit form of the Christoffel symbols (9.85), we note that the Riemann curvature tensor is a non-linear function of the metric tensor $g_{\mu\nu}$ (and its inverse $g^{\mu\nu}$), the partial derivatives $\partial_\rho g_{\mu\nu}$ (and $\partial_\rho g^{\mu\nu}$) and the second partial derivatives $\partial_\sigma \partial_\rho g_{\mu\nu}$.

The Riemann curvature tensor satisfies various properties which we shall now discuss. First of all, from (9.95) we immediately see that it is antisymmetric in the indices $\mu, \nu$, i.e.

$$R^{\beta}_{\rho\nu\mu} = -R^{\beta}_{\rho\mu\nu} \;. \tag{9.96}$$

Lowering the index $\beta$ according to $R_{\alpha\rho\nu\mu} := g_{\alpha\beta}\, R^{\beta}_{\rho\nu\mu}$, one can show that $R_{\alpha\rho\nu\mu}$ is antisymmetric in the indices $\alpha, \rho$, i.e.

$$R_{\alpha\rho\nu\mu} = -R_{\rho\alpha\nu\mu} \;, \tag{9.97}$$

and symmetric under exchange of index pairs $\alpha\rho$ and $\nu\mu$, i.e.

$$R_{\alpha\rho\nu\mu} = R_{\nu\mu\alpha\rho} \;. \tag{9.98}$$

The quickest way of proving these statements is by using the concept of *local inertial frames*, which we introduce in the next chapter. We will hence provide this proof later.

Two further important properties of the Riemann curvature tensor are the so-called *Bianchi identities*: The *first Bianchi identity* states that

$$R^{\beta}_{\rho\nu\mu} + R^{\beta}_{\nu\mu\rho} + R^{\beta}_{\mu\rho\nu} = 0 \;, \tag{9.99}$$

where we performed cyclic permutations of the lower indices $\rho\nu\mu$. The *second Bianchi identity* states that

$$\nabla_\sigma R^{\beta}_{\rho\nu\mu} + \nabla_\nu R^{\beta}_{\rho\mu\sigma} + \nabla_\mu R^{\beta}_{\rho\sigma\nu} = 0 \;, \tag{9.100}$$

where we performed cyclic permutations of the lower indices $\sigma\nu\mu$. Again, the quickest way to prove the Bianchi identities is to use local inertial frames, so we will provide this proof later.

EXAMPLE 9.15. The Riemann curvature tensor $R^{\beta}_{\rho\nu\mu}$ vanishes for the Minkowski spacetime. This follows from the fact that we may choose an inertial frame such that the Lorentzian metric is given by $\eta_{\mu\nu}$ and hence is constant. In this frame the Christoffel symbols (9.85) are zero and by (9.95) this implies that the Riemann curvature tensor vanishes. In contrast to the Christoffel symbols in Example 9.14, the Riemann curvature tensor vanishes for all choices of coordinate frames, because it is a tensor. $\triangledown$

CHAPTER 10

# Physics in Curved Spacetimes

## 10.1. Local Inertial Frames

Let $M$ be a spacetime with Lorentzian metric $g_{\mu\nu}$. We have already seen in (9.53) that, for any spacetime point $x_0 \in M$, we can find new coordinates such that the metric at this point is given by the Minkowski metric $g'_{\mu\nu}(x'_0) = \eta_{\mu\nu}$.

We shall now strengthen this statement as follows: For any point $x_0 \in M$, we can find new coordinates such that

$$g'_{\mu\nu}(x'_0) = \eta_{\mu\nu} \quad , \qquad \partial'_\rho \, g'_{\mu\nu}(x'_0) = 0 \, . \tag{10.1}$$

Such a coordinate frame is called a *local inertial frame* (at the point $x_0 \in M$). Notice that because of (9.85), the Christoffel symbols at $x_0$ in a local inertial frame are zero,

$$\Gamma'^{\rho}_{\mu\nu}(x'_0) = 0 \, . \tag{10.2}$$

Hence, in a local inertial frame the covariant derivative of any $(p, q)$-tensor field (9.81) reduces to the partial derivative

$$\nabla'_\mu A'^{\mu_1 \cdots \mu_p}_{\nu_1 \cdots \nu_q}(x'_0) = \partial'_\mu A'^{\mu_1 \cdots \mu_p}_{\nu_1 \cdots \nu_p}(x'_0) \tag{10.3}$$

at the point $x_0 \in M$.

Let us now show that local inertial frames always exist. Without loss of generality, we may assume that $x_0^\mu = 0$ is the origin of our coordinate system $x^\mu$ and that $g_{\mu\nu}$ already satisfies $g_{\mu\nu}(0) = \eta_{\mu\nu}$. Choose new coordinates $x'^\mu$ by any general coordinate transformation whose Taylor expansion around $0$ starts with

$$x'^\mu = x^\mu + \frac{1}{2} \, \Gamma^\mu_{\nu\rho}(0) \, x^\nu \, x^\rho + \cdots \, , \tag{10.4}$$

where the prefactor in front of the quadratic term is given by our Christoffel symbols (9.85) evaluated at our distinguished point $0$. The Jacobi matrix of this general coordinate transformation is

$$J^\mu{}_\nu(x) = \frac{\partial x'^\mu}{\partial x^\nu} = \delta^\mu_\nu + \Gamma^\mu_{\nu\rho}(0) \, x^\rho + \cdots \, , \tag{10.5}$$

where we used that the Christoffel symbols are symmetric in their lower indices. The metric tensor in our new coordinates is then given by

$$\begin{aligned} g'_{\mu\nu}(x') &= J^{-1\alpha}{}_\mu(x) \, J^{-1\beta}{}_\nu(x) \, g_{\alpha\beta}(x) \\ &= \left(\delta^\alpha_\mu - \Gamma^\alpha_{\mu\rho}(0) \, x'^\rho + \cdots\right) \left(\delta^\beta_\nu - \Gamma^\beta_{\nu\sigma}(0) \, x'^\sigma + \cdots\right) g_{\alpha\beta}(x) \, . \end{aligned} \tag{10.6}$$

Notice that our point $x_0^\mu = 0$ is transformed via (10.4) to $x'_0{}^\mu = 0$. Thus, we obtain

$$g'_{\mu\nu}(0) = \delta^\alpha_\mu \, \delta^\beta_\nu \, g_{\alpha\beta}(0) = \eta_{\mu\nu} \, , \tag{10.7}$$

where we used that by hypothesis $g_{\mu\nu}(0) = \eta_{\mu\nu}$, and

$$\partial'_\rho g'_{\mu\nu}(0) = \partial'_\rho g_{\mu\nu}(0) - \Gamma^\alpha_{\mu\rho}(0)\, g_{\alpha\nu}(0) - \Gamma^\alpha_{\nu\rho}(0)\, g_{\mu\alpha}(0)\ . \tag{10.8}$$

The first term of (10.8) can be simplified further as

$$\partial'_\rho g_{\mu\nu}(0) = \frac{\partial x^\alpha}{\partial x'^\rho}(0)\, \partial_\alpha g_{\mu\nu}(0) = \delta^\alpha_\rho\, \partial_\alpha g_{\mu\nu}(0) = \partial_\rho g_{\mu\nu}(0)\ . \tag{10.9}$$

Using this and the definition of Christoffel symbols (9.85), we find that (10.8) vanishes,

$$\begin{aligned}\partial'_\rho g'_{\mu\nu}(0) &= \partial_\rho g_{\mu\nu}(0) - \frac{1}{2}\Big(\partial_\mu g_{\nu\rho}(0) + \partial_\rho g_{\nu\mu}(0) - \partial_\nu g_{\mu\rho}(0)\Big) \\ &\quad - \frac{1}{2}\Big(\partial_\nu g_{\mu\rho}(0) + \partial_\rho g_{\mu\nu}(0) - \partial_\mu g_{\nu\rho}(0)\Big) \\ &= \partial_\rho g_{\mu\nu}(0) - \partial_\rho g_{\mu\nu}(0) = 0\ . \end{aligned} \tag{10.10}$$

Hence, our new coordinates given by (10.4) define a local inertial frame.

Summing up, we have shown that for any spacetime $M$ with Lorentzian metric $g_{\mu\nu}$ there exists a local inertial frame (10.1) for any chosen point $x_0 \in M$. In particular, in this coordinate system the metric tensor has a Taylor expansion of the form

$$g'_{\mu\nu}(x') = \eta_{\mu\nu} + d_{\mu\nu\rho\sigma}\,(x'^\rho - x_0'^{\,\rho})\,(x'^\sigma - x_0'^{\,\sigma}) + \cdots\ , \tag{10.11}$$

meaning that it is (up to quadratic corrections controlled by $d_{\mu\nu\rho\sigma}$) approximately the Minkowski metric in a small region around $x'_0$.

The physical relevance of local inertial frames is that they motivate a simple rule (or recipe) how to generalize physical laws from special relativity to general relativity. This is called *minimal coupling* or the *substitution rule* and will be discussed in the next section. Mathematically, local inertial frames are very useful for proving identities of tensor field equations, such as the identities for the Riemann curvature tensor (9.97), (9.98) and the Bianchi identities (9.99), (9.100) which we still have to prove.

**Proof of the identities of the Riemann curvature tensor:** The identities (9.97), (9.98), (9.99) and (9.100) for the Riemann curvature tensor are identities of tensor fields. In particular, they can be verified point-wise and if they hold true in one coordinate system they will hold true in all coordinate systems. Hence, we can use local inertial frames to verify our identities.

