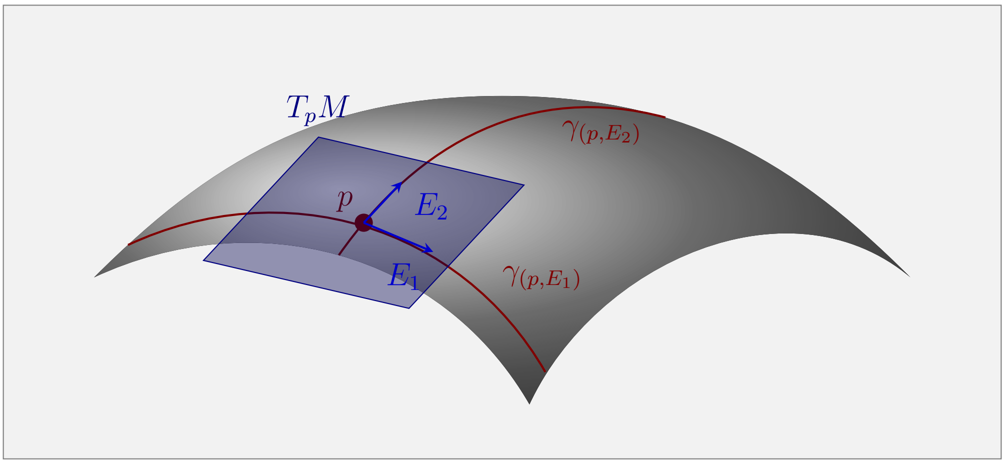
General Relativity

An introduction



Dominik Tschimmel

This page intentionally left blank.

Preface

The title might be a misnomer, as the focus of these notes is not the theory of general relativity, but rather the mathematics behind it. However, the basic concepts of a beginners course on general relativity are included. An exception is the chapter about special relativity, where the attempt was made to give a description that is as coordinate free as possible, with methods from (modern) differential geometry. The longest part, by far, is the appendix, covering the formal definition of tensors and Riemannian geometry amongst other things. Also a noteworthy chapter is the last one, introducing vector bundle valued differential forms almost from scratch.

Some parts of these notes were created, at the time I was learning the corresponding concepts. So be aware, that there may not only be the usual typos, but possibly wrong statements. In that sense, read with caution. However, nothing presented here is new, and usually well covered in textbooks.

This page intentionally left blank.

Contents

Preface i

1. Introduction 1
1.1. Galilean spacetime . . . 1
1.2. Newtonian gravitation . . . 2

2. Special relativity 5
2.1. Spacetime of special relativity . . . 5
2.1.1. Minkowski space and Poincaré group . . . 5
2.1.2. Lorentz boosts . . . 6
2.2. Minkowski diagrams . . . 9
2.3. Time dilation and Lorentz contraction . . . 10
2.4. Causality and Minkowski metric . . . 11
2.5. Dynamics in Minkowski spacetime (and four vectors) . . . 12
2.5.1. Mathematical background of four vectors . . . 12
2.5.2. Four vectors and kinematics . . . 14
2.5.3. Invariant description of velocity . . . 15
2.5.4. Four vectors and dynamics . . . 17
2.6. The geometry of proper time . . . 18

3. Einstein field equations 22
3.1. Physical motivation . . . 22
3.1.1. The postulates of general relativity . . . 22
3.1.2. Equivalence principle . . . 23
3.1.3. Equivalence principle and curvature . . . 25
3.2. Variational principle and energy-momentum tensor . . . 25
3.2.1. Einstein-Hilbert-Action review . . . 25
3.2.2. Energy-momentum tensor . . . 26

4. Gravitational waves in linear approximation 28
4.1. Linerized gravitation . . . 28
4.1.1. Linear limit . . . 28
4.1.2. Harmonic coordinates and trace reversed perturbation . . . 29
4.2. Propagation of free waves . . . 32
4.3. Radiation of gravitational waves . . . 34

5. Schwarzschild solution 36
5.1. The Schwarzschild metric . . . 36
5.2. Movement in a Schwarzschild system . . . 37

A. Tensors and Index-Notation 40
A.1. Tensor product . . . 40
A.1.1. Existence and uniqueness . . . 40
A.1.2. Tensors . . . 43
A.2. Ricci calculus . . . 47
A.2.1. Co-and contravarianz . . . 47
A.2.2. Tensors in Ricci calculus . . . 48

A.2.3. Raising and lowering indices 49

B. Variational calculus for fields **51**
B.1. Functional derivative and variation 51
B.2. Variation of Lagrange densities 53
B.3. Physical conventions and notations 56

C. Riemannian geometry **57**
C.1. Motivation: Geodesic equation 57
C.2. Connections 58
C.2.1. Vector bundles 58
C.2.2. Linear Connections 61
C.3. Levi-Civita-connection 68
C.3.1. Lie-derivative 68
C.3.2. Torsion and metric compatibility 74
C.4. Geodesics 77
C.4.1. Vector fields along curves 78
C.4.2. Geodesics and parallel transport 79
C.4.3. Riemannian normal coordinates 82
C.5. Curvature 85
C.5.1. Riemannian curvature 85
C.5.2. Symmetries of the Riemann tensor 86
C.5.3. Ricci and scalar curvature 88
C.6. Killing fields 89

D. Calculations in coordinates **92**
D.1. Determinant of the metric tensor 92
D.1.1. Definition: metric determinant 92
D.1.2. Derivatives of the metric determinant 93
D.2. Densities 95
D.2.1. Mathematical Density (short review) 95
D.2.2. Densities in Physics 95
D.2.3. Metric determinant and densities 96

E. Vector valued differential forms **99**
E.1. Natural bundle operations 99
E.2. Definition of vector bundle valued differential forms 101
E.3. Operations on vector forms 103
E.4. Connections and exterior derivative 106
E.5. Connection- and curvature forms 110

Index **114**

Bibliography **117**

1 Introduction

Before discussing the theory of general relativity, we outline the basic principles of the classical spacetime formulation, as can be found in [Zir15a]. Inspired by this, we briefly review Newtonian gravity from a differential-geometrical perspective.

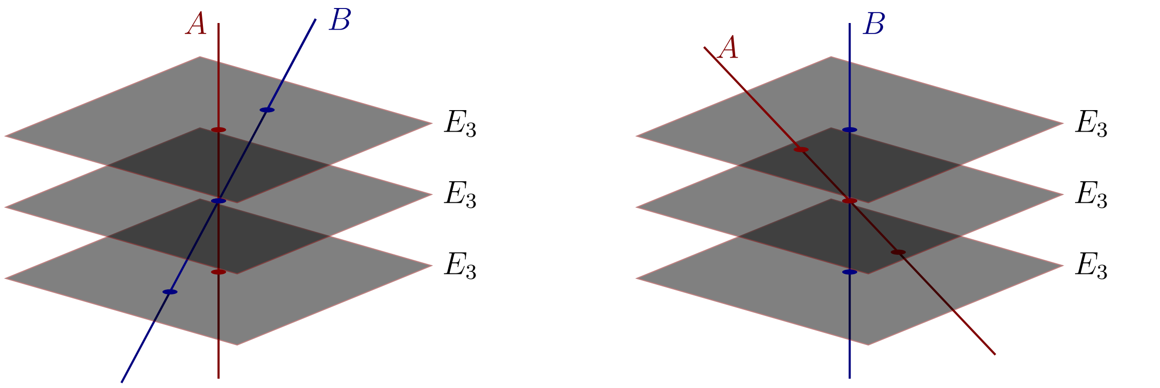
1.1. Galilean spacetime

The basis of all physics is the underlying concept of time and space. In classical mechanics the model for these concepts is the Galilean spacetime.

Definition 1.1.1.
The **Galilean spacetime** is a 4 dimensional affine space[1] $(M, V, +)$ with the following properties:

i) There is a unique linear form τ on V, called **absolute time**.
ii) The subspace of spacial translations $V_0 := \{v \in V | \tau(v) = 0\}$ is Euklidean.

Points in spacetime $p \in M$ are called **events**. Due to the absolute time, simultaneity has a well defined meaning, i.e. two events $p, q \in M$ are simultaneous if $\tau(p - q) = 0$. There is no absolute space however. That is, there is no unique construction of a vector space V_t with $V = V_t \oplus V_0$, splitting V in space an time directions. Thus the choice of V_t is arbitrary. Physically speaking, the choice of V_t is the choice of an inertial system.



(a) Being in system A, perceiving system B moving with velocity v.

(b) Being in system B, perceiving system A moving with velocity $-v$.

Figure 1.1.: There is no unique decomposition of V in $V_t \oplus V_0$. Both systems are equal according to the principle of relativity.

[1] More generally a 4-dim flat manifold M. The difference vector space V is the tangent space and the addition $+\colon M \times V \to M$ is the parallel transport.

There is a group of transformations, preserving the structure of Galilean spacetime, called **Galileo group** G. This group describes the symmetries of the classical spacetime.

Definition 1.1.2 (Properties of Galileo transformations)**.**
Let $g\colon M \to M$ be a Galileo transformation, then the following properties hold:

i) Preservation of affine structure:
$g(p) = g(\sigma) + D_\sigma g(p - \sigma) \quad \forall p \in M.$

ii) Preservation of absolute time: $\tau \circ D_\sigma g = \tau$.

iii) Preservation of metric:
$\langle D_\sigma g|_{V_0} v, D_\sigma g|_{V_0} v' \rangle = \langle v, v' \rangle \quad \forall v, v' \in V_0.$

The formulation of mechanical laws is unnecessarily complicated in M. Therefore, classical mechanics is usually described in an inertial system. This allows to reduce M to E_3.

1.2. Newtonian gravitation

Consider a mass point with mass M at position p_1. If there is a second mass point m at position p_2, then there is a force between the two mass points.According to Newton's theory of gravitation, the force that mass 1 exerts on mass 2 is given by

$$\vec{F}_{1 \text{ on } 2} = \frac{GMm}{\|p_1 - p_2\|^3}(p_1 - p_2) \ .$$

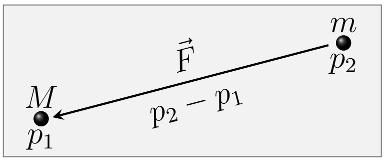


Figure 1.2.

If we fix the position of the first mass and move the second mass around, we find, that there is a force at every point in space. A more modern formulation of physics requires forces to be differential 1 forms.[2] It is easiest to use spherical coordinates centered at p_1:

$$\vec{F} = -\frac{GMm}{r^2}\vec{e_r} \qquad \Rightarrow \qquad F = -\frac{GMm}{r^2} dr \ .$$

In electrostatics there is an underlying Field E, creating the force on test charges. The exact value of the force depends on the charge q of the test charge, i.e. $F = qE$. Inspection of the force field created by mass points shows, that the mass m behaves like the charge:

$$F = m \cdot g \qquad \text{with} \qquad g = -\frac{GM}{r^2} dr \ .$$

The field g is called **Newtonian gravitational field**.

It can easily be seen, that there are functions f so that $df = g$ holds. Using the physical sign conventions and usual boundary conditions for potentials results in

$$\Phi(p) = -\frac{GM}{r(p)} \qquad \Rightarrow \qquad g = -d\Phi \ .$$

[2] The easiest way to see this, is the definition of work $W = \int_\gamma F$, where F is the force 1 form, the natural integrand over paths γ.

The equation $g = -d\Phi$ may be more familiar in classical vector calculus : $\vec{g} = -\vec{\nabla}\Phi$.

So far, we have only discussed idealized mass points. Nevertheless this treatment of gravity is justified. Before we can proof this, we need to find a description of finite mass distributions.

Definition 1.2.1.
The **mass density** ρ is a 3 form with the property, that $\int_U \rho$ is the mass of the described mass distribution in the region U.

Comparing the mass density and the sum of mass points shows, how to generalize the additivity of gravitational potentials:

$$\Phi(p) = -\sum_i \frac{GM_i}{r_{p_i}(p)} \qquad \xrightarrow{\text{continuous limit}} \qquad \Phi(p) = -\int_{E_3} \frac{G\rho}{r_p} \ .$$

The distance function r_p is defined by $r_p(q) = \|p - q\|$.

Lemma 1.2.2.
For any mass distribution there is a ***poisson equation****:*

$$\Delta\Phi = 4\pi G \star \rho \ .$$

Proof 1.2.3.
Contrary to electrostatics we do not have the differential equation connecting the mass density with a gravitational field yet. To obtain such a relation we integrate the 2 form $\star g$ over the surface ∂U of a volume U [3]enclosing the whole mass:

$$\begin{aligned}\int_{\partial U} \star g &= -\int_{\partial U} \frac{GM}{r_p^2} \star dr_p = -\int_{\partial U} \frac{GM}{r_p^2} r_p^2 \tau_p = -GM \int_{\partial U} \tau_p \\ &= -4\pi GM \ ,\end{aligned}$$

where we used the solid angle 2 form $\tau_p = \sin(\theta)[d\theta \wedge d\phi, R]$. With the definition of the mass density we find:

$$\int_{\partial U} \star g = -4\pi G \int_U \rho \ .$$

In 3 dimensions any 3 form is closed. By the lemma of Poincaré there is a 2 form $\mathfrak{g}$ such that $d\mathfrak{g} = \rho$. With Stoke's theorem we find:

$$\int_{\partial U} \star g = -4\pi G \int_U d\mathfrak{g} = \int_{\partial U} \mathfrak{g} \ .$$

Since U is chosen arbitrarily, the integrands are equal. Thus we found an equivalent to **Gauss' law for gravity**:

$$\boxed{d \star g = -4\pi G\rho} \ .$$

In odd dimensions the hodge star operator is self inverse $\star^{-1} = \star$. If we also use $-d\Phi = g$ we get:

$$-d \star d\Phi = -4\pi G\rho \qquad \Leftrightarrow \qquad \star d \star d\Phi = 4\pi G \star \rho \ .$$

The Laplacian for functions can be written as $\Delta = \star d \star d$, which concludes the proof. □

We close this subsection by proving the first part of Newton's shell theorem:

Theorem 1.2.4 (Newton's shell theorem)**.**
The external gravitational field of a spherically symmetric mass distribution is equal to the gravitational field of a mass point at the center of the mass distribution with the same total mass.

Proof 1.2.5.
The symmetry of the problem gives a radial ansatz for the gravitational field, $g = f(r_p)dr_p$. Integration over a ball $B_R(p)$ with radius R and the definition of the mass density yield by Stokes's theorem:

$$\begin{aligned} M &= \int\limits_{B_R(p)} \rho = -\frac{1}{4\pi G} \int\limits_{B_R(p)} d \star g = -\frac{1}{4\pi G} \int\limits_{\partial B_R(p)} \star g \\ &= -\frac{1}{4\pi G} \int\limits_{\partial B_R(p)} f(r_p) r_p^2 \tau_p = -\frac{R^2 f(R)}{4\pi G} \int\limits_{\partial B_R(p)} \tau_p = -\frac{R^2 f(R)}{G} \ , \\ &\Rightarrow \qquad f(R) = -\frac{GM}{R^2} \qquad \Rightarrow \qquad g = -\frac{GM}{r_p^2} dr_p \ . \end{aligned}$$

□

[3] The integration of the solid angle 2 form always yields 4π for these surfaces.

2

Special relativity

As special case of general relativity, the theory of special relativity shows already some of the peculiar features of spacetime. In addition, special relativity is the limit of general relativity in the absence of gravity and high velocities. We outline the important structures, principles as well as mathematical formulation.

2.1. Spacetime of special relativity

In 1905 Einstein published his insights about, what is now called special relativity. Though not being the first to notice discrepancies between the relativity principle and Maxwell's equations, he was the first to realize the meaning of this for spacetime. Special relativity is founded on two postulates (which are very well supported by experiment):

1) Principle of relativity:
Physical laws assume the same form in all inertial systems.

2) Constancy of the speed of light:
The speed of light has in all inertial systems the same finite value c.

These postulates give rise to the astonishing results of special relativity. One of them, being the relativity of simultaneity, contrasting Galilean spacetime.

2.1.1. Minkowski space and Poincaré group

As for classical mechanics, there is a mathematical description for the spacetime concept of special relativity. However a coordinate free description is rather strenuous for an introduction. For the next four sections we will follow [Zir15b] and [Zir98] mostly.

Definition 2.1.1.
The **Minkowski space** M is a 4 dimensional manifold with affine structure $(M, V, +)$, together with a pseudo Riemannian metric g, called **Minkowski metric**.

This definition of spacetime does not specify the structure of special relativity yet. To do so, we describe properties in coordinates of inertial systems:

Definition 2.1.2.
Inertial systems are affine coordinates $\{x^0, x^1, \ldots, x^3\}$ fulfilling the following conditions:

i) $g(\partial_{x_\nu}, \partial_{x_\mu}) = 0$ for $\nu \neq \mu$.
ii) $g(\partial_{x^0}, \partial_{x^0}) = -1$ and $(\partial_{x^i}, \partial_{x^i}) = 1$ for $i = 1, 2, 3$.

iii) The orientation is given by $\mathfrak{O} = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$.

In these coordinates the Minkowski metric can be written as $(0, 2)$-tensor field:

$$g = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dx^3 \otimes dx^3 \ .$$

By now we have laid down the mathematical structure of spacetime. It remains to connect this structure to the physical postulates. To do so, we consider transformations that preserve the structure of the Minkowski space, mapping inertial systems to inertial systems.

Definition 2.1.3.
The **Poincaré group** is the group of affine maps $f \colon M \to M$, that have the following properties:

i) f preserves the metric, i.e. $g(D_p f v, D_p f w) = g(v, w) \quad \forall v, w \in T_p M$.
ii) f preserves the orientation, i.e. $f^* \mathfrak{O} = \mathfrak{O}$.

Being an affine map means, that we can write f as a shift $f(0)$ plus a linear map Λ:

$$f(p) = f(0) + \Lambda(p - 0) \ .$$

The linear maps Λ are called **Lorentz transformations** and form a subgroup of the Poincaré group, called **Lorentz group**. The shift is, as in classical mechanics, just a static change of origin (4 parameters as in classical mechanics), so we will only analyze Λ.

One can easily verify $\Lambda = Df$ by differentiation. Thus the conditions of Poincaré group take the following form:

$$1) \quad g(\Lambda v, \Lambda w) = g(v, w) \ \forall v, w \in T_p M \qquad\qquad 2) \quad \det(\Lambda) = 1 \ .$$

Let us separate space and time for the moment. Although we do not know the exact coefficients yet, we can assume x^i for $i = 1, 2, 3$ to be cartesian spacial coordinates, leaving x^0 to be a coordinate describing time. Since the Minkowski metric is equal to the familiar Euclidean metric on the spacial subspace, we know, that transformations in that subspace are ordinary rotations. This allows us, to reduce the problem to inertial systems only differing in motion in one direction (e.g. x^1 direction). These transformations are called **boosts**.

Remark 2.1.4.
It should be mentioned that the previous (and following) is no rigorous derivation of the Lorentz group. One needs to show, that boosts in arbitrary directions and rotations span the whole group. Although that would be an interesting topic in Lie group theory, we focus on the physical meaning here.

2.1.2. Lorentz boosts

Before we can give an explicit transformation map, it is important to know, what kind of transformation we want to use. There are two equal transformations in physics:

passive transformation

$$\begin{array}{ccc} M & \xrightarrow{Id} & M \\ x\downarrow & & \downarrow y \\ \mathbb{R}^4 & \xrightarrow[\Psi]{} & \mathbb{R}^4 \end{array}$$

In passive transformations the spacetime remains the same. Instead one chooses different coordinate systems x, y for different inertial systems. There is a transition function between systems, given by:

$$\Psi = y \circ x^{-1}$$

active transformation

$$\begin{array}{ccc} M & \xrightarrow{\Phi} & M \\ x\downarrow & & \downarrow x \\ \mathbb{R}^4 & \xrightarrow[\widetilde{\Phi}]{} & \mathbb{R}^4 \end{array}$$

Active transformations on the other hand use a fixed coordinate system, yet transform spacetime accordingly:

$$\widetilde{\Phi} = (x \circ \Phi) \circ x^{-1}$$

The connection between both transformations is the following condition:

$$\Psi = \widetilde{\Phi} \qquad \Leftrightarrow \qquad y \circ x^{-1} = x \circ \Phi \circ x^{-1} \qquad \Leftrightarrow \qquad y = x \circ \Phi \ .$$

To follow the predominant part of the literature, we will use passive transformations here. There are various derivations of the Lorentz transformations in different mathematical complexities. We skip the derivation for that reason. Let (t, x^i) be an inertial system I and (t', x'^j) another inertial system I' moving in x^1 direction with velocity v. For simplicity assume, that both systems for $t = t' = 0$ were at the same position, with the same axis directions.[1] The relations between these coordinates are:

$$\boxed{\quad t' = \frac{t - \frac{v}{c^2}x^1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad\qquad x'^1 = \frac{x^1 - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad\qquad \begin{array}{l} x'^2 = x^2 \\ x'^3 = x^3 \end{array} \quad}$$

It is convenient to define a symbol for commonly appearing terms, e.g. the **$\boldsymbol{\gamma}$-factor**:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \Rightarrow \qquad t' = \gamma(t - \tfrac{v}{c^2}x) \quad \text{and} \quad x'^1 = \gamma(x^1 - vt) \ .$$

So far we have found, that x^0 is a time coordinate. Yet it could only be a function of time, i.e. $x^0 = f(t)$. However that is not the case. Lorentz boosts are linear maps, that map inertial systems to inertial systems. Thus, these maps should be contained in the Lorentz group for the theory to be useful.[2] For x^0 we make the ansatz $x^0 = c \cdot t$. We find (only considering t and x here):

$$-d(ct') \otimes d(ct') + dx' \otimes dx' = -c^2 d\left(\gamma(t - \tfrac{v}{c^2}x)\right) \otimes d\left(\gamma(t - \tfrac{v}{c^2}x)\right)$$

[1] In I' the system I moves with velocity $-v$.

[2] The implicit reason to consider it here.

$$\begin{aligned}
&+ d(\gamma(x - vt)) \otimes d(\gamma(x - vt)) \\
&= -c^2\gamma^2 \left(dt - \tfrac{v}{c^2}dx\right) \otimes \left(dt - \tfrac{v}{c^2}dx\right) + \gamma^2(dx - vdt) \otimes (dx - vdt) \\
&= \gamma^2(v^2 - c^2)dt \otimes dt + \gamma^2(1 - \tfrac{v^2}{c^2})dx \otimes dx \\
&\quad + (\gamma^2 v - \gamma^2 v)(dx \otimes dt + dt \otimes dx) \\
&= -c^2\gamma^2(1 - \tfrac{v^2}{c^2})dt \otimes dt + \gamma^2(1 - \tfrac{v^2}{c^2})dx \otimes dx \\
&= -d(ct) \otimes d(ct) + dx \otimes dx \ .
\end{aligned}$$

Thus $x^0 = c \cdot t$ is the wanted zeroth coordinate, such that Lorentz boosts preserve the Minkowski metric.

Lemma 2.1.5.
The Lorentz boosts can be written in terms of an angle θ:

$$t' = \cosh(\theta)\ t - \frac{\sinh(\theta)}{c}\ x \qquad \textit{and} \qquad x' = \cosh(\theta)\ x - c\ \sinh(\theta)\ t \ .$$

Proof 2.1.6.
In the real numbers it holds that $\gamma \in [1, \infty)$. Since $\cosh(\mathbb{R}) = [1, \infty)$ the following identification is possible:

$$\cosh(\theta) = \gamma.$$

The relation between the hyperbolic functions is $\cosh(\theta)^2 - \sinh(\theta)^2 = 1$. Thus one can show

$$\sinh(\theta) = \frac{v}{c}\gamma \ .$$

Applying the relation of cosh and sinh as well as the identification $\cosh(\theta) = \gamma$ shows after a little calculation the assertion. □

Corollary 2.1.7.
The quotient $\frac{v}{c}$ is connected to the angle θ, by $\tanh(\theta) = \frac{v}{c}$. This quantity is called ***rapidity****.*

Reintroducing x^0 allows to write the Lorentz boosts in a rather symmetrical form:

$$\begin{pmatrix} p'^0 \\ p'^1 \\ p'^2 \\ p'^3 \end{pmatrix} = \begin{pmatrix} \cosh(\theta) & -\sinh(\theta) & 0 & 0 \\ -\sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p^0 \\ p^1 \\ p^2 \\ p^3 \end{pmatrix}$$

Remark 2.1.8 (coordinates vs chart maps)**.**
It is common practice in the literature to identify the chart maps with the coordinates. To illustrate that, we consider a point $p \in M$. The coordinates of this point are $p^i = x^i(p)$. The identification of the literature is now to write $x^i(p) = x^i$. Usually this is no problem, yet one can easily confuse transformations.[3]

2.2. Minkowski diagrams

Reducing spacial dimensions to one allows to illustrate the consequences of special relativity. The standard representations are **Minkowski diagrams**, where x^0 and x^1 are plotted. The bisection describes a system moving with the speed of light[4], and has to be invariant in all coordinate systems. For that reason, the bisection gets a special name: **light cone**.

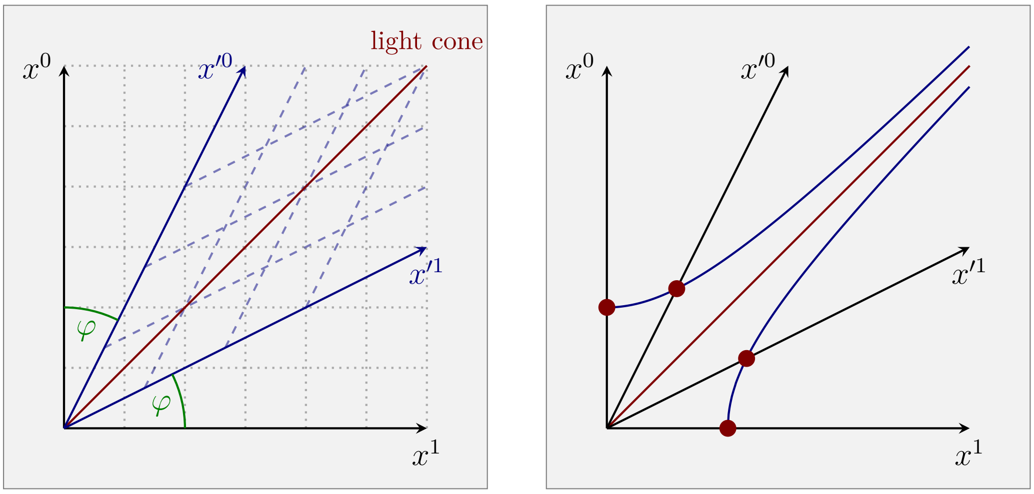


Figure 2.1.

The x'^1-axis of a second inertial system I' consists of points with $x'^0 = 0$. Likewise, the x'^0-axis consists of points with $x'^1 = 0$. To find the axes in terms of x^0 and x^1, one solves the following system of linear equations for the respective conditions:

$$\begin{pmatrix} x'^0 \\ x'^1 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} \\ -\gamma\frac{v}{c} & \gamma \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix}$$

One can show, that the angle φ between the x'^0- and x^0-axis is the same as the angle between the x'^1- and x^1-axis. Drawing the parallel lines to obtain the coordinate grid of I', one can easily see, that the light cone remains the bisection.

The reduced Minkowski metric is $g = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1$. As we have shown, this metric is an invariant of Lorentz boosts:

$$\begin{aligned} g &= -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 \\ &= -dx'^0 \otimes dx'^0 + dx'^1 \otimes dx'^1 \ . \end{aligned}$$

[3] Describing a particle with $x(t)$, one is tempted to transform this as $x'(t') = \gamma x(\gamma(t - \frac{v}{c^2}x)) - \gamma vt$. As one can see, an expression like $x(t, x)$ is meaningless. The correct transformation is:

$$\begin{pmatrix} s' \\ f(s)' \end{pmatrix} = \Lambda \begin{pmatrix} s \\ f(s) \end{pmatrix} = \begin{pmatrix} \gamma(s - \frac{v}{c^2} f(s)) \\ \gamma(f(s) - vs) \end{pmatrix} \ .$$

[4] $1 = \frac{x^1}{x^0} = \frac{x^1}{ct} = \frac{1}{c} v \qquad \Leftrightarrow \qquad v = c.$

A direct consequence is the constancy of the speed of light. Utilizing the affine structure of M allows to reframe the statement in a more obvious form. Any event $p \in M$ can be written as $p = 0 + v$ for $v \in T_0M$. Thus we can consider $x^i(p)$ as coefficients of a (tangent) vector. Calculating the negative length with $\|v\|^2 = g(v, v)$ yields:

$$-\|v\| = -g(v,v) = \left(x^0(p)\right)^2 - \left(x^1(p)\right)^2 = \left(x'^0(p)\right)^2 - \left(x'^1(p)\right)^2 \ .$$

A more common form of this equation is:

$$\boxed{t^2 - \frac{x^2}{c^2} = t'^2 - \frac{x'^2}{c^2}} \ .$$

This identity allows to find the unit points on the coordinate axes of I'. The unit point of the x'^1-axis is $(x'^0 = 0, x'^1 = 1)$. Inserting in the above equation results in $-1 = (x^0)^2 - (x^1)^2$. The equation of time-axis unit points is $1 = (x^0)^2 - (x^1)^2$. Rewriting in terms of $x^0(x^1)$ allows to plot these functions:

$$\begin{aligned} \text{time:} \quad x^0(x^1) &= \sqrt{1 + (x^1)^2} \ , \\ \text{space:} \quad x^0(x^1) &= \sqrt{-1 + (x^1)^2} \ . \end{aligned}$$

2.3. Time dilation and Lorentz contraction

The coordinate axes of I' are no longer orthogonal. Differences between points thus can get stretched or compressed. To describe these effects without creating confusion we need to introduce some vocabulary. The **rest frame** is an inertial system, that moves along the object one wants to describe (i.e. co-moving). The time coordinate of the rest frame is called **proper time**. A length measured with respect to coordinates of the rest frame is similarly called **proper length**.

The difference between two events are vectors in affine spaces (or tangent vectors on manifolds in the proper limit). Proving changes in length and time scales thus is best done in the context of (tangent) vectors. As a reminder: Let $\{y^\mu\}$ and $\{x^\nu\}$ be coordinate charts. The basis (tangent) vectors corresponding to those coordinates transform as follows:

$$\frac{\partial}{\partial y^\mu} = \sum_\nu \frac{\partial x^\nu}{\partial y^\mu} \frac{\partial}{\partial x^\nu} \qquad \text{also notice:} \qquad dy^\mu = \sum_\nu \frac{\partial y^\mu}{\partial x^\nu} dx^\nu \ .$$

Theorem 2.3.1.
Let I' be an inertial system moving with velocity v, $\mathfrak{L}$ the proper length of an object and $\mathfrak{T}$ the proper time of a time interval. A measurement of these quantities from system I yields

$$T = \gamma \mathfrak{T} \qquad \textit{and} \qquad L = \tfrac{1}{\gamma} \mathfrak{L} \ .$$

Proof 2.3.2.
A time interval of length is given by $\mathfrak{T}\frac{\partial}{\partial\tau}$. Expressing the (tangent) vector in the coordinates of system I (notice that $\tau = t'$) one finds:

$$\mathfrak{T}\frac{\partial}{\partial\tau} = \mathfrak{T}\left(\frac{\partial\gamma\cdot(t' + \frac{v}{c^2}x')}{\partial t'}\frac{\partial}{\partial t} + \frac{\partial\gamma\cdot(x' + vt')}{\partial t'}\frac{\partial}{\partial x}\right) = \gamma\mathfrak{T}\frac{\partial}{\partial t} + v\gamma\mathfrak{T}\frac{\partial}{\partial x} \ .$$

Notice, that we had to use the inverse Lorentz transformation. By calculation or physical reasoning one finds, that the substitution $v \to -v$ is enough. The interpretation of the equation above is as follows: An event that moves for an interval $\mathfrak{T}$ in the future in system I', moves the distance $v\gamma\mathfrak{T}$ in positive x-direction and $\gamma\mathfrak{T}$ into the future for observers in I, which is to say $T = \gamma\mathfrak{T}$. Since I' is moving relatively to I, it is hardly surprising, that the event moved.

To measure lengths, there is the additional hardship of simultaneity of measurements between starting and ending points (in one system). That is, we need to find a constant ξ, such that $\xi\frac{\partial}{\partial t'} + \frac{\partial}{\partial x'} \sim \frac{\partial}{\partial x}$ holds.

$$\xi\frac{\partial}{\partial t'} + \frac{\partial}{\partial x'} = \left(\xi\gamma + \gamma\frac{v}{c^2}\right)\frac{\partial}{\partial t} + (\xi v\gamma + \gamma)\frac{\partial}{\partial x} \overset{!}{=} (\xi v\gamma + \gamma)\frac{\partial}{\partial x}$$

$$\Rightarrow \qquad \xi\gamma + \gamma\frac{v}{c^2} = 0 \qquad \Leftrightarrow \qquad \xi = -\frac{v}{c^2}$$

$$\mathfrak{L}\left(\frac{\partial}{\partial t'} + \xi\frac{\partial}{\partial x'}\right) = \mathfrak{L}\left(-\gamma\frac{v^2}{c^2} + \gamma\right)\frac{\partial}{\partial x} = \mathfrak{L}\gamma\left(1 - \frac{v^2}{c^2}\right)\frac{\partial}{\partial x} = \frac{1}{\gamma}\mathfrak{L}\frac{\partial}{\partial x}$$

□

The theorem can be verbalized as follows: Moving clocks tick at faster rates (**time dilation**), and moving scales shrink (**Lorentz contraction**). It should be noticed, that these effects are reciprocal, expressing the symmetry between inertial systems.

2.4. Causality and Minkowski metric

We have commenced this section, by introducing the Minkowski metric and finding the right expression for x^0 in terms of classical coordinates. It remains to connect these definitions to the second postulate of special relativity, and to highlight some of the more unintuitive results. First of all, since the Minkowski spacetime has an affine structure, all points p can be written as sum of the origin 0 and a (tangent) vector $v(p)$. Hence it is justified to talk about $g(p, p)$ by meaning $g(v(p), v(p))$.

The Minkowski metric separates spacetime in three connected components. The speed of light is the limit for all movements of real particles/information. Thus particles can only reach the cone above the origin, and remain inside it. This region is the future, that information from the origin may influence. Points in that region (e.g. A) are called **time like**. The reason behind that naming convention is, that time like events are separated by time, but not necessarily by distance. That is, one can find inertial systems, in which these events lie parallel to the time axis. The past area, is similarly the union of all events, that could have caused the origin. As one can see easily, the condition for events p to be time like is $g(p, p) < 0$.

Conversely events with the property $g(p,p) > 0$ (e.g. C) are called **space like** events. One can find inertial systems such that space like events lie parallel to the space plane (3 coordinate axes). Thus space like events are simultaneous, i.e. the present.

Events that are located on the surface of the light cones (e.g. B) are called **light like** events. The condition $g(p,p) = 0$ is equal to a movement with the speed of light (which is the property of light, i.e. always moving with velocity c in all systems).

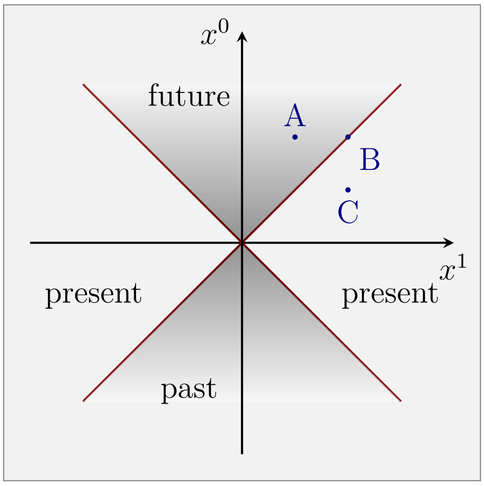


Figure 2.2.

2.5. Dynamics in Minkowski spacetime (and four vectors)

In classical mechanics time was an invariant quantity up to the choice of origin. Positions in spacetime were reduced to positions in space, depending on the time as parameter. In special relativity such a formulation is bothersome to hardly possible, since time coordinates have to be transformed. The solution of physics textbooks is called four vectors. In this section we try to connect the results of [Woo16, chapter 5] with differential geometrical methods.

2.5.1. Mathematical background of four vectors

Revisiting manifolds and charts will allow us to develop a proper notion of four vectors. So let M be a manifold. A chart x is a local diffeomorphism $x\colon U_x \subset M \to V_x \subset \mathbb{R}^n$. For simplicity we say $U_x = M$ and $V_x = \mathbb{R}^n$, as this holds for affine coordinates. Since $\mathbb{R}^n$ has a natural basis, one can define coordinate functions x^i by canonical projection: $x^i = \pi_i \circ x$. The canonical projection is exactly what one would expect:

$$\pi_i \left(\sum_i v^i e_i \right) = v^i \ .$$

Using passive transformations means to choose a second chart y. By definition there is a map connecting the coefficients of a point p:

$$p^i_y \overset{!}{=} \Psi^i(x(p)) \qquad \Leftrightarrow \qquad \pi_i \circ y = \pi_i \circ \Psi \circ x$$
$$\Leftrightarrow \qquad \Psi = y \circ x^{-1} \ .$$

For the context of Lorentz transformations we can assume Ψ to be linear, i.e. $y^i = \sum_j \Psi^i{}_j x^j$:

$$\frac{\partial}{\partial x^j} = \sum_i \frac{\partial y^i}{\partial x^j} \frac{\partial}{\partial y^i} = \sum_i \frac{\partial}{\partial x^j} \left(\sum_\ell \Psi^i{}_\ell y^\ell \right) \frac{\partial}{\partial y^i}$$

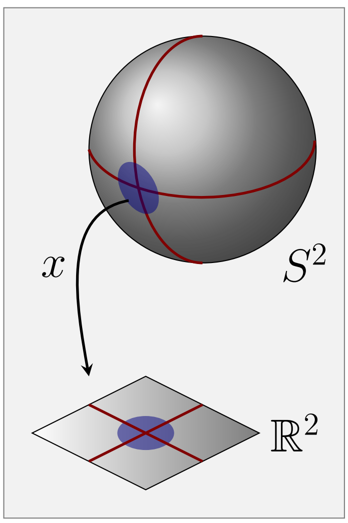


Figure 2.3.

$$= \sum_i \delta_{j\ell} {\Psi^i}_\ell \frac{\partial}{\partial y^i} = \sum_i {\Psi^i}_j \frac{\partial}{\partial y^i} \ .$$

Let v be a tangent vector with coefficients v_x^j, i.e. $v = \sum_j v_x^j \frac{\partial}{\partial x^j}$. In the case of linear Ψ these coefficients transform like the coefficients of points $p \in M$:

$$v = \sum_j v_x^j \frac{\partial}{\partial x^j} = \sum_{i,j} v_x^j {\Psi^i}_j \frac{\partial}{\partial y^i} = \sum_i \left(\sum_j {\Psi^i}_j v_x^j \right) \frac{\partial}{\partial y^i} \stackrel{!}{=} \sum_i v_y^i \frac{\partial}{\partial y^i} \ .$$

Thus, for linear transition functions Ψ, coordinate functions and coordinates of tangent vectors transform alike:

$$\boxed{y^i = \sum_j {\Psi^i}_j \, x^j \qquad \text{and} \qquad v_y^i = \sum_j {\Psi^i}_j \, v_x^j}$$

Remark 2.5.1.
Acceleration cannot be defined by a second derivation of parameters. One reason is, that the acceleration is not independent of the choice of coordinates. In physics, there is a way to circumvent that obstacle, by differentiation of basis vectors. This however is not possible in general (an affine structure is needed). The invariant definition (via connection) is the covariant derivative. In the case of affine coordinates and Minkowski metric, one can show, that the usual definition holds:[5]

$$\boxed{a^\mu(s) = \frac{d^2}{dt^2} \Gamma^\mu(s)}$$

[5]See remark C.4.4 for the proof.