Let the coordinates $x^\mu$ be a local inertial frame at $x_0 \in M$, i.e.

$$g_{\mu\nu}(x_0) = \eta_{\mu\nu} \quad , \qquad \partial_\rho g_{\mu\nu}(x_0) = 0\ . \tag{10.12}$$

It follows that $\Gamma^\mu_{\nu\rho}(x_0) = 0$ and hence we find for the Riemann curvature tensor (9.95) at $x_0$

$$R^\beta_{\rho\nu\mu}(x_0) = \partial_\nu \Gamma^\beta_{\mu\rho}(x_0) - \partial_\mu \Gamma^\beta_{\nu\rho}(x_0)\ . \tag{10.13}$$

Using (9.85), $g^{\mu\nu}(x_0) = \eta^{\mu\nu}$ and $\partial_\rho g^{\mu\nu}(x_0) = 0$, we can simplify this further and obtain

$$R^\beta_{\rho\nu\mu}(x_0) = \frac{\eta^{\beta\alpha}}{2}\Big(\partial_\nu\partial_\rho g_{\alpha\mu}(x_0) + \partial_\mu\partial_\alpha g_{\nu\rho}(x_0) - \partial_\nu\partial_\alpha g_{\mu\rho}(x_0) - \partial_\mu\partial_\rho g_{\alpha\nu}(x_0)\Big)\,. \tag{10.14}$$

Lowering the index $\beta$ then gives

$$R_{\alpha\rho\nu\mu}(x_0) = \frac{1}{2}\Big(\partial_\nu\partial_\rho g_{\alpha\mu}(x_0) + \partial_\mu\partial_\alpha g_{\nu\rho}(x_0) - \partial_\nu\partial_\alpha g_{\mu\rho}(x_0) - \partial_\mu\partial_\rho g_{\alpha\nu}(x_0)\Big)\,. \tag{10.15}$$

With these formulas it is easy to verify the identities (9.97) and (9.98) and also the first Bianchi identity (9.99). (That is a good exercise you should do!) Concerning the second Bianchi identity (9.100), we have to take the covariant derivative of the Riemann curvature tensor

$$\nabla_\sigma R^\beta_{\rho\nu\mu}\,. \tag{10.16}$$

In a local inertial frame at $x_0$, this expression simplifies to

$$\nabla_\sigma R^\beta_{\rho\nu\mu}(x_0) = \partial_\sigma\partial_\nu\Gamma^\beta_{\mu\rho}(x_0) - \partial_\sigma\partial_\mu\Gamma^\beta_{\nu\rho}(x_0)\,, \tag{10.17}$$

by using the Leibniz rule and $\Gamma^\mu_{\nu\rho}(x_0) = 0$. Hence,

$$\begin{aligned}
&\nabla_\sigma R^\beta_{\rho\nu\mu}(x_0) + \nabla_\nu R^\beta_{\rho\mu\sigma}(x_0) + \nabla_\mu R^\beta_{\rho\sigma\nu}(x_0)\\
&\quad = \partial_\sigma\partial_\nu\Gamma^\beta_{\mu\rho}(x_0) - \partial_\sigma\partial_\mu\Gamma^\beta_{\nu\rho}(x_0) + \partial_\nu\partial_\mu\Gamma^\beta_{\sigma\rho}(x_0) - \partial_\nu\partial_\sigma\Gamma^\beta_{\mu\rho}(x_0)\\
&\qquad + \partial_\mu\partial_\sigma\Gamma^\beta_{\nu\rho}(x_0) - \partial_\mu\partial_\nu\Gamma^\beta_{\sigma\rho}(x_0)\\
&\quad = 0\,,
\end{aligned} \tag{10.18}$$

where in the last step we used that the order of partial differentiation can be exchanged. This proves the second Bianchi identity.

## 10.2. Minimal Coupling/Substitution Rule

Let us assume that we are given a physical law of special relativity that is formulated as a tensor equation on the Minkowski spacetime ($g_{\mu\nu} = \eta_{\mu\nu}$) in any inertial frame for spacetime. A general relativistic generalization of this physical law can be obtained by the following prescription:

1.) Replace all Minkowski metrics $\eta_{\mu\nu}$ (or their inverses $\eta^{\mu\nu}$) by an arbitrary Lorentzian metric $g_{\mu\nu}(x)$ (or its inverse $g^{\mu\nu}(x)$).

2.) Replace all partial derivatives $\partial_\mu$ by covariant derivatives $\nabla_\mu$.

This prescription is called *minimal coupling* or the *substitution rule*.

A few important remarks are in order: First, notice that the physical laws obtained by the above prescription are automatically form-invariant under general coordinate transformations, because of the transformation properties of tensor fields and covariant derivatives. Second, if the Lorentzian metric $g_{\mu\nu}(x)$ is curved, i.e. there is a non-trivial gravitational force, we obtain a truly new physical law that describes the coupling of our special relativistic physical system to gravitation. Third, the minimal coupling prescription is compatible with the equivalence principle: In any local inertial frame, the new general relativistic law approximately reduces to the special relativistic law we started with, because locally the metric tensor is the Minkowski metric and the Christoffel symbols vanish, i.e. the covariant derivative reduces to the partial derivative.

A brief warning at this point: The minimal coupling prescription is unfortunately not unique. For example, we could add to a special relativistic tensor equation a term such as $\partial_\mu \partial_\nu B_\rho - \partial_\nu \partial_\mu B_\rho$, which is zero because partial derivatives commute with each other. Applying the substitution rule gives a term of the form $\nabla_\mu \nabla_\nu B_\rho - \nabla_\nu \nabla_\mu B_\rho$, which is in general not zero. (Notice that this term is proportional to the Riemann curvature tensor (9.95).) Hence, there are some ambiguities in applying the substitution rule. In practice, these ambiguities may be fixed by heuristic simplicity principles ("We do not want to add more terms than necessary.") or (much better!) by confronting the general relativistic laws with experiments that will decide whether such additional terms are present or not.

We will now apply the minimal coupling prescription to special relativistic point-particles (cf. Chapter 6) and energy-momentum tensors. This will lead to a general relativistic generalization of these physical theories which describes their coupling to gravitational fields.

## 10.3. General Relativistic Mechanics

Recall that the equation of motion of a massive special relativistic point-particle is given by

$$m_0 \, \frac{du^\mu(\tau)}{d\tau} = F^\mu \ , \tag{10.19}$$

where $m_0$ is the rest mass, $\tau$ is proper time defined by

$$\eta_{\mu\nu} \, dx^\mu \, dx^\nu = -c^2 \, (d\tau)^2 \ , \tag{10.20}$$

$u^\mu$ is the $4$-velocity

$$u^\mu(\tau) = \frac{dx^\mu(\tau)}{d\tau} \ , \tag{10.21}$$

and $F^\mu$ is the $4$-force. The $4$-force has to satisfy the orthogonality condition

$$\eta_{\mu\nu} \, u^\mu \, F^\nu = 0 \ , \tag{10.22}$$

and the $4$-velocity is normalized according to

$$\eta_{\mu\nu} \, u^\mu \, u^\nu = -c^2 \ . \tag{10.23}$$

The goal of this section is to generalize these equations to an arbitrary spacetime $M$ with Lorentzian metric $g_{\mu\nu}(x)$ using the minimal coupling prescription.

Let $M$ be a spacetime with Lorentzian metric $g_{\mu\nu}(x)$ and consider a time-like world-line

$$x^\mu : \mathbb{R} \longrightarrow M \ , \quad \lambda \longmapsto x^\mu(\lambda) \ , \tag{10.24}$$

where $\lambda$ is for the moment an arbitrary parametrization. Hence, by hypothesis, the tangent vector

$$\frac{dx^\mu(\lambda)}{d\lambda} \in T_{x(\lambda)} M \tag{10.25}$$

is time-like for all $\lambda$, i.e.

$$g_{\mu\nu}(x(\lambda))\,\frac{dx^\mu(\lambda)}{d\lambda}\,\frac{dx^\nu(\lambda)}{d\lambda} < 0\ ,\quad \forall\lambda\ . \tag{10.26}$$

It is always possible to find a parametrization by a new parameter $\tau$ such that the tangent vector satisfies the normalization condition

$$g_{\mu\nu}(x(\tau))\,\frac{dx^\mu(\tau)}{d\tau}\,\frac{dx^\nu(\tau)}{d\tau} = -c^2\ ,\quad \forall\tau\ . \tag{10.27}$$

Formally, we may also write this equation as

$$g_{\mu\nu}\,dx^\mu\,dx^\nu = -c^2\,(d\tau)^2\ , \tag{10.28}$$

i.e. our parameter $\tau$ plays the same role as proper time in special relativistic mechanics. Notice that (10.28) is obtained by applying our substitution rule to the corresponding Minkowski spacetime equation (10.20). From now on we shall always parametrize our world-line by proper time $\tau$ defined by (10.28). We call the corresponding tangent vector

$$u^\mu(\tau) := \frac{dx^\mu(\tau)}{d\tau} \in T_{x(\tau)}M \tag{10.29}$$

the $4$-velocity. By construction, it satisfies the normalization condition

$$g_{\mu\nu}(x(\tau))\,u^\mu(\tau)\,u^\nu(\tau) = -c^2 \tag{10.30}$$

which may also be obtained from the special relativistic condition (10.23) by our substitution rule.

Notice that the special relativistic equation of motion (10.19) involves a derivative of a tangent vector $u^\mu$ (i.e. the value of a vector field at a point). By our minimal coupling prescription, this derivative has to be replaced by a covariant derivative to obtain the general relativistic law. Notice however that the derivative $\frac{d}{d\tau}$ in (10.19) is *not* a partial derivative, but rather a parameter derivative. So the question arises how a covariant parameter derivative should look like? To obtain a formula for the *covariant parameter derivative* let us consider a $(1,0)$-tensor field $A^\mu(x)$ and evaluate it on the world-line $x(\tau)$, i.e. we consider $A^\mu(x(\tau))$. The ordinary parameter derivative then can be written as

$$\frac{dA^\mu(x(\tau))}{d\tau} = \frac{dx^\nu(\tau)}{d\tau}\,\partial_\nu A^\mu(x(\tau))\ , \tag{10.31}$$

where we used the chain rule for differentiation. To this expression we can apply our substitution rule and define the covariant parameter derivative as

$$\begin{aligned}\frac{\nabla A^\mu(x(\tau))}{d\tau} &:= \frac{dx^\nu(\tau)}{d\tau}\,\nabla_\nu A^\mu(x(\tau))\\ &= \frac{dx^\nu(\tau)}{d\tau}\Big(\partial_\nu A^\mu(x(\tau)) + \Gamma^\mu_{\nu\rho}(x(\tau))\,A^\rho(x(\tau))\Big)\\ &= \frac{dA^\mu(x(\tau))}{d\tau} + \Gamma^\mu_{\nu\rho}(x(\tau))\,\frac{dx^\nu(\tau)}{d\tau}\,A^\rho(x(\tau))\ .\end{aligned} \tag{10.32}$$

Applying this substitution rule to $\frac{du^\mu(\tau)}{d\tau}$ then results in

$$\frac{\nabla u^\mu(\tau)}{d\tau} = \frac{du^\mu(\tau)}{d\tau} + \Gamma^\mu_{\nu\rho}(x(\tau))\,u^\nu(\tau)\,u^\rho(\tau)\ . \tag{10.33}$$

The general relativistic generalization of the equation of motion (10.19) is therefore given by

$$m_0 \frac{\nabla u^\mu(\tau)}{d\tau} = F^\mu(x(\tau)) \,, \tag{10.34}$$

where $F^\mu(x)$ is a $(1,0)$-tensor field which we call 4-force. Defining the 4-momentum by

$$p^\mu(\tau) := m_0 \, u^\mu(\tau) \,, \tag{10.35}$$

the equation of motion (10.34) is equivalent to

$$\frac{\nabla p^\mu(\tau)}{d\tau} = F^\mu(x(\tau)) \,. \tag{10.36}$$

It is important to emphasize that the 4-force $F^\mu$ is used to describe *non-gravitational* external forces; gravitation, i.e. a curved metric tensor $g_{\mu\nu}$, is already included on the left-hand side of (10.34) in terms of the covariant derivative.

The orthogonality condition (10.22) generalizes to the general relativistic setting: Taking the parameter derivative $\frac{d}{d\tau}$ of the normalization condition (10.30), we obtain

$$\begin{aligned} 0 &= \frac{d}{d\tau}(-c^2) = \frac{d}{d\tau}\Big(g_{\mu\nu}(x(\tau))\, u^\mu(\tau)\, u^\nu(\tau)\Big) \\ &= \underbrace{\frac{\nabla g_{\mu\nu}(x(\tau))}{d\tau}}_{=0} u^\mu(\tau)\, u^\nu(\tau) + 2\, g_{\mu\nu}(x(\tau))\, u^\mu(\tau)\, \frac{\nabla u^\nu(\tau)}{d\tau} \\ &= \frac{2}{m_0}\, g_{\mu\nu}(x(\tau))\, u^\mu(\tau)\, F^\nu(x(\tau)) \,, \end{aligned} \tag{10.37}$$

where we used the Leibniz rule for covariant derivatives and the fact that the metric tensor satisfies $\nabla_\mu g_{\nu\rho} = 0$. Multiplying by $\frac{m_0}{2}$, this equation is equivalent to

$$g_{\mu\nu}(x(\tau))\, u^\mu(\tau)\, F^\nu(x(\tau)) = 0 \,, \tag{10.38}$$

which may also be obtained from (10.22) by our substitution rule.

Setting the external force $F^\mu = 0$ to zero allows us to study the propagation of particles subject to only gravitational fields. We obtain the equation of motion

$$\frac{\nabla u^\mu(\tau)}{d\tau} = 0 \,, \tag{10.39}$$

or more explicitly by using (10.33)

$$\frac{d^2 x^\mu(\tau)}{d\tau^2} + \Gamma^\mu_{\nu\rho}(x(\tau))\, \frac{dx^\nu(\tau)}{d\tau}\, \frac{dx^\rho(\tau)}{d\tau} = 0 \,. \tag{10.40}$$

In differential geometry, (10.40) is called the *geodesic equation*. It formalizes the condition that the tangent vectors $\frac{dx^\mu(\tau)}{d\tau}$ of the world-line $x^\mu(\tau)$ are transported parallel to themselves (with respect to the covariant derivative $\nabla_\mu$).

The geodesic equation (10.40) can also be obtained as the Euler-Lagrange equation corresponding to the action functional

$$S := \int ds := \int \sqrt{-g_{\mu\nu}(x(\lambda))\, \frac{dx^\mu(\lambda)}{d\lambda}\, \frac{dx^\nu(\lambda)}{d\lambda}}\, d\lambda \,, \tag{10.41}$$

for parametrized curves $x^\mu : \lambda \mapsto x^\mu(\lambda)$. (Here we use a general parametrization by $\lambda$ and not a parametrization by proper time.) Notice that this action functional determines the "length" of our world-line measured with respect to the Lorentzian metric $g_{\mu\nu}$. Hence, geodesics are those world-lines with extremal length.

Let us now work out the variation of the action functional (10.41) in order to confirm that its Euler-Lagrange equation is really the geodesic equation (10.40). To simplify notations for this calculation, we shall suppress all arguments and simply write

$$S = \int \sqrt{-g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}}\;d\lambda \tag{10.42}$$

for the action. By the principle of least action, we get

$$0 = \delta S = \int \delta\Big(\sqrt{-g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}}\Big)\,d\lambda = \int \frac{\delta\Big(-g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}\Big)}{2\sqrt{-g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}}}\,d\lambda \quad . \tag{10.43}$$

Hence, the extrema of our original action functional (10.41) are the same as the extrema of the alternative action functional

$$\widetilde{S} = \int g_{\mu\nu}(x(\lambda))\,\frac{dx^\mu(\lambda)}{d\lambda}\,\frac{dx^\nu(\lambda)}{d\lambda}\,d\lambda \quad , \tag{10.44}$$

which is simpler than (10.41) because it doesn't contain a square root. Let us therefore continue working with (10.44) instead of (10.41). We compute

$$\begin{aligned}
0 = \delta\widetilde{S} &= \int \delta\Big(g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}\Big)\,d\lambda\\
&= \int \Big((\delta g_{\mu\nu})\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda} + g_{\mu\nu}\,\frac{d\delta x^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda} + g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{d\delta x^\nu}{d\lambda}\Big)\,d\lambda\\
&= \int \Big(\partial_\alpha g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}\,\delta x^\alpha + 2g_{\mu\nu}\,\frac{d\delta x^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}\Big)\,d\lambda\\
&\overset{\text{P.I.}}{=} \int \Big(\partial_\alpha g_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\,\frac{dx^\nu}{d\lambda}\,\delta x^\alpha - 2\,\frac{d}{d\lambda}\Big(g_{\mu\nu}\,\frac{dx^\nu}{d\lambda}\Big)\,\delta x^\mu\Big)\,d\lambda\\
&= \int \Big(\partial_\mu g_{\alpha\nu}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\nu}{d\lambda} - 2\,\partial_\alpha g_{\mu\nu}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\nu}{d\lambda} - 2g_{\mu\nu}\,\frac{d^2x^\nu}{d\lambda^2}\Big)\,\delta x^\mu\,d\lambda\\
&= \int \Big(\partial_\mu g_{\alpha\nu}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\nu}{d\lambda} - \partial_\alpha g_{\mu\nu}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\nu}{d\lambda} - \partial_\nu g_{\mu\alpha}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\nu}{d\lambda} - 2g_{\mu\nu}\,\frac{d^2x^\nu}{d\lambda^2}\Big)\,\delta x^\mu\,d\lambda\\
&= \int -2g_{\mu\nu}\Big(\frac{d^2x^\nu}{d\lambda^2} + \Gamma^\nu_{\alpha\beta}\,\frac{dx^\alpha}{d\lambda}\,\frac{dx^\beta}{d\lambda}\Big)\,\delta x^\mu\,d\lambda \quad ,
\end{aligned} \tag{10.45}$$

where with "P.I." we indicated integration by parts. If we now choose proper time $\tau$ as a parametrization for the world-line, we recover the geodesic equation (10.40). Hence, the extrema of both action functionals (10.41) and (10.44) are determined by the geodesic equation.

## 10.4. General Relativistic Energy-Momentum Tensors

The concept of energy-momentum tensors we have introduced in Chapter 7 for the Minkowski spacetime immediately generalizes to arbitrary spacetimes $M$ with Lorentzian metrics $g_{\mu\nu}(x)$ by using our minimal coupling prescription/substitution rule. Recall from Chapter 7 that an energy-momentum tensor in an inertial frame on the Minkowski spacetime is a $(2, 0)$-tensor field $T^{\mu\nu}(x)$ that is symmetric, i.e. $T^{\mu\nu}(x) = T^{\nu\mu}(x)$, and satisfies the conservation law

$$\partial_\mu T^{\mu\nu}(x) = 0 \tag{10.46}$$

involving the partial derivative.

This generalizes immediately to arbitrary spacetimes $M$ with Lorentzian metrics $g_{\mu\nu}(x)$. A *general relativistic energy-momentum tensor* is a symmetric $(2, 0)$-tensor field $T^{\mu\nu}(x)$ on $M$ that satisfies the covariant conservation law

$$\nabla_\mu T^{\mu\nu}(x) = 0 \tag{10.47}$$

involving the covariant derivative. Physically, (10.47) describes the equation of motion of our matter model in presence of a gravitational field $g_{\mu\nu}(x)$.

As concrete examples, we find that the general relativistic energy-momentum tensor for dust (see (7.27) for the case of the Minkowski spacetime) is given by

$$T^{\mu\nu}_{\text{dust}}(x) = \rho(x)\, U^\mu(x)\, U^\nu(x) \ , \tag{10.48}$$

where $\rho$ is a scalar (i.e. $(0, 0)$-tensor) field describing the mass density and $U^\mu$ is a vector (i.e. $(1, 0)$-tensor) field describing the $4$-velocity of our dust matter model. The relevant equation of motion is $\nabla_\mu T^{\mu\nu}_{\text{dust}}(x) = 0$. Moreover, the general relativistic energy-momentum tensor for a fluid (see (7.36) for the case of the Minkowski spacetime) is given by our substitution rule by

$$T^{\mu\nu}_{\text{fluid}}(x) = \Big(\rho(x) + \frac{P(x)}{c^2}\Big)\, U^\mu(x)\, U^\nu(x) + P(x)\, g^{\mu\nu}(x) \ , \tag{10.49}$$

where $P$ is a scalar field describing the pressure density. Notice that to arrive from (7.36) at the general relativistic expression (10.49), we had to replace the inverse Minkowski metric $\eta^{\mu\nu}$ in (7.36) by $g^{\mu\nu}(x)$ according to our substitution rule. The relevant equation of motion is $\nabla_\mu T^{\mu\nu}_{\text{fluid}}(x) = 0$.

CHAPTER 11

# Einstein's Field Equation

So far we considered arbitrary spacetimes $M$ with Lorentzian metric $g_{\mu\nu}$ without asking the important question which equation we should use to determine the metric $g_{\mu\nu}$? As motivated in Chapter 8, the equation of motion for the metric should be a tensor field equation of the form

$$G^{\mu\nu} = \kappa\, T^{\mu\nu}\ , \tag{11.1}$$

where $T^{\mu\nu}$ is the energy-momentum tensor of the matter distribution throughout spacetime and $\kappa$ is a numerical constant that has to be fixed later. The symmetric $(2,0)$-tensor field $G^{\mu\nu}$ should be build up from the metric tensor $g_{\mu\nu}$ and its first and second derivatives ($\partial_\rho g_{\mu\nu}$ and $\partial_\sigma\partial_\rho g_{\mu\nu}$).

A natural candidate for a tensor field depending only on the metric tensor and its first and second derivatives is the Riemann curvature tensor

$$R^\beta_{\rho\nu\mu} = \partial_\nu\Gamma^\beta_{\mu\rho} - \partial_\mu\Gamma^\beta_{\nu\rho} + \Gamma^\sigma_{\mu\rho}\,\Gamma^\beta_{\nu\sigma} - \Gamma^\sigma_{\nu\rho}\,\Gamma^\beta_{\mu\sigma}\ , \tag{11.2}$$

where the Christoffel symbols are given by

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\, g^{\rho\sigma}\left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right)\ . \tag{11.3}$$

The problem is that the Riemann curvature tensor is a $(1,3)$-tensor field, while $G^{\mu\nu}$ in (11.1) has to be a $(2,0)$-tensor field. The way out is to use index contractions, so let us try all possibilities. For the $\beta,\rho$-contraction we find that

$$R^\beta_{\beta\nu\mu} = g^{\alpha\beta}\, R_{\alpha\beta\nu\mu} = 0\ , \tag{11.4}$$

because by (9.97) the tensor field $R_{\alpha\beta\nu\mu}$ is antisymmetric in $\alpha,\beta$, i.e. $R_{\alpha\beta\nu\mu} = -R_{\beta\alpha\nu\mu}$, while the inverse metric $g^{\alpha\beta}$ is symmetric. The $\beta,\nu$-contraction of the Riemann curvature tensor in general does not vanish, hence it is an interesting $(0,2)$-tensor field which is called the *Ricci tensor*. We denote the Ricci tensor by

$$R_{\rho\mu} := R^\beta_{\rho\beta\mu}\ , \tag{11.5}$$

and its corresponding $(2,0)$-tensor field which is obtained by raising the indices by

$$R^{\sigma\nu} = g^{\sigma\rho}\, g^{\nu\mu}\, R_{\rho\mu}\ . \tag{11.6}$$

Using that the Riemann curvature tensor $R^\beta_{\rho\nu\mu}$ is antisymmetric in the $\nu,\mu$-indices (9.96), i.e. $R^\beta_{\rho\nu\mu} = -R^\beta_{\rho\mu\nu}$, we find that the $\beta,\mu$-contraction of the Riemann curvature tensor is up to a minus sign the Ricci tensor,

$$R^\beta_{\rho\nu\beta} = -R^\beta_{\rho\beta\nu} = -R_{\rho\nu}\ , \tag{11.7}$$

so it leads to nothing new. Finally, using also the pair symmetry property $R_{\alpha\beta\nu\mu} = R_{\nu\mu\alpha\beta}$ from (9.98), we obtain that *all* possible contractions of two indices of the Riemann curvature tensor are either $0$, the Ricci tensor or minus the Ricci tensor.