2.5.2. Four vectors and kinematics

Definition 2.5.2.
Let q be an event in M and I an inertial system. The **four position** X of q is the coordinate representation of q: $X^\mu = x^\mu(q)$.

Since the components of four position are connected to inertial systems, there is a distinction between time and space components. The zeroth component is in time direction, whereas the latter three are the usual position in space. In that sense, it is convenient to write

$$X = \begin{pmatrix} x^0(p) \\ \vec{x}(p) \end{pmatrix} .$$

Note, that some authors write $\underline{X}$ for four position. In the mathematical digression, we discovered, that the vector of coordinates transforms like the coefficients of tangent vectors (at least for linear transformations). These objects are called **contravariant** and get upper indices. Dual objects, i.e. differential forms, which transform by the transposed matrix, are called **covariant** and get lower indices.[6] In case of covariance the notation of some authors is $\overline{X}$. We will not adopt that notation. To mark the tuple of spacial components, we will write $\vec{X}$ anyhow.

Most objects exist over some period of time, describing a curve in spacetime (not necessarily in space). Such a curve $\Gamma\colon I \subset \mathbb{R} \to M$ is called **world line** of that object. A parametrization defines a four position depending upon a parameter $s \mapsto X(s)$.

Our analysis of Lorentz boosts has been constrained to one dimension so far. Observing, that there are no spacial effects orthogonal to the boost velocity, separating $\vec{x}(s)$ in parallel $\vec{x}_\parallel(s)$ and orthogonal components $\vec{x}_\perp(s)$ allows to handle general Lorentz boosts:

$$\begin{pmatrix} x'^0(s) \\ \vec{x}\,'(s) \end{pmatrix} = \begin{pmatrix} \gamma \cdot (x^0(s) - \langle \vec{\beta}, \vec{x}_\parallel(s) \rangle) \\ \gamma \cdot (\vec{x}_\parallel(s) - \vec{\beta} x^0(s)) + \vec{x}_\perp(s) \end{pmatrix} \qquad \text{with} \qquad \begin{array}{l} \vec{x}_\parallel(s) = \frac{\langle \vec{x(s)}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}\, \vec{v} \\ \vec{x}_\perp(s) = \vec{x}(s) - \vec{x}_\parallel(s) \end{array} . \tag{2.1}$$

In the above transformation we also introduced the **$\boldsymbol{\beta}$-factor** $\vec{\beta} = \frac{\vec{v}}{c}$. Note, that:

$$\langle \vec{\beta}, \vec{x}_\parallel(s) \rangle = \frac{1}{c} \left\langle \vec{v}, \frac{\langle \vec{x(s)}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}\, \vec{v} \right\rangle = \frac{\langle \vec{x(s)}, \vec{v} \rangle}{c \cdot \langle \vec{v}, \vec{v} \rangle} \langle \vec{v}, \vec{v} \rangle = \frac{\langle \vec{x(s)}, \vec{v} \rangle}{c} .$$

Remark 2.5.3.
The curve parameter is invariant under Lorentz transformations, unlike the time coordinate $x^0(s)$. It is however possible for one inertial system, to identify the parameter as time, if $x^0(s) = s$. A natural, yet not necessary, choice for the parameter is the proper time τ.

[6] Informally, theories are also called covariant, when invariance (e.g. of equations) is meant.

Definition 2.5.4.
Let $\Gamma(\tau)$ be a curve in M parametrized by the proper time τ of the world line. The **four velocity** is the tangent vector of $\Gamma(\tau)$ with the following coordinate representation:

$$U^\mu(\tau) = dx^\mu \left(\frac{d}{d\tau} \Gamma(\tau) \right) .$$

Accordingly, the **four acceleration** is the second derivative of Γ with the following coordinate representation:

$$A^\mu(\tau) = dx^\mu \left(\frac{d^2}{d\tau^2} \Gamma(\tau) \right) = \frac{d}{d\tau} U^\mu(\tau) .$$

2.5.3. Invariant description of velocity

In the definition above, we used the textbook convention for the parameter to be the proper time. To develop a deeper understanding of the theory, we present the subject in a more general formulation.

A curve is a map $\Gamma \colon I \subset \mathbb{R} \to M$, mapping the parameter s to a point $\Gamma(s)$ on M. The tangent vector $\frac{d}{ds}\Gamma(s)$ is independent of coordinates. A coordinate parametrization defines functions $f_x^\mu \colon \mathbb{R} \to \mathbb{R}$ such that:

$$X^\mu(s) := x^\mu(\Gamma(s)) = f_x^\mu(s) .$$

Locally there are inverse functions g_x^μ such that $g_x^\mu \circ f_x^\mu = Id_{\mathbb{R}}$. In the physical context, we can assume f_x^0 to be bijective, that is time does always flow in one direction. Thus g_x^0 does always exist. In classical mechanics, the velocity is defined by $\vec{v}(t) = \frac{d}{dt}\vec{x}(t)$. In special relativity this translates to[7]:

$$\begin{aligned} \frac{v^i(s)}{c} := \frac{d}{dX^0(s)} X^i(s) &= \frac{d}{df_x^0(s)} f_x^i(g_x^0(f_x^0(s))) \\ &= \left(f_x^i\right)' (g_x^0(f_x^0(s))) \cdot \left. \frac{d}{dy} g_x^0(y) \right|_{y=f_x^0(s)} \\ &= \left(f_x^i\right)' (s) \cdot \left(g_x^0\right)' (f_x^0(s)) \\ &= \frac{d}{ds} X^i(s) \cdot \left(g_x^0\right)' (f_x^0(s)) \end{aligned}$$

Here we did not use more than the usual chain rule and plugging in our definitions. Yet this term does not appear initially in the tangent vector. However, we can simply divide by the right factor for that matter:

$$\boxed{\frac{d}{ds} X^i(s) = \frac{v^i(s)}{c} \cdot \frac{d}{ds} X^0(s)}$$

[7] Note, that $\frac{d}{dx^0} = \frac{1}{c}\frac{d}{dt}$ holds, by the same reasoning as below.

Proof 2.5.5.
To get the result we used the theorem about inverse functions and differentiations, which is:

$$\frac{d}{dx} f^{-1}(x) = \frac{1}{\frac{d}{dy} f(y)|_{y=f^{-1}(x)}} \ .$$

Using, that we defined $(f_x^0)^{-1} = g_x^0$ and $f_x^0 = X^0$ allows us to write:

$$\begin{aligned} \frac{1}{(g_x^0)'(f_x^0(s))} &= \frac{1}{\frac{d}{dy} g_x^0(y)|_{y=(g_x^0)^{-1}(s)}} = \frac{d}{ds}(g_x^0)^{-1}(s) = \frac{d}{ds} f_x^0(s) \\ &= \frac{d}{ds} X^0(s) \ . \end{aligned}$$

□

Remark 2.5.6 (about sloppy maths in textbooks)**.**
What we have proven so far, is written down in text books in a rather sloppy way, namely in expanding the differential fraction:

$$\frac{dX^i(s)}{ds} = \frac{dX^0(s)}{ds} \frac{dX^i(s)}{dX^0(s)} \ .$$

The problem of the usage of this formula (which is indeed correct) is, that one does not check, or even see, the conditions on the function X^0 to be a diffeomorphism (bijection that is differentiable with differentiable inverse).

To summarize our efforts, we state the results as a theorem:

Theorem 2.5.7.
The four velocity (with general parameter) has the following components:

$$\begin{pmatrix} U^0(s) \\ \vec{U}(s) \end{pmatrix} = \frac{d}{ds} \begin{pmatrix} X^0(s) \\ \vec{X}(s) \end{pmatrix} = \begin{pmatrix} \frac{d}{ds} X^0(s) \\ \frac{d}{ds} X^0(s) \cdot \frac{v^i(s)}{c} \end{pmatrix} ,$$

where $\vec{v}(s)$ is the classical three velocity, that would be measured in the inertial system.

As we have seen, the proper time is not necessary for the relativistic description. Nonetheless, using the proper time as parameter allows to write dynamical equations (involving mass and energy). The world line may describe accelerated motion. So there is no inertial system for the world line. To solve this problem one defines local inertial systems, for each point of the world line, called **instantaneous rest frames**. As illustrated in the figure, the world line always stays in the light cone. As tangent vector the, **four velocity is a time like vector**.

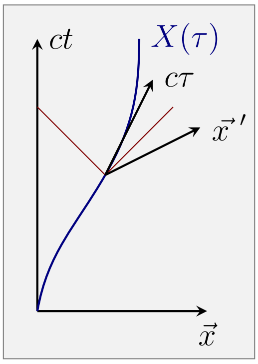


Figure 2.4.

To find the connection between four and three vectors, we need to find the relation between τ and t. Using theorem 2.3.1 we find:

$$\frac{dt(\tau)}{d\tau} = \gamma(\|\vec{v}(\tau)\|) \ .$$

It will also be important to know $\frac{d\tau}{dt}$. Using the theorem about inverse functions and differentiations yields:

$$\frac{d}{dt}\tau(t) = \frac{1}{\left.\frac{dt(\tau)}{d\tau}\right|_{\tau=t(\tau)}} = \frac{1}{\gamma(\|\vec{v}(t(\tau))\|)} =: \frac{1}{\gamma(\|\vec{v}(t)\|)} \ .$$

Lemma 2.5.8.
The components of the four velocity are $\begin{pmatrix} U^0(\tau) \\ \vec{U}(\tau) \end{pmatrix} = \begin{pmatrix} \gamma(\tau)c \\ \gamma(\tau)\vec{v}(\tau) \end{pmatrix}$.

Proof 2.5.9.
It is sufficient to notice that $X^0(\tau) = c \cdot t(\tau)$. The rest follows from theorem 2.5.7. □

One can show further results, like the three velocity addition theorem, or the components of four acceleration in terms of three acceleration. We skip this here, but mention, that $A(\tau) = (0, \vec{a}(\tau))$ holds in instantaneous rest frames.

Lemma 2.5.10.
For any four velocity $g(U,U) = -c^2$ holds. Also, the four velocity is directed in positive time direction.

Proof 2.5.11.

$$\begin{aligned} g(U,U) &= -(U^0)^2 + \sum_i (U^i)^2 = -\gamma^2 c^2 + \gamma^2 \|\vec{v}\|^2 = -c^2\gamma^2 \left(1 - \frac{\|\vec{v}\|^2}{c^2}\right) \\ &= -c^2 \ , \end{aligned}$$

$$c > 0 \ , \gamma > 0 \qquad \Rightarrow \qquad U^0 > 0 \ .$$

□

2.5.4. Four vectors and dynamics

Definition 2.5.12.
Let m be the mass of an object in it's rest frame, called **proper mass** or rest mass. The **four momentum** is defined as $P = m \cdot U$. The **four force** is

accordingly defined as $F = \frac{d}{d\tau}U$.

Indeed, up to a γ, the space component of the four momentum is the classical momentum. The zeroth component of P has the dimension Energy over velocity. Our assumption is, that the term describes the energy of the particle E/c. Additionally we define the **relativistic momentum** by $\vec{p}_\gamma = \gamma m \vec{v}$. For that reason, the term γm is called **relativistic mass**. It is the mass, associated to classical momentum, an observer in a different inertial system observes. Here we skip the derivation and focus on the consequences.

Theorem 2.5.13.
*From our assumptions $P^0 = \frac{E}{c}$ and $P^i = p^i_\gamma$, the **Energy-momentum-relation** follows directly:*

$$E^2 = m^2c^4 + \|\vec{p}_\gamma\|^2 c^2 \ .$$

Proof 2.5.14.
With the initial definition $P = mU$ we derive:

$$g(P,P) = -(m\gamma c)^2 + m^2\gamma^2\|\vec{v}\|^2 = m^2c^2\gamma^2\left(1 - \frac{\|\vec{v}\|^2}{c^2}\right) = m^2c^2 \ .$$

The assumptions on the other hand yield:

$$m^2c^2 = g(P,P) = -\frac{E^2}{c^2} + \|\vec{p}_\gamma\|^2 \qquad \Leftrightarrow \qquad E^2 = m^2c^4 + \|\vec{p}_\gamma\|^2c^2 \ .$$

□

2.6. The geometry of proper time

Although we have already defined the proper time, it is worthwhile to elaborate on the subtleties of the concept, especially in the context of accelerated motion. As defined before, the proper time τ is the time in a co-moving system. A more intuitive description would be, that the proper time is the time shown on a clock that is fixed to the object of interest. Note, that the system is not inertial in case of accelerated motion. The trick one uses is instantaneous rest frames. Fixing a specific value τ_0 defines a spacetime point as origin of the instantaneous inertial frame I' and a constant (for infinitesimal time) velocity $\vec{v}_{\tau_0} \equiv \vec{v}_0$ with respect to another valid inertial system I. More mathematically rigorous, we are interested in the coordinates of the tangent vector of the curve. The rest frame is characterized by the fact, that the tangent vector of the motion curve is always proportional to $\frac{\partial}{\partial x'^0}$. Graphically this can be seen in figure 2.4, where the $x'^0 = c\tau$-axis is the tangent vector of $X(\tau)$. With the help of Lemma 2.5.10 we can be more specific. Let $\Gamma(\tau)$ be a time like curve, then we have:

$$g(\dot{\Gamma}(\tau_0), \dot{\Gamma}(\tau_0)) = -c^2 \qquad \dot{\Gamma}(\tau_0) = h \cdot \partial_{x'^0}|_{\tau_0} \quad \text{and} \quad g(\partial_{x'^0}|_{\tau_0}, \partial_{x'^0}|_{\tau_0}) = -1 \ ,$$

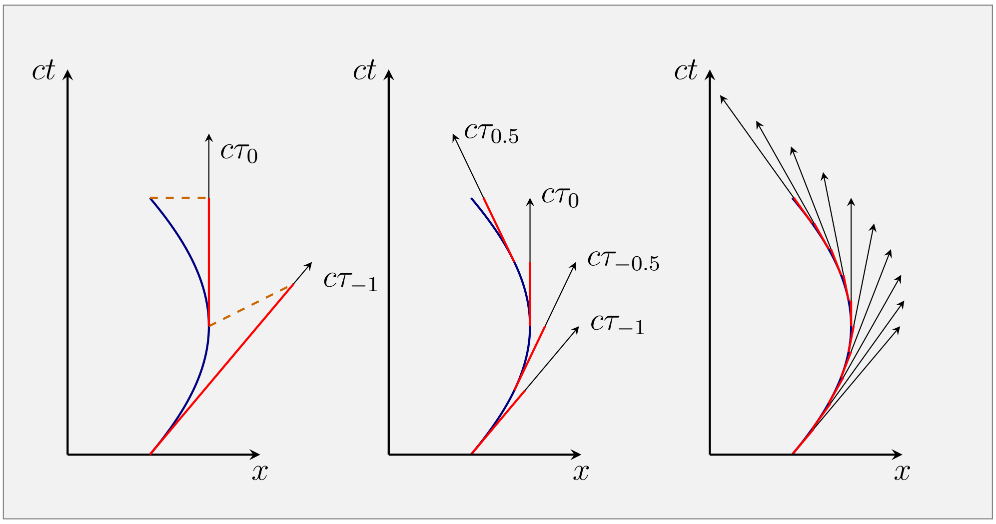


Figure 2.5.: Visualization of the arc-length formula for proper time. Choosing a discretization $\{\tau_j\}$ allows to approximate the proper time along the blue curve. To do so, the time coordinate of $\Gamma(\tau_j)$ in rest system $j-1$ is determined, by projection (orange dashed line). Adding up these time components, $\sum_j \Delta\tau_j$, approximates the proper time along the curve.

$$\Rightarrow \qquad -c^2 = g(\dot\Gamma(\tau_0), \dot\Gamma(\tau_0)) = h^2 g(\partial_{x'^0}|_{\tau_0}, \partial_{x'^0}|_{\tau_0}) \qquad \Leftrightarrow \qquad h = c\ .$$

The goal of this section is to find a formula to calculate the time that has passed on the co-moving clock between two spacetime events. Before passing to the proper limit, i.e. what one conceptually does, is to sum over a discretized set of instantaneous rest frames. So, let $\{\tau_j\}$ be a discretization of the proper time, defining points $\Gamma(\tau_i)$ on the curve. Let I_j be the instantaneous rest frames of these curves. The proper time between $\Gamma(\tau_{j-1})$ and $\Gamma(\tau_j)$, with the additional factor c of course, can be approximated by the x'^0-coordinate of $\Gamma(\tau_j)$ in the system I_{j-1}; (see figure 2.5 for an illustration). In the limit, we can write:

$$x'^0|_{j-1}(\Gamma(\tau_j)) \simeq c\cdot(\tau_j - \tau_{j-1}) = c\cdot\Delta\tau_j\ .$$

Indeed, using $\dot\Gamma(\tau) = c\cdot\partial_{x'^0}$ yields:

$$\left.\frac{d}{d\tau}\right|_{\tau_{j-1}} x'^0|_{j-1}(\Gamma(\tau_j)) = dx'^0|_{j-1}\left(\dot\Gamma(\tau_{j-1})\right) = dx'^0|_{j-1}\left(c\cdot\partial_{x'^0}|_{j-1}\right) = c\ .$$

Thus the proper time along the curve $\tau_{A,B}$ times c is:

$$\begin{aligned} c\cdot\tau_{A,B} &= \sum_j x'^0|_{j-1}(\Gamma(\tau_j)) \simeq \sum_j c\Delta\tau_j\cdot 1 = c\sum_j \Delta\tau_j\sqrt{-g(\partial_{x'^0}|_j, \partial_{x'^0}|_j)} \\ &= \sum_j \Delta\tau_j\sqrt{-g\left(\dot\Gamma(\tau_j), \dot\Gamma(\tau_j)\right)} \\ &\simeq \int_{\tau_A}^{\tau_B} \sqrt{-g\left(\dot\Gamma(\tau), \dot\Gamma(\tau)\right)}\, d\tau = L_{-g}(\Gamma, \tau_A, \tau_B)\ . \end{aligned}$$

With $L_{-g}(\Gamma, \tau_A, \tau_B)$ we denote the length of Γ between $\Gamma(\tau_A)$ and $\Gamma(\tau_B)$ with respect to the negative Minkowski metric $-g$.

Motivated by that equation, which could be written more rigorously using limits, we define:

Definition 2.6.1.
The **proper time interval** $\tau_{A,B}$, a clock moving along the curve $\Gamma\colon \mathbb{R} \to M$ measures between the spacetime events $\Gamma(\tau_A)$ and $\Gamma(\tau_B)$, is defined as

$$\tau_{A,B} = \frac{1}{c} L_{-g}(\Gamma, \tau_A, \tau_B) \ .$$

Lemma 2.6.2.
For the curve of (accelerated) motion, the inertial time differential dt and the ***proper time differential*** $d\tau$ *are related by*

$$d\tau(t) = \frac{1}{\gamma(t)} \ dt \ .$$

Proof 2.6.3.
In general the coordinate functions of inertial systems are related by affine transformations, that is:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} + X_0 \ .$$

For the coordinate differentials we get:

$$dx'^{\mu} = d\left(\Lambda^{\mu}{}_{\nu} x^{\nu} + X_0\right) = \Lambda^{\mu}{}_{\nu} dx^{\nu} \ .$$

From (2.1) we find the expression for $x'^0 = c\tau$, if we fix a time t, so that the rest frame can be pretended to be inertial:

$$c \ d\tau = \gamma(\|\vec{v}(t)\|) \left(c \ dt - \frac{1}{c} \sum_i v^i(t) \ dx^i(t) \right) \ .$$

We are only interested in the curve of motion (i.e. a one dimensional sub manifold). Thus the coordinates x^{μ} have to describe the position of the curve at the time t, that is $x^{\mu} \equiv x^i(t)$. At this point we also want to parametrize the proper time τ by the inertial frame time t. This can always be done since any point on the curve corresponds to a proper time value τ. Also for real motion, the inertial time value is the unique projection of the curve on the x^0-axis.

With the classical definition of velocity in an inertial system we find:

$$\begin{aligned} d\tau(t) &= \gamma(\|\vec{v}(t)\|) \left(dt - \frac{1}{c^2} \sum_i v^i(t) \ \partial_t x^i(t) \ dt \right) \\ &= \gamma(\|\vec{v}(t)\|) \left(1 - \frac{1}{c^2} \sum_i (v^i(t))^2 \right) \ dt \\ &= \gamma(\|\vec{v}(t)\|) \frac{1}{(\gamma(\|\vec{v}(t)\|))^2} \ dt = \frac{1}{\gamma(\|\vec{v}(t)\|)} \ dt \ . \end{aligned}$$

□

For the purpose of integration, we want to investigate the reparametrization of τ by t further. We begin, by noticing that the curve Γ and the inertial time axis T both are one dimensional manifolds. The map Ψ, that transforms the proper time into the inertial time, thus is, for physically reasonable motion, a diffeomorphism

$$\Psi \colon \Gamma \to T \ .$$

The reparametrized one form $d\tau(t)$ is then the pullback of $d\tau$ by Ψ^{-1}: $d\tau(t) = (\Psi^{-1})^* d\tau$.

The physical meaning of Ψ is far simpler than the mathematical formulation. Given a clock that moves in spacetime, for example. If the clock passes a specific position the time it shows will be noted, as well as the time on the clock of the observer. Let these be τ_0 for the moving clock and t_0 for the observer's clock. The relation between these times is $\Psi(\tau_0) = t_0$.

Theorem 2.6.4 (Formula for proper time)**.**
The proper time of a moving clock, between two spacetime events $A = \Gamma(\tau_A)$ and $B = \Gamma(\tau_B)$ is

$$\tau_{A,B} = \int_{\tau_A}^{\tau_B} d\tau = \int_{t_A}^{t_B} \sqrt{1 - \|\vec{v}(t)\|^2 / c^2} \ dt \ .$$

Proof 2.6.5.
The proper time interval is defined (Definition 2.6.1) to be the length of the curve divided by c. Let $X(\tau)$ be any parametrization of Γ over τ, with lemma 2.5.10 we get:

$$\begin{aligned}
\tau_{A,B} &= \frac{1}{c} L_{-g}(\Gamma, \tau_A, \tau_B) = \frac{1}{c} \int_{\tau_A}^{\tau_B} \sqrt{-g\left(\dot{X}(\tau), \dot{X}(\tau)\right)} \ d\tau \\
&= \frac{1}{c} \int_{\tau_A}^{\tau_B} \sqrt{c^2} \ d\tau = \int_{\tau_A}^{\tau_B} d\tau = \int_{\tau_A}^{\tau_B} \Psi^* (\Psi^{-1})^* d\tau = \int_{\Psi(\tau_A)}^{\Psi(\tau_B)} d\tau(t) \\
&= \int_{t_A}^{t_B} \sqrt{1 - \|\vec{v}(t)\|^2 / c^2} \ dt \ .
\end{aligned}$$

□

3

Einstein field equations

General Relativity is a theory that formulates gravity as curvature of spacetime. Behind this geometrical notion is a rather simple principle, that can be experienced every day, In this chapter we present the postulates of general relativity and give a brief outlook how the field equations can be derived form an action principle.

3.1. Physical motivation

As for any physical theory, there is no way to prove the postulates. All we can do, is to compare calculations to experiment, hoping for agreement. In that sense, we have some freedom, in where to start for the postulates. We could postulate principles, which lead to the field equations, as in [Car97]. On the other hand, we can start from the field equations, like in quantum mechanics with the Schrödinger equation, as its implications are tested very well and do agree with experiments.

Both approaches have their merits. The first, allowing for a deep intuition but using additional assumptions, that need not be physical principles, leading astray further research. The second, delivering ready equations, but overlooking deeper principles which could allow for an even more general theory.

In fact, while the equations are tested very well, the extension of the equivalence principle to non-gravitational effects is still under debate. For that reason, we will start with the postulates, that have proven to yield results in agreement with experiments

3.1.1. The postulates of general relativity

In this subsection we will follow the reasoning of [HE75, chapter 3].

Postulate 0: Configuration space

The configuration space, that is the space of all events, is taken to be a four dimensional pseudo Riemannian C^∞-manifold (M, g) with Lorentzian metric g. Being Lorentzian means, that there are local coordinates for each $p \in M$ such that $(g_{\mu\nu}(p)) = \operatorname{diag}(-1, 1, 1, 1)$.

By the meaning we ascribe points of M, all we can ever be aware of, has to be in the same connected component as ourselves. Hence we may assume the manifold to be at least connected.

Postulate 1: Locality

As special case of general relativity in the absence of all matter, the causality of special relativity, we have encountered in section 2.4, has to be present in general relativity as well. The postulate of locality is adopted as follow:

Let $U \subset M$ be a geodesically convex subset and $p, q \in U$ be two events. Information can only be transferred between these events, if there is a C^1 curve connecting p and q, such that every tangent vector of the curve is time-like or light-like.

The terminology above has exactly the same meaning as in special relativity. A tangent vector $v \in T_pM$ is time-like, if $g(v, v) < 0$, space-like if $g(v, v) > 0$ and light-like or called **null-vector** if $g(v, v) = 0$. A geodesic is called accordingly called time-like, space-like or null-geodesic, if every tangent vector is time-like, space-like or null-geodesic. Within this terminology we have already formulated, that light travels along null-geodesics (light-like geodesics).

Postulate 2: Local energy-momentum-conservation

The single most fundamental principle of all of physics is energy-conservation. Unless proven wrong by compelling experimental evidence in the future, no physical theory that violates energy conservation fundamentally is accepted today. Together with momentum conservation we thus have to postulate these conservations, at least locally.

To state the postulate more formal, we assume that to every matter field there exists a symmetric tensor of rank 2, called the **energy-momentum tensor** T, with the following properties:

i) $T \equiv 0$ on U, for $U \subset M$ open, if and only if the matter fields vanish on U.

ii) div $T \equiv 0$.

The first condition expresses that every matter field has a certain positive energy in the sense that there is no negative matter. This does not mean, that there cannot be chosen a gauge for potentials such that there are negative numbers, but that the equivalent matter to the energy is positive.[1] The second condition can be understood as the conservation law. In coordinates it reads ${T^{\mu\nu}}_{;\nu} \equiv 0$.

Postulate 3: Field equations

Gravitation is a geometrical effect rather than a force. Let $\mathcal{R}$ be the Ricci tensor, $\mathcal{S}$ be the scalar curvature and T be the energy-momentum tensor. In SI-units (Gravitation constant G and cosmological constant Λ) the **Einstein field equations** are: [Fli16, compare eq. (21.30)]

$$\boxed{\mathcal{R} - (\tfrac{1}{2}\mathcal{S} - \Lambda)g = \frac{8\pi G}{c^4}T\ .} \tag{3.1}$$

3.1.2. Equivalence principle

The postulates of the previous subsection demand that gravitation is a geometrical effect. Since all other physical interactions are described by fields, such a formulation is rather obscure at first. However, a simple observation Einstein made about the similarity between gravitation and acceleration can motivate the geometrical interpretation of gravity. The following explications is based on [Car97, chapter 4]:

[1] So far there has been no evidence of repulsive matter, as is the case for charges.

In the first chapter the concept of mass was treated as if it were a trivial one. However, comparing the classical definition of force with the gravitational force reveals, that there are three different concepts of mass:

$$m_I \cdot I_1 \left(\frac{d^2}{dt^2} \gamma(t) \right) = -\frac{GMm}{r_p^2(\gamma(t))} (dr_p)_{\gamma(t)} \ .$$

- The mass m_I, called **inertial mass**, connects the acceleration with the force
- The masses M and m, called **gravitational masses**, connect the concept of mass with gravitation. These gravitational masses can be distinguished further:
  - The gravitational mass M is called **active**, since it creates the gravitational field.
  - The gravitational mass m is called **passive**, since it reacts to the gravitational field.

The distinction between active and passive gravitational masses is not necessary in Newton's theory, as both masses can be exchanged. By the third law (actio = reactio) the force that m excerts on M is $F_{m \text{ on } M} = -F_{M \text{ on } m}$. Yet Newton's theory allows for different inertial gravitational masses. However, all experimental evidence hints to the equivalence of both masses so far, called **weak equivalence principal**.

In the theory of general relativity, Einstein extended the weak equivalence principle. He assumed that all dynamical laws in a small free falling system are equivalent to the laws in a system in gravity-free space. Also, accelerated small systems are equivalent to small systems in a gravitational field. These equivalences are called **equivalence principle**.

The auxiliary "small" for systems is important for gravitational systems, even in Newtonian gravity. For example consider two systems A and B in free fall. Each of these systems for itself is equivalent to a system in gravitation free space. However considering both systems at once, the distance between them decreases over time acceleratingly. There seems to be an force between the systems, called **tidal force**. If A is at position p and B at position $p + \vec{\chi}$, and we assume χ to be small, we can evaluate the tidal acceleration easily:

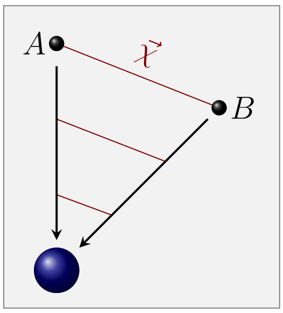


Figure 3.1.

1) $\frac{d^2}{dt^2} p_i = -\frac{\partial}{\partial x^i} \Phi(p)$

2) $\frac{d^2}{dt^2} (p + \vec{\chi})_i = -\frac{\partial}{\partial x^i} \Phi(p + \vec{\chi}) = -\frac{\partial}{\partial x^i} \Phi(p) - \sum_{j=1}^{3} \frac{\partial^2}{\partial x^i \partial x^j} \Phi(p) \chi^j + \mathcal{O}(\chi^2) \ .$

$$1) - 2) \quad \Rightarrow \quad \boxed{\frac{d^2}{dt^2} \chi_i = -\sum_{j=1}^{3} \frac{\partial^2}{\partial x^i \partial x^j} \Phi(p) \chi^j + \mathcal{O}(\chi^2)}$$

It should be mentioned that, for simplicity, we calculated in terms of coefficients here.[2]

[2] Notice that the positioning of the indices is already according to Ricci calculus. To compare forms with vectors we have also used $\partial_{x^i} = g_{ij} dx^j$ and $g_{ij} = \delta_{ij}$.

3.1.3. Equivalence principle and curvature

From the equivalence principle it might not be clear, how it connects to curved space. To illustrate that, we use an example, that is not necessarily a physical solution of the Einstein equations, but shows the concept clearly.

Assuming a huge mass at position P and a comparably small mass at position Q. Over time, the small mass will move towards the large mass, accelerating over time, while the large mass rests by assumption. Geometrically, the world line of the small mass is a straight line, defining a geodesic in a flat spacetime, whilst the world line of the small mass is curved. Taking the equivalence principle seriously, we have put no constraints on the masses, meaning they should be inertial systems (at least locally). This can be realized by choosing a metric g such that both world lines are Riemannian geodesics. Reformulated geometrically, we chose a curved space such that the world lines are locally as straight as possible, i.e. $\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0$, with respect to the Levi-Civita-connection. This point of view is visualized in figure 3.2.

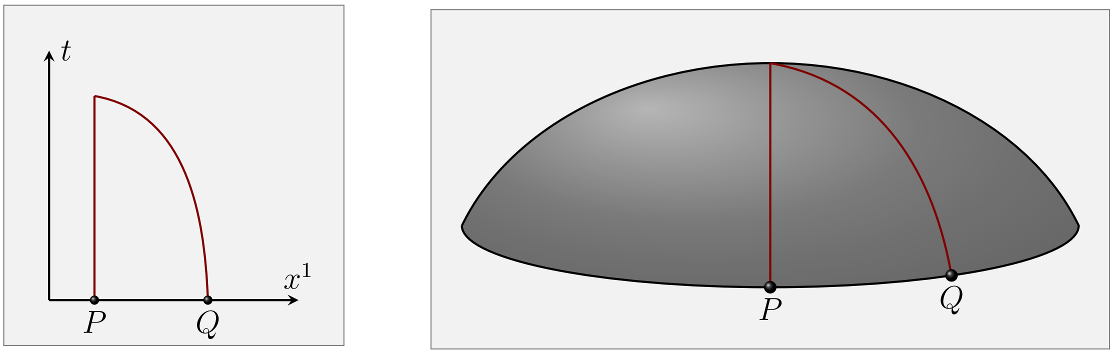


Figure 3.2.: Visualization of curvature instead of accelerated motion.

3.2. Variational principle and energy-momentum tensor

So far, we have not addressed, how to obtain energy-momentum tensors. Of course, in simple situations they can be guessed from classical physics by substituting partial derivatives with covariant derivatives, also called *comma-to-semicolon rule*. However the conditions imposed on the energy-momentum tensor may not be enough to uniquely determine it.

3.2.1. Einstein-Hilbert-Action review

The vacuum Einstein field equations can be derived from an action principle, the so called **Einstein-Hilbert action**, with a variation of the metric tensor coefficients $g^{\mu\nu}$. A derivation can be found in [Car97, section 4]. We will only present the results here:[3]

$$S_{EH} = \int_U \mathcal{L}_{EH}\sqrt{|g|}\; dx^4 \qquad \text{with} \quad \mathcal{L}_{EH} = \kappa(\mathcal{S} - 2\Lambda)\ .$$

[3]In the notation of this document adding information from [HE75, eq. (3.16)].

To obtain the field equations in SI-units, the constant is given by $\kappa = \left(\frac{16\pi G}{c^4}\right)^{-1}$. Assuming that variations on the boundary vanish, the functional derivative is given by

$$\frac{\delta S_{EH}}{\delta g^{\mu\nu}} = \sqrt{|g|}\ \kappa(\mathcal{R}_{\mu\nu} - \frac{1}{2}(\mathcal{S} - 2\Lambda)g_{\mu\nu})\ .$$

Of cause one could use the covariant Euler-Lagrange equations for a rigorous derivation, which would result in a lengthy calculation however, as the scalar curvature written in terms of the metric is a long expression.

Anyway, the condition $\delta S_{EH}(g_{\alpha\beta}; g^{\mu\nu}) = 0$ is equivalent to $\frac{\delta S_{EH}}{\delta g^{\mu\nu}} = 0$ resulting in the vacuum field equations:

$$\kappa(\mathcal{R}_{\mu\nu} - \tfrac{1}{2}(\mathcal{S} - 2\Lambda)g_{\mu\nu}) = 0\ .$$

3.2.2. Energy-momentum tensor

The Einstein-Hilbert Lagrange-density we used above differs from [HE75, eq. (3.16)], as we have ignored the matter Lagrange-density so far. As can be seen in [HE75, section 3.3] the energy-momentum tensor can be obtained from a variational principle, if the field equations correspond to that action. We have reversed the order compared to [HE75] for the following pedagogical reason: Noticing that the vacuum Einstein field equations can be derived from an action principle, which can be taken as an equivalent postulate, it can be assumed that there is an action principle for the full field equations. To see, how this is done, we will follow the reasoning of [HE75].

Assuming the matter/energy can be described by a field $\phi^{a\ldots b}{}_{c\ldots d} =: \phi^I_J$ with associate **matter Lagrange-density** $\mathcal{L}_M$ we define the **Einstein-Hilbert-Matter action** as follows:

$$S_{EHM}(g) = \int_U (\mathcal{L}_{EH} + \mathcal{L}_M)\sqrt{|g|}\ dx^4\ .$$

Using the linearity of the Fréchet-derivative and of integration, and assuming vanishing variations on the boundary, we see that

$$\delta(S_{EHM}, g^{\mu\nu}) = \delta_g \int_U \mathcal{L}_{EH}\sqrt{|g|}\ dx^4 + \delta_g \int_U \mathcal{L}_M\sqrt{|g|}\ dx^4 \overset{!}{=} 0\ ,$$

$$\Leftrightarrow \qquad \frac{1}{\sqrt{|g|}}\frac{\delta S_{EH}}{\delta g^{\mu\nu}} = \kappa(\mathcal{R}_{\mu\nu} - \frac{1}{2}(\mathcal{S} - 2\Lambda)g_{\mu\nu}) = T_{\mu\nu} = -\frac{1}{\sqrt{|g|}}\frac{\delta S_M}{\delta g^{\mu\nu}}$$

where δ_g denotes the variation of the integral-functional in direction $g^{\mu\nu}$. Thus, we have found an expression for the energy-momentum tensor:

$$\boxed{T_{\mu\nu} = -\frac{1}{\sqrt{|g|}}\frac{\delta S_M}{\delta g^{\mu\nu}}\ .}$$

A consequence of this definition of the energy momentum tensor is that it has to divergence free, if $\mathcal{R}_{\mu\nu} - \frac{1}{2}(\mathcal{S} - 2\Lambda)g_{\mu\nu}$ is divergence free.

Lemma 3.2.1.
The ***Einstein tensor*** $G = \mathcal{R} - \frac{1}{2}\mathcal{S}g$ *is divergence free.*

Proof 3.2.2.
Choosing an arbitrary chart allows to use the Contracted Bianchi identity in the easier form $\mathcal{R}_{\mu\nu;}{}^{\nu} = \frac{1}{2}\mathcal{S}_{;\mu}$. By the metric compatibility of the Levi-Civita-connection the metric is also divergence free: $g_{\mu\nu;\mu} = 0$. It follows that:

$$G_{\mu\nu;}{}^{\nu} = \mathcal{R}_{\mu\nu;}{}^{\nu} - \tfrac{1}{2}\mathcal{S}_{;}{}^{\nu} g_{\mu\nu} = \tfrac{1}{2}\mathcal{S}_{;\mu} - \tfrac{1}{2}\mathcal{S}_{;\mu} = 0 \ .$$

Since we have not used any special choice of coordinates, this result does not depend on coordinates.[4] □

As a result of this lemma, and since the metric is divergence free as well, the energy-momentum tensor has to be divergence free as well:

$$\kappa(\mathcal{R} - \frac{1}{2}(\mathcal{S} - 2\Lambda)g) = T \qquad \Rightarrow \qquad 0 = \kappa \cdot \mathrm{div}(\mathcal{R} - \frac{1}{2}(\mathcal{S} - 2\Lambda)g) = \mathrm{div}\ T \ .$$

[4] We could have used the coordinate free version of div from page 89 as well.

4

Gravitational waves in linear approximation

This chapter considers a linearization of gravitation and gravitational waves in this limit. This chapter follows [Car97, chapter 6] closely but is not complete.