This shows that the Ricci tensor plays a distinguished role. For later use, notice that the pair symmetry property (9.98) implies that the Ricci tensor is symmetric

$$R_{\rho\mu} = R^{\beta}_{\rho\beta\mu} = g^{\alpha\beta}\, R_{\alpha\rho\beta\mu} = g^{\alpha\beta}\, R_{\beta\mu\alpha\rho} = R^{\beta}_{\mu\beta\rho} = R_{\mu\rho}\ , \tag{11.8}$$

or equivalently that

$$R^{\rho\mu} = R^{\mu\rho} \tag{11.9}$$

for the associated $(2,0)$-tensor field with upper indices.

Setting $G^{\mu\nu} = R^{\mu\nu}$ in (11.1) is however the *wrong* choice of equation of motion for the metric tensor $g_{\mu\nu}$. The reason is that the right-hand side of (11.1) is an energy-momentum tensor, hence it has to satisfy the covariant conservation law

$$\nabla_\mu T^{\mu\nu} = 0\ . \tag{11.10}$$

The Ricci tensor (11.5) however satisfies

$$\nabla_\mu R^{\mu\nu} = \frac{1}{2}\,\nabla^\nu R = \frac{1}{2}\, g^{\nu\rho}\,\nabla_\rho R\ , \tag{11.11}$$

where $R$ is the scalar field defined by the index contraction

$$R := g^{\mu\nu}\, R_{\mu\nu} \tag{11.12}$$

of the Ricci tensor. We also call (11.12) the *scalar curvature* or *Ricci scalar*. Equation (11.11) can be proved by using the second Bianchi identity for the Riemann curvature tensor (9.100): In fact, contracting in the second Bianchi identity

$$\nabla_\sigma R^{\beta}_{\rho\nu\mu} + \nabla_\nu R^{\beta}_{\rho\mu\sigma} + \nabla_\mu R^{\beta}_{\rho\sigma\nu} = 0 \tag{11.13}$$

the indices $\beta, \sigma$ and also the indices $\nu$, $\rho$, we obtain

$$\begin{aligned} 0 &= g^{\nu\rho}\left(\nabla^\beta R_{\beta\rho\nu\mu} + \nabla_\nu R^{\beta}_{\rho\mu\beta} + \nabla_\mu R^{\beta}_{\rho\beta\nu}\right) \\ &= -\nabla^\beta R_{\beta\mu} - \nabla^\nu R_{\nu\mu} + \nabla_\mu R\ , \end{aligned} \tag{11.14}$$

which implies (11.11) by raising/lowering and relabeling the indices.

The property (11.11) motivates the definition of the following symmetric $(2,0)$-tensor field

$$G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}\, g^{\mu\nu}\, R\ , \tag{11.15}$$

which combines the Ricci tensor and the scalar curvature. Notice that (11.15) satisfies

$$\nabla_\mu G^{\mu\nu} = \nabla_\mu\Big(R^{\mu\nu} - \frac{1}{2}\, g^{\mu\nu}\, R\Big) = \nabla_\mu R^{\mu\nu} - \frac{1}{2}\underbrace{\nabla_\mu g^{\mu\nu}}_{=0}\, R - \frac{1}{2}\nabla^\nu R = 0\ , \tag{11.16}$$

where we used the Leibniz rule for the covariant derivative, metric compatibility $\nabla_\mu g^{\rho\sigma} = 0$ and (11.11). Hence, the equation

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}\, g^{\mu\nu}\, R = \kappa\, T^{\mu\nu} \tag{11.17}$$

is consistent with taking the covariant divergence on both sides.

The tensor field equation (11.17) with $G^{\mu\nu}$ defined by (11.15) has been proposed by Einstein as the fundamental equation of general relativity. It determines the metric tensor $g_{\mu\nu}$ in terms of the matter distribution described by the energy-momentum tensor $T^{\mu\nu}$. One also calls the tensor field $G^{\mu\nu}$ in (11.15) the *Einstein tensor* and (11.17) the *Einstein equation*. The constant $\kappa$ in the Einstein equation (11.17) can be fixed by studying its non-relativistic limit and comparing it to Newtonian gravitation. This will be the topic of the next chapter.

Finding solutions of the Einstein equation, together with the conservation equation $\nabla_\mu T^{\mu\nu} = 0$ of the energy-momentum tensor, is typically extremely hard. This is because the Einstein tensor is a highly non-linear function of the metric tensor $g_{\mu\nu}$, involving in particular also the inverse metric $g^{\mu\nu}$. We will later construct a special solution of the Einstein equation, which is called the *Schwarzschild solution* and which describes the exterior region of a gravitating star, planet or also a black hole. Notice that the Minkowski metric $g_{\mu\nu} = \eta_{\mu\nu}$ is a solution of Einstein's equation if and only if we assume that $T^{\mu\nu} = 0$, i.e. that there is no energy and momentum distribution throughout spacetime. The reason is that the Riemann curvature tensor $R^{\beta}_{\rho\nu\mu} = 0$ vanishes for the Minkowski metric (cf. Example 9.15) and hence so does the Ricci tensor and the curvature scalar. Thus, $G^{\mu\nu} = 0$ for the Minkowski metric.

REMARK 11.1 (Alternative form of Einstein's equation). The Einstein equation (11.17) can be rewritten in an alternative form, which is useful in some applications. Let us first consider the trace of (11.17), i.e. the index contraction of $\mu$ and $\nu$ by a metric field. We obtain

$$g_{\mu\nu}\Big(R^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}\,R\Big) = R - \frac{1}{2}\,\underbrace{\delta^{\mu}_{\mu}}_{=4}\,R = -R \stackrel{!}{=} \kappa\;g_{\mu\nu}\,T^{\mu\nu} =: \kappa\,T\;. \tag{11.18}$$

Inserting this back into (11.17) yields

$$R^{\mu\nu} + \frac{\kappa}{2}\,g^{\mu\nu}\,T = \kappa\,T^{\mu\nu} \tag{11.19}$$

or equivalently

$$R^{\mu\nu} = \kappa\,\Big(T^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}\,T\Big)\;. \tag{11.20}$$

This equation is equivalent to Einstein's equation (11.17). △

REMARK 11.2 (Cosmological constant term). Einstein also considered the possibility of adding a term of the form $\Lambda\,g^{\mu\nu}$ to the Einstein equation (11.17), where $\Lambda$ is a numerical constant called the *cosmological constant*. This leads to the modified equation

$$R^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}\,R + \Lambda\,g^{\mu\nu} = \kappa\,T^{\mu\nu}\;, \tag{11.21}$$

from which we of course recover (11.17) by setting $\Lambda = 0$. Notice that because of $\nabla_\mu g^{\nu\rho} = 0$, the covariant divergence of both sides of this equation is zero. In the more modern literature, one often prefers to regard the cosmological constant term as part of the right-hand side of Einstein's equation, i.e.

$$R^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}\,R = \kappa\,\Big(T^{\mu\nu} - \frac{\Lambda}{\kappa}\,g^{\mu\nu}\Big)\;. \tag{11.22}$$

The physical interpretation is that $-\frac{\Lambda}{\kappa}\,g^{\mu\nu}$ is a particular form of energy-momentum distribution throughout spacetime, which is called *dark energy* in cosmology. Dark energy is important to explain the expansion of our universe. △

CHAPTER 12

# The Newtonian Limit of General Relativity

In this chapter we study the non-relativistic limit of Einstein's equation (11.17) and the geodesic equation (10.40) in order to compare them with Newton's laws of gravitation, cf. Chapter 8. This is important for two very different reasons: 1.) It shows that general relativity can be approximated by Newtonian gravitation in certain regimes where the gravitational field is weak and the velocities involved are much smaller than the speed of light. Hence, it is a genuine generalization of Newtonian gravitation to strong gravitational fields and large velocities. 2.) It allows us to fix the unknown parameter $\kappa$ in the Einstein equation (11.17) in terms of Newton's gravitational constant $G$ and some numerical factors.

## 12.1. Linearization of the Einstein Tensor

Let us consider $M = \mathbb{R}^4$ with Lorentzian metric given by

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \; . \tag{12.1}$$

We assume that $h_{\mu\nu} = h_{\mu\nu}(x)$ is very small, i.e. that quadratic and higher-order terms in $h_{\mu\nu}$ are approximately zero. In technical jargon, one says that $h_{\mu\nu}$ is a *perturbation* of the Minkowski metric $\eta_{\mu\nu}$. As a consequence of this assumption, the inverse metric is approximately given by

$$g^{\mu\nu} \approx \eta^{\mu\nu} - \eta^{\mu\rho}\,\eta^{\nu\sigma}\,h_{\rho\sigma} =: \eta^{\mu\nu} - h^{\mu\nu} \; . \tag{12.2}$$

Here and in the following we use the Minkowski metric for raising and lowering indices. The Christoffel symbols (9.85) can be expanded to first order in the perturbation $h_{\mu\nu}$ and they read as

$$\begin{aligned}\Gamma^{\rho}_{\mu\nu} &\approx \frac{1}{2}\,\eta^{\rho\sigma}\left(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}\right) \\ &= \frac{1}{2}\left(\partial_\mu h^{\rho}_{\nu} + \partial_\nu h^{\rho}_{\mu} - \partial^{\rho} h_{\mu\nu}\right) \; .\end{aligned} \tag{12.3}$$

Concerning the Riemann curvature tensor (9.95), we observe that the quadratic terms in the Christoffel symbols are approximately zero because they involve only second or higher-order terms in $h_{\mu\nu}$. Hence, in our approximation we find

$$\begin{aligned}R^{\beta}_{\rho\nu\mu} &\approx \partial_\nu \Gamma^{\beta}_{\mu\rho} - \partial_\mu \Gamma^{\beta}_{\nu\rho} \\ &= \frac{1}{2}\left(\partial_\nu\partial_\rho h^{\beta}_{\mu} + \partial_\mu\partial^{\beta} h_{\nu\rho} - \partial_\mu\partial_\rho h^{\beta}_{\nu} - \partial_\nu\partial^{\beta} h_{\mu\rho}\right) \; .\end{aligned} \tag{12.4}$$

The Ricci tensor (11.5) in our approximation then reads as

$$R_{\rho\mu} = R^{\beta}_{\rho\beta\mu} \approx \frac{1}{2}\Big(\partial_\beta\partial_\rho h^{\beta}_{\mu} + \partial_\mu\partial^{\beta} h_{\beta\rho} - \partial_\mu\partial_\rho h^{\beta}_{\beta} - \partial_\beta\partial^{\beta} h_{\mu\rho}\Big) \; , \tag{12.5}$$

and the scalar curvature (11.12) as

$$R = g^{\rho\mu}\,R_{\rho\mu} \approx \eta^{\rho\mu}\,R_{\rho\mu} \approx \frac{1}{2}\Big(\partial_\beta\partial_\rho h^{\rho\beta} + \partial^\rho\partial^\beta h_{\beta\rho} - \partial^\rho\partial_\rho h^\beta_\beta - \partial_\beta\partial^\beta h^\rho_\rho\Big)$$
$$= \partial_\alpha\partial_\beta h^{\alpha\beta} - \partial^\alpha\partial_\alpha h^\beta_\beta\,, \tag{12.6}$$

where in the last equality we raised/lowered and relabeled the indices.

With these preparations, we find that the Einstein tensor (11.15) in our linear approximation reads as

$$G_{\rho\mu} \approx R_{\rho\mu} - \frac{1}{2}\eta_{\rho\mu}\,R \tag{12.7}$$
$$\approx \frac{1}{2}\Big(\partial_\beta\partial_\rho h^\beta_\mu + \partial_\mu\partial^\beta h_{\beta\rho} - \partial_\mu\partial_\rho h^\beta_\beta - \partial_\beta\partial^\beta h_{\mu\rho} - \eta_{\rho\mu}\,\partial_\alpha\partial_\beta h^{\alpha\beta} + \eta_{\rho\mu}\,\partial^\alpha\partial_\alpha h^\beta_\beta\Big)\,.$$

This expression looks horribly complicated! Notice that we can drastically simplify it by imposing the condition

$$\partial_\beta h^\beta_\mu = \frac{1}{2}\,\partial_\mu h^\beta_\beta \tag{12.8}$$

for the metric perturbation $h_{\mu\nu}$. Using more advanced techniques, which are unfortunately beyond the scope of this module, one can show that the condition (12.8) can always be fulfilled by applying infinitesimal coordinate transformations to (12.1). (In technical jargon, these are called *gauge transformations* and the condition (12.8) is called a *gauge fixing*.) In this module, we will take the condition (12.8) for granted.

Inserting our condition (12.8) into (12.7), we find after a short calculation

$$G_{\rho\mu} \overset{(12.8)}{\approx} -\frac{1}{2}\partial_\beta\partial^\beta h_{\rho\mu} + \frac{1}{4}\,\eta_{\rho\mu}\,\partial_\alpha\partial^\alpha h^\beta_\beta = -\frac{1}{2}\partial_\alpha\partial^\alpha\Big(h_{\rho\mu} - \frac{1}{2}\eta_{\rho\mu}\,h^\beta_\beta\Big)\,. \tag{12.9}$$

This expression can be simplified further by introducing the redefined perturbation

$$\overline{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\,\eta_{\mu\nu}\,h^\beta_\beta\,. \tag{12.10}$$

We obtain

$$G_{\rho\mu} \approx -\frac{1}{2}\partial_\alpha\partial^\alpha\,\overline{h}_{\rho\mu}\,, \tag{12.11}$$

which is indeed a much simpler expression than (12.7). As a side-remark, notice that $\partial_\alpha\partial^\alpha = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \triangle$ is the d'Alembert operator/wave operator. We will see in one of the problem sheets that linearized general relativity admits wave solutions, which are the famous gravitational waves.

## 12.2. Limit of the Einstein Equation

Using our final result (12.11) for the linearization of the Einstein tensor, Einstein's equation (11.17) is approximately given by

$$-\frac{1}{2}\partial_\alpha\partial^\alpha\,\overline{h}_{\mu\nu} \approx \kappa\,T_{\mu\nu}\,, \tag{12.12}$$

or equivalently

$$\partial_\alpha\partial^\alpha\,\overline{h}_{\mu\nu} \approx -2\,\kappa\,T_{\mu\nu}\,. \tag{12.13}$$

To study the non-relativistic limit, we make the following assumptions.

1.) The metric perturbation $h_{\mu\nu}$ (and hence also $\overline{h}_{\mu\nu}$ defined in (12.10)) changes just very slowly in time $x^0 = c\,t$, i.e. its partial derivative along $x^0$ is approximately zero,

$$\partial_0 \overline{h}_{\mu\nu} \approx 0 \; . \tag{12.14}$$

2.) The momentum density is negligible compared to the energy density, i.e. the energy-momentum tensor $T_{\mu\nu}$ is dominated by the 00-component $T_{00}$,

$$T_{\mu\nu} \approx \begin{pmatrix} T_{00} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \; . \tag{12.15}$$

3.) The redefined metric perturbation $\overline{h}_{\mu\nu}$ goes to zero at spatial infinity, i.e. $\overline{h}_{\mu\nu}(t, \mathbf{x}) \to 0$ for $|\mathbf{x}| \to \infty$.

Using assumption 1.), we find that the time derivatives drop out of (12.13) and hence this equation becomes

$$\partial_\alpha \partial^\alpha \, \overline{h}_{\mu\nu} \approx \sum_{i=1}^{3} \frac{\partial^2}{\partial x^{i2}} \, \overline{h}_{\mu\nu} = \Delta \overline{h}_{\mu\nu} \approx -2\,\kappa\, T_{\mu\nu} \; , \tag{12.16}$$

where $\Delta = \nabla^2$ is the Laplace operator. Using also assumption 2.), the 00-component of this equation reads as

$$\Delta \overline{h}_{00} = -2\,\kappa\, T_{00} \; , \tag{12.17}$$

the $0i$-component, with $i = 1, 2, 3$, as

$$\Delta \overline{h}_{0i} = 0 \; , \tag{12.18}$$

and the $ij$-component, with $i, j = 1, 2, 3$, as

$$\Delta \overline{h}_{ij} = 0 \; . \tag{12.19}$$

(Notice that we do not have to study the $i0$-component because $\overline{h}_{\mu\nu}$ is symmetric.) Comparing (12.17) with Newton's law of gravitation

$$\Delta \Phi = 4\pi\, G\, \rho \; , \tag{12.20}$$

where

$$\rho = \frac{1}{c^2} T_{00} \tag{12.21}$$

is the mass density (recall the formula $E = m\,c^2$ and that $T_{00}$ is the energy density), we find the relationship

$$\overline{h}_{00} = -\frac{\kappa\, c^2}{2\pi\, G}\, \Phi \tag{12.22}$$

between $\overline{h}_{00}$ and Newton's gravitational potential $\Phi$. Using assumption 3.), the unique solution of the other equations (12.18) and (12.19) is the zero solution, i.e. we find

$$\overline{h}_{0i} = 0 \quad , \qquad \overline{h}_{ij} = 0 \; . \tag{12.23}$$

Inverting our redefinition of the perturbation (12.10) via

$$h_{\mu\nu} = \overline{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\,\overline{h}_{\beta}^{\beta}\,, \tag{12.24}$$

we find for the components of our original perturbation $h_{\mu\nu}$ the following explicit expressions

$$h_{00} = \frac{1}{2}\overline{h}_{00} = -\frac{\kappa\, c^2}{4\pi\, G}\,\Phi\,, \tag{12.25a}$$
$$h_{0i} = 0\,, \tag{12.25b}$$
$$h_{ij} = \frac{1}{2}\,\delta_{ij}\,\overline{h}_{00} = -\delta_{ij}\,\frac{\kappa\, c^2}{4\pi\, G}\,\Phi\,. \tag{12.25c}$$

Inserting this into (12.1), we find

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} = \begin{pmatrix} -\left(1+\frac{\kappa\,c^2}{4\pi\,G}\,\Phi\right) & 0 & 0 & 0 \\ 0 & 1-\frac{\kappa\,c^2}{4\pi\,G}\,\Phi & 0 & 0 \\ 0 & 0 & 1-\frac{\kappa\,c^2}{4\pi\,G}\,\Phi & 0 \\ 0 & 0 & 0 & 1-\frac{\kappa\,c^2}{4\pi\,G}\,\Phi \end{pmatrix}, \tag{12.26}$$

which is our final expression to relate the metric tensor to Newton's gravitational potential $\Phi$ in the non-relativistic limit.

## 12.3. Limit of the Geodesic Equation

We shall now study how the geodesic equation (10.40) – describing the dynamics of massive point-particles propagating in a curved spacetime – looks like in our non-relativistic limit. In this section the metric tensor is always the one given by (12.26), where $\Phi$ is assumed to be very small and very slowly varying in time $\partial_0\Phi \approx 0$. (The latter is our assumption 1.) from the previous section.) This means that our metric is very close to the Minkowski metric. For a slowly moving particle in our particular spacetime (i.e. $\mathbf{v}(t)^2 \ll c^2$), proper time is approximately equal to coordinate time $\tau = t$ (cf. (6.6)) and the 4-velocity (cf. (6.18)) is approximately

$$u^\mu(\tau) = \frac{dx^\mu(\tau)}{d\tau} \approx \begin{pmatrix} c \\ \mathbf{v}(t) \end{pmatrix} = \begin{pmatrix} c \\ \frac{d\mathbf{x}(t)}{dt} \end{pmatrix}. \tag{12.27}$$

The 0-component $u^0 \approx c$ of the 4-velocity is thus the dominant term. This means that we can approximate the geodesic equation

$$\frac{d^2x^\mu(\tau)}{d\tau^2} + \Gamma^\mu_{\nu\rho}(x(\tau))\,\frac{dx^\nu(\tau)}{d\tau}\,\frac{dx^\rho(\tau)}{d\tau} = 0 \tag{12.28}$$

by

$$\frac{d^2x^\mu(t)}{dt^2} + \Gamma^\mu_{00}(t,\mathbf{x}(t))\,c^2 \approx 0\,. \tag{12.29}$$

Using our expression (12.3) for the approximation of the Christoffel symbols and (12.25), we find

$$\Gamma^{\mu}_{00} \approx \frac{1}{2}\eta^{\mu\sigma}\Big(\partial_0 h_{0\sigma} + \partial_0 h_{0\sigma} - \partial_\sigma h_{00}\Big) \approx \frac{\kappa\, c^2}{8\pi\, G}\, \partial^{\mu}\Phi\,. \tag{12.30}$$

In the last step we also used $\partial_0 h_{\mu\nu} \approx 0$, which is precisely our assumption 1.) from the previous section. Inserting this into (12.29) yields

$$\frac{d^2x^{\mu}(t)}{dt^2} + \frac{\kappa\, c^4}{8\pi\, G}\, \partial^{\mu}\Phi(t, \mathbf{x}(t)) \approx 0\,. \tag{12.31}$$

The 0-component of this equation is fulfilled in our approximation: By (12.27), the time derivative of $u^0 \approx c$ is approximately zero, and so is $\partial^0\Phi(t, \mathbf{x}(t)) \approx 0$ by our assumption 1.) from the previous section. The space components of (12.31) read in 3-vector notation as

$$\frac{d^2\mathbf{x}(t)}{dt^2} \approx -\frac{\kappa\, c^4}{8\pi\, G}\, \nabla\Phi(t, \mathbf{x}(t))\,. \tag{12.32}$$

This is precisely Newton's equation (8.8) for a particle in a gravitational field, provided that we fix $\kappa$ by

$$\frac{\kappa\, c^4}{8\pi\, G} = 1\,. \tag{12.33}$$

Hence, we have shown that the geodesic equation (10.40) reduces to Newton's equation in the non-relativistic limit.