4.1. Linerized gravitation

Gravitational waves are a general property of the field equations. Yet even in the linear approximation they do appear, and are also easier to understand. So we might as well restrict our attention to the linear limit we are to develop.

4.1.1. Linear limit

Linear limit means that the metric g can be written as:

$$g = \eta + h \ ,$$

where η is the flat Minkowski metric and h a perturbation. For the linearization it is assumed that the perturbation is small.

Remark 4.1.1.
We do not care about the precise meaning of small, but only want to neglect quadratic terms (and higher order terms) of h to get a linear theory. In fact we are going to use the coordinate expression $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, which need not be unique nor existing. Furthermore we will assume to have the freedom to choose the coordinates in a particular way. A part of these assumptions is addressed in in the classical literature by checking compatibility of the made assumptions and checking coordinate transformations (called gauge invariance in the literature). However, the very existence of the demanded coordinates remains to be proven, specifying also a chart domain. There do exist mathematical treatments of linearized gravity, see [SW12, section 7.6] for example. For that reason we will not address the aforementioned issues and assume everything to work out.

As a warning, in this chapter, raising and lowering indices is done by using $\boldsymbol{\eta}$. The reason will become clear momentarily. If we need the proper isomorphisms induced by $\boldsymbol{g}$, we underline the expression, e.g. $\underline{\boldsymbol{v}^{\boldsymbol{\mu}}}$.

We start by finding an expression for $g^{\mu\nu}$, which is defined to be the inverse matrix of $g_{\mu\nu}$, in first order of h. We make the guess, that $g^{\mu\nu} \approx \eta^{\mu\nu} - h^{\mu\nu}$ will work, and calculate:

$$(\eta_{\mu\nu} - h^{\mu\nu})g_{\nu\rho} = (\eta_{\mu\nu} h^{\mu\nu})(\eta_{\nu\rho} + h_{\nu\rho}) = \eta^{\mu\nu}\eta_{\nu\rho} + \eta^{\mu\nu}h_{\nu\rho} - h^{\mu\nu}\eta_{\nu\rho} + \mathcal{O}(h^2)$$
$$\approx \eta^{\mu\nu}\eta_{\nu\rho} + \eta^{\mu\nu}h_{\nu\rho} - h^{\mu\nu}\eta_{\nu\rho} = \delta^\mu_\rho + h^\mu{}_\rho - h^\mu{}_\rho = \delta^\mu_\rho \ .$$

Hence the guess of $g^{\mu\nu}$ was correct. We observe, that all quantities that are derived from h, i.e. contain only terms that are of linear order in h, can be transformed by η instead of g, if quadratic and higher order terms are to be neglected. Let $v = v(h)$ be a vector field derived from h:

$$\underline{v^\mu} = g^{\mu\nu}v_\nu = \eta^{\mu\nu}v_\nu + h^{\mu\nu}v_\nu = v^\mu + \mathcal{O}(h^2) \approx v^\mu \ .$$

Since the Minkowski metric is constant, the Christoffel symbols are (ignoring quadratic order):

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\eta^{\rho\lambda}(\partial_\mu h_{\nu\lambda} + \partial_\nu h_{\lambda\mu} - \partial_\lambda h_{\mu\nu})$$

We observe, that the Christoffel symbols are derived only from h, and thus all quantities concerned with curvature can be raised and lowered by η.

With corollary C.5.2 we observe the general structure of the coefficients of the Riemann tensor to be partial derivatives and products of the Christoffel symbols. Since products would be of order h^2 we obtain:

$$R_{ijk}{}^\ell = \partial_i \Gamma^\ell_{jk} - \partial_j \Gamma^\ell_{ik} \ .$$

It follows that:

$$R_{ijk\ell} = \frac{1}{2}(\partial_k\partial_i h_{\ell j} + \partial_\ell\partial_j h_{ik} - \partial_\ell\partial_i h_{jk} - \partial_k\partial_j h_{\ell i}) \ .$$

The Ricci tensor from subsection C.5.3 is the contraction of the first and last index (relabeling the indices to match [Car97]):

$$\mathcal{R}_{\mu\nu} = \frac{1}{2}(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\nu - \partial_\mu\partial_\nu h^\sigma{}_\sigma - \Box h_{\mu\nu}) \ ,$$

where $\Box = \partial_\sigma\partial^\sigma = \partial_x^2 + \partial_y^2 + \partial_z^2 - \frac{1}{c^2}\partial_t^2$ is the D'Alembert operator. For the scalar curvature one finds:

$$\mathcal{S} = \partial_\mu\partial_\nu h^{\mu\nu} - \Box h^\sigma{}_\sigma \ .$$

Putting all together in (3.1) (without cosmological constant) yields the **linearized Einstein field equations**:

$$\boxed{\frac{8\pi G}{c^4}T_{\mu\nu} = \frac{1}{2}\left(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\nu - \partial_\mu\partial_\nu h^\sigma{}_\sigma - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial_\mu\partial_\nu h^{\mu\nu} + \eta_{\mu\nu}\Box h^\sigma{}_\sigma\right)} \ . \tag{4.1}$$

4.1.2. Harmonic coordinates and trace reversed perturbation

The linear field equations can be simplified further by choosing a particular coordinate system. To define these coordinates we need a new operator.

Definition 4.1.2.
The **covariant Hessian** $\nabla^2 T$ of a tensor field T is defined by

$$\nabla^2 T(X,Y) = \nabla_Y(\nabla_X T)$$

for vector fields X, Y.

In the case of functions, using theorem C.2.24, we see that:

$$\nabla^2 f(X,Y) = \nabla_Y(\nabla_X f)\nabla_Y(X(f)) = Y(X(f)) - (\nabla_Y X)(f) \ .$$

In coordinates with $X = X^\nu \partial_\nu$ and $Y = Y^\mu \partial_\mu$ this reads:

$$\nabla^2 f(X,Y) = \left(\partial_\mu \partial \nu f - \Gamma^\sigma_{\mu\nu}\partial_\sigma f\right) Y^\mu X^\mu$$

leading to the local coordinate expression of $\nabla^2 f$:

$$\nabla^2 f = \left(\partial_\mu \partial_\nu f - \Gamma^\sigma_{\mu\nu}\partial_\sigma f\right) dx^\nu \otimes dx^\mu \ .$$

Definition 4.1.3.
The **Laplace-Beltrami operator** is defined by

$$\Delta f = \mathrm{tr}(\nabla^2 f) \ .$$

Remark 4.1.4.
For the Minkowski metric the Laplace-Beltrami operator is the D'Alembert operator. In case of the Euclidean metric on $\mathbb{R}^3$ it is the classical Laplace operator.

In coordinates the Laplace-Beltrami operator can be calculated as follows:

$$\begin{aligned}\Delta f &= \mathrm{tr}\left(\left(\partial_\mu \partial_\nu f - \Gamma^\sigma_{\mu\nu}\partial_\sigma f\right) dx^\nu \otimes dx^\mu\right) = \left(\partial_\mu \partial_\nu f - \Gamma^\sigma_{\mu\nu}\partial_\sigma f\right) dx^\nu (g^{\mu\rho}\partial_\rho) \\ &= g^{\mu\rho}\left(\partial_\mu \partial_\nu f - \Gamma^\sigma_{\mu\nu}\partial_\sigma f\right)\delta^\nu_\rho = g^{\mu\nu}\left(\partial_\mu \partial_\nu f - \Gamma^\sigma_{\mu\nu}\partial_\sigma f\right)\end{aligned}$$

$$\boxed{\Delta f = \partial^\mu \partial_\mu f - g^{\mu\nu}\Gamma^\sigma_{\mu\nu}\partial_\sigma f \ .}$$

Definition 4.1.5.
Coordinates (x^μ) that satisfy $\Delta x^\mu = 0$ are called **harmonic coordinates**.

Harmonic coordinates have a special consequence for the Christoffel symbols:

Corollary 4.1.6.
Coordinates (x^μ) are harmonic, if and only if $\Gamma^\mu_{\sigma\nu} = 0$.

Proof 4.1.7.

$$0 = \Delta x^\mu = \partial^\sigma \partial_\sigma x^\mu - g^{\sigma\nu}\Gamma^\rho_{\sigma\nu}\partial_\rho x^\mu = \partial^\mu \delta^\mu_\sigma - g^{\sigma\nu}\Gamma^\rho_{\sigma\nu}\delta^\mu_\rho = g^{\sigma\nu}\Gamma^\mu_{\sigma\nu}$$
$$\Leftrightarrow \qquad 0 = \Gamma^\mu_{\sigma\nu} \; .$$

□

Applying the choice of harmonic coordinates to the weak field approximation, since $\eta^{\rho\lambda}$ is an invertible matrix, this yields:

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\eta^{\rho\lambda}(\partial_\mu h_{\nu\lambda} + \partial_\nu h_{\lambda\mu} - \partial_\lambda h_{\mu\nu}) = 0$$
$$\Leftrightarrow \qquad \frac{1}{2}(\partial_\mu h_{\nu\lambda} + \partial_\nu h_{\lambda\mu} - \partial_\lambda h_{\mu\nu}) = 0$$
$$\Leftrightarrow \qquad \frac{1}{2}\eta^{\mu\nu}(\partial_\mu h_{\nu\lambda} + \partial_\nu h_{\lambda\mu} - \partial_\lambda h_{\mu\nu}) = \boxed{\partial_\mu h^\mu{}_\lambda - \frac{1}{2}\partial_\lambda h^\mu{}_\mu = 0} \; .$$

Calculating the Ricci coefficients in these coordinates yields:

$$\begin{aligned}\mathcal{R}_{\mu\nu} &= \frac{1}{2}(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\nu - \partial_\mu\partial_\nu h^\sigma{}_\sigma - \Box h_{\mu\nu}) \\ &= \frac{1}{2}(\tfrac{1}{2}\partial_\nu\partial_\mu h^\sigma{}_\sigma + \tfrac{1}{2}\partial_\mu\partial_\nu h^\sigma{}_\sigma - \partial_\mu\partial_\nu h^\sigma{}_\sigma - \Box h_{\mu\nu}) \\ &= -\Box h_{\mu\nu} \; .\end{aligned}$$

The scalar curvature becomes $\mathcal{S} = -\Box h^\sigma{}_\sigma$. Plugging in into the field equations yields:

$$\Box h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\Box h^\sigma{}_\sigma = -\frac{16\pi G}{c^4}T_{\mu\nu} \; .$$

Introducing a new form for the perturbation, $\overline{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^\sigma{}_\sigma$, which is called **trace reverse** leads to the simple form:

$$\boxed{\Box\overline{h}_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu} \; .} \tag{4.2}$$

With $\eta^\mu{}_\nu = \eta^{\mu\lambda}\eta_{\lambda\nu} = \delta^\mu_\nu$ and using the harmonic condition for h we calculate:

$$\begin{aligned}\partial_\mu\overline{h}^\mu{}_\nu &= \partial_\mu\eta^{\mu\lambda}\overline{h}_{\lambda\nu} = \partial_\mu\eta^{\mu\lambda}(h_{\lambda\nu} - \tfrac{1}{2}\eta_{\lambda\nu}h^\sigma{}_\sigma) = \partial_\mu h^\mu{}_\nu - \tfrac{1}{2}\eta^\mu{}_\nu\partial_\mu h^\sigma{}_\sigma \\ &= \partial_\mu h^\mu{}_\nu - \delta^\mu_\nu\partial_\sigma h^\sigma{}_\mu = \partial_\mu h^\mu{}_\nu - \partial_\mu h^\mu{}_\nu = 0 \; .\end{aligned}$$

Thus for $\overline{h}$ the harmonic condition reads

$$\partial_\mu\overline{h}^\mu{}_\nu = 0 \qquad \Leftrightarrow \qquad \partial_\mu\overline{h}^{\mu\nu} = 0 \; .$$

Remark 4.1.8.
The name trace reverse comes from the property, that $\overline{h}^\mu{}_\mu = -h^\mu{}_\mu$:

$$\overline{h}^\nu{}_\nu = \eta^{\mu\nu}h_{\mu\nu} - \tfrac{1}{2}\eta^{\mu\nu}\eta_{\mu\nu}h^\sigma{}_\sigma = h^\nu{}_\nu - \frac{1}{2}\left(\sum_{\nu=0}^{3}\delta^\nu_\nu\right)h^\nu{}_\nu = h^\nu{}_\nu - 2h^\nu{}_\nu$$

$$= -h^{\nu}{}_{\nu} \ .$$

This also allows to write down the inverse equation:

$$h_{\mu\nu} = \overline{h}_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\overline{h}^{\sigma}{}_{\sigma} \ .$$

4.2. Propagation of free waves

Free wave implies that there is no matter present. This means, that $T = 0$, such that the equation for the trace reversed perturbation becomes

$$\Box\overline{h}_{\mu\nu} = 0 \ .$$

As usual on chooses the ansatz

$$\overline{h}_{\mu\nu} = C_{\mu\nu}e^{ik_\sigma x^\sigma} \ ,$$

whew C is a constant symmetric $(0, 2)$ tensor. Plugging this ansatz into the wave equation shows, that the wave vector has to be light like, i.e. $k_\sigma k^\sigma = 0$:

$$0 = \Box\overline{h}_{\mu\nu} = \Box C_{\mu\nu}e^{ik_\sigma x^\sigma} = C_{\mu\nu}\eta^{\rho\sigma}\partial_\rho\partial_\sigma e^{ik_\sigma x^\sigma} = C_{\mu\nu}\eta^{\rho\sigma}k_\rho k_\sigma e^{ik_\sigma x^\sigma} = k_\sigma k^\sigma \overline{h}_{\mu\nu} \ .$$

$$\Rightarrow \qquad k_\sigma k^\sigma = 0 \ .$$

From the wave equation, as well as the fact that for non trivial perturbations the wave vectors have to be light like, we deduce that the perturbations propagates with the speed of light, hence a gravitational wave. From the four momentum $p = (E, p_1, \ldots, p_3)$ it we know that $k = (\omega, k_1, \ldots, k_3)$, such that the condition for k can be written:

$$\omega^2 = \sum_{j=1}^{3} k_j^2 \ .$$

Since we have derived the wave equation in harmonic coordinates, we have to apply the harmonic condition to $\overline{h}$:

$$0 = \partial_\mu\overline{h}^{\mu\nu} = \partial_\mu C^{\mu\nu}e^{ik_\sigma x^\sigma} C^{\mu\nu}k_\mu e^{ik_\sigma x^\sigma} \quad \Rightarrow \quad k_\mu C^{\mu\nu} = 0 \quad \Leftrightarrow \quad k^\mu C_{\mu\nu} = 0 \ .$$

To reduce the degrees of freedom, we note, that harmonic coordinates need not be unique. This can be seen by observing that the Laplace-Beltrami operator is linear. Let ξ^μ be harmonic functions, i.e. $\Delta\xi^\mu = 0$, and (x^μ) be harmonic coordinates, then $y^\mu = x^\mu + \xi^\mu$ are also harmonic coordinates.

We assume[1] that this freedom allows to find coordinates, such that

$$C^{\mu}{}_{\mu} = 0 \qquad \text{and} \qquad C_{0_\nu} = 0 \ .$$

[1] In [Car97, p. 149] it is tired to be verified. However the proof uses functions for which $\Box\xi^\mu = 0$ holds instead of $\Delta\xi^\mu = 0$. This can be seen by the ansatz for ξ^μ. The mistake might have happend by denoting both the Laplace-Beltrami operator and the D'Alembert operator with the same symbol. Thus the existence of the desired coordinate system remains to be verified.

Having applied all conditions for $C_{\mu\nu}$, we are ready to examine the degrees of freedom. The condition $C_{0\nu} = 0$ together with the symmetry reduce the free coefficients to a symmetric 3×3 matrix:

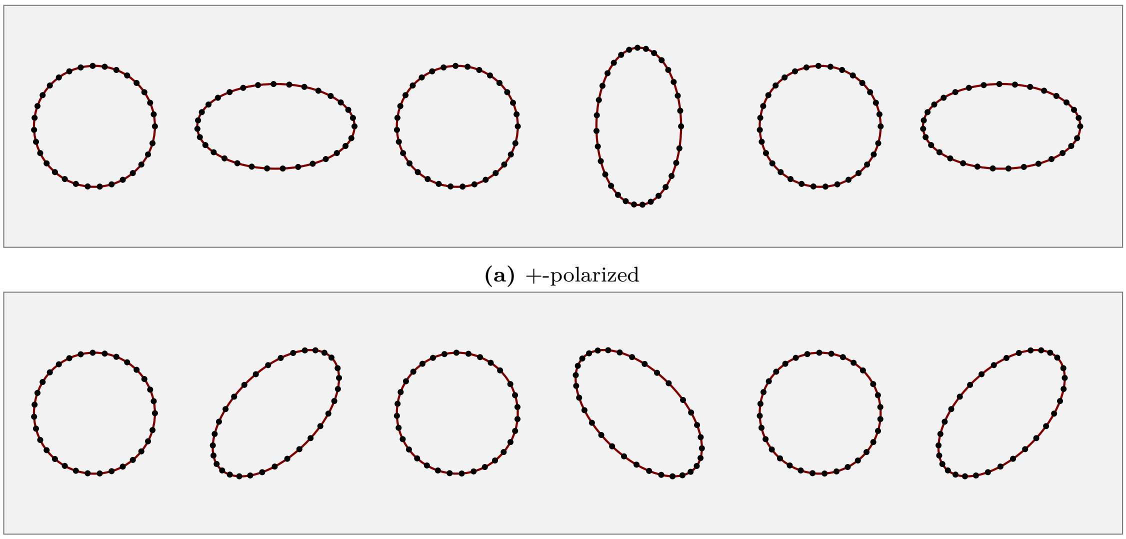
$$C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_{11} & C_{12} & C_{13} \\ 0 & C_{12} & C_{22} & C_{23} \\ 0 & C_{13} & C_{23} & C_{33} \end{pmatrix} .$$

As symmetric 3×3 matrix there are $\frac{3\cdot(3+1)}{2} = 6$ degrees of freedom. $k^\mu C_{\mu\nu} = 0$ are three equations (since the first row is zero) and $C^\mu{}_\mu = 0$ is one equation, such that there are only 2 degrees of freedom, that have a physical relevance. These two degrees of freedom are the polarizations of the gravitational waves.

The particular choice of coordinates we have made has another beneficial consequence. Since $C^\mu{}_\mu = 0$ it follows that $\overline{h}^\mu{}_\mu = C^\mu{}_\mu e^{ik_\sigma x^\sigma} = 0 = -h^\mu{}_\mu$. As a consequence it follows that:

$$h_{\mu\nu} = \overline{h}_{\mu\nu} .$$

This is only true in the special coordinates we have chosen to investigate gravitational waves, accounting for the name **radiation coordinates**.



(a) +-polarized

(b) ×-polarized

Figure 4.1.: Effect of linearly polarized gravitational waves on particles on a circle.

To investigate the properties of free gravitational waves further, we consider the spacial case of propagation in x^3 direction. In this case the wave vector is $k = (\omega, 0, 0, \omega)$, since $(k^0)^2 = \sum_i (k^i)^2 = (k^3)^2$, and thus $k^\flat = (-\omega, 0, 0, \omega)$. From $k^\mu C_{\mu\nu} = k^0 C_{0\nu} + k^3 C_{3\nu} = 0$ and $C_{0\nu} = 0$ it follows that $C_{3_\nu} = C_{\nu 3} = 0$. Thus, using symmetry and vanishing trace, we have:

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_{11} & C_{12} & 0 \\ 0 & C_{12} & -C_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

The two coefficients C_{11} and C_{12} are the two degrees of freedom, that can be chosen freely. Take for example two linearly polarized waves, defining C^+ by setting $C_{12} = 0$

and $C^{\times}$ by setting $C_{11} = 0$. Then the perturbations become

$$h^{+} = C_{11} e^{i\omega(x^3 - x^0)} (dx^1 \otimes dx^1 - dx^2 \otimes dx^2) \ ,$$

$$h^{\times} = C_{12} e^{i\omega(x^3 - x^0)} (dx^1 \otimes dx^2 + dx^2 \otimes dx^1) \ .$$

Fixing the coordinate x^3, the term $e^{i\omega(x^3 - x^0)}$ describes an oscillation in time, where x^0 is taken to be the time coordinate. Considering the full metrics $g^{+/\times} = \eta + h^{+/\times}$, the notation $+$ and $\times$ come from the axes along which distances are stretched and compressed, as can be seen in figure 4.1. As with light waves, there can be circular polarized waves:

$$C^R = \frac{1}{\sqrt{2}}(C^+ + iC^\times) \qquad \text{and} \qquad C^L = \frac{1}{\sqrt{2}}(C^+ - iC^\times) \ .$$

4.3. Radiation of gravitational waves

So far, we have only considered free waves, ignoring the source term. Mathematically speaking, we have only solved the homogeneous equation. Since the D'Alembert operator is a linear differential operator with constant coefficients and the background in the linearized theory is flat, the theory of fundamental solutions can be applied. A formal introduction would require distribution theory. Thus we will only give the results.

A green function[2] for the D'Alembert operator is a function $G(x, y)$, such that $\Box T_G = \delta_x$, where T_G is the regular distribution of the function $y \mapsto G(x, y)$ and δ_x the delta distribution with pole in $x \in \mathbb{R}^4$. In the more common notation of the textbooks with delta functions, the condition reads:

$$\Box_y G(x, y) = \delta(y - x).$$

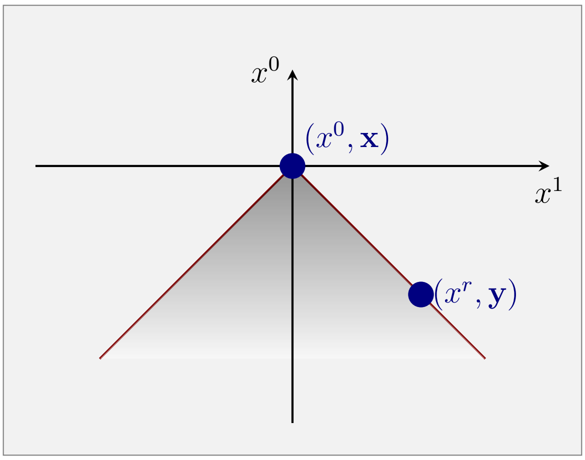


Figure 4.2.

A property of Green functions is, that a (weak[3]) solution of the differential equation $\Box u = f$ is given by

$$u(x) = \int_{\mathbb{R}^4} G(x, y) f(y) \ dy^4 \ .$$

Denoting the spacial parts of $x = (x^0, x^1, x^2, x^3)$ by $\mathbf{x} = (x^1, x^2, x^3)$, the Green function of the D'Alembert operator is:

$$G(x, y) = -\frac{1}{4\pi|\mathbf{x} - \mathbf{y}|} \delta\left(|\mathbf{x} - \mathbf{y}| - (x^0 - y^0)\right) \Theta(x^0 - y^0) \ ,$$

[2] The relation between Green functions $G(x, y)$ and regular fundamental solutions $g(y)$ for a linear differential operator with constant coefficients is $g(x - y) = G(x, y)$. This leads to the notation $G(x - y)$ in some texts.

[3] If f is a test function, then differentiability is ensured, and the solution is a proper solution. Otherwise it is only a solution w.r.t. weak differentiation.

where we have used the Heaviside function $\Theta(x) = 1$ for $x \geq 0$ and $\Theta(x) = 0$ for $x < 0$. A mathematical rigorous derivation takes some two to three pages, depending on the theoretical background, such that we have just copied it from [Car97].

Since the equation for the trace reversed perturbation (4.2) is coefficient wise, the Green function theory applies without changes:

$$\begin{aligned}\overline{h}_{\mu\nu}(x^0, \mathbf{x}) &= \frac{4G}{c^4} \int_{\mathbb{R}^4} \frac{1}{|\mathbf{x}-\mathbf{y}|} T_{\mu\nu}(y) \delta\left(|\mathbf{x}-\mathbf{y}| - (x^0 - y^0)\right) \Theta(x^0 - y^0)\ dy^4 \\ &= \frac{4G}{c^4} \int_{\mathbb{R}^3} \frac{1}{|\mathbf{x}-\mathbf{y}|} T_{\mu\nu}(x^0 - |\mathbf{x}-\mathbf{y}|, \mathbf{y})\ dy^3\ .\end{aligned}$$

In the second line, we have evaluated the time integral with help of the delta function. The expression $x^r := x^0 - |\mathbf{x}-\mathbf{y}|$ is called **retarded time**. To give the right interpretation, we choose $t = \frac{x^0}{c}$ as in special relativity. Then the retarded time is $t_r = t - \frac{|\mathbf{x}-\mathbf{y}|}{c}$. Fixing a point $(t, \mathbf{x})$ and choosing a space point $\mathbf{y}$, the retarded time t_r is the time point in the past (since $t_r \leq t$), when a light ray had to be emitted from $\mathbf{y}$ to reach $\mathbf{x}$ at the time t. This means that the point $(x^r, \mathbf{y})$ lies on the surface of the backwards directed light cone. In the integration this means, that only the surface of the light cone in the past contributes to $\overline{h}_{\mu\nu}(x^0, \mathbf{x})$.

5

Schwarzschild solution

The previous chapter presented solutions to the linearized field equations. However, from a theoretical point of view, one might ask, if this is possible for the full equations. Assuming certain symmetries, there do exist solutions. One such case is the Schwarzschild solution using spherical symmetry. This solution is discussed thoroughly in the literature and involves lots of tedious calculations. Here we do not want to copy them (as the reader can find them in any textbook of his liking) and will only consider results and concepts, choosing the simplest rather than the most complete approach.

5.1. The Schwarzschild metric

In this section we will motivate the Schwarzschild metric, following [Fli16, section 23]. This is not a formal derivation. However, there do exist derivations (e.g. in [HE75]) that show, that the Schwarzschild metric is the unique spherically symmetric vacuum solution solution. The condition of staticity, i.e. time independence is not needed in this case, as it follows from the Birkhoff theorem. Here we will make the assumption that the metric is static.

Assume, that the energy momentum tensor is static with compact support and spherically symmetric w.r.t. S^2 in space. This means, we are looking for a static metric in the vacuum (outside of the matter) that is spherical symmetric. Choose coordinates $(\tau, \rho, \theta, \phi)$, where (ρ, θ, ϕ) are spherical coordinates. Then we can assume:

$$\begin{aligned} g = &-f_1(\rho)c^2 d\tau \otimes d\tau + f_2(\rho) d\rho \otimes d\rho + f_3(\rho)(d\tau \otimes d\rho + d\rho \otimes d\tau) \\ &+ f_4(\rho)\rho^2 \left(d\theta \otimes d\theta + \sin(\theta)^2 d\phi \otimes d\phi\right) \ . \end{aligned}$$

Because of the spherical symmetry, the functions f_i can only depend on ρ. Also, cross terms involving the angles are not allowed, as $\theta \to -\theta$ would lead to $d\theta \to -d\theta$, violating the spherical symmetry. Choosing coordinates $r = \psi(\rho)$, such that $f_1(r) = 1$ and $t = \tau + \xi(r)$, such that the cross terms vanish, the metric can be brought in the so called standard form for spherically symmetric, isotropic systems:

$$g = -A(r)c^2 dt \otimes dt + B(r) dr \otimes dr + r^2 \left(d\theta \otimes d\theta + \sin(\theta)^2 d\phi \otimes d\phi\right) \ .$$

$$\Rightarrow \qquad g_{\mu\nu} = \mathrm{diag}(-A(r)c^2, B(r), r^2, r^2 \sin(\theta)^2) \ .$$

Furthermore, for large r, the spacetime should become flat, since in classical Newtonian gravity, the gravitational potential converges asymptotically to zero for $r \to 0$. This means, that for $r \to \infty$, it should hold that $g \to \eta$, where η is the Minkowski metric. Thus it has to hold that:

$$A(r) \longrightarrow 1 \quad \text{and} \quad B(r) \longrightarrow 1 \qquad \text{for} \qquad r \to \infty \ .$$

As assumed before, the energy momentum tensor has compact support. A solution of the Einstein field equations outside of the support thus is a vacuum solution. Plugging in the standard form, and going through a long calculation (Christoffel symbols → Curvature tensor coefficients → Ricci coefficients and Scalar curvature → solving to get expressions for A and B) yields the famous **Schwarzschild metric**:

$$g = -c^2 \left(1 - \frac{r_S}{r}\right) dt \otimes dt + \frac{1}{1 - \frac{r_S}{r}} dr \otimes dr + r^2 d\theta \otimes d\theta + r^2 \sin(\theta)^2 d\phi \otimes d\phi \ .$$

] Here r_S is the so called **Schwarzschild radius** $r_S = \frac{2GM}{c^2}$. If r approaches the Schwarzschild radius, the metric becomes singular. A priori, this singularity has no meaning, as it is only a coordinate singularity, that need not hold for all coordinates. More physically speaking: For all stars, there is a Schwarzschild radius. However, if it lies inside the star, the Schwarzschild metric is no longer a solution, as the inside of the star is not vacuum.

However, if the Schwarzschild radius is outside of the support of the energy momentum tensor, then $r = r_S$ is in the region of vacuum, for which the Schwarzschild metric is a solution. Here the coordinates apply, and the singularity has the physical meaning of the event horizon. A concept further discussed, when considering black holes.

5.2. Movement in a Schwarzschild system

Neglecting, that small and/or far away masses also curve spacetime, the solar system is an example of a Schwarzschild system, assuming that the sun is nearly spherically symmetric. Hence the movement in a Schwarzschild system have a practical application, apart from theoretical interest. Here we will qualitatively compare the relativistic case to the Newtonian case, using the results and ideas from [Car97, chapter 7].

To find geodesics, the equation $\nabla_{\dot\gamma}\dot\gamma = 0$ has to be solved. In local coordinates, this conditions reads (corollary C.4.3):

$$\ddot\gamma^\rho(s) = -\Gamma^\rho_{\mu\nu}(\gamma(s))\dot\gamma^\mu(s)\dot\gamma^\nu(s) \ .$$

Riemannian normal coordinates allow to find coordinates, such that $g(p) = \eta(p)$ for a point in space time (compare theorem C.4.17), such that lemma 2.5.10 also applies in general relativity. This means $g(\dot\gamma, \dot\gamma) = -c^2$ for massive particles and $g(\dot\gamma, \dot\gamma) = 0$, or in local coordinates:

$$g_{\mu\nu}\dot\gamma^\mu(s)\dot\gamma^\nu(s) = \varepsilon := \begin{cases} -c^2 & \text{, massive particles} \\ 0 & \text{, massless particles} \end{cases} \ . \tag{5.1}$$

Using the coordinates for the standard form of a spherically symmetric, static metric, these conditions give a set of equations. Again, we will not bother with the tedious calculations and present only the ideas and results:

- To keep the notation short we write $r \equiv \gamma^r$ etc.
- The first step is to observe that the spherical symmetry of the problem allows to arrange the coordinates, such that $\theta = \frac{\pi}{2}$.

- A direct calculation verifies that the Schwarzschild metric has the four Killing vector fields $\partial_t, \partial_r, \partial_\theta$ and ∂_ϕ.
- Invariance under time translations corresponds to energy, such that ∂_t is the Killing vector belonging to energy. Using theorem C.6.6 we find:

$$E = -c^2 \left(1 - \frac{r_S}{r}\right) \frac{d}{ds} t(s) \ ,$$

where E is the constant we associate with energy.

- By the choice of coordinates, invariance under rotations of ϕ correspond to the total angular momentum (per unit mass) ℓ. Hence:

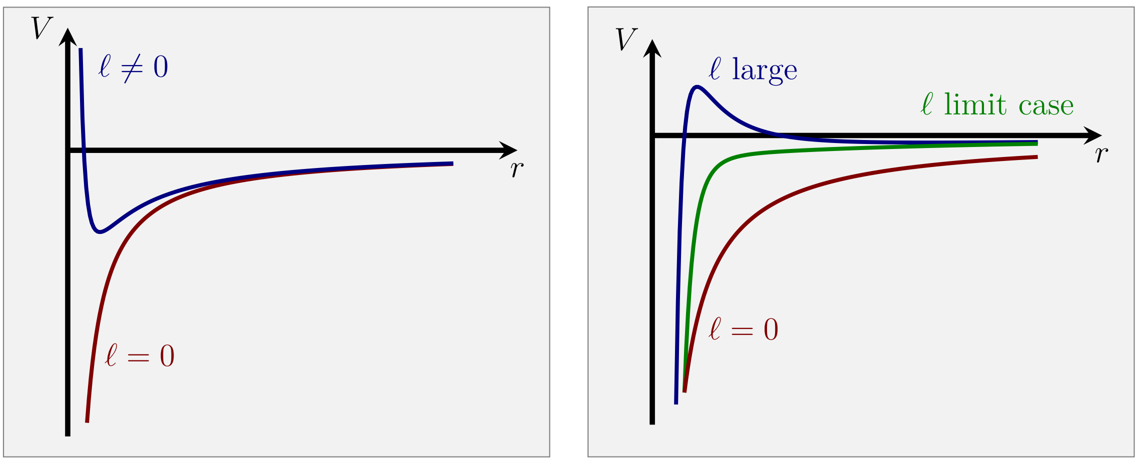
$$\ell = r^2 \frac{d}{ds} \phi(s) \ .$$

- Writing equation (5.1) explicitly for the Schwarzschild metric, plugging in the expressions for E and ℓ and calculating a bit longer yields:

$$\boxed{\frac{1}{2}\left(\frac{d}{ds} r\right)^2 + V(r) = \frac{1}{2} E^2}$$

with effective potential (according to [Fli16, eq. 25.27]):

$$V = \begin{cases} -\frac{GM}{r} + \frac{\ell^2}{2r^2} - \frac{GM\ell^2}{c^2r^3} & \text{, massive particles} \\ \frac{\ell^2}{2r^2} - \frac{GM\ell^2}{c^2r^3} & \text{, massless particles} \end{cases} \ .$$



(a) Classical case

(b) Relativistic case

Figure 5.1.: Sketches of effective potentials in the classical and relativistical case.

Although only a solution for the radial component it allows to find effects that distinguish general relativity from Newtonian gravity. For comparison, in the Kepler problem, for a massive particle, one obtains the following effective potential:

$$V(r) = -\frac{GM}{r} + \frac{\ell^2}{2r^2}.$$

In the classical theory, as long as the angular momentum is not zero, the potential will diverge to infinity for $r \to 0$, preventing a particle with angular momentum from falling into the star (center of mass). Also, if the angular momentum is non-zero, there is a stable equilibrium.

However, in the relativistic case, the extra term $-\frac{GM\ell^2}{c^2r^3}$ has the potential diverge to $-\infty$ for $r \to 0$. This means, however large the angular momentum may be, there is a radius, that if a particle falls below it, it will fall into the star. In the case $r \to \infty$, the relativity term will become comparatively small, such that the large r behavior is as the classical one. For the angular momentum, there can be three cases (quantitatively by calculating local extrema). For large ℓ, the effective potential has a local maximum followed by a local minimum for increasing r. Hence, there is an instable and a stable equilibrium. If ℓ decreases, the equlibria will move towards each other, until they coincide in the limit case. Finally, for smaller angular momenta, the effective potential will be exclusively attractive, as the classical $\ell = 0$ case.

A

Tensors and Index-Notation

Tensors are a vital concept in physics. This chapter tries to introduce the underlying algebraic concept of tensor products following [RW05] and [HO07]. As result, the universal property is the starting point, allowing to prove existence and the well known properties for calculations. We conclude this chapter with an introduction to Ricci-calculus, as given in [Jän05], observing the invariant isomorphisms behind index manipulations.

A.1. Tensor product

Usually (at least in the physical literature) tensors are defined by transformation behavior or calculational properties. Though this approach delivers a ready introduction for calculations, the concepts remain non-transparent. The most transparent way, yet sadly also the most abstract way, is the definition by universal properties commonly used in abstract algebra. However, the fundamental behavior, the transformation behavior, follows as direct corollary. An important structure left behind in the approach of transformation behavior, is the tensor product, that will be the first object to investigate here.

A.1.1. Existence and uniqueness

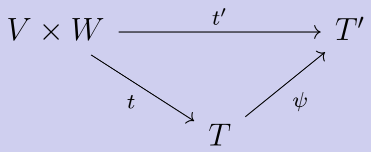
The following section is rather technical and can be skipped, if one is only interested in the behavior of tensors.

Definition A.1.1 (Universal property)**.**
Let V and W be $\mathbb{K}$-vector spaces. Also let (T, t) be a tupel, consisting of a vector space T and a bilinear map $t\colon V \times W \to T$. The tuple is called **tensor product** if the following universal property is fulfilled:
Let U be another vector space and $f\colon V \times W \to U$ be a linear map. Then there exists a linear map $\varphi_f\colon T \to U$, such that the following diagram commutes:

$$\begin{array}{ccc} V \times W & \xrightarrow{\quad f \quad} & U \\ & {\scriptstyle t}\searrow \quad \nearrow{\scriptstyle \varphi} & \\ & T & \end{array}$$

Theorem A.1.2 (Existence and uniqueness)**.**
For any two $\mathbb{K}$- vector spaces V and W, there exists always a tensor product (T, t). This tensor product is unique up to isomorphy. That is, if (T', t') is a second tensor product, then there exists a defined isomorphism $\Psi\colon T \to T'$ such

that the following diagram commutes:



Proof A.1.3.
Preparations: Let $\mathbb{K}^M$ denote the set of maps from M in the field $\mathbb{K}$. The set $\mathbb{K}^M$ is a vector space with point wise addition. A special class of maps in this vector space are *Kronecker-deltas*

$$\delta_m \colon M \longrightarrow \mathbb{K} \ , \quad x \longmapsto \delta_m(x) := \begin{cases} 1 & , x = m \\ 0 & , \text{sonst} \end{cases} \ .$$

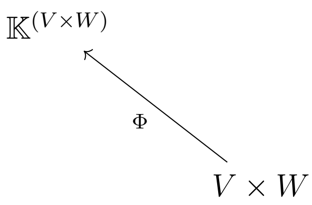
This gives rise to an embedding Φ of M in $\mathbb{K}^M$:

$$\Phi \colon M \hookrightarrow \mathbb{K}^M \ , \quad m \mapsto \delta_m \ .$$

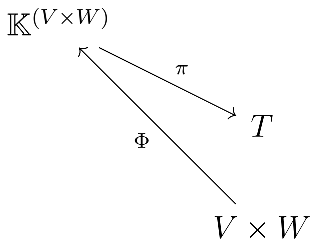
An important subspace of $\mathbb{K}^M$ for this proof is the space of maps with $f(x) = 0$ for all but finitely many $x \in M$. This subspace will be denoted by $\mathbb{K}^{(M)}$. Every element of this space can be written as linear combination of finitely many deltas δ_m. By definition the deltas are linear independent and thus forming a basis of $\mathbb{K}^{(M)}$:

$$\mathbb{K}^{(M)} = \text{span}_{\mathbb{K}}(im(\Phi)) \ .$$

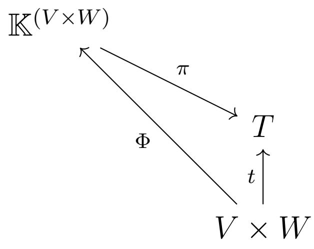
Existence:
The existence follows from inspecting a commutative diagram:



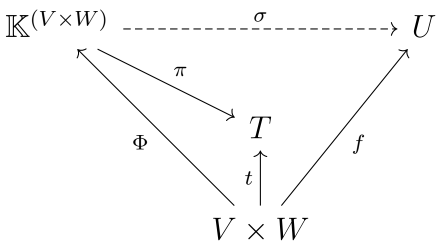
(a)



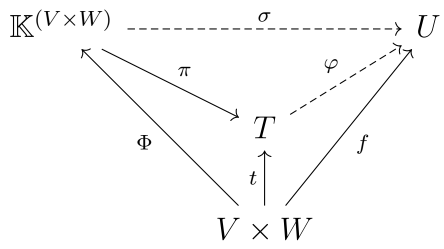
(b)



(c)



(d)



(e)

(a) As we have seen, there is an embedding, that is, an injective map $\Phi\colon V \times W \to \mathbb{K}^{(V\times W)}$.