Using (12.33), we can make now explicit the relation (12.26) between the metric tensor and Newton's gravitational potential. We obtain

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} = \begin{pmatrix} -\Big(1 + \frac{2\,\Phi}{c^2}\Big) & 0 & 0 & 0 \\ 0 & 1 - \frac{2\,\Phi}{c^2} & 0 & 0 \\ 0 & 0 & 1 - \frac{2\,\Phi}{c^2} & 0 \\ 0 & 0 & 0 & 1 - \frac{2\,\Phi}{c^2} \end{pmatrix}, \tag{12.34}$$

where now all numerical constants are known. For completeness, let us also mention that the line element corresponding to (12.34) reads as

$$ds^2_{\text{Newton}} = -\Big(1 + \tfrac{2\,\Phi}{c^2}\Big)\, c^2\, (dt)^2 + \Big(1 - \tfrac{2\,\Phi}{c^2}\Big)\Big((dx^1)^2 + (dx^2)^2 + (dx^3)^2\Big)\,. \tag{12.35}$$

This expression illustrates very nicely the way in which a Newtonian gravitational potential $\Phi$ may be interpreted as a curved spacetime geometry.

## 12.4. The Constant $\kappa$ in Einstein's Equation

Our discussion in this chapter has shown that general relativity reduces to Newtonian gravitation in the non-relativistic limit, provided that $\kappa$ is fixed by (12.33). Explicitly,

$$\kappa = \frac{8\pi\, G}{c^4}\,, \tag{12.36}$$

where we recall that $G$ is Newton's constant

$$G = 6.67408 \times 10^{-11}\ \mathrm{m^3\, kg^{-1}\, s^{-2}}\,. \tag{12.37}$$

The Einstein equation (11.17) then explicitly reads as

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}\, g^{\mu\nu}\, R = \frac{8\pi\, G}{c^4}\, T^{\mu\nu} \,, \tag{12.38}$$

where now all numerical constants are known.

CHAPTER 13

# Schwarzschild Spacetime

## 13.1. Introductory Remarks

So far we know only one solution of the Einstein equation (12.38): The spacetime $M = \mathbb{R}^4$ with the Minkowski metric $\eta_{\mu\nu}$ is a solution of (12.38) for $T^{\mu\nu} = 0$. The physical scenario described by this solution is however rather boring, because $T^{\mu\nu} = 0$ means that the whole universe is 'empty' in the sense that there is no matter distributed in it. The goal of this chapter is to find another, physically more interesting, solution of the Einstein equation. The physical scenario we would like to describe is as follows (see also Figure 13.1 for a graphical illustration): Using Cartesian coordinates $x^0 = ct$, $x^1$, $x^2$ and $x^3$ on $\mathbb{R}^4$, we consider a spherically symmetric and time-independent mass density centered at the origin $\mathbf{x} = \mathbf{0}$ of our coordinate frame, i.e. $\rho(x) = \rho(r)$ with $r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$ the radial coordinate. We further assume that this mass density is concentrated in some finite region of space, i.e. that $\rho(r) = 0$ for $r > r_{\max}$ some maximal radius $r_{\max} > 0$. Physical examples of this scenario are astrophysical objects such as stars and planets; in this case $r_{\max}$ describes the radius of the star/planet.

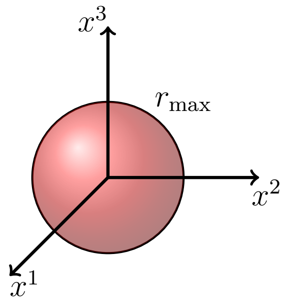


FIGURE 13.1. Visualization of a spherically symmetric and time-independent mass density $\rho(x) = \rho(r)$ centered at the origin $\mathbf{x} = \mathbf{0}$ of the spatial coordinate system. The mass is assumed to be localized inside a ball of radius $r_{\max}$, i.e. $\rho(r) = 0$ for $r > r_{\max}$. You can think of the red ball as (an idealization of) a planet or a star.

To analyze this scenario, it is definitely beneficial to work in spherical coordinates $ct$, $r$, $\theta$ and $\varphi$. Notice that there is no obstruction in doing so because general relativity is by construction invariant under general coordinate transformations! In order to find a reasonable *Ansatz* for the Lorentzian metric $g_{\mu\nu}$ modeling our physical scenario, it is instructive to find out how the Minkowski metric looks like in spherical coordinates. Recalling Example 9.2, the coordinate transformation from spherical to

Cartesian coordinates is given by

$$\begin{aligned}
x'^0 &= x^0 = c\,t\;, \\
x'^1 &= r\;\cos\varphi\;\sin\theta\;, \\
x'^2 &= r\;\sin\varphi\;\sin\theta\;, \\
x'^3 &= r\;\cos\theta\;,
\end{aligned} \tag{13.1}$$

and the associated Jacobi matrix reads as

$$J^\mu{}_\nu(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi\;\sin\theta & r\;\cos\varphi\;\cos\theta & -r\;\sin\varphi\;\sin\theta \\ 0 & \sin\varphi\;\sin\theta & r\;\sin\varphi\;\cos\theta & r\;\cos\varphi\;\sin\theta \\ 0 & \cos\theta & -r\;\sin\theta & 0 \end{pmatrix} . \tag{13.2}$$

Notice that we denote by $x'^\mu$ the Cartesian coordinates and by $x^\mu$ the spherical coordinates. From the transformation law for metric tensors

$$g'_{\mu\nu}(x') = J^{-1\rho}{}_\mu(x)\,J^{-1\sigma}{}_\nu(x)\,g_{\rho\sigma}(x)\;, \tag{13.3}$$

we obtain by inverting the Jacobi matrices the transformation rule

$$g_{\mu\nu}(x) = J^\rho{}_\mu(x)\,J^\sigma{}_\nu(x)\,\eta_{\rho\sigma} \tag{13.4}$$

for the Minkowski metric from Cartesian coordinates $x'^\mu$ to spherical coordinates $x^\mu$. Writing (13.4) as a matrix equation, i.e.

$$g_{\mu\nu}(x) = (J^{\mathrm{T}}\,\eta\,J)_{\mu\nu}(x) \tag{13.5}$$

with $J^{\mathrm{T}}$ the transposed Jacobi matrix, one finds after a short calculation

$$g_{\mu\nu}(x) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\,(\sin\theta)^2 \end{pmatrix} . \tag{13.6}$$

In this calculation one frequently has to use trigonometric identities of the form $(\cos\varphi)^2 + (\sin\varphi)^2 = 1$. The line element corresponding to $g_{\mu\nu}$ then has the simple form

$$ds^2 = -c^2\,dt^2 + dr^2 + r^2\,d\theta^2 + r^2\,(\sin\theta)^2 d\varphi^2\;. \tag{13.7}$$

For the rest of this chapter we will always use spherical coordinates $x^\mu$, i.e. $x^0 = c\,t$, $x^1 = r$, $x^2 = \theta$ and $x^3 = \varphi$.

Our spherically symmetric and time-independent mass density $\rho(x) = \rho(r)$ will curve the Minkowski metric (given in spherical coordinates by (13.6)) in a way that is also spherically symmetric and time-independent. To model this gravitational effect, we consider the following *Ansatz* for the metric in spherical coordinates

$$g_{\mu\nu}(x) = \begin{pmatrix} -e^{a(r)} & 0 & 0 & 0 \\ 0 & e^{b(r)} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\,(\sin\theta)^2 \end{pmatrix} , \tag{13.8}$$

where $a$ and $b$ are functions of only the radial coordinate $r$ that we will determine later from the Einstein equation. The line element corresponding to (13.8) reads as

$$ds^2 = -e^{a(r)}\,c^2\,dt^2 + e^{b(r)}\,dr^2 + r^2\,d\theta^2 + r^2\,(\sin\theta)^2 d\varphi^2 \, . \tag{13.9}$$

Of course, we recover the Minkowski metric (and line element) in spherical coordinates by setting $a(r) = b(r) = 0$.

In this chapter we will solve Einstein's equation

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}\,R = \frac{8\pi\,G}{c^4}\,T^{\mu\nu} \tag{13.10}$$

by inserting our *Ansatz* (13.8) for the metric tensor and using as energy-momentum tensor our spherically symmetric and time-independent mass density, i.e.

$$T^{\mu\nu}(x) = \begin{pmatrix} \rho(r)\,c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{13.11}$$

Recalling our assumption that $\rho(r) = 0$, for $r > r_{\text{max}}$, we observe that the energy-momentum tensor vanishes

$$T^{\mu\nu}(x) = 0 \qquad (\text{for } r > r_{\text{max}}) \, , \tag{13.12}$$

in the exterior region $r > r_{\text{max}}$ of our mass density. Hence, in the exterior region Einstein's equation simplifies to

$$R^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}\,R = 0 \qquad (\text{for } r > r_{\text{max}}) \, . \tag{13.13}$$

Using our equivalent form of the Einstein equation (11.20), we obtain that (13.13) is equivalent to

$$R_{\mu\nu} = 0 \qquad (\text{for } r > r_{\text{max}}) \, , \tag{13.14}$$

i.e. that the Ricci tensor vanishes in the exterior region.

You might wonder how it is possible that our mass density $\rho(r)$ influences the solution of Einstein's equation (13.14) in the exterior region $r > r_{\text{max}}$, where it completely drops out. The key point is that solutions to (13.14) depend on a free parameter, which we shall relate to the *total mass*

$$M_{\text{tot}} := \int_{\mathbb{R}^3} \rho(\mathbf{x})\,d^3\mathbf{x} = 4\pi \int_0^\infty \rho(r)\,r^2\,dr \tag{13.15}$$

of our mass density $\rho(r)$. In particular, we will see that the Riemann curvature tensor of solutions to (13.14) will be non-vanishing in the exterior region $r > r_{\max}$ for $M_{\mathrm{tot}} \neq 0$, meaning that spacetime is curved.