(b) We define the subspace $X \subset \mathbb{K}^{(V\times W)}$. Let $v, v' \in V$, $w, w' \in W$ and $a \in \mathbb{K}$, then X shall be defined as linear span of the following elements

$$\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)} \ , \quad \delta_{(v,w+w')} - \delta_{(v,w)} - \delta_{(v,w')} \ ,$$

$$\delta_{(av,w)} - \delta_{a(v,w)} \quad \text{and} \quad \delta_{(v,aw)} - \delta_{a(v,w)} \ .$$

The space T can now be defined as quotient $T \coloneqq \mathbb{K}^{(V\times W)}/X$. Hence T is a space of equivalence classes with the following equivalence relation:

$$h \sim h' \quad \Leftrightarrow \quad \exists\, x \in X\colon x' = h + x \ .$$

Let $\pi\colon \mathbb{K}^{(V\times W)} \to T$ be the canonical projection, i.e. the surjective map assigning every $h \in \mathbb{K}^{(V\times W)}$ its equivalence class $\pi(h) = [h] \in \mathbb{K}^{(V\times W)}/X$.

(c) The map $t\colon V \times W \to T$ will be defined by $t = \pi \circ \Phi$. Due to the choice of X the map is bilinear:

$$t((v + v', w) - (v, w) - (v', w)) = \pi(\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)}) = [0]$$

$$\Rightarrow \quad [\delta_{(v+v',w)}] = [\delta_{(v,w)}] + [\delta_{(v',w)}] \quad \Leftrightarrow \quad t((v+v', w)) = t((v, w)) + t((v', w)) \ .$$

The remaining properties can be shown similarly.

(d) It remains to show, that (T, t) satisfies the universal property. So let U be a $\mathbb{K}$-vector space and $f\colon V \times W \to U$ a bilinear map. The image of Φ defines a basis of $\mathbb{K}^{(V\times W)}$. Define the map σ by

$$\sigma(\Phi(v, w)) = f(v, w) \ .$$

By linear completion σ is a linear map $\mathbb{K}^{(V\times W)} \to U$. From the bilinearity of f follows the bilinearity of σ:

$$\sigma(\Phi(v + v', w)) = f((v + v', w)) = f(v, w) + f(v', w) = \sigma(\Phi(v, w)) + \sigma(\Phi(v', w)) \ .$$

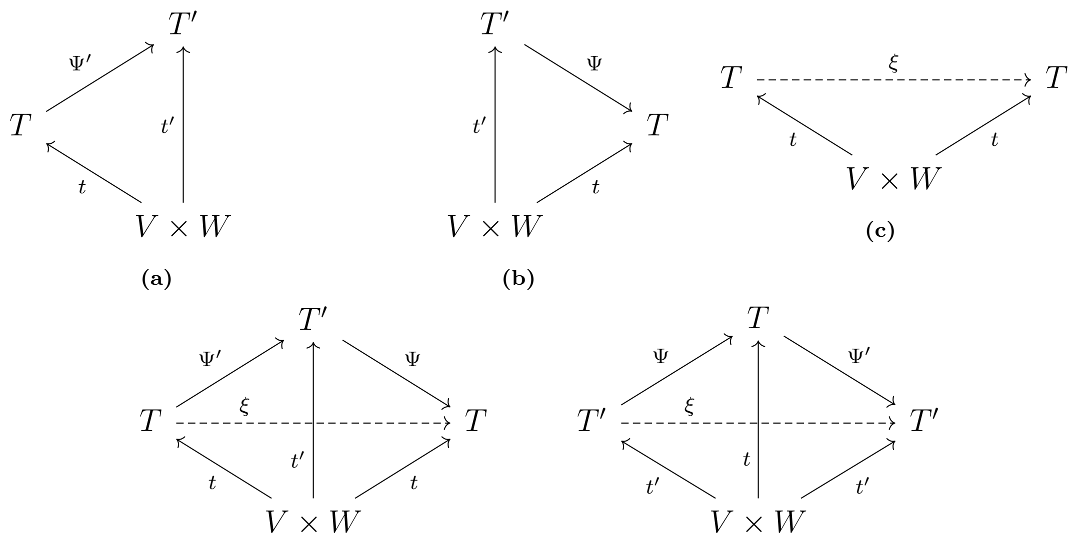
Hence $U \subseteq \ker(\sigma)$.

(e) The fundamental theorem on homomorphisms states, that there exists a unique linear map φ, completely defined by σ, such that the upper triangle of the diagram commutes. Since σ is uniquely defined by f, so is φ, proving the universal property.

Uniqueness:
To prove uniqueness we use a second tensor space T' for U:

(a) Since we have already proven the universal property, we know that there is a unique map $\Psi'\colon T \to T'$, defined by t'.

(b) Similarly there is a unique map $\Psi\colon T' \to T$ defined by t.



(a) (b) (c) (d) (e)

(c) Also there is a unique linear map $\xi \in \mathrm{End}(T)$, such that

$$\xi \circ t = t$$

holds. Furthermore, $Id_T \circ t = t$ holds. Since ξ was unique, $\xi = Id_T$.

(d) Combining the diagrams, which commute by definition, shows that the whole diagram does, too. Hence:

$$Id_T = \xi = \Psi' \circ \Psi \ .$$

(e) Similarly it follows that $Id_{T'} = \Psi \circ \Psi'$. Due to the uniqueness of Ψ, being determined by t and the uniqueness of Ψ', being determined by t', theses maps coincide in both diagrams. Thus finally $\Psi' = \Psi^{-1}$, i.e. it is an isomorphism.

□

A.1.2. Tensors

The tensor product is unique up to isomorphy, so it is common to speak about the tensor product. The usual notation for (T, t) is $(V \otimes W, \otimes)$. As long as the field $\mathbb{K}$ is understood, one can write $\otimes$, otherwise one needs to specify the field, e.g. $\otimes_{\mathbb{K}}$. This is important, since $\otimes$ is bilinear only with respect to $\mathbb{K}$:

$$(v + v') \otimes w = v \otimes w + v' \otimes w \ ,$$

$$v \otimes (w + w') = v \otimes w + v \otimes v' \ ,$$

$$(\alpha v) \otimes w = \alpha(v \otimes w) = v \otimes (\alpha w) \ .$$

Definition A.1.4.
Tensors of the form $v \otimes w \in V \otimes W$ with $v \in V$ and $w \in W$ are called **pure tensors**

By the construction of $V \otimes W$ it follows that every tensor can be written as sum of pure tensors.

Theorem A.1.5.
Let $\{v_i\}_{i\in I}$ be a basis of V and $\{w_j\}_{j\in J}$ a basis of W for finite dimensional vector spaces. Then $\{v_i \otimes v_j\}_{i\in I, j\in J}$ is a basis of $V \otimes W$ and it follows that:

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W) \ .$$

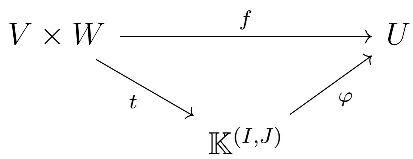
Proof A.1.6.
We consider the following map:

$$t' \colon V \times W \longrightarrow \mathbb{K}^{(I,J)} \ , \qquad \left(\sum_{i\in I} x_i v_i, \sum_{j\in J} y_j w_j \right) \longmapsto x_i y_j \cdot \delta_{(i,j)} \ .$$

Here I and J are the index sets of the bases of V and W. The map t' is bilinear and maps elements (v_i, w_j) to the basis elements $\delta_{(i,j)}$ of $\mathbb{K}^{(I,J)}$. By linear completion there is a unique linear map $\varphi \colon \mathbb{K}^{(I,J)} \to U$ for every bilinear map $f \colon V \times W \to U$, defined by

$$f(v,w) = \varphi(t'(v,w)) \qquad \forall\, v \in V,\ w \in W \ .$$

Thus the following diagram commutes:



Hence the tuple $(\mathbb{K}^{(I,J)}, t')$ satisfies the universal property. With theorem A.1.2 we find $\mathbb{K}^{(I,J)} \cong V \otimes W$. The isomorphism Ψ between those vector spaces has the property $\Psi \circ t' = \otimes$, mapping $t'(v_i, w_j) = \delta_{(i,j)}$ to $v_i \otimes w_j$. Since $\{\delta_{(i,j)}\}$ is a basis of $\mathbb{K}^{(I,J)}$, so is $\{v_i \otimes w_j\}$ a basis of $V \otimes W$. $\square$

With the universal property we can prove the following isomorphies:

Lemma A.1.7 (Ismorphisms of tensor spaces)**.**
Let V, W and U be $\mathbb{K}$-vector spaces, then the following isomorphies are unique for the stated conditions:

(i) $V \otimes W \cong W \otimes V, \quad v \otimes w \mapsto w \otimes v.$

(ii) $(U \otimes V) \otimes W \simeq U \otimes (V \otimes W), \quad (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w).$

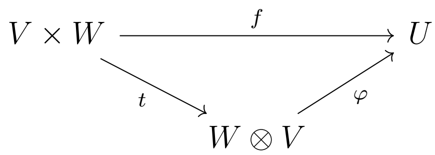
(iii) $(U \bigoplus V) \otimes W \simeq (U \otimes W) \bigoplus (V \otimes W) \quad (u,v) \otimes w \mapsto (u \otimes w, v \otimes w).$

(iv) $\mathbb{K} \otimes_{\mathbb{K}} V \cong V, \quad a \otimes v \mapsto a \cdot v.$

Proof A.1.8 (Œ for (i))**.**
Let $t\colon V \times W \to W \otimes V$ be defined by $t(v,w) = w \otimes v$. Every bilinear map $f\colon V \times W \to U$ determines $\varphi\colon W \times V \to U$ uniquely by

$$\varphi(w \otimes v) = f(v,w) \ ,$$

such that the following diagram commutes:



Thus $(W \otimes V, t)$ satisfies the universal property and is isomorphic to $V \otimes W$ due to theorem A.1.2. The remaining isomorphisms can be proven similarly.[1] □

The last isomorphism (iv) is only valid, if the field used as vector space is the same as the field used to define the tensor product. If $\mathbb{K}' \supset \mathbb{K}$ is a field containing $\mathbb{K}$ as subfield (e.g. $\mathbb{C}$ and $\mathbb{R}$), then statement (iv) fails: $\mathbb{K}' \otimes_{\mathbb{K}} V \ncong V$. Yet $\mathbb{K}' \otimes_{\mathbb{K}} V$ becomes a $\mathbb{K}'$-vector space. That is, the scalar range of V is extended to $\mathbb{K}'$, by

$$a' \cdot (b' \otimes v) = (a' \cdot b') \otimes v \ .$$

Definition A.1.9.
Let $\mathbb{K}$ be a subfield of $\mathbb{C}$, i.e. $\mathbb{K} = \mathbb{Q}$ or $\mathbb{R}$, and let V be a $\mathbb{K}$-vector space. The tensor product $V^{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{K}} V$ is called **complexification** of V.

It can be shown, that if $\{v_j\}_{j \in J}$ is a basis of V, so is $\{1 \oplus_{\mathbb{K}} v_j\}_{j \in J}$ of $V^{\mathbb{K}'}$.

Lemma A.1.10.
Let V_1, V_2, W_1, W_2 be vector spaces, and $\varphi_1\colon V_1 \to W_1$ as well as $\varphi_2\colon V_2 \to W_2$ be linear maps. Then there is a unique linear map

$$\varphi_1 \otimes \varphi_2\colon V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2 \ , \qquad (\varphi_1 \otimes \varphi_2)(v_1 \otimes v_2) = \varphi_1(v_1) \otimes \varphi_2(v)_2 \ .$$

This tensor product of linear maps has the following properties:

1. $Id_{V_1} \otimes Id_{V_2} = Id_{V \otimes V_2}$.
2. *For two additional linear maps $\varphi_1'\colon V_1 \to W_1$ and $\varphi_2'\colon V_2 \to W_2$, it holds that:*
$$(\varphi_1 \otimes \varphi_2) \circ (\varphi_1' \otimes \varphi_2') = (\varphi_1 \circ \varphi_1') \otimes (\varphi_2 \circ \varphi_2') \ .$$

[1] For example see [RW05, p. 4-5] for (iv).

3. *The map*

$$\begin{aligned}\mathrm{Hom}(V_1,W_1)\times\mathrm{Hom}(V_2,W_2)&\hookrightarrow\mathrm{Hom}(V_1\otimes V_2,W_1\otimes W_2)\\(\varphi_1,\varphi_2)&\mapsto\varphi_1\otimes\varphi_2\end{aligned}$$

is bilinear and injective. For finite-dimensional vector spaces it is an isomorphism.

From the definition of the algebraic dual space $V^* = \mathrm{Hom}(V,\mathbb{K})$ and the isomorphy $\mathbb{K}\otimes_{\mathbb{K}} V \cong V$ the injection

$$V_1^*\otimes V_2^* \hookrightarrow (V_1\otimes V_2)^* \ ,$$

follows as direct result from the previous lemma. Accordingly it is an isomorphism for finite-dimensional vector spaces.

Lemma A.1.11.
There is an embedding $V_2\otimes V_1^*$ into $\mathrm{Hom}(V_1,V_2)$, defined by the following injective linear map

$$V_2\otimes V_1^* \hookrightarrow \mathrm{Hom}(V_1,V_2)\ , \quad v_2\otimes\vartheta_1\mapsto \ell_{v_2,\vartheta_1}\ ,$$

where $\ell_{v_2,\vartheta_1}(v)=\vartheta_1(v)\cdot v_2$. This map can be extended linearly for $V_2\otimes V_1^*$.

As before, the embedding becomes an isomorphy in the finite-dimensional case.

Corollary A.1.12.
Let $\{e_i\}_i$ be a basis of a finite-dimensional vector space V and $\{\vartheta_j\}_j$ the dual basis. Then, every linear operator $L\in\mathrm{Hom}(V,W)$ can be written as tensor from $W\otimes V^*$:

$$L=\sum_i (Le_i)\otimes\vartheta_i \ .$$

Although the tensor product is defined for infinite-dimensional vector spaces, too, the tensor product of two Hilbert spaces does not need to be one as well. The missing property is completeness here.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. One can define a scalar product on $\mathcal{H}_1\otimes\mathcal{H}_2$ by

$$\langle v_1\otimes v_2|w_1\otimes v_2\rangle = \langle v_1|w_1\rangle\cdot\langle v_2|w_2\rangle \ .$$

To obtain a Hilbert space, one can take the metric completion with respect to that scalar product.

We have spoken about bases so far, even in the infinite-demensional case. A basis allows to linear combine every element of the space with finitely many basis elements.[2] Theorem A.1.5 can be extended to Hilbert-bases.

[2]Contrasting Hilbert-bases, where the sum is infinite. In infinite spaces, a true basis may even become uncountable.

A.2. Ricci calculus

In Ricci calculus, tensors are characterized by their coefficients (becoming a list of numbers). There are three important principles we need to understand:

1) Components define objects
2) Position of indices determines transformation behavior.
3) Summation convention.

A.2.1. Co-and contravarianz

Let V be a vector space, $\{e_i\}_{i=1,\dots,n}$ a basis and V^* be the dual space with dual basis $\{\vartheta^i\}_{i=1,\dots,n}$. A vector $v \in V$ is called **contravariant** and is described by coefficients with upper indices:

$$v = \sum_{i=1}^{n} v^i e_i \equiv v^i e_i \ .$$

A dual vector $\varphi \in V^*$ is called **covariant** and is described by coefficients with lower indices

$$\varphi = \varphi_i \vartheta^i \ .$$

In the last equation we have already used the summation convention. Over same indices, one upper and one lower, will always be summed (without having to write the summation symbol).

Changing the basis $e_i \to \tilde{e}_i$ does not change the element v, but its components:

$$v = v^i e_i = \tilde{v}^k \tilde{e}_k \ .$$

By definition there are coefficients, such that $e_i = A^j{}_i \tilde{e}_j$ can be written. Plugging in yields the connection between v^i and $\tilde{v}^k$:

$$v = \tilde{v}^k \tilde{e}_k = v^i e_i = v^i A^j{}_i \tilde{e}_j \qquad \Rightarrow \qquad \tilde{v}^k = A^k{}_i v^i \ .$$

For the coefficients of the matrix A, we have already used the Ricci convention. Still, the first index describes the row and the second index the columns, independent if it is an upper or lower index.

Remark A.2.1 (Composition of linear maps)**.**
Let $A, B \in \mathrm{End}(V)$ for a finite dimensional vector space V. Due to the isomorphy $\mathrm{End}(V) \simeq V \otimes V^*$, these maps can be written as tensors:

$$A = Ae_i \otimes \vartheta^i = A^j{}_i e_j \otimes \vartheta^i \quad \text{and} \quad B = B^k{}_\ell e_k \otimes \vartheta^\ell \ .$$

Evaluating $(A \circ B)(v)$ for an arbitrary vector $v \in V$ yields

$$\begin{aligned}(A \circ B)(v) &= A^j{}_i e_j \cdot \vartheta^i \left(B^k{}_\ell e_k \cdot \vartheta^\ell(v)\right) = A^j{}_i B^k{}_\ell \vartheta^\ell(v) \vartheta^i(e_k) \cdot e_j \\ &= A^j{}_i B^k{}_\ell \vartheta^\ell(v) \delta^i_k \cdot e_j = A^j{}_i B^i{}_\ell \vartheta^\ell(v) \cdot e_j \\ &= \left(\left(A^j{}_i B^i{}_\ell\right) e_j \otimes \vartheta^\ell\right)(v) \ .\end{aligned}$$

From the last equality we can read off the coefficient behavior under composition:

$$\boxed{(A \circ B)^j{}_\ell = A^j{}_i B^i{}_\ell}$$

If A is a basis change matrix, there is an inverse A^{-1}. From the above remark we know, that this can be expressed by $A^i{}_j (A^{-1})^j{}_k = \delta^i_k$. Dual vectors are linear and defined by $\vartheta^i(e_j) = \delta^i_j$ and respectively $\tilde{\vartheta}^i(\tilde{e}_j) = \delta^i_j$, thus we find:

$$\delta^i_j = \vartheta_i(e_j) = \vartheta_i(A^k{}_j \tilde{e}_k) = A^k{}_j \vartheta_i(\tilde{e}_k) \ .$$

Since the dual vector space is also a vector space, there are coefficients such that $\vartheta^i = M^i{}_\ell \tilde{\vartheta}^\ell$. Plugging in results in the transformation behavior of covectors:

$$\delta^i_j = A^k{}_j \vartheta^i(\tilde{e}_k) = A^k{}_j M^i{}_\ell \tilde{\vartheta}^\ell(\tilde{e}_k) = A^k{}_j M^i{}_\ell \delta^k_\ell = A^k{}_j M^i{}_k = M^i{}_k A^k{}_j \quad \Rightarrow \quad M = A^{-1} \ .$$

$$\Rightarrow \qquad \varphi = \tilde{\varphi}_j \tilde{\vartheta}^j = \varphi_i \vartheta^i = \varphi_i (A^{-1})^i{}_k \tilde{\vartheta}^k \qquad \Rightarrow \qquad \tilde{\varphi}_j = (A^{-1})^i{}_j \varphi_i \ .$$

By definition, basis change matrices are orthogonal/unitary. That is $A^{-1} = A^\dagger$ / $A^{-1} = A^T$. Summing up our findings:

	coefficients	**transformation basis vectors**
contravariant	$\tilde{v}^k = A^k{}_i v^i$	$\tilde{e}_k = (A^T)^i{}_k e_i$
covariant	$\tilde{\varphi}_k = (A^T)^i{}_k \varphi_i$	$\tilde{\vartheta}^k = A^k{}_i \vartheta^i$

A.2.2. Tensors in Ricci calculus

After we have seen the foundations of Ricci calculus we can use this formulation on tensors:

Definition A.2.2.
A tensor, consisting of r vectors and s covectors

$$T \in \underbrace{V \otimes ... \otimes V}_{r \text{ times}} \otimes \underbrace{V^* \otimes ... \otimes V^*}_{s \text{ times}}$$

is called **tensor of type (r, s)**. The number $r + s$ is called the **rank**, also for a general order of vectors and covectors.

A tensor of type (r, s) can be expanded as follows:

$$T = T^{i_1 \dots i_r}{}_{j_1 \dots j_s} \, e_{i_1} \otimes ... \otimes e_{i_r} \otimes \vartheta^{j_1} \otimes ... \otimes \vartheta^{j_s} \ .$$

In Ricci calculus one agrees upon the following identification:

$$T = T^{i_1 \dots i_r}{}_{j_1 \dots j_s} \ .$$

A change of basis results in the following transformation behavior:

$$\tilde{T}^{n_1 \dots n_r}{}_{m_1 \dots m_s} = A^{n_1}{}_{i_1} \dots A^{n_r}{}_{i_r} (A^{-1})^{m_1}{}_{j_1} \dots (A^{-1})^{m_s}{}_{j_s} \, T^{i_1 \dots i_r}{}_{j_1 \dots j_s}$$

This behavior is used to define tensors in the physical literature.

A.2.3. Raising and lowering indices

In case of a Riemannian manifold there is a fourth principle, induced by the Riemannian metric:

4) Raising and lowering indices

To understand the invariant meaning behind these manipulations, instead of just defining them, it is best to use the coordinate free formulation first.

Let g be a scalar product (or Riemannian metric). Then there is an isomorphism

$$I_1 \colon \text{vector} \longrightarrow \text{covector} \qquad v \longmapsto I_1(v) = g(v, \cdot)\ .$$

Remark A.2.3.
In the literature, the isomorphism I_1 and its inverse I_1^{-1} are called **flat-** and **sharp isomorphism** respectively. The usual notation is

$$I_1(v) = v^\flat \qquad \text{and} \qquad I_1^{-1}(\omega) = \omega^\# \ .$$

These isomorphisms can be applied to individual parts of the tensor, still defining an isomorphism between tensor spaces. For example, a $(1, 1)$-tensor becomes a $(0, 2)$-tensor if $I_1 \otimes \mathbb{1}$ is applied, and a $(2, 0)$-tensor, if $\mathbb{1} \otimes I_1^{-1}$ is applied.

Let ∂_μ be a tangent basis and dx^ν be the dual basis of the cotangent space. Defining the coefficients of the Riemannian metric by $g_{\mu\nu} = g(\partial_\mu, \partial_\nu)$, there is an inverse matrix (list of numbers) $g^{\mu\nu}$. By definition of scalar products the matrices are symmetric: $g_{\mu\nu} = g_{\nu\mu}$ and hence $g^{\mu\nu} = g^{\nu\mu}$. For the isomorphisms I_1 and I_1^{-1} it follows that:

$$I_1(\partial_\mu) = g(\partial_\mu, \cdot) = g_{\mu\nu} dx^\nu \qquad \text{and thus}^{3} \qquad I_1^{-1}(dx^\mu) = g^{\mu\nu}\partial_\nu\ .$$

The coefficients transform as follows:

$$I_1(v^\mu \partial_\mu) = g_{\mu\nu} v^\mu dx^\nu =: v_\nu dx^\nu \qquad \text{and} \qquad I_1^{-1}(u_\mu dx^\mu) = g^{\mu\nu} u_\mu \partial_\nu =: u^\nu \partial_\nu\ .$$

Remark A.2.4.
A contravariant vector v^μ becomes a covariant vector v_ν by lowering the index:

$$v_\nu = g_{\mu\nu} v^\mu\ .$$

Raising an index on the other hand, transforms a covector into a vector:

$$u^\nu = g^{\mu\nu} u_\mu\ .$$

The raising and lowering can be applied for indices of tensors separately:

$$g_{\mu\nu} A^{\cdots\mu\cdots}{}_{\cdots} = A^{\cdots}{}_{\mu}{}^{\cdots}{}_{\cdots}\ .$$

[3] $\partial_\mu = I_1^{-1}(g_{\mu\nu} dx^\mu) = g_{\mu\nu} I_1^{-1}(dx^\mu) \quad \Rightarrow \quad g^{\mu\nu} g_{\mu\nu} I_1^{-1}(dx^\mu) = I_1^{-1}(dx^\mu) = g^{\mu\nu}\partial_\mu$.

Remark A.2.5.
The scalar product of two vectors $u^\mu \partial_\mu$ and $v^\nu \partial_\nu$ can be written as composition:

$$g(u^\mu \partial_\mu, v^\nu \partial_\nu) = u^\mu v^\nu g_{\mu\nu} = u_\nu v^\nu = v_\mu u^\mu \ .$$

B
Variational calculus for fields

In this chapter we cover the fundamentals of variational calculus. We introduce the mathematical notion of functionals and the concepts, as well as properties, of their derivatives, following [Wer11, chapter III.5]. In the second section we focus on the important class of functionals that can be expressed by Lagrange-densities, deriving the Euler-Lagrange-equations. We conclude this chapter by discussing notations of the physical literature, used in [AS10] for example.

B.1. Functional derivative and variation

In variational calculus (e.g. in the context of least action) the goal is to find functions, that minimize or maximize a given functional. In general a **functional** is a map from a normed space X into its number field. Recalling, that for functions $f\colon \mathbb{R} \to \mathbb{R}$ a vanishing first derivative is a necessary condition for extrema, we are looking for a similar computational tool in the case of functionals. We will see, that there is a concept of derivative, that is not only connected to the concept of variations, but also the computational tool we are looking for.

Definition B.1.1.
Let $F\colon U(x_0) \subset X \to \mathbb{R}$ be a functional defined on a neighborhood containing x_0. The ***n*-th variation** $\delta^n F(x_0; h)$ of F in direction of $h \in X$ is defined, if it exists, by

$$\delta^n F(x_0; h) = \frac{d^n}{dt^n} F(x_0 + th)\Big|_{t=0} .$$

The variation resembles the directional derivative of multivariable calculus. This leads to the following definition:

Definition B.1.2.
Let F be a functional as before. If $\delta F(x_0; h)$ exists for all $h \in X$ and there is a linear continuous functional $L\colon X \to \mathbb{R}$, such that

$$L(h) = \delta F(x_0; h) \quad \forall\ h \in X\ ,$$

the functional is called **Gâteaux-differentiable** in x_0. The linear continuous functional is called the **Gâteaux-derivative** $\widehat{G}_{x_0} F$ of F in x_0. Accordingly F is called Gâteaux-differentiable on X if it is Gâteaux-differentiable in all $x \in X$.

As in multivariable calculus, we are not only interested in the directional derivative, but also for the derivative itself. Even in the finite-dimensional case, the existence of the directional derivative in every direction is not enough to prove the existence of the total derivative. Writing the definition of the Gâteaux-derivative slightly different reveals the generalization necessary to define an analogue to the total differential:

$$\lim_{t\to 0} \frac{|F(x_0 + th) - F(x_0) - t\widehat{G}_{x_0} F(h)|}{t} = 0 \ .$$

In this form one can see easily the similarity to the directional derivative. In the finite-dimensional case one demands the sequence to be uniformly convergent to define the total differential, motivating the following definition:

Definition B.1.3.
Let $F\colon U(x_0) \subset X \to \mathbb{R}$ be a functional on a neighborhood around x_0. The functional is said to be **Fréchet-differentiable** in x_0, if there is a continuous linear functional $L\colon X \to \mathbb{R}$ such that

$$\lim_{X\ni h\to 0} \frac{|F(x_0 + h) - F(x_0) - L(h)|}{\|h\|_X} = 0 \ .$$

The continuous linear functional L is called the **Fréchet-derivative** $D_{x_0}F$ in x_0 of F.

From the definitions it is clear, that if a functional is Fréchet-differentiable in x_0 is is also Gâteaux-differentiable in x_0 and

$$D_{x_0}F(h) = \widehat{G}_{x_0}F(h) = \delta F(x_0; h) \ .$$

The condition to be Fréchet-differentiable can be reinterpreted as linearization of the functional. The existence of the limit is equivalent to

$$F(x_0 + h) - F(x_0) - D_{x_0}F(h) = r(h) \qquad \text{with} \qquad \lim_{X\ni h\to 0} \frac{r(h)}{\|h\|_X} = 0 \ .$$

In introductory texts about analytical mechanics the concept of **functional derivative** is introduced as follows: If the change of the functional $F(x_0 + h) - F(x_0)$ can be written as sum of a part $L(x_0, h)$ that is linear in h and a part $R(x_0, h)$ that decreases faster than $\|h\|_X$, i.e.

$$F(x_0 + h) - F(x_0) = L(x_0, h) + R(x_0, h) = D_{x_0}F(h) + r(h) \ ,$$

the functional L is called functional derivative of F in x_0. With the knowledge about Fréchet-derivatives, one can clearly see, that theses texts introduce the Fréchet-derivative, which can be defined in even more general cases. The Fréchet-derivative of an operator $F\colon X \to Y$ can be defined as before, using $\|\cdot\|_Y$ instead of $|\cdot|$.

Theorem B.1.4 (Properties of the Fréchet-derivative)**.**

i) The continuous linear operator L defining the Fréchet-derivative is unique.

ii) If F and G are Fréchet-differentiable in x_0, so are $F + G$ and λF for all

$\lambda \in \mathbb{R}$, and it holds, that

$$D_{x_0}(F+G) = D_{x_0}F + D_{x_0}G \qquad \text{and} \qquad D_{x_0}(\lambda F) = \lambda D_{x_0}F \ .$$

iii) If F and G are Fréchet-differentiable in x_0, then $F \cdot G$ is Fréchet-differentiable and the Fréchet-derivative satisfies the product rule:

$$D_{x_0}(F \cdot G) = G(x_0) D_{x_0}F + F(x_0) D_{x_0}G \ .$$

iv) Let X, Y, Z be normed spaces, and $F\colon D(F) \subset X \to Y$, as well as $G\colon D(G) \subset Y \to Z$ be Fréchet-differentiable operators with $F(D(F)) \subset D(G)$. Then $G \circ F$ is Fréchet-differentiable and the chain-rule applies:

$$D_{x_0}(G \circ F) = D_{F(x_0)}G \circ D_{x_0}F \ .$$

The first property allows to speak about "the" Fréchet-derivative. Also it allows to prove the remaining properties in the same way one does in the finite-dimensional case, by using the uniqueness of the linearization. If Fréchet-differentiability is given, the same rules apply to the Gâteaux-derivative and hence also for the first variation. Yet the opposite direction is not true.

By the similarities we have encountered so far, it is hardly a surprise that a vanishing Fréchet-derivative and thus vanishing variations for all h, is a necessary condition for extrema. Sufficient conditions, unlike in the finite-dimensional case, are however more complicated. In most cases, at least for physical theories, it is not important if a solution is minimizing or maximizing. Otherwise, usually there is an easy way to determine what kind of extremum it is, by physical reasoning.

B.2. Variation of Lagrange densities

In most physical applications the functionals of interest can be written as integration over a Lagrange density. A **Lagrange-density** $\mathcal{L}$ is an object, that maps functions to functions and thus defining itself a function of functions (and its variables):

$$\mathcal{L}(f(x), g(x), \ldots, x) = \ell(x) \ ,$$

where $\ell\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is a proper function. Furthermore, the functionals $S(\phi)$ of interest often have Lagrange densities that do only depend on the components of the fields ${\phi^{a\ldots b}}_{c\ldots d}$ and their partial derivatives ${\phi^{a\ldots b}}_{c\ldots d,\mu}$:

$$\boxed{S(\phi) = \int_U \mathcal{L}({\phi^{a\ldots b}}_{c\ldots d}(x), {\phi^{a\ldots b}}_{c\ldots d,\mu}(x), x)\ d^n x.}$$

Here U has to be a compact subset of $\mathbb{R}^n$. In the case of invariant theories on Riemannian manifolds one can also use covariant derivative instead of partial derivatives. We shorten the notation of tensor components in the following, using $\phi^I_J \equiv {\phi^{a\ldots b}}_{c\ldots d}$.

Theorem B.2.1.
Let $S(\phi) = \int_U \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)\ d^n x$ be a Fréchet-differentiable functional with two times continuously differentiable Lagrangian. Then the first variation over fields, vanishing at the boundary ∂U, can be written as follows:

$$\delta S(\phi;\psi) = \int_U \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \sum_{\mu=1}^{n} \frac{\partial}{\partial \mu} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) \psi^I_J\ d^n x \ .$$

Proof B.2.2.
Assuming Fréchet-differentiability we know that the Gâteaux-derivative exists and can simply calculate the first variation for the proof:

$$S(\phi + t\psi) - S(\phi) = \int_U \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) - \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)\ d^n x \ .$$

The difference of the Lagrange-densities can be Taylor-expanded for small t, which is given, when we take the limit.

$$\begin{aligned} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) = \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x) \\ + t \left.\frac{d}{dt}\right|_{t=0} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) \\ + \mathcal{O}(t^2) \ . \end{aligned}$$

Thus the integrand becomes:

$$\begin{aligned} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) - \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x) = t \left.\frac{d}{dt}\right|_{t=0} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) \\ + \mathcal{O}(t^2) \ . \end{aligned}$$

Evaluating the parameter derivative, using the parameter chain rule[1], we find

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) = \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} \psi^I_J + \sum_{\mu=1}^{n} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \psi^I_{J,\mu} \ .$$

Also we observe that

$$\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \psi^I_{J,\mu} = \frac{\partial}{\partial \mu} \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \psi^I_J \right) - \left(\frac{\partial}{\partial \mu} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) \psi^I_J \ .$$

Inserting our findings in the expression for $S(\phi + t\psi) - S(\phi)$ we get:

$$\begin{aligned} S(\phi + t\psi) - S(\phi) = t \int_U \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \sum_{\mu=1}^{n} \frac{\partial}{\partial \mu} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) \psi^I_J\ d^n x \\ + t \int_U \frac{\partial}{\partial \mu} \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \psi^I_J \right)\ d^n x + \mathcal{O}(t^2) \ . \end{aligned}$$

The integral in the second line can be rewritten using Stokes theorem (better known as Gauss-theorem or divergence theorem here). By assumption $\psi^I_J = 0$ on ∂U we find:

$$\int_U \frac{\partial}{\partial\mu}\left(\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_{J,\mu}}\psi^I_J\right)\, d^n x = \int_{\partial U} \frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_{J,\mu}}\psi^I_J \, d\Omega = 0 \ .$$

Finally with $\delta S(\phi;\psi) = \lim_{t\to 0} \frac{1}{t} S(\phi + t\psi) - S(\phi)$ we get an expression for the first variation:

$$\delta S(\phi;\psi) = \int_U \left(\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_J} - \sum_{\mu=1}^n \frac{\partial}{\partial\mu}\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_{J,\mu}}\right)\psi^I_J \, d^n x \ .$$

□

It is common practice to imply the sum over μ by using ∂_μ instead of $\frac{\partial}{\partial\mu}$ in the sense of Ricci-calculus.

Lemma B.2.3 (Euler-Lagrange-equations)**.**
A smooth field ϕ extremizes the functional $S(\phi) = \int_U \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)\, d^n x$ with the boundary condition $\phi^I_J|_{\partial U} \equiv 0$ if it satisfies the ***Euler-Lagrange-equations*** *:*

$$\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_J} - \partial_\mu \frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_{J,\mu}} = 0 \ .$$

Proof B.2.4.
All we need to do, is to show that $\delta S(\phi;\psi) = 0$, is equivalent to the Euler-Lagrange-equations. From the last theorem we know that

$$\delta S(\phi;\psi) = \int_U \left(\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_J} - \partial_\mu \frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial\phi^I_{J,\mu}}\right)\psi^I_J \, d^n x \ .$$

If the Euler-Lagrange-equations are satisfied, we are integrating over the zero function, such that $\delta S(\phi;\psi) = 0$. The opposite direction is harder to show. Yet, there is an important theorem, called **Fundamental lemma of calculus of variations**, which states that an integrable function $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is identically zero, if

$$\int_\Omega f(x) g(x)\, d^n x = 0$$

for all $g \in C^\infty$ with compact support. Since we assumed U to be compact ψ^I_J has a compact support. □

[1] $\frac{d}{dt} f(x_1(t), \ldots, x_m(t)) = \sum_{i=1}^m \frac{\partial f(x_1(t), \ldots, x_m(t))}{\partial x_i}\frac{d}{dt} x_i(t).$

B.3. Physical conventions and notations

In the physical literature there are different notations and conventions for variational calculus. For example, arguments of functionals get square brackets $S[\phi]$. Most notably however, is the notation and concept of the functional derivative. As we have seen, the functional derivative (Fréchet-derivative) of a functional S defined by an integration over a Lagrange-density, can be written as integral operator

$$D_{\phi^I_J} S = \int_U d^n x \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \partial_\mu \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) .$$

The domain of the functional is the set of smooth tensor fields on a compact subset of $\mathbb{R}^n$. Also the Lagrange-density had to satisfy differentiability and dependence conditions. We notice, that the integral-operator defines a distribution over a function, commonly denoted by $\frac{\delta S}{\delta \phi^I_J}$:

$$\boxed{\frac{\delta S}{\delta \phi^I_J} = \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \partial_\mu \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} .}$$

The term $\frac{\delta S}{\delta \phi^I_J}$ is also called **functional derivative**. This is possible, as a function defines a regular distribution by integration. However, we had to restrict the domain of the functional, that can be defined on a larger set. However, if one does so, the term $\frac{\delta S}{\delta \phi^I_J}$ defines no longer the Fréchet-derivative. We have seen so in the proof, where we could ignore boundary terms, which enter in a general setting. The physical jargon of this is: *the field variations vanish at the boundary.*

However, even with relaxed conditions, $\frac{\delta S}{\delta \phi^I_J}$ is called functional derivative in the physical literature. The reasoning behind this is, that although it does no longer define the Fréchet-derivative, it still defines the variation with respect to fields ψ, that have vanishing variations on the boundary:

$$\delta S(\phi, \psi) = \int_U \frac{\delta S}{\delta \phi^I_J} \psi^I_J \, d^n x .$$

The field ϕ is not necessary an extrmum anymore, since only variations in special directions vanish.

Also, the condition of U to be compact can be relaxed, as long as the variation fields ψ have a compact support.

Remark B.3.1.
More carefully one should be with concepts involving the delta function. An example for this would be the definition of functional derivatives:

$$\frac{\delta S}{\delta \phi^I_J(y)} = \lim_{t \to 0} \frac{1}{t} (S[\phi + t\delta(x - y)] - S[\phi]) .$$

As a means of notation, this works for Lagrange-densities. However, the delta-distribution is not regular. Hence, the above definition for functional derivative fails for more general functionals.

C
Riemannian geometry

Riemannian geometry is the most fundamental mathematical field for general relativity. The field equations, that determine the physics of gravity, link the matter/energy to the curvature and metric of the space. These objects (metric and curvature) are the central objects of Riemannian geometry. In the following we cover the basics of Riemannian geometry, including: bundles, connections, Lie-derivatives, geodesics, curvature and Killing fields. Most of this chapter is based on [Lee97], which will be followed very closely. The parts about Lie-derivatives and Killing fields are based on [Mor01] and [Win07, chapter 1].

C.1. Motivation: Geodesic equation

In the physical literature, Riemannian geometry is mostly discussed in coordinates, hiding the bigger picture and mathematical symmetry. The reason for the physical approach is, partly due to simplicity, and partly due to the readiness of equations that can be used to calculate. The mathematical approach is, on the other hand rather abstract. To motivate this chapter, we show, how the Christoffel symbols arise in the physical context (using Ricci notation).

In addition to time dilation, caused by accelerated movement, we expect a gravitational time dilation due to the equivalence principle. As we have seen in section 2.6 the proper time and the length of the world line with respect to $-g$ correspond to each other. One can argue, that inertial movement creates curves of maximal length, equivalent to the observation that time passes slower in accelerated systems. We assume this to be true generally, finding the following variational principle:

$$\delta \int_{\lambda_1}^{\lambda_2} L(x, \dot{x})\ d\lambda = \delta \int_{\lambda_1}^{\lambda_2} \sqrt{-g_{\alpha\beta}(x(\lambda))\dot{x}^\alpha(\lambda)\dot{x}^\beta(\lambda)}\ d\lambda = 0\ .$$

One recognizes, that L is a Lagrange function, where $\dot{x}^\mu = \frac{d}{d\lambda}x^\mu(\lambda)$. To solve the variational problem, we can use the Euler-Lagrange equations.

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0$$

$$\Leftrightarrow \qquad g_{\alpha\mu}(x)\ddot{x}^\alpha + \frac{1}{2}\left(\frac{\partial}{\partial x^\beta}g_{\mu\alpha}(x) + \frac{\partial}{\partial x^\alpha}g_{\beta\mu}(x) - \frac{\partial}{\partial x^\mu}g_{\alpha\beta}(x)\right)\dot{x}^\alpha\dot{x}^\beta - g_{\alpha\mu}(x)\frac{d}{d\lambda}L(x,\dot{x})\dot{x}^\alpha = 0$$

Using the inverse metric $g^{\mu\nu}(x)$ we get the geodesic equation:

$$\ddot{x}^\alpha + \Gamma^\nu_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta - \frac{d}{d\lambda}L(x,\dot{x})\dot{x}^\alpha = 0\ ,$$

where we defined the **Christoffel symbols** (which define a so called connection)

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

C.2. Connections

The geodesic equation, from the last chapter, describes the geodesics of the pseudo Riemannian spacetime manifold. Geodesics play a central role in Riemannian geometry in general. These special curves are defined with the help of the Levi-Civita-connection.

C.2.1. Vector bundles

Connections, in our context, are generally defined on vector bundles, a concept underlying many structures used in physics.

Definition C.2.1.
A smooth k-dimensional real **vector bundle** is the triple (E, M, π) consisting of C^{∞}-manifolds E, called **total space**, M, called **base space** and a surjective map $\pi \colon E \to M$, called **projection**, such that the following properties hold:

i) Every **fiber** $E_p := \pi^{-1}(p)$ is a k-dimensional vector space.

ii) For every $p \in M$ there exists a neighborhood U and a diffeomorphism $\varphi \colon \pi^{-1}(U) \to U \times \mathbb{R}^k$, called **local trivialization**, such that the following diagram is commutative:

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\ \varphi\ } & U \times \mathbb{R}^k \\ {\scriptstyle \pi}\big\downarrow & & \big\downarrow{\scriptstyle \pi_1} \\ U & \xrightarrow[Id_U]{\qquad} & U \end{array}$$

where $\pi_1 \colon U \times \mathbb{R}^k \to U$ is the canonical projection on U.

iii) The restriction of φ to a fiber $\varphi|_{E_p} \colon \pi^{-1}(p) \to \{p\} \times \mathbb{R}^k$ is a vector space isomorphism.

A very well known example for vector bundles is the tangent bundle TM, i.e. the disjoint union of the tangent spaces T_pM:

$$TM := \{(p, \alpha) \mid p \in M,\ \alpha \in T_pM\} \ .$$

In the same way one defines the cotangent bundle T^*M. This construction can be extended to tensor product spaces of tangent and cotangent spaces. Let $\bigotimes^{(r,s)} T_pM$ be the tensor space of rank (r, s)-tensors:

$$\bigotimes^{(r,s)}(T_pM) = \underbrace{T_pM \otimes \ldots \otimes T_pM}_{r-\text{times}} \otimes \underbrace{T_p^*M \otimes \ldots \otimes T_p^*M}_{s-\text{times}} \ .$$

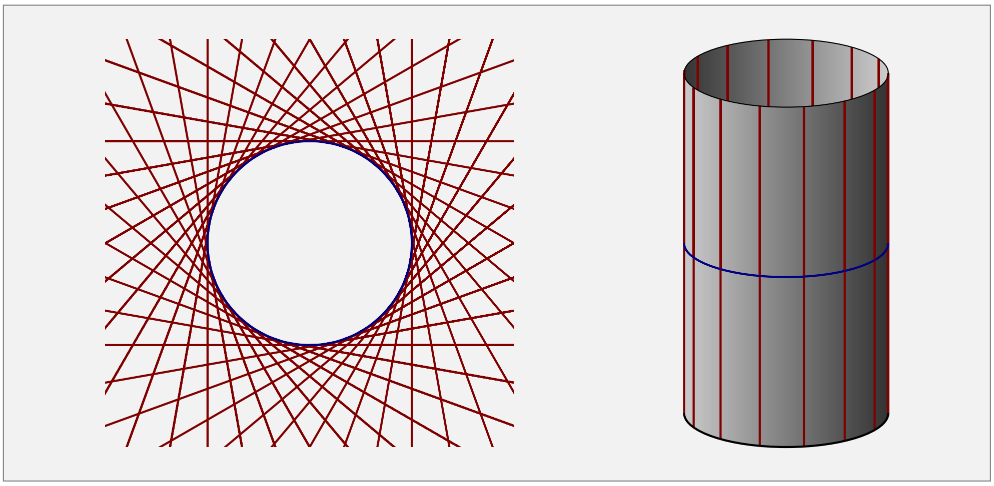


Figure C.1.: Illustration of the tangent bundle. On the left, the tangent lines of S^1. On the right, the corresponding tangent bundle, with base space S^1 defining the total space $[0, 2\pi) \times \mathbb{R}$, which is a smooth manifold.

The (r, s)-tensor bundle is defined by:

$$\bigotimes^{(r,s)}(TM) := \left\{ (p, T) \,\middle|\, p \in M,\ T \in \bigotimes^{(r,s)}(T_pM) \right\} .$$

Definition C.2.2.
Let (E, M, π) be a vector bundle. A **section** is continuous map $s\colon M \to E$ such that $s(p) \in E_p$ $(\Leftrightarrow \pi \circ s = Id_M)$ holds.

Since sections are maps between differentiable manifolds, one can define sections of class C^r as usual. The vector space of r-times continuously differentiable sections will be denoted by $C^r(E, M)$ in the following. The vector space of smooth sections is commonly denoted by $\Gamma(E, M)$ or $\Gamma(E)$.

Example C.2.3.
A smooth (r, s)-tensor field is a section in $\Gamma\left(\bigotimes^{(r,s)}(TM), M\right)$.

After introducing sections, there is an easy way to define Riemannian metrics.

Definition C.2.4.
A **pseudo Riemannian metric** is a smooth symmetric $(0, 2)$-tensor field $g \in \Gamma\left(\bigotimes^{(0,2)}(TM), M\right)$. If the tensor field is positively definite, that is

$$\begin{aligned} &g(v, v)|_p =: g_p(v|_p, v|_p) \geq 0 \\ &g_p(v|_p, v|_p) = 0 \ \Leftrightarrow \ v|_p = 0 \quad \forall\, v \in C^\infty(TM, M)\ \forall p \in M \ , \end{aligned}$$

it is called **Riemannian metric**. A (pseudo) Riemannian manifold (M, g) thus is a manifold M together with a (pseudo) Riemannian metric.

Remark C.2.5.
Let $\{E_1, \dots, E_n\}$ be a local smooth frame and $\{\vartheta_1, \dots, \vartheta_n\}$ the dual frame. The metric g can be written in terms of these frames by

$$g = g_{ij}\vartheta^i \otimes \vartheta^j \ , \qquad \text{with } g_{ij} = g(E_i, E_j) \ .$$

In particular, for a chart we have

$$g = g_{ij} dx^i \otimes dx^j \ , \qquad \text{with } g_{ij} = g(\partial_{x^i}, \partial_{x^j}) \ .$$

Still, the coefficients g_{ij} are functions $M \to \mathbb{R}$. To obtain a coordinate representation of the metric tensor one uses the chart induced pullback:

$$(x^{-1})^* g = \left(g_{ij} \circ x^{-1}\right) (x^{-1})^* dx^i \otimes (x^{-1})^* dx^j \ .$$

The function $(g_{ij} \circ x^{-1})$ is a map from $U \subset \mathbb{R}^n \to \mathbb{R}$ commonly denoted by $g_{ij}(x)$. For the pullback of the coordinate differentials, abusing the notation, one often writes dx^i although e_i^* is meant.

So far,vector bundles are a global construction, that fulfills local conditions. However, the bundle can be reconstructed from local information. To motivate this, we consider two local trivializations $\varphi_{\alpha,\beta} \colon \pi^{-1}(U_{\alpha,\beta}) \to U_{\alpha,\beta} \times \mathbb{R}^k$, where the intersection $U_\alpha \cap U_\beta$ is nonempty. Since $\varphi_{\alpha,\beta}$ have to be diffeomorphisms and thus have to be invertible, the following map a well defined diffeomorphism:

$$\varphi_\alpha \circ \varphi_\beta^{-1} \colon \ (U_\alpha \cap U_\beta) \times \mathbb{R}^k \longrightarrow (U_\alpha \cap U_\beta) \times \mathbb{R}^k \ .$$

The condition on local trivializations (commutativity of diagram) shows, that the above defined map acts as follows:

$$(p, v) \xrightarrow{\varphi_\alpha \circ \varphi_\beta^{-1}} (p, f_{\alpha\beta}(p) v) \ .$$

The restriction to a fiber has to be a vector space isomorphism, hence $f_{\alpha\beta}$, called **transition function**, maps points to invertible linear maps: $f_{\alpha\beta} \colon U_\alpha \cap U_\beta \to \mathrm{GL}(\mathbb{R}^k)$.

Remark C.2.6.
In the fiber bundle context, the Lie-group $\mathrm{GL}(\mathbb{R}^k)$ would be called **structure group** of the bundle. Indeed, different vector bundles can have the transition functions being only in a subgroup of $\mathrm{GL}(\mathbb{R}^k)$, and thus giving rise to special symmetries, e.g. orientation line bundles.

Theorem C.2.7 (vector bundle construction theorem)**.**
Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of a C^∞-manifold M, and $\{f_{\alpha\beta}\}$ be a set of differentiable functions, such that the ***cocycle condition***

$$f_{\alpha\beta}(p) f_{\beta\gamma}(p) = f_{\alpha\gamma}(p) \qquad \forall p \in U_\alpha \cap U_\beta \cap U_\gamma$$

is fulfilled. Then there exists a vector bundle over M, that has $\{f_{\alpha\beta}\}$ as transition *functions.*

We do not prove the theorem here[1], since we will only be concerned about the tangent bundle and the Levi-Civita-connection.[2]

Definition C.2.8.
Let (E, M, π) and (E', M', π') be two vector bundles. A **bundle map** is a pair $(\widetilde{f}, f)$ of continuous maps $\widetilde{f}\colon E \to E'$ and $f\colon M \to M'$ such that the following diagram commutes:

$$\begin{array}{ccc} E & \xrightarrow{\widetilde{f}} & E' \\ {\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\pi'} \\ M & \xrightarrow[f]{} & M' \end{array}$$

The vector bundles are called **isomorphic**, if the have the same base space M and there is a bundle map of the form $(\widetilde{f}, Id_M)$.

C.2.2. Linear Connections

In the first section of this chapter we introduced the geodesic equation. The reasoning is, that one wants to have an equivalent of straight lines in curved spaces. As can be seen easily, it is not as simple as taking a second derivative, as long as the Christoffel symbols do not vanish. Indeed, while the first derivative of curves is well defined and coordinate independent, the second derivative with respect to a parameter is not. A simple example is a circle in $\mathbb{R}^2$ in Cartesian coordinates and in polar coordinates:

$$\begin{aligned} r(t) &= \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} & \qquad r(t) &= \begin{pmatrix} r_0 \\ t \end{pmatrix} , \\ \ddot{r}(t) &= -r(t) \neq 0 & \qquad \ddot{r}(t) &= 0 \ . \end{aligned}$$

The problem arises, for there is no way to construct a coordinate independent difference quotient. This is, because the tangent vectors in the difference quotient are in different vector spaces. Hence we need an additional concept: connections:

Definition C.2.9.
Let (E, M, π) be a vector bundle. A **connection** ∇ is a map

$$\nabla\colon \Gamma(TM) \times \Gamma(E) \longrightarrow \Gamma(E) \ , \quad (v, X) \longmapsto \nabla_v X \ ,$$

satisfying the following properties:

i) ∇ is linear over $C^\infty(M)$ in $\Gamma(TM)$:

$$\nabla_{fv_1 + gv_2} X = f \nabla_{v_1} X + g \nabla_{v_2} X \qquad \forall\ f, g \in C^\infty(M) \ .$$

[1]See [Mor01, poof of proposition 6.2] for a proof in the more general case of fiber bundles.
[2]Still, the theorem deserves to be mentioned, especially because some authors take it as quick definition for vector bundles in order to study topological properties.

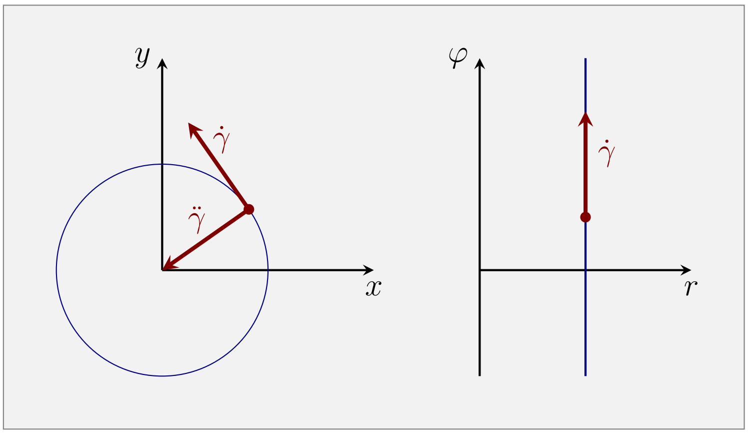


Figure C.2.: Illustration of the difference of second order derivatives in different coordinates (Left: Cartesian coordinates. Right: polar coordinates).

ii) ∇ is liner over $\mathbb{R}$ in $\Gamma(E)$:

$$\nabla_v(aX_1 + bX_2) = a\nabla_v X_1 + b\nabla_v X_2 \ .$$

iii) ∇ satisfies the product rule:

$$\nabla_v(fX) = f\nabla_v X + v(f) \cdot X \qquad \forall\ f \in C^\infty(M)$$

From this definition, a connection seems to be a global object. However, on can show, that the evaluation depends only on local information:

Lemma C.2.10.
The result of $\nabla_v X|_p$ only depends on the behavior of X in an arbitrary small neighborhood U around p and the value of $v|_p$. Thus one can write:

$$\nabla_v X|_p = \nabla_{v(p)} X \ .$$

Proof C.2.11.
Let φ be a test function with support in a neighborhood U around p and $\varphi \leq 1$, $\varphi(p) = 1$. Consider $\widetilde{Y}$ with $\widetilde{X}|_U = X_U$. It is sufficient to show that $\nabla_v(X - \widetilde{X})|_p = 0$ on a small neighborhood around p. We will show that $\nabla_v\varphi \cdot (X - \widetilde{X})|_p$ is zero and find as consequence that $\nabla_v(X - \widetilde{X})|_p$ is zero. We start by noticing that $X - \widetilde{X}$ is zero on the support of φ and thus on the support of $v(\varphi) = \mathcal{L}_v\varphi$. By linearity, the product rule, and the construction of $\widetilde{X}$ we get:

$$\begin{aligned}
\nabla_v\varphi \cdot (X - \widetilde{X})|_p &= \nabla_v\varphi \cdot (X|_U - \widetilde{X}|_U)|_p = 0 \\
&= v(\varphi) \cdot (X - \widetilde{X})|_p + \varphi\nabla_v(X - \widetilde{X})|_p \\
&= \varphi\nabla_v(X - \widetilde{X})|_p \ ,
\end{aligned}$$

which is, that $\varphi\nabla_v(X - \widetilde{X})|_p$ is zero on the support of φ.

The same argument can be used for v:

$$0 = \nabla_{\varphi\cdot(v|_U - \widetilde{v}|_U)} X|_p = \nabla_{\varphi\cdot(v-\widetilde{v})} X|_p = \varphi \cdot \nabla_{v-\widetilde{v}} X|_p \ .$$

This allows us to express v in a basis ∂_i representation $X = v^i\partial_i$, after choosing U small enough such that we find a chart containing U:

$$\nabla_v X|_p = \nabla_{v^i\partial_i} X|_p = (v^i \nabla_{\partial_i} X)|_p = v^i(p)\nabla_{\partial_i} X|_p \ .$$

As before, by linearity, we only need to show, that $\nabla_{v-\widetilde{v}} X|_p$ is zero, if $v|_p = \widetilde{v}|_p$:

$$\nabla_{v-\widetilde{v}} X|_p = (v^i(p) - \widetilde{v}^i(p))\nabla_{\partial_i} X|_p = 0\nabla_{\partial_i} X|_p = 0 \ .$$

□

Morally speaking, the term $\nabla_{v(p)} X$ is like some kind of directional derivative of X in direction of $v(p)$ at the point p. It allows to compare the changing of the vector field X around p with respect to a fixed tangent vector $v(p)$. In that sense it connects the fibers locally. Hence the name connection.

Definition C.2.12.
A **linear connection** is a connection over $\Gamma(TM)$, i.e.

$$\nabla \colon \Gamma(TM) \times \Gamma(TM) \longrightarrow \Gamma(TM) \ .$$

One is tempted to interpret ∇ as (2,1) tensor field in this case. However, a tensor field is linear in both arguments over $C^\infty(M)$, while ∇ satisfies the product rule in the second argument. That is, **a connection is no tensor field**.

Let $\{E_i\}$ be a local frame, then a linear connection defines n^3 functions, called **Christoffel symbols**, by

$$\nabla_{E_i} E_j = \Gamma^k_{ij} E_k \ .$$

These coefficients of the connection do not transform as tensor coefficients would, since linear connections are no tensors, as explained above. Explicit calculations show, that additional terms appear in transformations, related to the product rule.

Lemma C.2.13 (Frame representation of linear connections)**.**
Let ∇ be a linear connection and $\{E_i\}$ a local frame. Let $v = v^i E_i$ and $u = u^i E_i$ be vector field expanded in terms of the local frame, then it holds that:

$$\nabla_v u = \left(v(u^j) + v^i u^j \Gamma^k_{ij} \right) E_k \ .$$

Proof C.2.14.
This is a straightforward calculation using the defining properties of connections.

□

Corollary C.2.15.
Using a coordinate frame one finds:

$$\nabla_v u = \left(v^i \partial_i(u^j) + v^i u^j \Gamma^k_{ij}\right) \partial_k \ .$$

Remark C.2.16 (Notation in physics)**.**
In the physical literature one often encounters terms like $\nabla_\mu v^\nu$, called **covariant derivative**. This is a shorthand notation for $\nabla_{\partial_\mu}(v^\nu \partial_\nu)$. Using the linearity of the connection, we find the term usually used in the physical literature to define $\nabla_\mu v^\nu$:

$$\nabla_\mu v^\nu \equiv \nabla_{\partial_\mu}(v^\nu \partial_\nu) = \partial_\mu(v^\mu) + \Gamma^k_{\mu\nu} v^\nu \partial_k \equiv \partial_\mu v^\nu + \Gamma^k_{\mu\nu} v^\nu \ .$$

It is common to shorten the notation further, using

$$\partial_\mu v^\nu \equiv v^\nu{}_{,\mu} \qquad \text{and} \qquad \nabla_\mu v^\nu \equiv v^\nu{}_{;\mu} \ .$$

Definition and lemma C.2.17 (induced connection on tensors)**.**
A linear connection ∇ induces a connection $\widetilde{\nabla}$ on every tensor bundle $\bigotimes^{(r,s)}(TM)$ by the following conditions:

i) $\widetilde{\nabla}$ and ∇ are equal on $\Gamma(TM)$, i.e. $\nabla_v u = \widetilde{\nabla}_v u$.

ii) For functions we define $\widetilde{\nabla}_v f = v(f)$.

iii) For 1-forms we define $(\widetilde{\nabla}_v \omega)(u) = \widetilde{\nabla}_v \omega(u) - \omega(\widetilde{\nabla}_v u)$.

iv) The connection satisfies the product rule for tensor products: $\widetilde{\nabla}_v(X \otimes Y) = (\widetilde{\nabla}_v X) \otimes Y + X \otimes \widetilde{\nabla}_v Y$.

Proof C.2.18.
To proof the lemma, we need to check, that $\widetilde{\nabla}$ is a proper connection on the respective vector bundle. This is a straightforward calculation, using, that tangent vectors are derivatives. In the last case, one uses the bilinearity of tensor products and demands linearity of $\widetilde{\nabla}_v$ concerning sums of tensors. □

Theorem C.2.19.
The ∇-induced connection $\widetilde{\nabla}$ is unique.

Proof C.2.20.
Let $\widehat{\nabla}$ be another induced connection. We want to show, that $\widetilde{\nabla}_v T = \widehat{\nabla}_v T$ for every tensor T and every vector field v. By definition we have:

$$\widehat{\nabla}_v u = \nabla_v u = \widetilde{\nabla}_v u \qquad \text{and} \qquad \widehat{\nabla}_v f = v(f) = \widetilde{\nabla}_v f \ .$$

From this, we also find $\widehat{\nabla}_v\omega = \widetilde{\nabla}_v\omega$ for any 1-form. The rest follows from linearity and the product rule. □

Due to uniqueness we do not bother to mark the induced connection, and write ∇ in all cases.

Lemma C.2.21.
Let x be a chart and $\{\Gamma^i_{jk}\}$ the corresponding Christoffel symbols, then the covariant derivative of a 1-form is given by:

$$\nabla_v\omega = (v^i\partial_i(\omega_k) - v^i\omega_j\Gamma^j_{ik})dx^k \ .$$

Proof C.2.22.

$$\nabla_v\omega = \nabla_v\omega_k dx^k = v(\omega_k)dx^k + \omega_k\nabla_v dx^k = v^i\partial_i(\omega_k)dx^k + v^i\omega_j\nabla_{\partial_i}dx^j \ .$$

Using the definition of $\nabla_{\partial_i}dx^j(u)$ we want to find an expression for $\nabla_{\partial_i}dx^j$:

$$\begin{aligned}\nabla_{\partial_i}dx^j(u) &= \nabla_{\partial_i}dx^j(u) - dx^j(\nabla_{\partial_i}u^k\partial_k)\\ &= \partial_i(u^j) - dx^j\left(\partial_i(u^k)\cdot\partial_k + u^k\Gamma^\ell_{ik}\partial_\ell\right)\\ &= \partial_i(u^j) - \partial_i(u^k)\delta^j_k - u^k\Gamma^\ell_{ik}\delta^j_\ell = -u^k\Gamma^j_{ik} = -\Gamma^j_{ik}dx^k(u) \ .\end{aligned}$$

Thus we have found $\nabla_{\partial_i}dx^j = -\Gamma^j_{ik}dx^k$ and hence proven the lemma. □

Corollary C.2.23 (Characteristic of Christoffel symbols on basis elements)**.**
The covariant derivative of chart induced basis elements of tangent and cotangent spaces are related to the Christoffel symbols by:

$$\nabla_{\partial i}\partial_j = \Gamma^k_{ij}\partial_k \qquad \textit{and} \qquad \nabla_{\partial_i}dx^j = -\Gamma^j_{ik}dx^k \ .$$

This corollary allows to find the rule for the coordinates of general (r,s)-tensor fields. Before we state how to calculate the covariant derivative of tensor fields in terms of coordinates, we investigate the covariant derivative of tensors as multilinear maps.[3]

Theorem C.2.24.
Let T be a tensor field of rank (r,s), $\omega_1,\dots,\omega_r$ 1-forms and $v_1,\dots,v_s$ vector fields, then the covariant derivative of T along u can be calculated as follows:

$$\begin{aligned}(\nabla_u T)(\omega_1,\dots,\omega_r,v_1,\dots,v_s) = {}& u(T(\omega_1,\dots,\omega_r,v_1,\dots,v_s))\\ &-\sum_{\ell=1}^{r} T(\omega_1,\dots,\nabla_u\omega_\ell,\dots,\omega_r,v_1,\dots,v_s)\\ &-\sum_{\ell=1}^{s} T(\omega_1,\dots,\omega_r,v_1,\dots,\nabla_u v_\ell,\dots,v_s) \ .\end{aligned}$$

[3] Which is the natural definition of tensors.

Proof C.2.25 (sketch of proof).
By definition we have $(\nabla_u\omega)(v) = \nabla_u\omega(v) - \omega(\nabla_u v)$ for 1-forms, that are essentially rank $(0, 1)$ tensors. To find an analogue for vector fields, we use the isomorphy of finite dimensional vector spaces to their double dual spaces:

$$(\nabla_u v)(\omega) \equiv \omega(\nabla_u v) = \nabla_u\omega(v) - (\nabla_u\omega)(v) \equiv \nabla_u v(\omega) - v(\nabla_u\omega) \ .$$

Any tensor can be written as sum of pure tensors by construction. Due to linearity of connections we may assume without loss of generality, that T has the form:

$$T = X_1 \otimes \ldots \otimes X_r \otimes Y_1 \otimes \ldots \otimes Y_s \ ,$$

where X_j are vector fields and Y_j are 1-forms. To keep things short, yet illustrate the general calculation we use $T = X \otimes Y$:

$$\begin{aligned}
(\nabla_u T)(\omega, v) &= (\nabla_v[X \otimes Y])(\omega, v) = ((\nabla_u X) \otimes Y + X \otimes \nabla_u Y)(\omega, v)\\
&= ((\nabla_u X) \otimes Y)(\omega, v) + (X \otimes \nabla_u Y)(\omega, v)\\
&= (\nabla_u X)(\omega) \cdot Y(v) + X(\omega) \cdot (\nabla_u Y)(v)\\
&= (\nabla_u X(\omega)) \cdot Y(v) - X(\nabla_u\omega) \cdot Y(v)\\
&\qquad + X(\omega) \cdot \nabla_u Y(v) - X(\omega) \cdot Y(\nabla_u v)\\
&= \nabla_u(X(\omega) \cdot Y(v)) - X(\nabla_u\omega) \cdot Y(v) - X(\omega) \cdot Y(\nabla_u v)\\
&= u((X \otimes Y)(\omega, v)) - (X \otimes Y)(\nabla_u\omega, v) - (X \otimes Y)(\omega, \nabla_u v)\\
&= u(T(\omega, v)) - T(\nabla_u\omega, v) - T(\omega, \nabla_u v) \ .
\end{aligned}$$

□

One can use corollary C.2.23 to find the coordinate expression for covariant derivatives of tensor fields:

Lemma C.2.26 (Components of covariant derivative).
The components of the covariant derivative in direction ∂_μ of a tensor field are:

$$\begin{aligned}
(\nabla_\mu T)^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} =& \partial_\mu \left(T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}\right)\\
&+ \Gamma^{i_1}_{\eta\mu} T^{\eta \ldots i_r}{}_{j_1 \ldots j_s} + \ldots + \Gamma^{i_r}_{\eta\mu} T^{i_1 \ldots \eta}{}_{j_1 \ldots j_s}\\
&- \Gamma^{\eta}_{j_1\mu} T^{i_1 \ldots i_r}{}_{\eta \ldots j_s} - \ldots - \Gamma^{\eta}_{j_s\mu} T^{i_1 \ldots i_r}{}_{j_1 \ldots \eta} \ .
\end{aligned}$$

Remark C.2.27.
There are further notations for the covariant derivative of tensors:

$$(\nabla_\mu T)^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} = T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s;\mu} = \nabla_\mu T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \ .$$

These notations may cause ambiguity, especially the last one. To avoid misunderstandings, we will write $\nabla_{\partial_\mu} T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}$ if we mean the covariant derivative to act on the components only, preserving ∇_μ for the action on the whole object.

The last thing we need to take care about are contractions and how they behave in the presence of covariant derivatives.

Definition C.2.28.
Let $T = X_1 \otimes \ldots \otimes X_r \otimes \Omega_1 \otimes \ldots \otimes \Omega_s$ be a tensor field of rank (r, s). The **contraction** of the i-th contravariant component with the j-th covariant component is the following $(r-1, s-1)$ tensor field:

$$\mathrm{tr}_{ij}(T) = \Omega_j(X_i) X_1 \otimes \ldots \otimes \widehat{X_i} \otimes \ldots \otimes X_r \otimes \Omega_1 \otimes \ldots \otimes \widehat{\Omega}_j \otimes \ldots \otimes \Omega_s \ ,$$

where the elements with hat are left out. The contraction is defined to be linear.

With $\Omega_j(X_i) = (\Omega_j)^\mu (X_i)_\mu$ and due to linearity, contractions become relatively easy to be written down in Ricci-notation:

$$(\mathrm{tr}_{ij}(T))^{\mu_1\ldots\mu_r}{}_{\nu_1\ldots\nu_s} = T^{\mu_1\ldots\alpha\ldots\mu_r}{}_{\nu_1\ldots\alpha\ldots\nu_s}$$

The same index α in the lower and upper part indicates summation as usual. In the presence of a covariant derivative, using the semicolon notation together with contractions would usually need a convention about order. What does $T^{\mu_1\ldots\alpha\ldots\mu_r}{}_{\nu_1\ldots\alpha\ldots\nu_s;\rho}$ mean? We could define either of the following possibilities:

$$\begin{aligned} & T^{\mu_1\ldots\alpha\ldots\mu_r}{}_{\nu_1\ldots\alpha\ldots\nu_s;\rho} \equiv (T^{\mu_1\ldots\alpha\ldots\mu_r}{}_{\nu_1\ldots\alpha\ldots\nu_s})_{;\rho} \\ \text{or} \quad & T^{\mu_1\ldots\alpha\ldots\mu_r}{}_{\nu_1\ldots\alpha\ldots\nu_s;\rho} \equiv (\nabla_\mu T)^{\mu_1\ldots\alpha\ldots\mu_r}{}_{\nu_1\ldots\alpha\ldots\nu_s} \ . \end{aligned}$$

However, such a choice is not necessary. The next theorem shows, that both definitions are the same. That is, contractions and covariant derivatives commute.

Theorem C.2.29.
The covariant derivative commutes with any contraction:

$$\nabla_v tr_{ij}(T) = tr_{ij}(\nabla_v T)$$

Proof C.2.30.
By linearity we are free to choose $T = X_1 \otimes \ldots \otimes X_r \otimes \Omega_1 \otimes \ldots \otimes \Omega_s$ and denote the reduced tensor by

$$\begin{aligned} \widetilde{T} &= X_1 \otimes \ldots \otimes \widehat{X_i} \otimes \ldots \otimes X_r \otimes \Omega_1 \otimes \ldots \otimes \widehat{\Omega}_j \otimes \ldots \otimes \Omega_s \\ \Rightarrow \qquad & \mathrm{tr}_{ij}(T) = \Omega_j(X_1) \cdot \widetilde{T} \ . \end{aligned}$$

The covariant derivative will create a number of terms in a sum, acting only on one (co)-vector at a time due to the product rule. Hence it will be convenient to prove $\nabla_v \mathrm{tr}_{ij}(T) - \mathrm{tr}_{ij}(\nabla_v T) = 0$, which is equivalent to the claim.

$$\begin{aligned} \nabla_v \mathrm{tr}_{ij}(T) - \mathrm{tr}_{ij}(\nabla_v T) &= \ldots = (\nabla_v \Omega_j(X_i)) \cdot \widetilde{T} - \Big(\Omega_j(\nabla_v X_i) \cdot \widetilde{T} + (\nabla_v \Omega_j)(X_i)\Big) \cdot \widetilde{T} \\ &= (\nabla_v \Omega_j(X_i) - \Omega_j(\nabla_v X_i) - (\nabla_v \Omega_j)(X_i)) \cdot \widetilde{T} = 0 \cdot \widetilde{T} = 0 \ . \end{aligned}$$

The fact, that $\nabla_v \Omega_j(X_i) - \Omega_j(\nabla_v X_i) - (\nabla_v \Omega_j)(X_i) = 0$ holds, follows from definition C.2.17. □

The same statement can be shown for partial derivatives in local coordinates. Thus both index notations (comma and semicolon) behave well with contractions.

Remark C.2.31.
In the following we will use a generalized version of contraction, allowing for contractions of the form $\mathrm{tr}_{12}(X_1 \otimes X_2)$, by using the flat and sharp isomorphisms:

$$\mathrm{tr}_{12}(X_1 \otimes X_2) = X_1(X_2^\flat) = X_2^\flat(X_1) = g(X_1, X_2) \ .$$

Furthermore, in this special case, we may drop the indices of the trace operator, as there is no ambiguity.
In general, this trace and covariant derivatives do not commute. However, if the connection is compatible with the metric (see definition C.3.17), all contractions commute with covariant derivatives. The structure of the proof remains the same. All one has to see is, that

$$\begin{aligned} \nabla_v \mathrm{tr}(X_1 \otimes X_2) &= \nabla_v g(X_1, X_2) = g(\nabla_v X_1, X_2) + g(X_1, \nabla_v X_2) \\ &= \mathrm{tr}(\nabla_v X_1 \otimes X_2) + \mathrm{tr}(X_1 \otimes \nabla_v X_2) \\ &= \mathrm{tr}(\nabla_v X_1 \otimes X_2 + X_1 \otimes \nabla_v X_2) \\ &= \mathrm{tr}(\nabla_v (X_1 \otimes X_2)) \ . \end{aligned}$$

C.3. Levi-Civita-connection

So far we were investigating the concept of general linear connections. These were not related to the metric at all. Indeed, choosing a set of Christoffel symbols allows to define a connection without knowledge about the metric at all. As we will see in this chapter however, there is a unique torsion free connection, related to the metric, called the *Levi-Civita-connection*.

C.3.1. Lie-derivative

To define the torsion of a connection we need a more general concept of directional derivative, called *Lie-Derivative*. There are different ways to introduce this operator. Here we choose the geometric definition using local flows:

Definition C.3.1.
Let $v \in \Gamma(TM)$ be a vector field on M. A curve $\gamma\colon (a, b) \to M$ is called **integral curve** of v, if for all $t \in (a, b)$ the tangent vector of γ coincides with v:

$$T_{\gamma(t)}M \ni \dot{\gamma}(t) = v|_{\gamma(t)} \qquad \forall\, t \in (a, b) \ .$$

The theorem about existence and uniqueness of solutions to first order differential equations assures the existence of maximal integral curves for every initial value problem of a smooth vector field. That is, for all $p \in M$ and $t_0 \in \mathbb{R}$ there exists an integral curve $\gamma_p\colon (a_p, b_p) \to M$, such that

$$\dot{\gamma}_p(t) = v|_{\gamma_p(t)} \qquad \forall\, t \in (a_p, b_p) \qquad \text{and} \qquad \gamma_p(t_0) = p$$

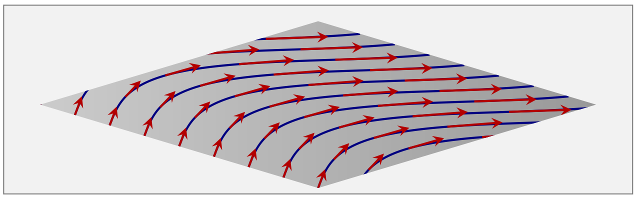


Figure C.3.: Sketch of a vector field (red) with its flow (blue).

hold. These integral curves define a map $\Phi\colon (a_p, b_p) \times M \to M$ by $\Phi_t(p) = \gamma_p(t)$. One can show, that Φ_t is a **local flow**:

$$\Phi_{s+t} = \Phi_s \circ \Phi_t \qquad \text{and} \qquad \Phi_0 = Id_M \ .$$

Definition C.3.2 (Lie-derivative).
Let v be a vector field, ω a k-form and f a function. The Lie-derivative in direction of the vector field u with local flux Φ_t of these objects is defined by:

$$i)\ \mathcal{L}_u f = \left.\frac{d}{dt}\right|_{t=0} \Phi_t^* f \qquad ii)\ \mathcal{L}_u \omega = \left.\frac{d}{dt}\right|_{t=0} \Phi_t^* \omega$$

$$iii)\ \mathcal{L}_u v = \left.\frac{d}{dt}\right|_{t=0} (\Phi_{-t})_* v \ ,$$

where Φ_t^* denotes the pullback and $(\Phi_t)_*$ the push forward of the map Φ_t.

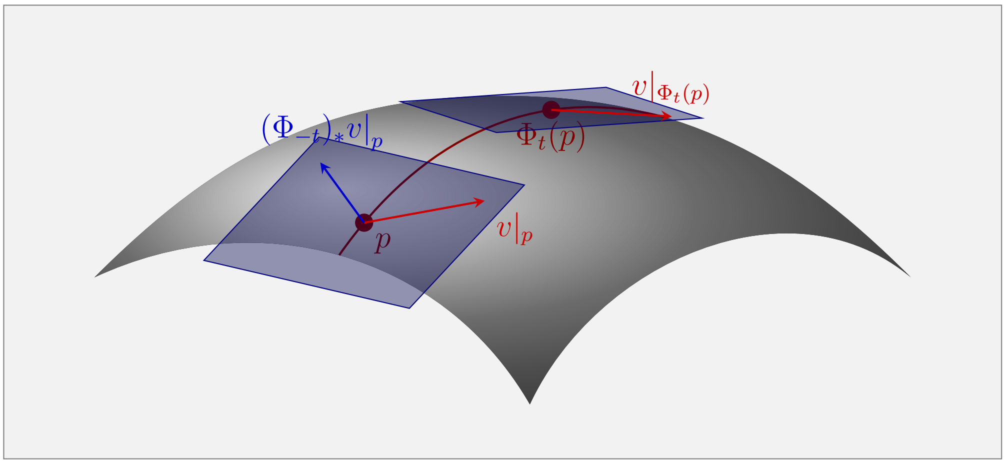


Figure C.4.: Illustration of the transport of tangent vectors by a local flow.