After this detailed motivation of what we would like to do in this chapter, we now have to work quite hard to do the necessary calculations. This is an excellent opportunity to practice and apply our techniques we learned in this module.

### 13.2. Christoffel Symbols, Riemann Curvature Tensor and Ricci Tensor

We shall always work in spherical coordinates ($x^0 = ct$, $x^1 = r$, $x^2 = \theta$ and $x^3 = \varphi$) and consider the class of metric tensors defined by (13.8) or equivalently (13.9). The functions $a(r)$ and $b(r)$ are at the moment arbitrary. Notice that the inverse metric is given by

$$g^{\mu\nu}(x) = \begin{pmatrix} -e^{-a(r)} & 0 & 0 & 0 \\ 0 & e^{-b(r)} & 0 & 0 \\ 0 & 0 & r^{-2} & 0 \\ 0 & 0 & 0 & r^{-2}\,(\sin\theta)^{-2} \end{pmatrix} . \tag{13.16}$$

For the sake of a convenient notation, we denote the components of the partial derivatives $\partial_\mu$, for the space directions $\mu = 1, 2, 3$, by the coordinates they represent, i.e.

$$\partial_1 = \frac{\partial}{\partial r} = \partial_r \ , \quad \partial_2 = \frac{\partial}{\partial \theta} = \partial_\theta \ , \quad \partial_3 = \frac{\partial}{\partial \varphi} = \partial_\varphi \ . \tag{13.17}$$

We use the same convention for the components of tensor fields. For example,

$$C^0_{\varphi r} = C^0_{31} \quad , \qquad C^\theta_{r\varphi} = C^2_{13} \ . \tag{13.18}$$

This might be a bit unusual at first sight, but you will get quickly used to that.

Let us first compute the Christoffel symbols

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\, g^{\rho\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right) \tag{13.19}$$

corresponding to our metric tensor (13.8). For $\rho = 0$, we obtain

$$\begin{aligned} \Gamma^0_{\mu\nu} &= -\frac{1}{2}\, e^{-a}\Big(\partial_\mu g_{\nu 0} + \partial_\nu g_{\mu 0} - \partial_0 g_{\mu\nu}\Big) \\ &= -\frac{1}{2}\, e^{-a}\Big(\partial_\mu g_{\nu 0} + \partial_\nu g_{\mu 0}\Big) \ , \end{aligned} \tag{13.20}$$

where in the second step we used that the metric is time-independent, i.e. $\partial_0 g_{\mu\nu} = 0$. Inserting (13.8) into this expression, we find that the non-zero Christoffel symbols $\Gamma^0_{\mu\nu}$ with upper index 0 are

$$\Gamma^0_{0r} = \Gamma^0_{r0} = \frac{1}{2}\, \partial_r a \ . \tag{13.21a}$$

By a similar calculation, we obtain that the non-zero Christoffel symbols $\Gamma^r_{\mu\nu}$ with upper index $r$ are

$$\begin{aligned} \Gamma^r_{00} &= \frac{1}{2}\, e^{a-b}\, \partial_r a \quad , \qquad \Gamma^r_{rr} = \frac{1}{2}\, \partial_r b \quad , \\ \Gamma^r_{\theta\theta} &= -e^{-b}\, r \quad , \qquad \Gamma^r_{\varphi\varphi} = -e^{-b}\, r\, (\sin\theta)^2 \; . \end{aligned} \tag{13.21b}$$

The non-zero Christoffel symbols $\Gamma^\theta_{\mu\nu}$ with upper index $\theta$ are

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = r^{-1} \quad , \qquad \Gamma^\theta_{\varphi\varphi} = -\sin\theta\, \cos\theta \; . \tag{13.21c}$$

Finally, the non-zero Christoffel symbols $\Gamma^\varphi_{\mu\nu}$ with upper index $\varphi$ are

$$\Gamma^\varphi_{r\varphi} = \Gamma^\varphi_{\varphi r} = r^{-1} \quad , \qquad \Gamma^\varphi_{\theta\varphi} = \Gamma^\varphi_{\varphi\theta} = \frac{\cos\theta}{\sin\theta} \; . \tag{13.21d}$$

Hence, we see that from the 40 independent Christoffel symbols $\Gamma^\rho_{\mu\nu}$ only 9 are non-vanishing for our metric *Ansatz* (13.8).

Computing the Riemann curvature tensor

$$R^\beta_{\rho\nu\mu} = \partial_\nu \Gamma^\beta_{\mu\rho} - \partial_\mu \Gamma^\beta_{\nu\rho} + \Gamma^\sigma_{\mu\rho}\, \Gamma^\beta_{\nu\sigma} - \Gamma^\sigma_{\nu\rho}\, \Gamma^\beta_{\mu\sigma} \tag{13.22}$$

associated to our Christoffel symbols (13.21) is a rather tedious task. This task can be simplified considerably by using the (anti)symmetry properties (9.96), (9.97) and (9.98) of this tensor. It turns out that the non-vanishing components of $R^\beta_{\rho\nu\mu}$ are

$$R^0_{r0r}\;, \quad R^0_{\theta 0\theta}\;, \quad R^0_{\varphi 0\varphi}\;, \quad R^r_{\theta r\theta}\;, \quad R^r_{\varphi r\varphi}\;, \quad R^\theta_{\varphi\theta\varphi}\;, \tag{13.23}$$

and of course also the ones fixed by the (anti)symmetry properties (9.96), (9.97) and (9.98). Let us now compute as an example the component $R^0_{r0r}$. Using (13.21), we find

$$\begin{aligned} R^0_{r0r} &= \partial_0 \Gamma^0_{rr} - \partial_r \Gamma^0_{0r} + \Gamma^\sigma_{rr}\, \Gamma^0_{0\sigma} - \Gamma^\sigma_{0r}\, \Gamma^0_{r\sigma} \\ &= -\partial_r \Gamma^0_{0r} + \Gamma^r_{rr}\, \Gamma^0_{0r} - (\Gamma^0_{0r})^2 \\ &= -\frac{1}{2}\, \partial_r^2 a + \frac{1}{4}\, (\partial_r b - \partial_r a)\, \partial_r a \; . \end{aligned} \tag{13.24}$$

By similar computations one obtains also the other components (13.23). The final result is given by

$$R^0_{r0r} = -\frac{1}{2}\, \partial_r^2 a + \frac{1}{4}\, \big(\partial_r b - \partial_r a\big)\, \partial_r a \;, \tag{13.25a}$$

$$R^0_{\theta 0\theta} = -\frac{1}{2}\, e^{-b}\, r\, \partial_r a \;, \tag{13.25b}$$

$$R^0_{\varphi 0\varphi} = -\frac{1}{2}\, e^{-b}\, r\, (\sin\theta)^2 \partial_r a \;, \tag{13.25c}$$

$$R^r_{\theta r\theta} = \frac{1}{2}\, e^{-b}\, r\, \partial_r b \;, \tag{13.25d}$$

$$R^r_{\varphi r\varphi} = \frac{1}{2}\, e^{-b}\, r\, (\sin\theta)^2\, \partial_r b \;, \tag{13.25e}$$

$$R^\theta_{\varphi\theta\varphi} = (1 - e^{-b})\, (\sin\theta)^2 \; . \tag{13.25f}$$

For the Ricci tensor

$$R_{\mu\nu} = R^\beta_{\mu\beta\nu} \tag{13.26}$$

one finds after a calculation

$$R_{00} = e^{a-b}\left(\frac{1}{2}\,\partial_r^2 a + \frac{1}{4}\,(\partial_r a)^2 - \frac{1}{4}\,\partial_r a\,\partial_r b + r^{-1}\,\partial_r a\right)\;, \tag{13.27a}$$

$$R_{rr} = -\frac{1}{2}\,\partial_r^2 a - \frac{1}{4}\,(\partial_r a)^2 + \frac{1}{4}\,\partial_r a\,\partial_r b + r^{-1}\,\partial_r b\;, \tag{13.27b}$$

$$R_{\theta\theta} = 1 - e^{-b}\left(1 + \frac{r}{2}\big(\partial_r a - \partial_r b\big)\right)\;, \tag{13.27c}$$

$$R_{\varphi\varphi} = (\sin\theta)^2\,R_{\theta\theta}\;, \tag{13.27d}$$

while all other components vanish. In this calculation one has to use carefully (13.25) together with the antisymmetry properties (9.96) and (9.97). We exemplify this by computing explicitly $R_{rr}$ as an example,

$$\begin{aligned}
R_{rr} = R^{\beta}_{\;r\beta r} &= R^{0}_{\;r0r} + \underbrace{R^{r}_{\;rrr}}_{=0 \text{ by } (9.96)} + R^{\theta}_{\;r\theta r} + R^{\varphi}_{\;r\varphi r} \\
&= R^{0}_{\;r0r} + g^{\theta\theta}\,R_{\theta r\theta r} + g^{\varphi\varphi}\,R_{\varphi r\varphi r} \\
&\overset{(9.96)\,\&\,(9.97)}{=} R^{0}_{\;r0r} + g^{\theta\theta}\,R_{r\theta r\theta} + g^{\varphi\varphi}\,R_{r\varphi r\varphi} \\
&= R^{0}_{\;r0r} + g^{\theta\theta}\,g_{rr}\,R^{r}_{\;\theta r\theta} + g^{\varphi\varphi}\,g_{rr}\,R^{r}_{\;\varphi r\varphi} \\
&\overset{(13.25)}{=} -\frac{1}{2}\,\partial_r^2 a - \frac{1}{4}\,(\partial_r a)^2 + \frac{1}{4}\,\partial_r a\,\partial_r b + r^{-1}\,\partial_r b\;,
\end{aligned} \tag{13.28}$$

where we also have used frequently that the metric and inverse metric are diagonal.