The idea behind Lie-derivatives is, as in the case of connections, to provide an alternative to the directional derivative of $\mathbb{R}^n$. As before, there is no natural way to compare tangent vectors from different tangent spaces. In the case of connections, one choses arbitrary functions[4], the Christoffel symbols, that define a connection. In the case of Lie-derivatives, the existence of local flows is used to compare tangent vectors of different tangent spaces. In figure C.4 this is illustrated before the limit of the derivative

[4]or the ones of the Levi-Civita-connection, if the additional structure of a metric is available.

is taken. The tangent vector $v|_p \in T_pM$ of the vector field $v \in \Gamma(TM)$ can be compared to the tangent vector $v|_{\Phi_t(p)} \in T_{\Phi_t(p)}M$ by comparing it to the push forwarded[5] vector $(\Phi_{-t})_* v|_p \in T_pM$.

It is not hard to show, that a Lie-derivative is no connection. One finds ready examples which show, that $\mathcal{L}_v$ is not linear in v over $C^\infty(M)$, but only over $\mathbb{R}$ or $\mathbb{C}$. By the way we constructed, this is hardly surprising. To have the existence theorem of local flows apply, one needs to provide a vector field in at least a small neighborhood. A tangent vector is not enough, i.e. $\mathcal{L}_v w|_p$ cannot be reduced to $\mathcal{L}_{v(p)} w|_p$. This will be the reason, we will not define geodesics by Lie-derivatives.

Corollary C.3.3.
Let f be a function and u be a vector field, then the Lie-derivative coincides with the application of u as derivation:

$$\mathcal{L}_u f = u(f) \ .$$

Proof C.3.4.

$$(\mathcal{L}_u f)(p) = \left(\left.\frac{d}{dt}\right|_{t=0} \Phi_t^* f \right)(p) = \left.\frac{d}{dt}\right|_{t=0} f(\Phi_t(p)) = \left.\frac{d}{dt}\right|_{t=0} f(\gamma_p(t))$$
$$= df_p(\dot\gamma_p(t)) = df_p(u(p)) = u(f)|_p \ .$$

□

This corollary gives a simple formula to calculate the Lie-derivative of a function in terms of coordinates:

$$\mathcal{L}_u(f) = u(f) = u^\mu \partial_\mu(f) = u^\mu f_{,\mu} \ .$$

One can show that the Lie-derivative of differential forms and vector fields has the following alternative expressions:

Theorem C.3.5 (Commutator and Cartan-formula)**.**
Let u, v be vector fields and ω a k-form. Then it holds that:

$$\mathcal{L}_u v = [u, v] \qquad \text{and} \qquad \mathcal{L}_u \omega = (u\lrcorner \circ d + d \circ u\lrcorner)\omega \ .$$

*The latter is known as **Cartan-formula***

Proof C.3.6 (only for vector fields and functions)**.**
Using the limit definition for time derivatives and the definition of tangent vectors

[5] With the inverse map. Some authors define that as pullback of vector fields.

as derivations, we find:

$$\begin{aligned}(\mathcal{L}_u v)(f) &= \lim_{t\to 0}\frac{(\Phi_{-t})_* v - (\Phi_0)_* v}{t}(f) = \lim_{t\to 0}\frac{(\Phi_{-t})_* v - v}{t}(f) \\ &= \lim_{t\to 0}\frac{1}{t}\left(\{(\Phi_{-t})_* v\}(f) - v(f)\right) = \lim_{t\to 0}\frac{1}{t}(v(f\circ\Phi_{-t}) - v(f)) \\ &= \lim_{t\to 0}\frac{1}{t}(v(f\circ\Phi_{-t}) - v(f) + \{\Phi_t^* v(f\circ\Phi_{-t})\} - \{\Phi_t^* v(f\circ\Phi_{-t})\}) \\ &= \lim_{t\to 0}\Phi_t^*\left\{v\left(\frac{f\circ\Phi_{-t} - f}{t}\right)\right\} + \lim_{t\to 0}\frac{\Phi_{-t}^*\{v(f)\} - v(f)}{t} \\ &= \lim_{t\to 0} v\left(\frac{f\circ\Phi_{-t} - f}{t}\right) + \lim_{t\to 0}\frac{\Phi_{-t}^*\{v(f)\} - v(f)}{t} \\ &= v(\mathcal{L}_{-u} f) + \mathcal{L}_u(v(f)) = u(v(f)) - v(u(f)) = [u,v](f)\end{aligned}$$

In the last steps we used corollary C.3.3 and the fact, that $\Phi_0 = Id_M$.

The proof for k-forms is rather involved, thus we only show the case for $k = 0$, i.e. functions here:

$$u \lrcorner df + d(u \lrcorner f) = u \lrcorner df = df(u) = u(f) = \mathcal{L}_u(f)\ .$$

□

These alternative ways of calculating the Lie-derivative allow to find the coordinate expressions for vector fields and 1-forms, rather easily:

Corollary C.3.7.
In coordinates the Lie-derivative in direction u of a vector field v and a 1-form ω are:

$$\mathcal{L}_u v = \left(u^\mu \partial_\mu(v^\nu) - v^\mu \partial_\mu(u^\nu)\right)\partial_\nu\ ,$$

$$\mathcal{L}_u \omega = \left(u^\mu \partial_\mu(\omega_\nu) + \omega_\mu \partial_\nu(u^\mu)\right) dx^\nu\ .$$

Proof C.3.8.
For the definition of tangent vector fields to be point wise derivations, we demand f to be $C^\infty(M)$. Thus partial derivatives commute.

$$\begin{aligned}(\mathcal{L}_u v)(f) &= [u,v](f) = u(v(f)) - v(u(f)) = u^\mu\partial_\mu(v^\nu\partial_\nu(f)) - v^\nu\partial_\nu(u^\mu\partial_\mu(f)) \\ &= u^\mu\partial_\mu(v^\nu)\partial_\nu(f) - u^\mu v^\nu\partial_\mu\partial_\nu(f) - v^\nu\partial_\nu(u^\mu)\partial_\mu(f) + u^\mu v^\nu\partial_\nu\partial_\mu(f) \\ &= u^\mu\partial_\mu(v^\nu)\partial_\nu(f) - v^\nu\partial_\nu(u^\mu)\partial_\mu(f) \\ &= \left(u^\mu\partial_\mu(v^\nu)\partial_\nu - v^\nu\partial_\nu(u^\mu)\partial_\mu\right)(f) = \left(u^\mu\partial_\mu(v^\nu)\partial_\nu - v^\mu\partial_\mu(u^\nu)\partial_\nu\right)(f)\ .\end{aligned}$$

For the 1-form we use, that $u\lrcorner$ is a linear antiderivation:

$$\begin{aligned}\mathcal{L}_u\omega &= u\lrcorner d\omega + d(u\lrcorner\omega) = u^\mu\partial_\mu\lrcorner d(\omega_\nu dx^\nu) + d(\omega_\nu u^\mu\ \partial_\nu\lrcorner dx^\nu) \\ &= u^\mu\partial_\eta(\omega_\nu)\cdot\partial_\mu\lrcorner(dx^\eta\wedge dx^\nu) + d(\omega_\nu u^\nu) \\ &= u^\mu\partial_\eta(\omega_\nu)(\delta_\mu^\eta dx^\nu - \delta_\mu^\nu dx^\eta) + (\partial_\mu(\omega_\nu)u^\nu + \omega_\nu\partial_\mu(u^\nu))dx^\mu\end{aligned}$$

$$= u^\mu \partial_\mu(\omega_\nu)dx^\nu - u^\mu \partial_\eta(\omega_\mu)dx^\eta + \partial_\mu(\omega_\nu)u^\nu dx^\mu + \omega_\nu \partial_\mu(u^\nu)dx^\mu$$
$$= u^\mu \partial_\mu(\omega_\nu)dx^\nu + \omega_\nu \partial_\mu(u^\nu)dx^\mu = u^\mu \partial_\mu(\omega_\nu)dx^\nu + \omega_\mu \partial_\nu(u^\mu)dx^\nu \ .$$

□

Remark C.3.9.
A result of the last corollary is, that the components of Lie-derivatives only utilize partial derivatives of functions. With definition C.2.17 we may rewrite these in terms of covariant derivatives, i.e.

$$\mathcal{L}_u v = \left(u^\mu \nabla_{\partial_\mu}(v^\nu) - v^\mu \nabla_{\partial_\mu}(u^\nu)\right)\partial_\nu \ ,$$

$$\mathcal{L}_u \omega = \left(u^\mu \nabla_{\partial_\mu}(\omega_\nu) + \omega_\mu \nabla_{\partial_\nu}(u^\mu)\right)dx^\nu \ ,$$

which can come handy in some calculations. (Here we have used our convention from remark C.2.27).

Lemma C.3.10.
The Lie-derivative can be written with covariant derivatives, if the corresponding Christoffel symbols are symmetric in the lower indices for all coordinate-frames:

$$\mathcal{L}_X Y = \nabla_X Y - \nabla_Y X \ .$$

Proof C.3.11.

$$\mathcal{L}_X Y = [X,Y] = X^i \partial_i(Y^j)\partial_j - Y^j \partial_j(X^i)\partial_j = X^i \nabla_{\partial_i}(Y^j)\partial_j - Y^j \nabla_{\partial_j}(X^i)\partial_j$$
$$= \nabla_X(Y) - X^i Y^j \Gamma^k_{ij}\partial_k - \nabla_Y(X) + X^i Y^j \Gamma^k_{ji}\partial_k = \nabla_X Y - \nabla_Y X \ .$$

We now know how to compute the Lie-derivative for vector fields and 1-forms (in fact k-forms). If we demand[6] for the Lie-derivative to satisfy the product rule for tensor products:

$$\mathcal{L}_u(S \otimes T) := (\mathcal{L}_u S) \otimes T + S \otimes \mathcal{L}_u T \ ,$$

we are able to differentiate, in the Lie-sense, arbitrary Tensors.

Example C.3.12.
Let u be a vector field and $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$ a metric tensor. The Lie-derivative in direction u is:

$$\mathcal{L}_u g = \mathcal{L}_u(g_{\mu\nu}dx^\mu) \otimes dx^\nu + g_{\mu\nu}dx^\mu \otimes \mathcal{L}_u dx^\nu$$
$$= (u^\eta \partial_\eta(g_{\mu\nu})dx^\mu + g_{\eta\nu}\partial_\mu(u^\eta)dx^\mu) \otimes dx^\nu + g_{\mu\nu}dx^\mu \otimes \partial_\eta(u^\nu)dx^\eta$$
$$= (u^\eta \partial_\eta(g_{\mu\nu}) + g_{\eta\nu}\partial_\mu(u^\eta))\, dx^\mu \otimes dx^\nu + g_{\mu\nu}\partial_\eta(u^\nu)dx^\mu \otimes dx^\eta$$

[6] Either that, or one defines pullbacks of tensor fields. In that case the product rule appears naturally.

$$= (u^\eta \partial_\eta (g_{\mu\nu}) + g_{\eta\nu}\partial_\mu(u^\eta))\, dx^\mu \otimes dx^\nu + g_{\mu\eta}\partial_\nu(u^\eta) dx^\eta \otimes dx^\nu$$

In the last step we relabeled the summation indices (which is always possible) to find the index rule of the physical literature:

$$\mathcal{L}_u g_{\mu\nu} = u^\eta g_{\mu\nu,\eta} + g_{\eta\nu} u^\eta{}_{,\mu} + g_{\mu\eta} u^\eta{}_{,\nu}$$

From corollary C.3.7 and the behavior of the product rule together with the multilinearity of tensor products one can derive the coordinate expression for Lie-derivatives of arbitrary tensors:

$$\begin{aligned}(\mathcal{L}_u T)^{i_1\ldots i_r}{}_{j_1\ldots j_s} = {}& u^\eta T^{i_1\ldots i_r}{}_{j_1\ldots j_s,\eta} \\ & - T^{\eta\ldots i_r}{}_{j_1\ldots j_s} u^{i_1}{}_{,\eta} - \ldots - T^{i_1\ldots \eta}{}_{j_1\ldots j_s} u^{i_r}{}_{,\eta} \\ & + T^{i_1\ldots i_r}{}_{\eta\ldots j_s} u^\eta{}_{,j_1} + \ldots + T^{i_1\ldots i_r}{}_{j_1\ldots \eta} u^\eta{}_{,j_s} \ .\end{aligned}$$

To conclude this section we consider Tensors again as multilinear maps and show the behavior of Lie-derivatives acting on these maps. We also show again, that Lie-derivatives, like covariant derivatives, commute with contractions.

Theorem C.3.13.
Let T be a tensor field of rank (r, s), $\omega_1, \ldots, \omega_r$ 1-forms and $v_1, \ldots, v_s$ vector fields, then the Lie-derivative of T along u can be calculated as follows:

$$\begin{aligned}(\mathcal{L}_u T)(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s) = {}& \mathcal{L}_u(T(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s)) \\ & - \sum_{\ell=1}^{r} T(\omega_1, \ldots, \mathcal{L}_u\omega_\ell, \ldots, \omega_r, v_1, \ldots, v_s) \\ & - \sum_{\ell=1}^{s} T(\omega_1, \ldots, \omega_r, v_1, \ldots, \mathcal{L}_u v_\ell, \ldots, v_s) \ .\end{aligned}$$

Proof C.3.14.
We can copy the proof of theorem C.2.24, if we can proof a similar behavior for the Lie-derivative acting on 1-forms and vectors. The remaining properties that were used in the proof are linearity and the product rule, which apply to Lie-derivatives as well. There is a coordinate free way to show the following, involving techniques also used to prove Cartan-formula. We can however use arbitrary coordinates, as long as we do not use special properties of a chosen system:

$$\mathcal{L}_u\, \omega(v) = u(\omega(v)) = u^\alpha \partial_\alpha(\omega_\eta v^\eta) = u^\alpha(\omega_{\eta,\alpha} v^\eta + \omega_\eta v^\eta{}_{,\alpha}) \ .$$

On the other hand we have, using corollary C.3.7:

$$\begin{aligned}(\mathcal{L}_u\omega)(v) + \omega(\mathcal{L}_u v) &= \omega_{\eta,\alpha} u^\alpha v^\eta + \omega_\alpha u^\alpha{}_{,\eta} v^\eta + \omega_\eta u^\alpha v^\eta{}_{,\alpha} - \omega_\eta v^\alpha u^\eta{}_{,\alpha} \\ &= \omega_{\eta,\alpha} u^\alpha v^\eta + \omega_\eta u^\alpha v^\eta{}_{,\alpha} = u^\alpha(\omega_{\eta,\alpha} v^\eta + \omega_\eta v^\eta{}_{,\alpha}) = \mathcal{L}_u\, \omega(v) \ .\end{aligned}$$

Subtraction of $\omega(\mathcal{L}_u v)$ gives the desired result:

$$(\mathcal{L}_u\omega)(v) = \mathcal{L}_u\,\omega(v) - \omega(\mathcal{L}_u v)\ .$$

At this point we can copy the proof of theorem C.2.24. □

Theorem C.3.15.
The Lie-derivative commutes with any contraction:

$$\mathcal{L}_v\,tr_{ij}(T) = tr_{ij}(\mathcal{L}_v T)$$

Proof C.3.16.
The proof of theorem C.2.29 carries over without changes. □

C.3.2. Torsion and metric compatibility

The goal of this chapter was to connect the connection with the metric. We took a rather long detour, exploring the details of the Lie-derivative. However, that will not have been a waste of time, for Lie-derivatives play a central role in the context of symmetries and Killing fields.

Definition C.3.17.
Let g be a Riemannian metric and ∇ a linear connection. The connection is said to be **metric/compatible with g** if for all $u, v, w \in \Gamma(TM)$ the product rule is fulfilled:

$$\nabla_u g(v,w) = g(\nabla_u v, w) + g(v, \nabla_u w)\ .$$

Lemma C.3.18.
Compatibility of a connection ∇ with a metric g is equivalent to $\nabla_\bullet g \equiv 0$, where $\nabla_\bullet g$ is a map defined in the natural way:

$$\begin{aligned} \nabla_\bullet g \colon \Gamma(TM)\times\Gamma(TM)\times\Gamma(TM) &\longrightarrow \Gamma(TM) \\ (u,v,w) &\longmapsto (\nabla_u g)(v,w)\ . \end{aligned}$$

Proof C.3.19.
The lemma follows from theorem C.2.24:

$$(\nabla_u g)(v,w) = \nabla_u g(v,w) - g(\nabla_u v, w) - g(v, \nabla_u w)$$

$$\begin{aligned} \nabla_u g \equiv 0 \quad &\Leftrightarrow \quad 0 = \nabla_u g(v,w) - g(\nabla_u v, w) - g(v, \nabla_u w) \\ &\Leftrightarrow \quad \nabla_u g(v,w) = g(\nabla_u v, w) + g(v, \nabla_u w)\ . \end{aligned}$$

□

Remark C.3.20.
The condition $\nabla g = 0$ can be written in coordinates as $\nabla_\eta g_{\mu\nu} = 0$.

Compatibility withe the metric relates the connection the the metric, yet is not enough to determine the connection uniquely. Additionally requiring the connection to be torsion free does however, allowing to calculate the Christoffel symbols in terms of the metric.

Definition C.3.21.
The **Torsion tensor field** T of a linear connection is a rank $(2,1)$ tensor field $T\colon \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ defined by

$$T(v,w) = \nabla_v w - \nabla_w v - [v,w] \ .$$

A linear connection is called **symmetric**, if the torsion vanishes identically, i.e.

$$\nabla_v w - \nabla_w v = [v,w] \qquad \Leftrightarrow \qquad T(v,w) \equiv 0 \ .$$

Lemma C.3.22.
A connection is symmetric, if and only if the Christoffel symbols for coordinate frames are symmetric in the lower indices, i.e. $\Gamma^k_{ij} = \Gamma^k_{ji}$.[7]

Proof C.3.23.
Let $\{\partial_\ell\}$ be a coordinate frame. By linearity we have the following equivalence from the definition of torsion:

$$\nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i = [\partial_i, \partial_j] \qquad \Leftrightarrow \qquad T \equiv 0 \ .$$

With corollary C.3.7 we see, that Lie-derivatives of the form $[\partial_i, \partial_j]$ vanish. Also using the definition of the Christoffel symbols we find:

$$T \equiv 0 \qquad \Leftrightarrow \qquad 0 = \nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i = (\Gamma^k_{ij} - \Gamma^k_{ji})\partial_k \qquad \Leftrightarrow \qquad \Gamma^k_{ij} = \Gamma^k_{ji} \ .$$

□

The torsion is an interesting property in itself, describing the drill of tangent vectors along a geodesic. Since torsion plays a minor role in non quantum gravity, we will not investigate this object any further, and only use torsion to determine a unique connection that is linked to the metric.

[7] This is not necessarily true for other local frames.

Theorem C.3.24 (fundamental lemma of Riemannian geometry).
Let (M, g) be a (pseudo) Riemannian manifold. There exists a unique linear connection on M, that is symmetric and compatible with the metric, called the ***Levi-Civita-connection****.*

Proof C.3.25 (identical in [Lee97]).

1) We show uniqueness first, since it will give rise to a simpler construction method to prove existence. Let u, v, w be arbitrary vector fields. Writing down the metric compatibility condition for cyclic permutations of u, v, w and using the symmetry of the connection yields:

$$\begin{aligned}
\nabla_u g(v,w) &= g(\nabla_u v, w) + g(v, \nabla_u w) = g(\nabla_u v, w) + g(v, \nabla_w u) + g(v, [u,w]) \\
\nabla_v g(w,u) &= g(\nabla_v w, u) + g(w, \nabla_v u) = g(\nabla_v w, u) + g(w, \nabla_u v) + g(w, [v,u]) \\
\nabla_w g(u,v) &= g(\nabla_w u, v) + g(u, \nabla_w v) = g(\nabla_w u, v) + g(u, \nabla_v w) + g(u, [w,v]) \ .
\end{aligned}$$

Adding the first two equations and subtracting the third one, we get after solving for $g(\nabla_u v, w)$ the following equation:

$$\begin{aligned}
g(\nabla_u v, w) &= \frac{1}{2}\Big(\nabla_u g(v,w) + \nabla_v g(w,u) - \nabla_w g(u,v) \\
&\qquad - g(v,[u,w]) - g(w,[v,u]) - g(u,[w,v])\Big) \\
&= \frac{1}{2}\Big(u(g(v,w)) + v(g(w,u)) - w(g(u,v)) \\
&\qquad - g(v,[u,w]) - g(w,[v,u]) - g(u,[w,v])\Big) \ .
\end{aligned} \tag{C.1}$$

We see, that the right-hand side does not depend on the connection. Thus, if there is another symmetric connection $\widetilde{\nabla}$, that is compatible with the metric, we find:

$$g(\nabla_u v - \widetilde{\nabla}_u v, w) = 0 \qquad \forall\ u, v, w \ ,$$

which is equivalent to $\nabla_u v - \widetilde{\nabla}_u v = 0$ and respectively $\nabla_u v = \widetilde{\nabla}_u v$. This proves uniqueness.

2) If there is a symmetric connection that is compatible with the metric, it is unique. If we can construct such a connection locally in coordinates, uniqueness guarantees agreement on overlaps. Let (U, x) be a chart of M. We want to define ∇ using (C.1). With corollary C.3.7 we see, that Lie-derivatives of the form $[\partial_i, \partial_j]$ vanish, and find:

$$g(\nabla_{\partial_i}\partial_j, \partial_k) = \frac{1}{2}\Big(\partial_i(g(\partial_j,\partial_k)) + \partial_j(g(\partial_k,\partial_i)) - \partial_k(g(\partial_i,\partial_j))\Big) \ .$$

Using the Christoffel symbols and the notational convention $\partial_i(g(\partial_j,\partial_k)) = g_{jk,i}$ we find a formula for the Christoffel symbols:

$$\Gamma^{\ell}_{ij} g_{\ell k} = \frac{1}{2}(g_{jk,i} + g_{ki,j} - g_{ij,k}) \ .$$

With $g_{\ell k} g^{km} = \delta_\ell^m$ we recover the formula most prevalent in the literature:

$$\Gamma_{ij}^m = \frac{1}{2} g^{mk} (g_{jk,i} + g_{ki,j} - g_{ij,k}) \ .$$

The Christoffel symbols define the connection and vice versa. We can see, that the Christoffel symbols defined above are symmetric in the lower indices. By lemma C.3.22 the corresponding connection is symmetric. A short calculation shows, that the connection is also compatible with the metric, i.e. $\nabla g = 0$, by lemma C.3.18. □

Corollary C.3.26.
The Christoffel symbols of the Levi-Civita-connection can be calculated as follows:

$$\Gamma_{ij}^k = \frac{1}{2} g^{km} (g_{jm,i} + g_{mi,j} - g_{ij,m}) \ .$$

Theorem C.3.27 (Naturality of the Levi-Civita-connection)**.**
Let $\varphi \colon M \to \widetilde{M}$ be an ***isometry*** *between two Riemannian manifolds (M, g) and $(\widetilde{M}, \widetilde{g})$, i.e. a diffeomorphism such that $\varphi^* \widetilde{g} = g$. Also let ∇ be the Levi-Civita-connection of M and $\widetilde{\nabla}$ the Levi-Civita-connection of $\widetilde{M}$, then it holds that*

$$\varphi_*(\nabla_X Y) = \widetilde{\nabla}_{\varphi_* X} (\varphi_* Y) \ .$$

Proof C.3.28 (Idea of the proof)**.**
Define the pullback connection $\varphi^* \widetilde{\nabla}$ on M by

$$(\varphi^* \widetilde{\nabla})_X Y = \varphi_*^{-1} (\widetilde{\nabla}_{\varphi_* X} (\varphi_* Y)) \ .$$

It remains to show, that $\varphi^* \widetilde{\nabla}$ is a connection on M that is compatible with g and torsion free. If so, $\varphi^* \widetilde{\nabla} = \nabla$ due to the uniqueness of the Levi-Civita-connection, and the claim follows:

$$\varphi_*((\varphi^* \widetilde{\nabla})_X Y) = \varphi_*(\nabla_X Y) = (\widetilde{\nabla}_{\varphi_* X} (\varphi_* Y)) \ .$$

C.4. Geodesics

In Euklidean spaces there exists a special set of curves, the straight lines. These give rise to affine structures. More physically speaking, they define what is called acceleration free movement. To define acceleration free movement on general (pseudo) Riemannian manifolds a similar concept is needed. The motivation to introduce covariant derivatives is also the reason why we cannot extend the classical definition directly. The second derivative of curves depends on the choice of coordinates.[8] To define a class of curves similar to straight lines, called geodesics, we will need a generalization of second derivatives. In the following, if not specified otherwise, a curve means a smooth map $\gamma\colon I \to M$, where I is an interval of $\mathbb{R}$.

C.4.1. Vector fields along curves

Let $\gamma\colon I \to M$ be a curve. A **vector field along curve** is a map $V\colon I \to TM$ with $V(t) \in T_{\gamma(t)}M$ for every $t \in I$. We will denote the set of vector fields along the curve γ in the style of sections with $\Gamma(\gamma)$. The vector field V is called **extendible** if there are a neighborhood around $\gamma(I)$ and a smooth vector field $\widetilde{V}$ in that neighborhood, such that $V(t) = \widetilde{V}|_{\gamma(t)}$ holds for all $t \in I$. The vector field $\widetilde{V}$ is called an **extension** of V. To be extendible is no trivial property. For example the tangent field $\dot{\gamma}(t)$ is not extendible if there are self intersections, as can be seen in figure C.5. On the other hand, it can be shown, that the restriction of a smooth vector field to a curve is a smooth vector field along the curve.

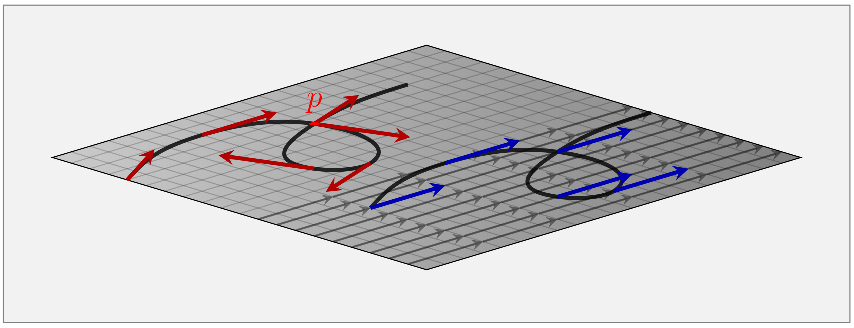


Figure C.5.: Illustration of an extendible vector field (blue) and a non-extendible vector field (red). The problem of the red vector field is, that any extension would have to assign two different tangent vectors to the point p.

Definition and Lemma C.4.1.
Let ∇ be a linear connection. For each curve there is a unique operator $D_t^\nabla\colon \Gamma(\gamma) \to \Gamma(\gamma)$ induced by ∇, that satisfies the following properties:

i) Linearity over $\mathbb{R}$:

$$D_t^\nabla(aV + bW) = aD_t^\nabla(V) + bD_t^\nabla(W)\ .$$

ii) Product rule:

$$D_t^\nabla(fV) = \dot{f}V + fD_t^\nabla(V) \quad \forall\, f \in C^\infty(I)\ .$$

iii) If V is extendible, then for any extension $\widetilde{V}$ it holds that $D_t^\nabla(V(t)) = \nabla_{\dot{\gamma}(t)}\widetilde{V}$.

[8]See figure C.2 for an example.

Proof C.4.2.
It is best to show uniqueness first. So suppose D_t is such an operator. Similar to proof C.2.11 one can show, that $D_t V|_{t_0}$ only depends on the values of V in an arbitrary small interval $(t_0 - \varphi, t_0 + \varepsilon)$. We may choose coordinates in a small neighborhood around $\gamma(t_0)$ and write[9]

$$V(t) = V^j(t)\partial_j \ .$$

Since ∂_j are extendible, and by the assumed properties of D_t we have:

$$\begin{aligned} D_t V|_{t_0} &= \dot{V}^j(t_0)\partial_j + V^j(t_0)\nabla_{\dot{\gamma}(t_0)}\partial_j \\ &= \left(\dot{V}^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0))\right)\partial_k \ . \end{aligned}$$

The right hand side does only depend on ∇, proving the uniqueness $D_t = D_t^\nabla$. Existence may now be proven by defining D_t^∇ by the above equation and showing that the properties are satisfied. Due to uniqueness this definition agrees on any overlap of charts proving well-definiedness. □

Corollary C.4.3.
*The operator D_t^∇, called **covariant curve derivative** can be calculated in terms of coordinates as follows:*

$$D_t^\nabla V|_{t_0} = \left(\dot{V}^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0))\right)\partial_k \ .$$

In the spacial case of $V(t) = \dot{\gamma}(t)$ one finds the definition of the physical literature:

$$D_t^\nabla \dot{\gamma}|_{t_0} = \left(\ddot{\gamma}^k(t_0) + \dot{\gamma}^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0))\right)\partial_k \ .$$

Remark C.4.4.
In the case of the Levi-Civita-connection the Christoffel symbols of affine coordinates vanish. Thus the corollary proves the claim of remark 2.5.1.

In the following, if the connection is understood, we will not bother to write D_t^∇ any longer and simply use D_t. In the literature there are differing notations, including $\frac{\nabla}{dt}$, $\frac{D}{Dt}$, $\frac{d}{Dt}$ or $\frac{D}{dt}$.

C.4.2. Geodesics and parallel transport

Straight lines in $\mathbb{R}^n$ are those curves, whose second derivative $\frac{d^2}{dt^2}\gamma(t)$ vanishes. Physically speaking, the velocity $\dot{\gamma}(t)$ does not change along the curve. The way we constructed the connection allows to generalize such a statement. Loosely speaking $\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t)$ would mean, that around $\gamma(t)$ the vector field $\dot{\gamma}(t)$ does not change. This is the property, that

[9] This does not require extendibility yet, since $V(t_0)$ is uniquely extendible even on a self intersection, if the interval around t_0 is chosen small enough.

generalizations of straight lines should satisfy. However, $\dot{\gamma}(t)$ is only defined on $\gamma(I)$, being no proper smooth vector field on M. For that reason we defined D_t, leading to the following definition:

Definition C.4.5.
Let ∇ be a linear connection. A **geodesic** corresponding to ∇ is a curve γ such that $D_t\dot{\gamma}(t)$ vanishes identically:

$$D_t\dot{\gamma}(t) = 0 \quad \forall t \in I \ .$$

With corollary C.4.3 the **geodesic equation** $D_t\dot{\gamma}(t) = 0$ is exactly the one from the motivation at the beginning of this chapter, yet in a more general context of linear connections:

$$\ddot{\gamma}^k(t) + \dot{\gamma}^j(t)\dot{\gamma}^i(t)\Gamma^k_{ij}(\gamma(t)) = 0$$

By the definition it is still questionable, if geodesics exist at all. With the theorem about existence and uniqueness of solutions to first order differential equations, using the coordinate form of the geodesic equation, one can prove the following lemma:

Lemma C.4.6 (Existence of geodesics)**.**
For every pair $(p, v) \in TM$ there exists a unique geodesic $\gamma\colon I \to M$ with maximal interval, such that $\gamma(0) = p$ and $\dot{\gamma}(0) = v$ hold.

It is convenient to refer to such a geodesic with $\gamma_{(p,v)}$. Reparametrization allows to choose t_0 instead of 0 if needed.

We defined geodesics such that $\dot{\gamma}$ does not change locally. We realized that be demanding $D_t\dot{\gamma}(t) = 0$. In Euklidean space, a vector field that does not change, i.e. a constant vector field, is said to consist of parallel vectors. This motivates the next definition:

Definition C.4.7.
A vector field V along a curve γ is called **parallel along $\boldsymbol{\gamma}$**, if $D_tV(t) = 0$ holds for all t.

Again, writing the condition $D_tV(t) = 0$ in coordinates, using the theorem about differential equations and inspecting the intersection of chart domains results in the following lemma:

Lemma C.4.8.
*Let γ be a curve and V_0 be a tangent vector of $T_{\gamma(t_0)}M$. There exists a unique parallel vector field V along γ, called the **parallel translate** of V_0, such that $V(t_0) = V_0$.*

Using this lemma, we can define an operator that parallel translates vectors along a curve, called **parallel transport**:

$$P^\gamma_{t_0t_1} : T_{\gamma(t_0)}M \longrightarrow T_{\gamma(t_1)}M, \quad v = V(t_0) \longmapsto P^\gamma_{t_0t_1}v = V(t_1) \ .$$

One can show, that the parallel transport is a linear isomorphism.

Lemma C.4.9.
The following statements are equivalent:

i) The connection ∇ is compatible with the metric g.

ii) For all vector fields V, W along a curve γ the following equation holds:

$$\frac{d}{dt}g(V(t), W(t)) = g(D_tV(t), W(t)) + g(V(t), D_tW(t)) \ .$$

Proof C.4.10.
Choose a chart around $\gamma(t)$. Then, as in proof C.4.2 we can write the vector fields in the tangent basis:

$$\begin{aligned}\frac{d}{dt}g(V(t), W(t)) = \frac{d}{dt}V^i(t)W^j(t)g_{\gamma(t)}(\partial_i, \partial_j) &= \frac{d}{dt}V^i(t)W^j(t)g_{ij}|_{(\gamma(t))} \\ &= g(\dot{V}^i(t)\partial_i, W^j(t)\partial_j) + g(V^i(t)\partial_i, \dot{W}^j(t)\partial_j) \\ &\quad + V^i(t)W^j(t) \cdot dg_{ij}(\dot{\gamma}(t))|_{\gamma(t)} \\ &= g(\dot{V}^i(t)\partial_i, W^j(t)\partial_j) + g(V^i(t)\partial_i, \dot{W}^j(t)\partial_j) \\ &\quad + V^i(t)W^j(t) \cdot \dot{\gamma}^k(t)\partial_k(g_{ij})(\gamma(t)) \ .\end{aligned}$$

The right hand side of the equation from claim ii) can be expanded as well:

$$\begin{aligned}&g_{\gamma(t)}(D_tV(t), W(t)) + g_{\gamma(t)}(V(t), D_tW(t)) \\ &= g(\dot{V}^i(t)\partial_i, W^j(t)\partial_j) + g(V^i(t)\partial_i, \dot{W}^j(t)\partial_j) \\ &\quad + g(\dot{V}^j(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))\partial_k, W^j(t)\partial_j) \\ &\quad + g(V^i(t)\partial_i, \dot{W}^j(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))\partial_k) \ .\end{aligned}$$

Comparing both equations with the claim shows:

$$\frac{d}{dt}g(V(t), W(t)) = g(D_tV(t), W(t)) + g(V(t), D_tW(t)) \ ,$$

$$\begin{aligned}\Leftrightarrow \quad V^i(t)W^j(t)\dot{\gamma}^k(t)\partial_k(g_{ij})(\gamma(t)) &= g(\dot{V}^j(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))\partial_k, W^j(t)\partial_j) \\ &\quad + g(V^i(t)\partial_i, \dot{W}^j(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))\partial_k) \ ,\end{aligned}$$

$$\Leftrightarrow \quad \partial_k g(\partial_i, \partial_j) = \nabla_k g(\partial_i, \partial_j) = g(\Gamma^k_{ij}\partial_k, \partial_j) + g(\partial_i, \Gamma^k_{ij}\partial_k) \ .$$

The last equation is the condition for ∇ to be compatible with the metric, written in coordinates. □

Corollary C.4.11.
If the connection is compatible with the metric, then the following properties follow immediately:

i) Let V, W be parallel along the curve, then $g(V, W)$ is constant along the curve.

ii) The parallel transport is an isometry.

(On can even show, that these claims are equivalent to each other and those of the last lemma.)

Geodesics with respect to the Levi-Civita-connection are called **Riemannian geodesics**. As long as we will be concerned with the Levi-Civita-connection, we will simply use the term geodesic.

C.4.3. Riemannian normal coordinates

Although geodesics exist for all $(p, v) \in TM$, the maximal interval for $\gamma_{(p,v)}$ need not be large enough to define $\gamma_{(p,v)}(1)$. We define the set of pairs $(p, v) \in TM$, that allow $\gamma_{(p,v)}(1)$:

$$\mathcal{E} := \left\{ (p, v) \in TM \,\middle|\, \gamma_{(p,v)} \text{ is defined on an interval containing } [0, 1] \right\} .$$

Definition C.4.12.
The **exponential map** $\exp\colon \mathcal{E} \to M$ is defined by $\exp(p, v) \equiv \exp_p(v) \equiv \gamma_{(p,v)}(1)$.

In the following we give some properties of the exponential map without proof (see [Lee97] for the proofs):

Proposition C.4.13.

i) The set $\mathcal{E}$ is an open subset of TM containing $(0, 0)$. The restriction $\mathcal{E}_p$ is star shaped.

ii) The exponential map is smooth.

iii) The geodesic $\gamma_{(p,v)}$ is given by $\gamma_{(p,v)}(t) = \exp_p(tv)$ for all t such that both sides are defined.

*iv) **Rescaling property:** For all $(p, v) \in TM$ and $c, t \in \mathbb{R}$ it holds that*

$$\gamma_{(p,cv)}(t) = \gamma_{(p,v)}(ct) \ ,$$

whenever one side is defined.

Lemma C.4.14.
For all $p \in M$, there are neighborhoods U around the origin of T_pM and V around p, such that $\exp_p\colon U \to V$ is a diffeomorphism.

Proof C.4.15.
By the theorem of inverse functions we only need to show, that $(\exp_p)_*$ is invertible

at $0 \in T_pM$. Identifying $T_0(T_pM) \cong T_pM$ we calculate $(\exp_p)_* v$:

$$(\exp_p)_* v = \frac{d}{dt}\bigg|_{t=0} \exp_p(tv) = \gamma_{(p,v)}(t) = v \quad \Rightarrow \quad (\exp_p)_* = Id_{T_pM} \ .$$

□

If U is star shaped, then V is called **normal neighborhood**. There is a technical proof, that for every $p \in M$ there is a normal neighborhood.[10] The last lemma allows to find a diffeomorphism $V \to \mathbb{R}^n$, that is a chart and hence a coordinate system, whenever V is a normal neighborhood.

Let V be a normal neighborhood around $p \in M$ and $\{E_i\}$ be an orthonormal basis of T_pM. The basis defines an isomprhism

$$E\colon \mathbb{R}^n \to T_pM \quad \text{by} \quad \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \longmapsto x^i E_i \ .$$

Since $\exp_p$ is an isomorphism, there exists $\exp_p^{-1}\colon V \to U \subset T_pM$. Thus $\varphi := E^{-1} \circ \exp_p^{-1}\colon V \to \mathbb{R}^n$ is a diffeomorphism, and by definition (V, φ) is a chart of M.

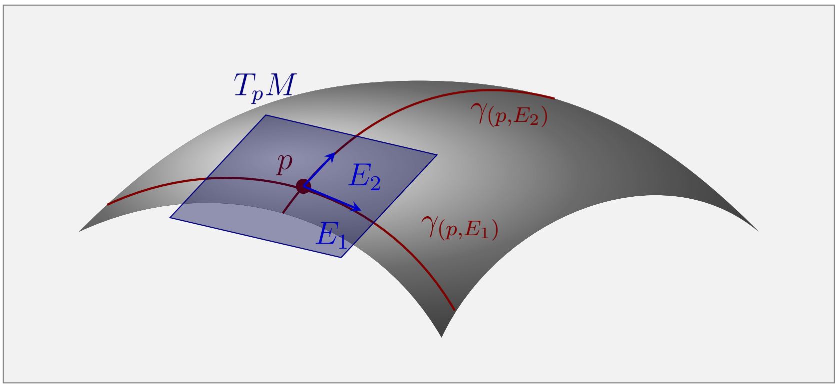


Figure C.6.: The Idea behind Riemannian normal coordinates is as follows: Choose an orthonormal basis (here (E_1, E_2)) of the tangent space T_pM. To any vector exists a geodesic, especially to E_1 and E_2. To a point q near p on M we assign the special tangent vector V of T_pM which defines a geodesic $\gamma_{(p,V)}$ such that $\gamma_{(p,V)}(0) = p$ and $\gamma_{(p,V)}(1) = q$ hold. The coordinates of q are then the coordinates of V with respect to the initially chosen orthonormal basis.

Definition C.4.16.
Any coordinate system (V, φ) defined as above is called **Riemannian normal coordinate system**.

Theorem C.4.17 (Properties of Riemannian normal coordinates)**.**

i) Let $v = v^i \partial_i$ be an element of T_pM. Any geodesic $\gamma_{(p,v)}$ passing through p

[10] Even more, the proof shows, that there is a uniformly normal neighborhood. A concept we will not use here.

has the following particular easy form in Riemannian normal coordinates:

$$\gamma_{(p,v)}(t) = (tv^1, \ldots, tv^n).$$

ii) The coordinates of p are $(0, \ldots, 0)$.

iii) The first partial derivatives of g_{ij} and the Christoffel symbols of the Levi-Civita-connection vanish at p.

Proof C.4.18.
By construction, the normal coordinates φ^i are given by $\varphi^i = (E^{-1} \circ \exp_p^{-1})^i = \vartheta^i \circ E^{-1} \circ \exp_p^{-1}$. From proposition C.4.13 we know, that any geodesic $\gamma_{(p,v)}(t)$ can be written as $\gamma_{(p,v)}(t) = \exp_p(tv)$. Thus we have:

$$\varphi^i(\gamma_{(p,t)}(t)) = \left(\vartheta^i \circ E^{-1} \circ \exp_p^{-1} \circ \exp_p\right)(tv) = \vartheta^i(E^{-1}(tv)) = t\vartheta^i(v^j e_j) = tv^i \ ,$$

proving the first claim. Since $\gamma_{(p,v)}(0) = p$, the second claim follows immediately from i).

Although vanishing Christoffel symbols follow directly from vanishing partial derivatives of the components of the metric, it is easier to prove the claim about the Christoffel symbols first. Using the coordinate expression of D_t, and normal coordinates for the geodesic $\gamma_{(p,v)}(t)$ we find:

$$\begin{aligned} 0 = D_t\gamma_{(p,v)}(t) &= \frac{d^2}{dt^2}\varphi^k(\gamma_{(p,v)}(t)) + \Gamma^k_{ij}(\gamma_{(p,v)}(t))\varphi^i(\dot\gamma_{(p,v)}(t))\varphi^j(\dot\gamma_{(p,v)}(t)) \\ &= \frac{d^2}{dt^2}tv^k + \Gamma^k_{ij}(\gamma_{(p,v)}(t))v^i v^j = \Gamma^k_{ij}(\gamma_{(p,v)}(t))v^i v^j \ . \end{aligned}$$

For all $v \in T_pM$ all geodesics $\gamma_{(p,v)}(t)$ satisfy $\gamma_{(p,v)}(0) = p$.[11]Then we get (by evaluating $D_t\gamma_{(p,t)}(0)$):

$$\Gamma^k_{ij}(p)v^i v^j = 0 \qquad \Rightarrow \qquad \Gamma^k_{ij}(p) + \Gamma^k_{ji}(p) = 0 \ .$$

Since Christoffel symbols of a torsion free connection (e.g. Levi-Civita) are symmetric in the lower indices, we found $\Gamma^k_{ij}(p) = 0$. Using metric compatibility in coordinate form (see end of proof C.4.10) we find the first part of the third claim:

$$\begin{aligned} \partial_k(g_{ij}(p)) = \partial_k g(\partial_i, \partial_j)|_p &= g(\Gamma^k_{ij}\partial_k, \partial_j)|_p + g(\partial_i, \Gamma^k_{ij}\partial_k)|_p \\ &= \Gamma^k_{ij}(p)g_{kj}(p) + \Gamma^k_{ij}(p)g_{ik}(p) = 0 \ . \end{aligned}$$

□

[11] It is important to evaluate at the point $t = 0$. Only there, the Christoffel symbols do not depend on v, i.e. are constant with respect to v.

C.5. Curvature

The meaning of intrinsic curvature differs from the intuition we have about curved surfaces of curves. The goal of this section is to define a curvature, that can be measured without relying on an embedding. An example for the difference between intrinsic and extrinsic curvature can be found using a sphere and a cylinder. It is possible to draw a triangle of geodesics with a total angle sum larger than π on a sphere but not on a cylinder. A more intuitive way to see the difference is the following: Without stretching overlapping and creasing, it is possible ro obtain a cylinder from a piece of paper. Yet that is not possible with a sphere. The reason is, that the cylinder has no intrinsic, but only extrinsic curvature.

C.5.1. Riemannian curvature

A way to define a computational tool to determine intrinsic curvature is to observe the following behavior of covariant derivatives with respect to the Levi-Civita-connection for Euklidean spaces (e.g. $\mathbb{R}^n$ with standard metric):

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = (X(Y(Z^k)) - Y(X(Z^k)))\partial_k = \nabla_{[X,Y]} Z \ .$$

This condition, called **flatness condition**, does not only hold for Euklidean spaces, which we understand as the prime example of flat spaces, but for all manifolds that are locally isometric to $\mathbb{R}^n$ due to the naturality of the Levi-Civita-connection. That motivates the definition of the Riemannian curvature endomorphism:

Definition C.5.1.
The **Riemannian curvature endomorphism** is the map $R\colon \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$, defined by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z = [\nabla_X, \nabla_Y] Z - \nabla_{[X,Y]} Z \ .$$

Loosely speaking, the curvature endomorphism measures how much the flatness condition fails, and hence defining an **intrinsic curvature**. The commutator of partial derivatives vanishes, leading to a simple basis expression for R:

$$R(\partial_i, \partial_j)\partial_k = [\nabla_{\partial_i}, \nabla_{\partial_j}]\partial_k \ .$$

Showing that R is multilinear over $C^\infty(M)$ shows, that the curvature endomorphism is a tensor, called the **Riemann (curvature) tensor**:

$$R = R_{ijk}{}^{\ell} dx^i \otimes dx^j \otimes dx^k \otimes \partial_\ell \quad \text{with} \quad R(\partial_i, \partial_j)\partial_k = R_{ijk}{}^{\ell} \partial_\ell \ .$$

Corollary C.5.2.
The coefficients of the Riemann tensor in terms of the Christoffel symbols are:

$$R_{ijk}{}^{\ell} = \partial_i \Gamma^\ell_{jk} - \partial_j \Gamma^\ell_{ik} + \Gamma^\eta_{jk}\Gamma^\ell_{i\eta} - \Gamma^\eta_{ik}\Gamma^\ell_{j\eta} \ .$$

Proof C.5.3.

$$\begin{aligned} R_{ijk}{}^{\ell}\partial_\ell &= R(\partial_i,\partial_j)\partial_k = [\nabla_{\partial_i},\nabla_{\partial_j}]\partial_k = \nabla_{\partial_i}\nabla_{\partial_j}\partial_k - \nabla_{\partial_j}\nabla_{\partial_i}\partial_k \\ &= \nabla_{\partial_i}\Gamma^{\ell}_{jk}\partial_\ell - \nabla_{\partial_j}\Gamma^{\ell}_{ik}\partial_\ell \\ &= \left(\partial_i\Gamma^{\ell}_{jk}\right)\partial_\ell + \Gamma^{\ell}_{jk}\Gamma^{\eta}_{i\ell}\partial_\eta - \left(\partial_j\Gamma^{\ell}_{ik}\right)\partial_\ell - \Gamma^{\ell}_{ik}\Gamma^{\eta}_{j\ell}\partial_\eta \\ &= \left(\left(\partial_i\Gamma^{\ell}_{jk}\right) + \Gamma^{\eta}_{jk}\Gamma^{\ell}_{i\eta} - \left(\partial_j\Gamma^{\ell}_{ik}\right) - \Gamma^{\eta}_{ik}\Gamma^{\ell}_{j\eta}\right)\partial_\ell \ . \end{aligned}$$

In the last line we have relabeled dummy indices. □

A Riemannian manifold is called **flat** if for every point $p \in M$ there are an open neighborhood U around p and an isometry into an open subset of $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, $\phi\colon U \to \phi(U) \subset \mathbb{R}$. A lengthy proof (see [Lee97]) shows, that the manifold is flat, if and only if $R \equiv 0$.

Remark C.5.4.
There are different sign conventions for the curvature tensor in the literature, e.g. :

$$R(\partial_i,\partial_j)\partial_k = -R_{ijk}{}^{\ell}\partial_\ell \ .$$

With the symmetries stated in the next subsection, one finds that

$$-R_{ijk}{}^{\ell} = -R^{\ell}{}_{kji} = R^{\ell}{}_{kij} \ .$$

C.5.2. Symmetries of the Riemann tensor

To find the symmetries of the Riemann tensor it is convenient to define an isomorphic tensor field, that is completely covariant, i.e. has only lower indices.

Definition C.5.5.
The **covariant Riemann tensor** $R^\flat$ is defined by lowering the last index of R.

For the coordinates we get

$$R_{ijkm} = g_{\ell m}R_{ijk}{}^{\ell} \ .$$

In the coordinate free language $R^\flat$ is defined by

$$R^\flat(X,Y,Z,W) = g(R(X,Y)Z,W) \ .$$

Theorem C.5.6.
Let X,Y,Z,W be vector fields. The covariant Riemann tensor has the following symmetries:

i) $R^\flat(X,Y,Z,W) = -R^\flat(Y,X,Z,W)$

ii) $R^\flat(X,Y,Z,W) = -R^\flat(X,Y,W,Z)$

iii) $R^\flat(X,Y,Z,W) = R^\flat(Z,W,X,Y)$

iv) $R^\flat(X,Y,Z,W) + R^\flat(Y,Z,X,W) + R^\flat(Z,X,Y,W) = 0$

*The last symmetry is called **algebraic Bianchi identity**.*

The proof, that is more or less a series of simple steps in the right order, can be found in [Lee97].

Corollary C.5.7.
For the coefficients of $R^\flat$ the symmetries have the following form:

i) $R_{ijk\ell} = -R_{jik\ell}$

ii) $R_{ijk\ell} = -R_{ij\ell k}$

iii) $R_{ijk\ell} = R_{k\ell ij}$

iv) $R_{ijk\ell} + R_{jki\ell} + R_{kij\ell} = 0$

Theorem C.5.8 (Differential Bianchi identity)**.**
*The covariant Riemann tensor satisfies the **differential Bianchi identity**:*

$$(\nabla_W R^\flat)(X,Y,Z,V) + (\nabla_Z R^\flat)(X,Y,V,W) + (\nabla_V R^\flat)(X,Y,W,Z) = 0 \ .$$

In coordinates this reads

$$R_{ijk\ell;m} + R_{ij\ell m;k} + R_{ijmk;\ell} = 0 \ .$$

Proof C.5.9.
We include the proof, not because it is particularly helpful in understanding the relation or where it comes from, but because it shows a standard technique of Riemannian geometry.

To prove the claimed identity for a single (but not specified) point p, it is enough to prove the identity for basis vectors (i.e. the coordinate version), because of the multilinearity in all five vector fields. The main idea to simplify the calculations is, that we are free to choose the coordinates (also because of multilinearity). Here, a good choice would be coordinates, such that $\nabla_{\partial_i}\partial_j|_p = \Gamma^k_{ij}(p)\partial_k$ vanish. Indeed such coordinates exist for all p, for example Riemannian normal coordinates do (see theorem C.4.17). Also $[\partial_i, \partial_j] = 0$ for all coordinates.

We begin by observing, with the help of theorem C.2.24, that the symmetries of the covariant Riemann tensor may be used in the presence of a covariant derivative. Thus the differential Bianchi identity is equivalent to

$$(\nabla_W R^\flat)(Z,V,X,Y) + (\nabla_Z R^\flat)(V,W,X,Y) + (\nabla_V R^\flat)(W,Z,X,Y) = 0 \ .$$

As we have argued, it is enough to show

$$(\nabla_m R^\flat)(\partial_k,\partial_\ell,\partial_i,\partial_j)|_p + (\nabla_k R^\flat)(\partial_\ell,\partial_m,\partial_i,\partial_j)|_p + (\nabla_\ell R^\flat)(\partial_m,\partial_k,\partial_i,\partial_j)|_p = 0 \ .$$

With the metric compatibility of ∇, theorem C.2.24 and our special choice, such that $\Gamma^k_{ij}(p) = 0$, we get:

$$\begin{aligned}(\nabla_m R^\flat)(\partial_k,\partial_\ell,\partial_i,\partial_j)|_p &= \nabla_m R^\flat(\partial_k,\partial_\ell,\partial_i,\partial_j)|_p = \nabla_m g(R(\partial_k,\partial_\ell)\partial_i,\partial_j)|_p \\ &= g(\nabla_m\nabla_k\nabla_\ell\partial_i - \nabla_m\nabla_\ell\nabla_k\partial_i, \partial_j)|_p \ .\end{aligned}$$

Although $\nabla_m \partial_j$ has vanished, this does not happen to $\nabla_m \nabla_k \nabla_\ell \partial_i$ in general, similarly to zeros of first derivatives not necessarily being zeros of higher derivatives.

$$\begin{aligned}(\nabla_m R^\flat)(\partial_k, \partial_\ell, \partial_i, \partial_j)|_p + (\nabla_k R^\flat)(\partial_\ell, \partial_m, \partial_i, \partial_j)|_p + (\nabla_\ell R^\flat)(\partial_m, \partial_k, \partial_i, \partial_j)|_p \\ = g(\nabla_m \nabla_k \nabla_\ell \partial_i - \nabla_m \nabla_\ell \nabla_k \partial_i + \nabla_k \nabla_\ell \nabla_m \partial_i - \nabla_k \nabla_m \nabla_\ell \partial_i \\ + \nabla_\ell \nabla_m \nabla_k \partial_i - \nabla_\ell \nabla_k \nabla_m \partial_i, \partial_j)|_p \\ = g(R(\partial_m, \partial_k)\nabla_\ell \partial_i + R(\partial_k, \partial_\ell)\nabla_m \partial_i + R(\partial_\ell, \partial_m)\nabla_k \partial_i, \partial_j)|_p \ . = 0\end{aligned}$$

In the last line we used, that all covariant derivatives are of first order of and thus vanish at p. □

C.5.3. Ricci and scalar curvature

Tensors with 4 indices are rather complicated objects one wants to avoid when possible. A way to do so, is to encode some of the information of R in tensors of lower rank, which leads to the following definition:

Definition C.5.10.
The **Ricci (curvature) tensor** $\mathcal{R}$ and the **scalar curvature** $\mathcal{S}$ are defined as contractions of the Riemann tensor:

$$\mathcal{R} = \mathrm{tr}_{14}(R) \qquad \text{and} \qquad \mathcal{S} = \mathrm{tr}_{12}(\mathcal{R}) \ .$$

In coordinates this reads

$$\mathcal{R}_{ij} = R_{kij}{}^k \qquad \text{and} \qquad \mathcal{S} = \mathcal{R}_i{}^i = g^{ij}\mathcal{R}_{ij} \ .$$

The scalar curvature depends on the Ricci tensor, we defined to be a contraction of the Riemann tensor. One wants to have a well defined meaning of positive or negative scalar curvature, yet there are different sign conventions for the Riemann tensor. This leads to a different definition of the Ricci tensor, if one chooses a different sign convention (changing index position for trace or an additional minus sign).

Corollary C.5.11.
The Ricci tensor is symmetric, i.e. $\mathcal{R}_{ij} = \mathcal{R}_{ji}$.

Proof C.5.12.

$$\mathcal{R}_{ij} = R_{kij}{}^k = -R^k{}_{jki} = R^k{}_{jik} = R_{kji}{}^k = \mathcal{R}_{ji} \ .$$

□

Lemma C.5.13 (Contracted Bicanchi identity)**.**
The covariant derivative of the scalar curvature is twice the covariant divergence

of the Ricci tensor:

$$\mathcal{R}_{ij;}{}^{j} = \frac{1}{2} S_{;i} \ .$$

Proof C.5.14.
We start from the differential Bianchi identity:

$$R_{ijk\ell;m} + R_{ij\ell m;k} + R_{ijmk;\ell} = 0 \ .$$

Contracting with $g^{i\ell}$ (since $g^{i\ell}{}_{;\bullet}$ is zero) and using the symmetries results in:

$$0 = R_{ijk}{}^{i}{}_{;m} - R_{ijm}{}^{i}{}_{;k} + R_{ijmk;}{}^{i} = \mathcal{R}_{jk;m} - \mathcal{R}_{jm;k} + R_{ijmk;}{}^{i} \ .$$

Again, contracting with g^{jk} gives

$$0 = \mathcal{R}_{j}{}^{j}{}_{;m} - \mathcal{R}_{jm;}{}^{j} - R_{jim}{}^{j}{}_{;}{}^{i} = \mathcal{S}_{;m} - \mathcal{R}_{jm;}{}^{j} - \mathcal{R}_{im;}{}^{i} \ ,$$

$$\Leftrightarrow \qquad \frac{1}{2}\mathcal{S}_{;m} = \mathcal{R}_{jm;}{}^{j} = \mathcal{R}_{mj;}{}^{j} \ .$$

□

Since we have emphasized the coordinate free approach, we want to state the last lemma without relying on coordinates. To do so, we need to define the divergence of a symmetric covariant tensor of rank 2:

$$\mathrm{div}\mathcal{R} := \mathrm{tr}_{01}(\nabla\mathcal{R}) \ .$$

The zeroth index corresponds to the vector of the covariant derivative. Since $\mathcal{R}$ is symmetric, we could have chosen tr_{02}, as well. So there is no ambiguity. The contracted Bianchi identity reads in the coordinate free version:

$$\mathrm{div}\mathcal{R} = \frac{1}{2}\nabla\mathcal{S} \ .$$

C.6. Killing fields

In classical physics and quantum mechanics, conserved quantities correspond to symmetries, by Noether's theorem. In general relativity, conserved quantities can be constructed from special vector fields, called *Killing fields*, that can be understood as symmetries of the metric.

Definition C.6.1.
Let (M, g) be a (pseudo) Riemannian manifold. A vector field $X \in \Gamma(TM)$ is called **Killing field**, if $\mathcal{L}_X g = 0$ on M.

The condition to be a Killing field is a global condition and it is not clear, if non-trivial Killing fields exists at all. In fact, one can construct Riemannian manifolds that do not allow for non-trivial Killing fields.

Lemma C.6.2.
*A vector field $X \in \Gamma(TM)$ is a Killing field, if and only if it satisfies the **Killing equation**:*

$$g(\nabla_v X, w) + g(v, \nabla_w X) = 0 \ ,$$

for all vector fields v, w, where ∇ is the Levi-Civita connection.

Proof C.6.3.
We begin with the Killing condition:

$$0 = (\mathcal{L}_X g)(v, w) = \mathcal{L}_X g(v, w) - g(\mathcal{L}_X v, w) - g(v, \mathcal{L}_X w)$$

Since $\mathcal{L}_X g(v, w) = X(g(v, w)) = \nabla_X g(v, w)$ it follows, that

$$\nabla_X g(v, w) = g(\mathcal{L}_X v, w) + g(v, \mathcal{L}_X w) \ .$$

On the other hand:

$$\nabla_X g(v, w) = -(\nabla_X g)(v, w) + g(\nabla_X v, w) + g(v, \nabla_X w) = g(\nabla_X v, w) + g(v, \nabla_X w) \ .$$

In the last step we used the compatibility with the metric $\nabla_X g \equiv 0$ of the Levi-Civita-connection. Combining the equations yields:

$$g(\nabla_X v, w) + g(v, \nabla_X w) = g(\mathcal{L}_X v, w) + g(v, \mathcal{L}_X w)$$

The Lie-derivative can be written with covariant derivatives (see lemma C.3.10), hence

$$\begin{aligned} g(\nabla_X v, w) + g(v, \nabla_X w) &= g(\mathcal{L}_X v, w) + g(v, \mathcal{L}_X w) \\ &= g(\nabla_X v - \nabla_v X, w) + g(v, \nabla_X w - \nabla_w X) \\ &= g(\nabla_X v, w) + g(v, \nabla_X w) - g(\nabla_v X, w) - g(v, \nabla_w X) \ , \end{aligned}$$

$$\Leftrightarrow \qquad 0 = g(\nabla_v X, w) + g(v, \nabla_w X) \ .$$

□

Corollary C.6.4.
In coordinates, the Killing equations reads $X_{\mu;\nu} + X_{\nu;\mu} = 0$.

Proof C.6.5.
Using the flat isomorphism, and noticing, that v and w are arbitrary, we can write:

$$\begin{aligned} 0 &= g(\nabla_v X, w) + g(v, \nabla_w X) = (\nabla_v X)^\flat(w) + (\nabla_w X)^\flat(v) \\ &= v^\mu(\nabla_{\partial_\mu} X)^\flat(w) + w^\nu(\nabla_{\partial_\nu} X)^\flat(v) \end{aligned}$$

$$
\begin{aligned}
&= v^\mu (\nabla_{\partial_\mu} X)_\nu dx^\nu(w) + w^\nu (\nabla_{\partial_\nu} X)_\mu dx^\mu(v) \\
&= v^\mu w^\nu (\nabla_{\partial_\mu} X)_\nu + v^\mu w^\nu (\nabla_{\partial_\nu} X)_\mu \\
&= v^\mu w^\nu ((\nabla_{\partial_\mu} X)_\nu + (\nabla_{\partial_\nu} X)_\mu) \equiv v^\mu w^\nu (X_{\nu;\mu} + X_{\mu;\nu}) \ ,
\end{aligned}
$$

$$\Leftrightarrow \quad 0 = X_{\mu;\nu} + X_{\nu;\mu} \ .$$

□

The importance of Killing fields for general relativity comes from the following theorem:

Theorem C.6.6.
Let γ be a geodesic and ∇ a metric connection. If X is a Killing field, then $g(\dot\gamma, X) = $ const. along the curve.

Proof C.6.7.
To show that $g(\dot\gamma, X)$ is constant along the curve it is enough to show that $\nabla_{\dot\gamma} g(\dot\gamma, X) = 0$. In the following, we use that $\nabla_{\dot\gamma}\dot\gamma = 0$ for geodesics and $g(\nabla_v X, v) = -g(v, \nabla_v X) = 0$ for Killing fields, because of the Killing equation. Using the metric compatibility:

$$\nabla_{\dot\gamma} g(\dot\gamma, X) = g(\nabla_{\dot\gamma}\dot\gamma, X) + g(\dot\gamma, \nabla_{\dot\gamma} X) = 0 \ .$$

□

D
Calculations in coordinates

Although it is the author's opinion, that physical theories should be formulated coordinate free, since physics does not depend on man-made coordinates, comparing experiment with theory will end in coordinate dependent calculations. Also, historical, and partly also contemporary literature prefer the coordinate approach. Hence, in this chapter, some useful calculational tools are presented.

D.1. Determinant of the metric tensor

In many coordinate expressions in Riemannian geometry and general relativity one encounters the determinant of the metric tensor. In the following we show some useful calculations and mention mathematical subtleties.

D.1.1. Definition: metric determinant

Let $A_p \in \mathrm{End}(T_pM)$ be an endomorphism for all $p \in U \subset M$ and ω be an n-form on M, where $\dim(M) = n$. The determinant of A_p in p is defined by

$$\omega(A_p v_1, \ldots, A_p v_n) = \det(A_p)\omega(v_1, \ldots, v_n) \qquad \forall v_1, \ldots, v_n \in T_pM \ .$$

With this definition one can see easily, that the determinant does not depend on the choice of coordinates. Hence $\det(A_p)\colon U \to \mathbb{R}$ defines a function on U. Choosing coordinates, such that A_p can be expressed as matrix ${A^\mu}_\nu(p)$, it can be shown that

$$\det(A_p) = \sum_{\sigma \in \Sigma_n} \left(\mathrm{sgn}(\sigma) \prod_{\mu=1}^{n} {A^\mu}_{\sigma(\mu)}(p) \right) \ .$$

Definition D.1.1.
Let (M, g) be a (pseudo) Riemannian manifold. For a chart (x, U) define the **metric determinant** $\det(g_{\mu\nu})$ by

$$\det(g_{\mu\nu}) = \sum_{\sigma \in \Sigma_n} \left(\mathrm{sgn}(\sigma) \prod_{\mu=1}^{n} g_{\mu\sigma(\mu)} \right) \ .$$

The formula for the metric determinant seems to be innocent enough. In contrast to the determinant of endomorphisms, the metric determinant has no coordinate free definition. In fact, the **metric determinant depends on the choice of coordinates**. The reason is, that endomorphisms, for finite-dimensional vector spaces, are isomorphic to (1,1)-tensors. The metric tensor is a (0,2) tensor.[1]. Hence, the metric determinant is

[1] One might attempt to use the flat or sharp isomorphisms to fix that. However, these isomorphisms use the metric tensor, changing the determinant

defined with respect to a chart. Also notice that the metric determinant defines no function on the manifold!

D.1.2. Derivatives of the metric determinant

There is a very useful formula for derivatives involving determinants, called **Jacobi's formula**:

Theorem D.1.2 (Jacobi's formula)**.**
Let $A(t)$ be an invertible matrix. Then the parameter derivative of the determinant is given as follows:

$$\frac{d}{dt}\det(A(t)) = \det(A(t))\mathrm{tr}\left(A^{-1}(t)\frac{d}{dt}A(t)\right) \ .$$

We skip the proof, which can readily be found, and focus on its implications for the metric determinant.

Lemma D.1.3.
The derivative the determinant with respect to one coefficient is:

$$\frac{d}{dA_{\mu\nu}}\det(A) = \det(A)(A^{-1})_{\nu\mu} \ .$$

Proof D.1.4.
Using Jacobi's formula it remains to show that

$$(A^{-1})_{\nu\mu} = \mathrm{tr}\left(A^{-1}\frac{d}{dA_{\mu\nu}}A\right) \ .$$

The matrix $B = \frac{d}{dA_{\mu\nu}}A$ is a matrix with $B_{\eta\rho} = 0$ for all $\eta\rho \neq \mu\nu$. The coefficient $B_{\mu\nu}$ is equal to 1. matrix multiplication shows, that all diagonal elements of $A^{-1}\frac{d}{dA_{\mu\nu}}A$ are zero but one, being equal to $(A^{-1})_{\nu\mu}$. □

Jacobi's formula and the last lemma allow to calculate some common derivatives of the metric determinant.

Lemma D.1.5.
Let (M, g) be a (pseudo) Riemannian manifold. Then for all positions x, for which $\det(g_{\mu\nu}(x)) \neq 0$ holds, the following derivatives of $\sqrt{|\det(g_{\mu\nu})|}$ apply:

i) $$\frac{d}{dg_{\mu\nu}}\sqrt{|\det(g_{\mu\nu})|} = \frac{1}{2}\sqrt{|\det(g_{\mu\nu})|}g^{\mu\nu} \ ,$$

ii) $$\frac{d}{dg^{\mu\nu}}\sqrt{|\det(g_{\mu\nu})|} = -\frac{1}{2}\sqrt{|\det(g_{\mu\nu})|}g_{\mu\nu} \ ,$$

iii) $\frac{\partial}{\partial x^\eta}\sqrt{|\det(g_{\mu\nu})|} = \sqrt{|\det(g_{\mu\nu})|}\Gamma^\mu_{\mu\eta}$.

Proof D.1.6.
By assumption we are at positions where the absolute value is a differentiable function with derivative (for $t \neq 0$) $\frac{d}{dt}|t| = \frac{t}{|t|} = \text{sign}(t)$. Let $\ominus$ denote $\text{sign}(\det(g_{\mu\nu}))$ in the following:

$$\begin{aligned}\frac{d}{dg_{\mu\nu}}\sqrt{|\det(g_{\mu\nu})|} &= \frac{1}{2\sqrt{|\det(g_{\mu\nu})|}}\frac{d}{dg_{\mu\nu}}|\det(g_{\mu\nu})| \\ &= \ominus\frac{1}{2\sqrt{|\det(g_{\mu\nu})|}}\frac{d}{dg_{\mu\nu}}\det(g_{\mu\nu}) \\ &= \ominus\frac{\det(g_{\mu\nu})}{2\sqrt{|\det(g_{\mu\nu})|}}g^{\nu\mu} = \frac{|\det(g_{\mu\nu})|}{2\sqrt{|\det(g_{\mu\nu})|}}g^{\mu\nu} \\ &= \frac{1}{2}\sqrt{|\det(g_{\mu\nu})|}g^{\mu\nu} \ .\end{aligned}$$

With $\det(g_{\mu\nu}) = \det(g^{\mu\nu})^{-1}$, since $(g_{\mu\nu})^{-1} = (g^{\mu\nu})$ it follows that:

$$\begin{aligned}\frac{d}{dg^{\mu\nu}}\det(g_{\mu\nu}) &= \frac{d}{dg^{\mu\nu}}\frac{1}{\det(g^{\mu\nu})} = -\frac{1}{\det(g^{\mu\nu})^2}\frac{d}{dg^{\mu\nu}}\det(g^{\mu\nu}) \\ &= -\frac{1}{\det(g^{\mu\nu})}g_{\mu\nu} = -\det(g_{\mu\nu})g_{\mu\nu} \ .\end{aligned}$$

With the first calculation we obtain:

$$\frac{d}{dg_{\mu\nu}}\sqrt{|\det(g_{\mu\nu})|} = -\frac{1}{2}\sqrt{|\det(g_{\mu\nu})|}g_{\mu\nu} \ .$$

Before we can prove the last equation, we need to show how the trace of the partial derivatives of the metric $g^{\mu\rho}g_{\mu\rho,\eta}$ relate to the Christoffel symbols:

$$2\Gamma^\mu_{\mu\eta} = g^{\mu\rho}(g_{\eta\rho,\mu} + g_{\rho\mu,\eta} - g_{\mu\eta,\rho}) = g_{\eta\rho,}{}^{\rho} - g_{\eta\mu,}{}^{\mu} + g^{\mu\rho}g_{\mu\rho,\eta} = g^{\mu\rho}g_{\mu\rho,\eta} \ .$$

With this , using Jacobi's formula it follows, that:

$$\begin{aligned}\frac{\partial}{\partial x^\eta}\sqrt{|\det(g_{\mu\nu})|} &= \frac{1}{2\sqrt{|\det(g_{\mu\nu})|}}\frac{\partial}{\partial x^\eta}|\det(g_{\mu\nu})| = \\ &\ominus\frac{1}{2\sqrt{|\det(g_{\mu\nu})|}}\frac{\partial}{\partial x^\eta}\det(g_{\mu\nu}) \\ &= \ominus\frac{\det(g_{\mu\nu})}{2\sqrt{|\det(g_{\mu\nu})|}}\text{tr}(g^{\mu\nu}g_{\nu\rho,\eta}) \\ &= \frac{|\det(g_{\mu\nu})|}{2\sqrt{|\det(g_{\mu\nu})|}}g^{\mu\nu}g_{\mu\nu,\eta} \\ &= \sqrt{|\det(g_{\mu\nu})|}\Gamma^\mu_{\mu\eta} \ .\end{aligned}$$

□

Corollary D.1.7.
In the last proof, we have seen, that:

$$g^{\mu\nu} g_{\mu\nu,\eta} = 2\Gamma^{\mu}_{\mu\eta} \ .$$

D.2. Densities

There is much confusion about the meaning of densities in the physical and mathematical literature, due to different usage of the term. Even worse, the proper meaning of quantities like the current-density is a mathematical density not a physical density.

D.2.1. Mathematical Density (short review)

The necessity of densities arise in the context of integration over non-orientable manifolds. In the case of oriented manifolds one naturally uses top-forms to integrate. There are no problems, as long as one restricts oneself to oriented charts. However, since there are two possible orientations, there are two different oriented atlases. Choosing a chart with different orientation results in an additional minus sign in integration. This inconvenience reveals, that top-forms are not the perfect integrands. Loosely speaking, **densities** in mathematics are differential forms that account for orientation changes with an additions sign change, such that integration remains invariant. In fact, densities allow for integration on non-orientable manifolds. More formally, densities are sections of $\bigwedge(T^*M) \otimes \mathcal{O}(M)$, where $\mathcal{O}(M)$ is the orientation line bundle of M. In the orientable case, densities and differential forms are related by a sign change, allowing to define densities as equivalence classes.[Jän05]

D.2.2. Densities in Physics

The meaning of densities in the physical literature comes form the need to integrate on manifolds, at a time when differential forms were not available. We can motivate the definition with the integration of a top form/density over a manifold, similarly to [Car97, chapter 2].

Given an n-form ω, the integration $\int_U \omega$ for $U \subset M$ is well defined. If U can be covered with a single chart (U_x, x), then

$$\int_U \omega = \int_{x(U)} \omega^{(x)}_{1,\ldots,n} \, dx^1 \wedge \ldots \wedge dx^n \ .$$

Here $\omega^{(x)}_{1,\ldots,n} \, dx^1 \wedge \ldots \wedge dx^n$ is the pulled back form $(x^{-1})^*\omega$ on $\mathbb{R}^n$. Choosing a second chart (U_y, y), that covers U, we obtain a different integration region $y(U)$, but more importantly for this motivation section, we obtain a different form on $\mathbb{R}^n$:

$$(y^{-1})^*\omega = \omega^{(y)}_{1,\ldots,n} \, dy^1 \wedge \ldots \wedge dy^n \ .$$

Denoting the Jacobi determinant $\det\left(\frac{\partial x^i}{\partial y^i}\right)$ with J^x_y, we can write

$$dx^1 \wedge \ldots \wedge dx^n = J^x_y dy^1 \wedge \ldots \wedge dy^n \ .$$

Hence the transformation for the coefficients is

$$\omega^{(y)}_{1,\dots,n} = J^x_y \omega^{(x)}_{1,\dots,n} \ .$$

If however ω were a density (in the mathematical sense), the transformation would read

$$\omega^{(y)}_{1,\dots,n} = |J^x_y| \omega^{(x)}_{1,\dots,n} \ .$$

The space of top forms/densities is one dimensional, as is the space of functions, for each $p \in M$.[2] In fact, the functions and top forms/densities are isomorphic on oriented manifolds. However, functions transform different than the coefficients of top forms/densities. Motivated by this, we define:

Definition D.2.1.
A **tensor density** $T^{a\dots b}{}_{c\dots d}$ transforms as a tensor with an additional $|J^x_y|$:

$$\left(T^{(y)}\right)^{a\dots b}{}_{c\dots d} = |J^x_y| \frac{\partial y^a}{\partial x^\alpha} \cdots \frac{\partial y^b}{\partial x^\beta} \frac{\partial x^\gamma}{\partial y^c} \cdots \frac{\partial x^\delta}{\partial y^d} \left(T^{(x)}\right)^{\alpha\dots\beta}{}_{\gamma\dots\delta} \ .$$

A **pseudo tensor density** transforms as a tensor with an additional J^x_y:

$$\left(T^{(y)}\right)^{a\dots b}{}_{c\dots d} = J^x_y \frac{\partial y^a}{\partial x^\alpha} \cdots \frac{\partial y^b}{\partial x^\beta} \frac{\partial x^\gamma}{\partial y^c} \cdots \frac{\partial x^\delta}{\partial y^d} \left(T^{(x)}\right)^{\alpha\dots\beta}{}_{\gamma\dots\delta} \ .$$

In the same way a (pseudo) tensor density of weight w transforms as a tensor with an additional $|J^x_y|^w$ and $((J^x_y)^w$ for pseudo densities).[3]

One can see, that tensor densities are related to mathematical densities. But be aware, that an antisymmetric tensor density is no mathematical density.

Remark D.2.2.
To avoid confusion, we use the term density on its own only to describe mathematical densities. Otherwise we write physical density, with the exception of (pseudo) tensor density.

D.2.3. Metric determinant and densities

We stated above, that the metric determinant is no proper function on a manifold. In fact it is a scalar density. To see that, we begin with the transformation behavior of the metric tensor coefficients:

$$g^{(y)}_{ab} = \frac{\partial x^\alpha}{\partial y^a} \frac{\partial x^\beta}{\partial y^b} g^{(x)}_{\alpha\beta} \ .$$

[2] Of course over M it becomes infinite dimensional.

[3] In the literature pseudo tensor densities of weight w transform with $\mathrm{sign}(J^x_y)|J^x_y|^w$. However, from our motivation $(J^x_y)^w$ is the more natural extension.

Taking the metric determinant and observing that the contraction can be understand as matrix multiplication in this context, we find:

$$\det(g^{(y)}_{ab}) = \det\left(\frac{\partial x^\alpha}{\partial y^a}\right)\det\left(\frac{\partial x^\beta}{\partial y^b}\right)\det(g^{(x)}_{\alpha\beta}) = (J^x_y)^2\det(g^{(x)}_{\alpha\beta}) \ .$$

Finally we see that $\sqrt{|\det(g_{\mu\nu})|}$ is a scalar density:

$$\sqrt{|\det(g^{(y)}_{\mu\nu})|} = |J^x_y|\sqrt{|\det(g^{(x)}_{\mu\nu})|} \ .$$

For the remainder we use the convention, that $|g| = |\det(g^{(x)}_{\mu\nu})|$, since $|g|$ has no further meanings here anyway. Thus:

$$\boxed{\sqrt{|g^{(y)}|} = |J^x_y|\sqrt{|g^{(x)}|} \ .}$$

Constructing tensor fields by components, we immediately find:

Theorem D.2.3.
Let $T^{a\ldots b}{}_{c\ldots d}$ be a tensor field, then $\mathfrak{T}^{a\ldots b}{}_{c\ldots d} = \sqrt{|g|}^{\,w} T^{a\ldots b}{}_{c\ldots d}$ is a tensor density of weight w. Conversely, if $\mathfrak{T}^{a\ldots b}{}_{c\ldots d}$ is a tensor density of weight w, then $T^{a\ldots b}{}_{c\ldots d} = \frac{1}{\sqrt{|g|}^{\,w}}\mathfrak{T}^{a\ldots b}{}_{c\ldots d}$ is a regular tensor field.