## 13.3. Solving Einstein's Equation

In the exterior region $r > r_{\max}$ where our mass density $\rho(r)$ is zero, Einstein's equation is given by (13.14), i.e. vanishing of the Ricci tensor. Hence, all non-vanishing components of the Ricci tensor (13.27) have to be set to zero. This leads to the following three equations

$$\frac{1}{2}\,\partial_r^2 a + \frac{1}{4}\,(\partial_r a)^2 - \frac{1}{4}\,\partial_r a\,\partial_r b + r^{-1}\,\partial_r a = 0\;, \tag{13.29a}$$

$$-\frac{1}{2}\,\partial_r^2 a - \frac{1}{4}\,(\partial_r a)^2 + \frac{1}{4}\,\partial_r a\,\partial_r b + r^{-1}\,\partial_r b = 0\;, \tag{13.29b}$$

$$1 - e^{-b}\left(1 + \frac{r}{2}\big(\partial_r a - \partial_r b\big)\right) = 0\;. \tag{13.29c}$$

Adding Eqns. (13.29a) and (13.29b), we obtain

$$\partial_r a = -\partial_r b\;, \tag{13.30}$$

which yields by integration[1]

$$a = -b\;. \tag{13.31}$$

[1] A non-zero integration constant, i.e. $a = -b + k$ with $k$ constant, can always be absorbed by rescaling our time coordinate. Explicitly, for the line element (13.9), $ds^2 = -e^{-b(r)+k}\,c^2\,dt^2 + e^{b(r)}\,dr^2 + r^2\,d\theta^2 + r^2\,(\sin\theta)^2 d\varphi^2 = -e^{-b(r)}\,c^2\,d(e^{k/2}t)^2 + e^{b(r)}\,dr^2 + r^2\,d\theta^2 + r^2\,(\sin\theta)^2 d\varphi^2 = -e^{-b(r)}\,c^2\,dt'^2 + e^{b(r)}\,dr^2 + r^2\,d\theta^2 + r^2\,(\sin\theta)^2 d\varphi^2$, with the new time coordinate $t' = e^{k/2}t$. Hence, we can set without loss of generality $k = 0$.

Inserting this into (13.29c) we find

$$1 - e^a \left(1 + r\,\partial_r a\right) = 0 \qquad \Longleftrightarrow \qquad \partial_r e^a + r^{-1}\, e^a = r^{-1}\,. \tag{13.32}$$

Taking the derivative $\partial_r$ of this equation, we obtain by using also (13.31) the difference between Eqns. (13.29a) and (13.29b), hence (13.32) is the only remaining equation we have to solve. The general solution of (13.32) is

$$e^a = 1 - \frac{r_S}{r}\,, \tag{13.33}$$

where $r_S$ is an integration constant called the *Schwarzschild radius*. By (13.31), we also find

$$e^b = e^{-a} = \frac{1}{e^a} = \frac{1}{1 - \frac{r_S}{r}}\,. \tag{13.34}$$

Inserting this into our metric *Ansatz* (13.8), we find

$$g_{\mu\nu}(x) = \begin{pmatrix} -\left(1 - \frac{r_S}{r}\right) & 0 & 0 & 0 \\ 0 & \frac{1}{1-\frac{r_S}{r}} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\,(\sin\theta)^2 \end{pmatrix}, \tag{13.35}$$

or equivalently for the line element (13.9) we find

$$ds^2 = -\left(1 - \frac{r_S}{r}\right) c^2\, dt^2 + \frac{1}{1 - \frac{r_S}{r}}\, dr^2 + r^2\, d\theta^2 + r^2\,(\sin\theta)^2 d\varphi^2\,. \tag{13.36}$$

This metric is called the *Schwarzschild solution* of Einstein's equation.

## 13.4. Properties

For very large $r \gg r_S$, the Schwarzschild metric (13.35) can be approximated by

$$g_{\mu\nu}(x) \approx \begin{pmatrix} -\left(1 - \frac{r_S}{r}\right) & 0 & 0 & 0 \\ 0 & \left(1 + \frac{r_S}{r}\right) & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\,(\sin\theta)^2 \end{pmatrix} \qquad (\text{for } r \to \infty)\,. \tag{13.37}$$

In particular, it tends to the Minkowski metric (in spherical coordinates) (13.4) in the limit $r \to \infty$. In technical jargon, one says that the Schwarzschild metric is *asymptotically flat*. Comparing this to the metric in the Newtonian limit (12.34), we see that for $r \gg r_S$ the Schwarzschild metric is described by the Newtonian gravitational potential

$$\Phi = -\frac{c^2\, r_S}{2\, r}\,. \tag{13.38}$$

From Newtonian gravitation, we know that a mass density $\rho(r)$ produces a gravitational potential of the form

$$\Phi = -\frac{G\, M_{\text{tot}}}{r} \qquad (\text{for } r > r_{\text{max}})\,, \tag{13.39}$$

where $M_{\rm tot}$ is the total mass (13.15) and $G$ is Newton's constant. Comparing these two expressions, we find that the Schwarzschild radius is related to the total mass via

$$r_S = \frac{2\,G\,M_{\rm tot}}{c^2}\,. \tag{13.40}$$

This provides us with a physical interpretation of the Schwarzschild radius.

One can compute the Riemann curvature tensor (13.25) for the Schwarzschild solution. It is non-vanishing for $r_S \neq 0$ (i.e. $M_{\rm tot} \neq 0$), hence the Schwarzschild solution is a curved spacetime. In particular, we note (without going into the details of the calculation) that the quantity

$$R_{\mu\nu\rho\sigma}\,R^{\mu\nu\rho\sigma} = 12\,\frac{r_S^2}{r^6} \tag{13.41}$$

called *Kretschmann scalar* is non-zero.

We conclude this chapter with a few remarks about the Schwarzschild radius $r_S$. Notice that the metric tensor (13.35) becomes singular for $r \to r_S$. Because our solution was derived under the assumption that $r > r_{\rm max}$, where $r_{\rm max}$ is the maximal radius in which the mass density is localized, there will be no problem for the case where $r_S < r_{\rm max}$. Using (13.40), we can compute the Schwarzschild radius for our earth and sun:

Earth: $r_S = 8.87 \times 10^{-3}$ m

Sun: $r_S = 2.95 \times 10^{3}$ m

This is much smaller than $r_{\rm max}$ of these astrophysical objects, which is given by the radius of earth/sun:

Earth: $r_{\rm max} = 6.37 \times 10^{6}$ m

Sun: $r_{\rm max} = 6.96 \times 10^{8}$ m

'Normal' astrophysical objects, such as planets and stars, always satisfy $r_S < r_{\rm max}$, so there will occur no problems with the metric singularity at $r = r_S$.

In the opposite case where $r_S > r_{\rm max}$ one speaks of a *black hole*. The singularity at $r = r_S$ turns out to be just a coordinate singularity in the sense that one can find another set of 'better' coordinates to study the Schwarzschild solution also inside the region $r < r_S$. However, the surface defined by $r = r_S$ – called the *event horizon* – has very interesting features: For any particle or light ray inside the event horizon it is impossible to cross it, independently of how much force we apply. This is where the name *black* in black hole comes from: Once inside the event horizon, not even light can escape the black hole anymore, so it looks black! Due to time constraints, we cannot provide a solid treatment of black holes in this module. In case you would like to understand the details of what I said above, you should take the module *MATH4016 Black Holes* offered by the University of Nottingham!

CHAPTER 14

# Gravitational Time Dilation

Gravitational time dilation is the general relativistic physical effect that clocks within a gravitational field tick slower than clocks outside of gravitational fields. We shall now provide a derivation of this effect.

Let us consider as an example the Schwarzschild spacetime, i.e. the line element

$$ds^2 = -\Big(1 - \frac{r_S}{r}\Big)\, c^2\, dt^2 + \frac{1}{1 - \frac{r_S}{r}}\, dr^2 + r^2\, d\theta^2 + r^2\, (\sin\theta)^2 d\varphi^2\ . \qquad (14.1)$$

Let $x^\mu : \mathbb{R} \to M\ ,\ \tau \mapsto x^\mu(\tau)$ be the world-line of a particle which is at rest at some fixed position with radius $r > r_S$. (Due to spherical symmetry, we can choose any fixed value $\theta$ and $\varphi$ for the angles at which the particle sits at rest.) Recalling the definition of proper time $\tau$, i.e.

$$-c^2\, d\tau^2 = ds^2\ , \qquad (14.2)$$

we obtain

$$\begin{aligned} -c^2\, d\tau^2 &= \left(-\Big(1 - \frac{r_S}{r}\Big)\, c^2 + \frac{1}{1 - \frac{r_S}{r}}\Big(\frac{dr}{dt}\Big)^2 + r^2\Big(\frac{d\theta}{dt}\Big)^2 + r^2\, (\sin\theta)^2\Big(\frac{d\varphi}{dt}\Big)^2\right) dt^2 \\ &= -\Big(1 - \frac{r_S}{r}\Big)\, c^2\, dt^2\ . \end{aligned} \qquad (14.3)$$

In the second step we used that the particle is at rest, i.e. the time derivatives of $r$, $\theta$ and $\varphi$ vanish. This provides us with the relation

$$\Delta\tau(r) = \sqrt{1 - \frac{r_S}{r}}\ \Delta t \qquad (14.4)$$

between a time interval $\Delta t$ of coordinate time and the corresponding proper time interval $\Delta\tau$. Notice that this relation depends on the radial position $r$ of our particle. For $r \to \infty$, the proper time interval coincides with the coordinate time interval, i.e.

$$\Delta\tau(\infty) := \lim_{r\to\infty} \Delta\tau(r) = \Delta t\ . \qquad (14.5)$$

Hence, one should physically interpret the coordinate time $t$ in the Schwarzschild spacetime as the time measured by an observer that is at rest very, very far away from the mass density.

Let us now consider the ratio

$$\frac{\Delta\tau(r)}{\Delta\tau(\infty)} = \sqrt{1 - \frac{r_S}{r}} \qquad (14.6)$$

of the time measured by an observer at $r$ and the time measured by an observer at $r \to \infty$. Plotting this as a function of $r$, we obtain the behavior visualized in Figure 14.1. We notice the following feature: The time interval $\Delta\tau(r)$ increases monotonically in $r$ and approaches (of course) $\Delta\tau(\infty)$ in the limit $r \to \infty$. The physical interpretation

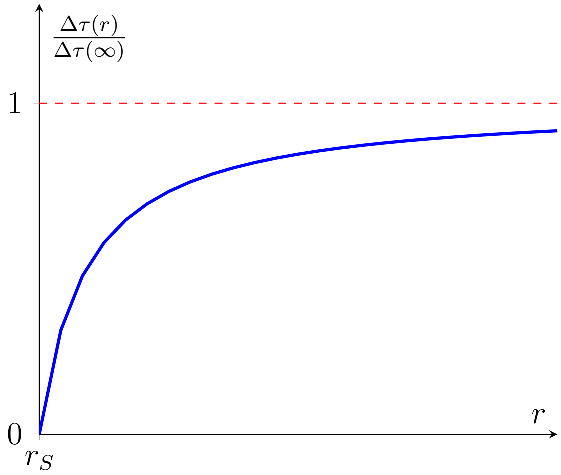


FIGURE 14.1. Radius dependence of the ratio $\frac{\Delta\tau(r)}{\Delta\tau(\infty)}$ of the time measured by an observer at $r$ and an observer at $r \to \infty$ for the Schwarzschild spacetime.

is as follows: The stronger the gravitational field, i.e. the smaller $r$, the slower the clock is ticking; this means there is *gravitational time dilation*.

Gravitational time dilation is a generic physical effect of general relativity and not only a particular feature of the Schwarzschild spacetime. We just used the Schwarzschild spacetime as an example where the relevant calculations are very easy to do. The effect of gravitational time dilation has been confirmed experimentally by using atomic clocks in gravitational fields.