There is a very useful relation between the metric determinant and covariant derivatives, stated in the following lemma:

Lemma D.2.4.
*Let V^μ be a tensor field, then the **covariant divergence** can be calculated as follows:*

$$V^\mu{}_{;\mu} = \frac{1}{\sqrt{|g|}}\left(\sqrt{|g|}V^\mu\right)_{,\mu} \ .$$

Proof D.2.5.
Beginning from the right side, using lemma D.1.5, we calculate

$$\begin{aligned}\left(\sqrt{|g|}V^\mu\right)_{,\mu} &= (\partial_\mu\sqrt{|g|})V^\mu + \sqrt{|g|}\partial_\mu V^\mu = \sqrt{|g|}(\Gamma^\eta_{\eta\mu}V^\mu + V^\mu{}_{,\mu}) \\ &= \sqrt{|g|}V^\mu{}_{;\mu} \ .\end{aligned}$$

□

The last Lemma can extended to higher rank tensors, if interpreted right, to find covariant Euler-Lagrange-equations, according to [HE75, section 3.3]:

Corollary D.2.6.
Understanding $T^{a\ldots\mu\ldots b}{}_{c\ldots d} = V^\mu$ as the coefficients of a vector, we can write

$$\nabla_{\partial_\mu} T^{a\ldots\mu\ldots b}{}_{c\ldots d} = \frac{1}{\sqrt{|g|}}\partial_\mu\left(\sqrt{|g|}T^{a\ldots\mu\ldots b}{}_{c\ldots d}\right) \ ,$$

using a slightly different notation than before, to make visible that we do not take the covariant derivative over the full tensor. Form $g^{ab}{}_{;\mu} = 0$ we see that we can also write

$$\nabla_{\partial_\mu} T^{a\ldots b}{}_{c\ldots\mu\ldots d} = \frac{1}{\sqrt{|g|}}\partial_\mu\left(\sqrt{|g|}T^{a\ldots b}{}_{c\ldots\mu\ldots d}\right) \ .$$

The preceding corollary can become rather confusing, if it is not mentioned explicitly how to understand the covariant derivative there.

Theorem D.2.7.
Let $\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)$ be a Lagrange density. Then the first variation $\delta S[\phi,\psi]$ of the functional $S[\phi] = \int_U \mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)\ \sqrt{|g|}d^n x$ vanishes for all $\psi \equiv 0$ on ∂U, if and only if the ***covariant Euler-Lagrange-equations*** *are satisfied:*

$$\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_J} - \nabla_\mu \frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_{J;\mu}} = 0 \ .$$

Proof D.2.8.
The proof of theorem B.2.1 carries over due to the last corollary. For the adaptions necessary we observe, that:

$$\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_{J;\mu}}\psi^I_{J;\mu} = \nabla_\mu\left(\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_{J;\mu}}\psi^I_J\right) - \left(\nabla_\mu\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_{J;\mu}}\right)\psi^I_J \ .$$

Also, to use the same argument (Stokes theorem) as before, we write

$$\int_U \nabla_\mu\left(\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_{J;\mu}}\psi^I_J\right)\ d^n x = \int_U \frac{1}{\sqrt{|g|}}\partial_\mu\left(\sqrt{|g|}\frac{\partial\mathcal{L}(\phi^I_J, \phi^I_{J;\mu}, x)}{\partial\phi^I_{J;\mu}}\psi^I_J\right)\ d^n x \ .$$

□

E
Vector valued differential forms

Differential forms have been used throughout this document without introduction. Since they are a well covered topic in mathematics, we refer to the literature for a thorough introduction. However, to alleviate the lack of introduction, we will cover Vector valued differential forms more detailed than usually done. Part of the reason for the mostly concise introductions is, that many definitions and theorems carry over without changes. This chapter is mostly based on [Bau14], [Mor01] and [RS18].

E.1. Natural bundle operations

Without proof here, some operations on vector bundles need to be introduced. For that reason, we recall the definition of transition functions (see page 60), and quote a vector bundle construction lemma:

Lemma E.1.1 ([Lee97, lemma 2.2]).
Let M be a smooth manifold, E a set and $\pi\colon E \to M$ a surjective map. Let $\{U_\alpha, \varphi_\alpha\}$ be an open cover of M with maps $\varphi\colon \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^n$ that satisfy $\pi_1 \circ \varphi_\alpha = \pi$, where π_1 is the projection on the first component. If $\varphi_\alpha \circ \varphi_\beta^{-1}$ define maps

$$\varphi_\alpha \circ \varphi_\beta^{-1}\colon\ (U_\alpha \cap U_\beta) \times \mathbb{R}^k \longrightarrow (U_\alpha \cap U_\beta) \times \mathbb{R}^k \ , \quad (p, v) \longmapsto (p, f_{\alpha\beta}(p)v) \ .$$

such that $f_{\alpha\beta}\colon U_\alpha \cap U_\beta \to GL(\mathbb{R}^k)$, then (E, M, π) is a vector bundle with local trivializations $\{\varphi_\alpha\}$.

This lemma proves, that the tensor bundles we introduced in subsection C.2.1 are indeed vector bundles. Here, we use it to define some important vector bundles, two of which we are already familiar with, following [Bau14, section 2.4]:

The dual bundle

Let (E, M, π) be a vector bundle with local trivializations $\{U_\alpha, \varphi_\alpha\}$ and define

$$E^* := \bigcup_{p \in M} E_p^* \ ,$$

where E_p^* is the dual vector space of the fiber $E_p = \pi^{-1}(\{p\})$. Define further $\pi^*\colon \varphi_p \in E_p^* \to p \in M$. Then (E^*, M, π^*) is a vector bundle, called **dual bundle**, together with the maps

$$\varphi_\alpha^*\colon U_\alpha \longrightarrow U_\alpha \times (\mathbb{R}^n)^* \ , \quad \vartheta_p \in E_p^* \longmapsto (p, \vartheta_p \circ \varphi_\beta^{-1}) \ ,$$

for all β, such that $p \in U_\beta$.

The tensor bundle

Let (E, M, π) and $(\widetilde{E}, M, \widetilde{\pi})$ be two vector bundles with local trivializations $\{U_\alpha, \varphi_\alpha\}$ and $\{\widetilde{U_\alpha}, \widetilde{\varphi_\alpha}\}$. The **tensor bundle** $(E \otimes \widetilde{E}, M, \pi_\otimes)$ is defined by

$$E \otimes \widetilde{E} \coloneqq \bigcup_{p\in M} E_p \otimes \widetilde{E}_p \ , \quad \pi_\otimes \colon (v_p, w_p) \in E_p \otimes \widetilde{E}_p \to p \in M \ ,$$

$$\varphi_\alpha(v_p, w_p) = (p, \varphi'_\alpha(v_p) \otimes \widetilde{\varphi_\alpha}'(w_p)) \ ,$$

where φ'_α denotes $\pi_2 \circ \varphi_\alpha$ with $\pi_2 \colon U \times \mathbb{R}^n \to \mathbb{R}^n$ and $\widetilde{\varphi_\alpha}'$ respectively.

The Whitney bundle

Let (E, M, π) and $(\widetilde{E}, M, \widetilde{\pi})$ be as before. The **Whitney bundle** $(E \oplus \widetilde{E}, M, \pi_\oplus)$ is defined by

$$E \oplus \widetilde{E} \coloneqq \bigcup_{p\in M} E_p \oplus \widetilde{E}_p \ , \quad \pi_\oplus \colon (v_p, w_p) \in E_p \oplus \widetilde{E}_p \to p \in M \ ,$$

$$\varphi_\alpha(v_p, w_p) = (p, \varphi'_\alpha(v_p) \oplus \widetilde{\varphi_\alpha}'(w_p)) \ .$$

The Homomorphism bundle

Let (E, M, π) and $(\widetilde{E}, M, \widetilde{\pi})$ be as before. The **Homomorphism bundle** is defined by

$$\mathrm{Hom}(E, \widetilde{E}) \coloneqq \bigcup_{p\in M} \mathrm{Hom}(E_p, \widetilde{E}_p) \ , \quad \pi_H \colon L_p \in \mathrm{Hom}(E_p, \widetilde{E}_p) \to p \in M \ ,$$

$$\varphi_\alpha(L_p) = (p, \widetilde{\varphi}_\alpha \circ L_p \circ \varphi_\alpha^{-1}) \ .$$

The Pullback bundle

Let (E, M, π) be a vector bundle and $\psi \colon N \to M$ a diffeomorphism. The **pullback bundle** $(\psi^* E, N, \overline{\pi})$ is a vector bundle defined by

$$\psi^* N \coloneqq \{(p, v) \in N \times E \mid f(p) = \pi(v)\} \quad \text{and} \quad \overline{\pi}(p, v) \coloneqq p \ .$$

Let $\{U_\alpha, \varphi_\alpha\}$ be local trivializations, then $\{\psi(U_\alpha), \phi_\alpha\}$ with

$$\phi_\alpha(p, v) = (p, \varphi'_\alpha(v))$$

are local trivializations of $(\psi^* E, N, \overline{\pi})$.

Definition E.1.2.
We keep the notation introduced in this section for the rest of this chapter, that is:

$$\varphi_\alpha = (\pi, \varphi'_\alpha) \qquad \Rightarrow \qquad \varphi_\alpha(\omega_p) = (p, \varphi'_\alpha(\omega_p)) \ .$$

E.2. Definition of vector bundle valued differential forms

Usual differential k-forms are sections of $(\bigwedge^k(T^*M), M, \pi)$, i.e $\omega \in \Gamma\left(\bigwedge^k(T^*M)\right)$. That means:

$$\omega_p \colon \underbrace{T_pM \times \ldots \times T_pM}_{k\text{-times}} \longrightarrow \mathbb{R}$$

is an alternating multilinear map, that varies smoothly in $p \in M$. Vector bundle valued k-forms are alternating linear maps into the fibers E_p of an vector bundle (E, M, π):

$$\Omega_p \colon \underbrace{T_pM \times \ldots \times T_pM}_{k\text{-times}} \longrightarrow E_p \ .$$

Recalling the isomorphy of tensor spaces and the space of linear maps (see lemma A.1.11) in the case of finite dimensional vector spaces, this means:

$$\Omega_p \in E_p \otimes \left(\bigwedge^k(T_pM)\right)^* \cong E_p \otimes \bigwedge^k(T_p^*M) \cong \bigwedge^k(T_p^*M) \otimes E_p \ .$$

Using the fiber wise construction of the tensor bundle, we obtain the formal definition of vector bundle valued differential forms:

Definition E.2.1.
A **vector bundle valued differential k-form** with values in the vector bundle (E, M, π) is a smooth section of the vector bundle $(\bigwedge^k(T^*M) \otimes E, M, \pi_k)$. The vector space of these vector valued k-forms is denoted by $\Omega^k(M, E)$.

Remark E.2.2.
For short, we will call vector bundle valued differential k-form simply **vector k-form**.

Indeed, this definition is a straightforward generalization of k-forms, since $\bigwedge^k(T_p^*M) \otimes_{\mathbb{R}} \mathbb{R} \cong \bigwedge^k(T_p^*M)$ and thus

$$\Omega^k(M) = \Gamma\left(\bigwedge^k(TM), M\right) \cong \Gamma\left(\bigwedge^k(TM) \otimes \mathbb{R}, M\right) \ .$$

Lemma E.2.3.
Let $\hat{\otimes}$ denote a tensor product linear in $C^\infty(M)$-functions, i.e. $\hat{\otimes} = \otimes_{C^\infty(M)}$, then for every $\eta \in \Omega^k(M)$ and $v \in \Gamma(E, M)$ it holds that $\eta \hat{\otimes} v \in \Omega^k(M, E)$, with the definition:

$$(\eta \hat{\otimes} v)_p := \eta_p \otimes v_p \ .$$

Proof E.2.4.
First, we notice that $\eta_p \otimes v_p$ is well defined by the construction of tensor bundles. Also, $\eta_p \otimes v_p$ defines an alternating multilinear map

$$\eta_p \otimes v_p \colon \underbrace{T_pM \times \ldots \times T_pM}_{k\text{-times}} \longrightarrow E_p \ ,$$

by setting

$$(\eta_p \otimes v_p)(X_1, \ldots, X_k) = \eta_p(X_1, \ldots, X_k) \cdot v_p \in E_p \ .$$

Hence $(\eta \widehat{\otimes} v)_p$ is such an alternating multilinear map. We need to show, that $(\eta \widehat{\otimes} v)_p$ is smooth in p.

This can be done as follows: Let (U_α, ϕ_α) be a local trivialization of $(\bigwedge^k(T^*M) \otimes E, M, \pi_k)$ as defined for tensor bundles:

$$\phi_\alpha \colon \pi_k^{-1}(U_\alpha) \longrightarrow U_\alpha \times \left(\bigwedge\nolimits^k(\mathbb{R}^n) \otimes \mathbb{R}^m\right) \ .$$

Let $\varphi_\alpha(\eta_p) = (p, \eta(p))$ and $\widetilde{\varphi}_\alpha(v_p) = (p, v(p))$. Then, by definition, $\eta(p)$ is a smooth vector field $U_\alpha \to \bigwedge^k(\mathbb{R}^n)$, $v(p)$ is a smooth vector field $U_\alpha \to \mathbb{R}^m$, and also $\eta(p) \otimes v(p)$ is a smooth vector field on $\bigwedge^k(\mathbb{R}^n) \otimes \mathbb{R}^m$. Hence,

$$\phi_\alpha^{-1}(p, \eta(p) \otimes v(p)) =: \rho$$

defines an element $\rho \in \Omega^k(M, E)$, i.e. a smooth section. But for all $p \in U_\alpha$ we have:

$$\phi_\alpha(\rho_p) = (p, \eta(p) \otimes v(p)) = \phi_\alpha(\eta_p \otimes v_p) = \phi_\alpha((\eta \widehat{\otimes} v)_p) \ .$$

Thus $\eta \widehat{\otimes} v$ is smooth on U_α and by repetition with different U_β on the whole of M.

To be well defined, we finally need to show, that the linearity over $C^\infty(M)$ leads to no contradictions. This can be done fiber-wise:

$$\left((f \cdot \eta) \widehat{\otimes} v\right)_p = f(p) \cdot \eta_p \otimes v_p = \eta_p \otimes f(p) v_p = \left(\eta \widehat{\otimes} f \cdot v\right)_p \ .$$

□

As a corollary of the last proof, we also have

Corollary E.2.5.
Let (U_α, ϕ_α) be a local trivialization of $(\bigwedge^k(T^*M) \otimes E, M, \pi_k)$. On U_α any vector k-form $\omega \in \Omega^k(M, E)$ can be written as

$$\omega = \omega_\alpha{}^\beta \eta^\alpha \otimes_{C^\infty(U_\alpha)} v_\beta \ ,$$

with local frames $\{\eta^\alpha\}$ of $(\bigwedge^k(T^*M), M, \pi)$ and $\{e_\beta\}$ of (E, M, π).

In fact, some sources[1] state that even $\Omega^k(M, E) \simeq \Omega^k(M) \oplus_{C^\infty(M)} \Gamma(E)$ holds.

[1] For example [Wik18], which refers to a forum, that itself refers to the book “Differentiable Manifolds” from Lawrence Colon.

Remark E.2.6.
As we have done throughout the whole document, we will drop the special notation and simply write $\otimes$ for $\widehat{\otimes}$ and $\otimes_{C^\infty(U_\alpha)}$ etc., yet being aware of its special meaning on bundles.

To avoid misunderstandings in the next section, we mention the natural meaning of $\omega \in \Omega^k(M, E)$ as map:

$$\omega\colon \Gamma(TM, M) \times \ldots \times \Gamma(TM, M) \longrightarrow \Gamma(E, M) \ ,$$

that is alternating and multilinear over $C^\infty(M)$.

E.3. Operations on vector forms

Definition E.3.1.
Let $Q\colon E_1 \oplus E_2 \to E_3$ be a fibre wise non-degenerate bilinear map of vector bundles (E_i, M, π_i) that is smooth, then the **wedge product** $\wedge_Q\colon \Omega^k(M, E_1) \times \Omega^\ell(M, E_2) \to \Omega^{k+\ell}(M, E_3)$ is defined by

$$(\omega \wedge_Q \eta)(X_1, \ldots, X_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in \Sigma_{k+\ell}} \operatorname{sgn}(\sigma) Q\Big(\omega(X_{\sigma(1)}, \ldots, X_{\sigma(k)}); \\ \eta(X_{\sigma(k+1)}, \ldots, X_{\sigma(k+\ell)})\Big) \ ,$$

for tangent vector fields $X_1, \ldots, X_{k+\ell}$.

It should be mentioned, that different normalization coefficients than $\frac{1}{k!\ell!}$ are used by some authors in some situations. In the case of usual differential forms, the natural choice of a bilinear map is the field multiplication $\cdot\colon \mathbb{R} \oplus \mathbb{R} \to \mathbb{R}$.

Definition and corollary E.3.2.
Differential k-forms operate on vector ℓ-forms by the wedge product $\wedge\colon \Omega^k(M) \times \Omega^\ell(M, E) \to \Omega^{k+\ell}(M, E)$. The natural bilinear map is the product, fiber wise defined by $\cdot\colon \mathbb{R} \times E_p \to E_p$.

Proof E.3.3.
From

$$\bigwedge\nolimits^k(TM) \cong \bigwedge\nolimits^k(TM) \otimes \mathbb{R} \quad \text{and} \quad \bigwedge\nolimits^k(TM) \otimes E \cong \bigwedge\nolimits^k(TM) \otimes E \otimes \mathbb{R} \ ,$$

the statement is immediate. $\square$

In the following, we will introduce operators, known from usual differential forms, that carry over without any changes in the definitions. We will also see, that they keep their behavior.

Definition E.3.4.
Let $\omega \in \Omega^k(M, E_1)$ and $\eta \in \Omega^k(N, E_2)$ be vector forms, $X_i \in \Gamma(TN, N)$ vector fields and $\psi\colon N \to M$ a diffeomorphism.

- The **pulled back vector form** $\psi^*\omega \in \Omega^k(N, \psi^*E_1)$ is defined by

$$(\psi^*\omega)_p(X_1, \dots, X_k)_p := \omega_{\psi(p)}(D_p\psi(X_1), \dots, D_p\psi(X_k)) \ .$$

- The **interior product** of X_0 with η is defined by:

$$X_0 \lrcorner\colon \Omega^k(N, E_2) \longrightarrow \Omega^{k-1}(N, E_2) \ ,$$

$$(X_0 \lrcorner \eta)(X_1, \dots, X_{k-1}) = \eta(X_0, X_1, \dots, X_{k-1}) \ .$$

Lemma E.3.5.
Let $\omega \in \Omega^k(M, E)$ and $\eta \in \Omega^\ell(M, E)$ be vector forms, $X \in \Gamma(TM, M)$ be a vector field and $\psi\colon M \to N$ a diffeomorphism, then:

i) The wedge product $\wedge_Q$ is bilinear.

ii) If Q is symmetric: $\omega \wedge_Q \eta = (-1)^{k\ell} \eta \wedge_Q \omega$.

iii) ψ^* is a linear map $\Omega^k(M, E) \to \Omega^k(\psi^*E, N)$.

iv) $\psi^*(\omega \wedge_Q \eta) = \psi^*\omega \wedge_{\psi^*Q} \psi^*\eta$.

v) $X\lrcorner$ is a linear operator.

vi) $X \lrcorner(\omega \wedge_Q \eta) = (X\lrcorner\omega) \wedge_Q \eta + (-1)^k \omega \wedge_Q X \lrcorner \eta$.

Proof E.3.6.

i) This is immediate form the bilinearity of Q.

ii) For the notation, let $\widehat{X}$ denote the omission of X, and assume $\ell + 1 = k + j$:

$$\begin{aligned}
(\omega \wedge_Q \eta)(X_1, \dots, X_{k+\ell}) &= (-1)^\ell (\omega \wedge_Q \eta)(X_{k+j}, X_1, \dots, \widehat{X}_{k+j}, \dots, X_{k+\ell}) \\
&= (-1)^{k\ell} (\omega \wedge_Q \eta)(X_{k+j}, \dots, X_{k+\ell}, X_1, \dots, X_\ell) \\
&= (-1)^{k\ell} \frac{1}{k!\ell!} \sum_{\sigma \in \Sigma_{k+\ell}} \operatorname{sgn}(\sigma) Q\Big(\omega(X_{\sigma(k+j)}, \dots, X_{\sigma(k+\ell)}) \ ; \\
&\qquad \eta(X_{\sigma(1)}, \dots, X_{\sigma(\ell)})\Big) \\
&= (-1)^{k\ell} \frac{1}{k!\ell!} \sum_{\sigma \in \Sigma_{k+\ell}} \operatorname{sgn}(\sigma) Q\Big(\eta(X_{\sigma(1)}, \dots, X_{\sigma(\ell)}) \ ; \\
&\qquad \omega(X_{\sigma(k+j)}, \dots, X_{\sigma(k+\ell)})\Big) \\
&= (-1)^{k\ell} (\eta \wedge_Q \omega)(X_1, \dots, X_{k+\ell}) \ .
\end{aligned}$$

iii) This follows from the fact, that the fibers E_p are vector spaces.

iv) Plugging in the definition for both sides shows the equality.

v) Immediate from definitions.

vi) To simplify the proof, we notice that

$$\begin{aligned}(\omega \wedge_Q \eta)(X_1, \dots, X_{k+\ell}) &= \frac{1}{k!\ell!} \sum_{\sigma \in \Sigma_{k+\ell}} \operatorname{sgn}(\sigma) Q \left(\omega(X_{\sigma(1)}, \dots, X_{\sigma(k)}) \; ; \right. \\ &\qquad \left. \eta(X_{\sigma(k+1)}, \dots, X_{\sigma(k+\ell)}) \right) \\ &= \sum_{\sigma \in \bigwedge_{k+\ell}} \operatorname{sgn}(\sigma) Q \left(\omega(X_{\sigma(1)}, \dots, X_{\sigma(k)}); \right. \\ &\qquad \left. \eta(X_{\sigma(k+1)}, \dots, X_{\sigma(k+\ell)}) \right)\end{aligned}$$

Where $\sigma \in \bigwedge_{k+\ell}$ means $\sigma(1) < \ldots < \sigma(k)$ and $\sigma(k+1) < \ldots < \sigma(k+\ell)$. This can be done, since any transposition creates a minus sign in $\operatorname{sgn}(\sigma)$ as well as in ω or η. Thus there are $k!$ equal terms in ω and $\ell!$ equal terms in η in the set of full permutations, explaining the normalization $\frac{1}{k!\ell!}$. The set of permutation $\sigma \in \bigwedge_{k+\ell}$ consist of those that satisfy $\sigma(1) = 1$ and those that satisfy $\sigma(k+1) = 1$. This covers all allowed permutations, since $\sigma(1) < \ldots < \sigma(k)$ and $\sigma(k+1) < \ldots < \sigma(k+\ell)$.

$$\begin{aligned}&(\omega \wedge_Q \eta)(X_1, \dots, X_{k+\ell}) = \\ &\quad = \sum_{\sigma \in \bigwedge_{k+\ell}, \sigma(1)=1} \operatorname{sgn}(\sigma) Q \left(\omega(X_{\sigma(1)}, \dots, X_{\sigma(k)}); \eta(X_{\sigma(k+1)}, \dots, X_{\sigma(k+\ell)}) \right) \\ &\qquad + \sum_{\sigma \in \bigwedge_{k+\ell}, \sigma(k+1)=1} \operatorname{sgn}(\sigma) Q \left(\omega(X_{\sigma(1)}, \dots, X_{\sigma(k)}); \eta(X_{\sigma(k+1)}, \dots, X_{\sigma(k+\ell)}) \right)\end{aligned}$$

Let τ be a permutation in $\bigwedge_{k-1+\ell}$, that acts on $\{2, \dots, k+\ell\}$. Then, in the first term $\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)$ if $\tau(i) = \sigma(i)$ for $i \in \{2, \dots, k+\ell\}$. For the second term, we use the permutation

$$\rho = \begin{pmatrix} 1 & 2 & \dots & k & k+1 & \dots \\ 2 & 3 & \dots & k+1 & 1 & \dots \end{pmatrix} \equiv (1\;2 \dots k+1) \;,$$

which has a signum[2] of $(-1)^(k+1) - 1 = (-1)^k$. It follows that

$$\begin{aligned}&(\omega \wedge_Q \eta)(X_1, \dots, X_{k+\ell}) = \\ &\quad = \sum_{\tau \in \bigwedge_{k-1+\ell}} \operatorname{sgn}(\sigma) Q \left(\omega(X_1, \dots, X_{\tau(k)}); \eta(X_{\tau(k+1)}, \dots, X_{\tau(k+\ell)}) \right) \\ &\qquad + \sum_{\tau \in \bigwedge_{k-1+\ell}} \operatorname{sgn}(\tau \circ \rho) Q \left(\omega(X_{\tau(\rho(1))}, \dots, X_{\tau(\rho(k))}); \eta(X_{\tau(\rho(k+1))}, \dots, X_{\tau(\rho(k+\ell))}) \right) \\ &\quad = \sum_{\tau \in \bigwedge_{k-1+\ell}} \operatorname{sgn}(\sigma) Q \left([X_1 \lrcorner \omega](X_{\tau(2)}, \dots, X_{\tau(k)}); \eta(X_{\tau(k+1)}, \dots, X_{\tau(k+\ell)}) \right) \\ &\qquad + (-1)^k \sum_{\tau \in \bigwedge_{k-1+\ell}} \operatorname{sgn}(\tau) Q \left(\omega(X_{\tau(2)}, \dots, X_{\tau(k+1)}); [X_1 \lrcorner \eta](X_{\tau(k+2)}, \dots, X_{\tau(k+\ell)}) \right) \\ &\quad = ([X_1 \lrcorner \omega] \wedge_Q \eta + (-1)^k \omega \wedge_Q [X_1 \lrcorner \eta])(X_2, \dots, X_{k+\ell}) \;.\end{aligned}$$

□

Corollary E.3.7.
Let $\omega \in \Omega^k(M,E)$ and $\omega \in \Omega^\ell(M,E)$ be vector forms with local frame representation $\omega = {\omega_\alpha}^\beta \eta^\alpha \otimes e_\beta$ and $\rho = {\rho_\gamma}^\delta \xi^\gamma \otimes f_\delta$. Then, the wedge product and the interior product have the following form:

$$\omega \wedge_Q \rho = {\omega_\alpha}^\beta {\rho_\gamma}^\delta \left(\eta^\alpha \wedge \xi^\delta\right) \otimes Q(e_\beta, f_\delta) \ ,$$

$$X \lrcorner \left({\omega_\alpha}^\beta \eta^\alpha \otimes e_\beta\right) = {\omega_\alpha}^\beta \left(X \lrcorner \eta^\alpha\right) \otimes e_\beta \ .$$

Proof E.3.8.
Inserting the definitions of the operators, together with the last lemma proves the claims. □

E.4. Connections and exterior derivative

We recall the definition of a connection on an arbitrary vector bundle (definition C.2.9). Such a connection $\nabla\colon \Gamma(TM,M) \times \Gamma(E,M) \to \Gamma(E,M)$ can be understood as a map $d_\nabla\colon \Gamma(E,M) \to \Omega^1(M,E)$ by setting

$$d_\nabla X \equiv \omega := \nabla_\bullet X \qquad (d_\nabla X)(v) = \nabla_v X \ .$$

Indeed, vector 1-forms are $C^\infty(M)$-linear, agreeing with the definition of Connections. Noticing, that $\Gamma(E,M)$ has the meaning of E-valued functions, i.e.

$$\Gamma(E,M) \cong \Gamma(\mathbb{R} \otimes E, M) = \Gamma\left(\bigwedge\nolimits^0 (T^*M) \otimes E, M\right) = \Omega^0(M,E) \ ,$$

we may restate the definition of connections in the language of vector forms.

Definition E.4.1.
A **connection** on a vector bundle (E,M,π) is a linear map

$$d_\nabla\colon \Omega^0(M,E) \longrightarrow \Omega^1(M,E)$$

that satisfies the product rule:

$$d_\nabla(f \cdot X) = f \cdot d_\nabla X + df \otimes X \ .$$

Proof E.4.2.
Indeed, the product rule is the same as in definition C.2.9. Let $v \in \Gamma(TM,M)$,

[2] Let ℓ be the length of a cycle, then the cycle has the signum $(-1)^{\ell-1}$.

then, with $df(v) = v(f)$:

$$d_\nabla(fX)(v) = f d_\nabla X(v) + (df \otimes X)(v) = f\nabla_v X + v(f)X = \nabla_v(fX) \ .$$

□

Corollary E.4.3.
To every connection $d_\nabla \colon \Omega^0(M,E) \longrightarrow \Omega^1(M,E)$ *belongs a unique* $\nabla \colon \Gamma(TM,M) \times \Gamma(E,M) \to \Gamma(E,M)$ *and vice versa, by the definition*

$$d_\nabla X \equiv \nabla_\bullet X \qquad (d_\nabla X)(v) = \nabla_v X \ .$$

For d_∇ it is the same as for ∇. There are many possible connections in general. Hence the operator d_∇ is not uniquely determined a priory. However, one can fix a unique choice, e.g. the Levi-Civita connection on a Riemannian manifold.

Example E.4.4.
The usual Cartan-derivative $d \colon \Omega^0(M) \to \Omega^1(M)$ is a connection on the trivial bundle $M \otimes \mathbb{R}$.

Before defining the exterior covariant derivatives, that belong to a connection, we recall the invariant definition of the k-th order exterior derivative:

$$\begin{aligned}\left(d^k\omega\right)(X_0, X_1, \dots, X_k) \coloneqq &\sum_{i=0}^{k}(-1)^i X_i \left(\omega(X_0, \dots, \widehat{X_i}, \dots, X_k)\right) \\ &+ \sum_{0 \le i < j \le k} (-1)^{i+j}\omega([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k) \ .\end{aligned}$$

For the case of $d \colon \Omega^0(M) \to \Omega^1(M)$ the exterior derivative is defined by $df(v) = v(f)$, which allows to write

$$\begin{aligned}\left(d^k\omega\right)(X_0, X_1, \dots, X_k) \coloneqq &\sum_{i=0}^{k}(-1)^i d \left(\omega(X_0, \dots, \widehat{X_i}, \dots, X_k)\right)(X_i) \\ &+ \sum_{0 \le i < j \le k} (-1)^{i+j}\omega([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k) \ .\end{aligned}$$

Definition E.4.5 (See [KM13, section 11.13])**.**
Let d_∇ be a connection. The **exterior covariant derivative** of order k is a map

$$d_\nabla^k \colon \Omega^k(M,E) \longrightarrow \Omega^{k+1}(M,E) \ ,$$

defined by

$$\begin{aligned}\left(d_\nabla^k\omega\right)(X_0, X_1, \dots, X_k) \coloneqq &\sum_{i=0}^{k}(-1)^i d_\nabla \left(\omega(X_0, \dots, \widehat{X_i}, \dots, X_k)\right)(X_i) \\ &+ \sum_{0 \le i < j \le k} (-1)^{i+j}\omega([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k) \ .\end{aligned}$$

It is common practice, to drop the index k, simply writing d_∇.

Theorem E.4.6 (See [KM13, section 11.13]).
The exterior covariant derivative has the following properties:

i) For a decomposition $\omega = \eta \otimes v$ it holds, that: $d_\nabla(\eta \otimes v) = d\eta \otimes v + (-1)^{\deg \eta} \eta \wedge d_\nabla v$.

ii) For $\omega \in \Omega^k(M)$, $\eta \in \Omega^\ell(M, E)$ it holds, that: $d_\nabla(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^k \omega \wedge d_\nabla \eta$.

Proof E.4.7.

i) Writing out the definition of d_∇ for $\omega = \eta \otimes v$ and using the product rule for $d_\nabla \colon \Omega^0(M, E) \to \Omega^1(M, E)$ yields:

$$
\begin{aligned}
&(d_\nabla(\eta \otimes v))(X_0, \dots, X_k) = \\
&\quad = \sum_{i=0}^{k} (-1)^i d_\nabla \left(\eta(X_0, \dots, \widehat{X_i}, \dots, X_k) \cdot v \right) (X_i) \\
&\qquad + \sum_{0 \le i < j \le k} (-1)^{i+j} \eta([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k) \cdot v \\
&\quad = \sum_{i=0}^{k} (-1)^i d \left(\eta(X_0, \dots, \widehat{X_i}, \dots, X_k) \right) (X_i) \cdot v + \eta(X_0, \dots, \widehat{X_i}, \dots, X_k) \cdot (d_\nabla v)(X_i) \\
&\qquad + \sum_{0 \le i < j \le k} (-1)^{i+j} \eta([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k) \cdot v \\
&\quad = (d\eta \otimes v)(X_0, \dots, X_k) + \sum_{i=0}^{k} (-1)^i \eta(X_0, \dots, \widehat{X_i}, \dots, X_k) \cdot (d_\nabla v)(X_i) \ .
\end{aligned}
$$

To see, that the last term is exactly, what the theorem claims, we start from the other side, using the reasoning of proof E.3.6:

$$
(\eta \wedge d_\nabla v)(X_0, \dots, X_k) = \sum_{\sigma \in \bigwedge_{k,1}} \operatorname{sgn}(\sigma) \eta(X_{\sigma(0)}, \dots, X_{\sigma(k-1)}) \cdot (d_\nabla v)(X_{\sigma(k)})
$$

With the condition $\sigma(0) < \ldots < \sigma(k-1)$, we see that

$$
\begin{aligned}
&\sum_{\sigma \in \bigwedge_{k,1}} \operatorname{sgn}(\sigma) \eta(X_{\sigma(0)}, \dots, X_{\sigma(k-1)}) \cdot (d_\nabla v)(X_{\sigma(k)}) \\
&\qquad = (-1)^{\deg \eta} \sum_{i=0}^{k} (-1)^i \eta(X_0, \dots, \widehat{X_i}, \dots, X_k) \cdot (d_\nabla v)(X_i) \ ,
\end{aligned}
$$

which proves the claim.

ii) Choosing local frames $\{\xi^\alpha\}$, $\{e_\alpha\}$ and using corollary E.3.7, together with the last claim yields:

$$\begin{aligned} d_\nabla(\omega\wedge\eta) &= d_\nabla\left({\eta_\alpha}^\beta(\omega\wedge\xi^\alpha)\otimes e_\beta\right) = d\left(\omega\wedge{\eta_\alpha}^\beta\xi^\alpha\right)\otimes e_\beta \\ &\quad + (-1)^{k+\ell}\left(\omega\wedge{\eta_\alpha}^\beta\xi^\alpha\right)\wedge d_\nabla e_\beta \\ &= d\omega\wedge{\eta_\alpha}^\beta\xi^\alpha\otimes e_\beta + (-1)^k\omega\wedge d\left({\eta_\alpha}^\beta\xi^\alpha\right)\otimes e_\beta \\ &\quad + (-1)^{k+\ell}\left(\omega\wedge{\eta_\alpha}^\beta\xi^\alpha\right)\wedge d_\nabla e_\beta \\ &= d\omega\wedge\eta + (-1)^k\omega\wedge\left(d\left({\eta_\alpha}^\beta\xi^\alpha\right)\otimes e_\beta + (-1)^\ell{\eta_\alpha}^\beta\xi^\alpha\wedge d_\nabla e_\beta\right) \\ &= d\omega\wedge\eta + (-1)^k\omega\wedge d_\nabla\eta\ . \end{aligned}$$

□

The theory of vector forms allows us to reinterpret the meaning of curvature. Recalling definition C.5.1, a curvature endomorphism w.r.t a connection ∇ on an arbitrary vector bundle is a map $R\colon \Gamma(E,M)\times\Gamma(E,M)\times\Gamma(E,M)\to\Gamma(E,M)$, defined by

$$R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z = [\nabla_X,\nabla_Y]Z - \nabla_{[X,Y]}Z\ .$$

We observe, that for fixed X and Y, $R(X,Y)$ defines an endomorphism $\mathrm{End}(E)\equiv\mathrm{Hom}(E,E)$. Also, it is easy to check, that $R(X,Y)$ is multilinear and alternating in X and Y. This is reason to define the following:

Definition E.4.8.
The **curvature form** $F^\nabla\in\Omega^2(M,\mathrm{End}(E))$ with respect to a connection ∇ is defined by

$$F^\nabla(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}\ .$$

Theorem E.4.9.
Let (E,M,π) be a vector bundle and ∇ a connection, the it holds that:

$$\left((d_\nabla\circ d_\nabla)Z\right)(X,Y) = F^\nabla(X,Y)Z \quad \forall\ X,Y,Z\in\Gamma(E,M)\ .$$

Proof E.4.10.
With $d_\nabla Z = \nabla_\bullet Z$ and the definition of d^1_∇ we see that:

$$\begin{aligned} (d_\nabla(d_\nabla Z))(X,Y) &= d_\nabla(d_\nabla Z(Y))(X) - d_\nabla(d_\nabla Z(X))(Y) - d_\nabla Z([X,Y]) \\ &= d_\nabla(\nabla_Y Z)(X) - d_\nabla(\nabla_X Z)(Y) - \nabla_{[X,Y]}Z \\ &= \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z \\ &= F^\nabla(X,Y)Z\ . \end{aligned}$$

□

Example E.4.11.
The trivial bundle $M \oplus \mathbb{R}$ is flat with respect to the naturally induced connection of the exterior derivative, i.e. $d \circ d = 0$.

E.5. Connection- and curvature forms

This section follows [Mor01, section 5.3 (c)] and is inspired by [RS18, sections 8.1 and 8.2] and [Wik18].

Lemma E.5.1.
*Let (E, M, π) be a vector bundle with local trivializations $\{U_\alpha, \varphi_\alpha\}$. Let $\{e_\beta\}$ be a local frame on U_α. For every connection ∇ there are differential forms $\omega^i_j \in \Omega^1(U_\alpha)$, called **connection differential forms for the frame $\{e_\beta\}$** such that*

$$d_\nabla e_j = \omega^i_j \otimes e_i \ .$$

Remark E.5.2.
So far, objects like ${\omega^i}_j$ have been functions. Here however, ω^i_j are proper differential 1-forms. To highlight that, we do not use the tensor notation of indented indices. But still, we keep the summation convention. It should also be mentioned, that the differential forms depend on the choice of local trivialization and the choice of the local frame.

Proof E.5.3.
Let $X \in \Gamma(E, M)$ be arbitrary. Since a connection ∇_X defines a fiber wise linear map, we can write

$$\nabla_X e_j = \sum_i f^i_j(X) \cdot e_i \ ,$$

where $f^i_j(X) \in C^\infty(U)$ are functions. To be connection however, for $g \cdot X$ the objects f^i_j have to be $C^\infty(U)$-linear, i.e. $f^i_j(g \cdot X) = g \cdot f^i_j(X)$. Thus f^i_j are linear maps $\Gamma(E|_U, U)$ and hence local differential 1-forms. Renaming $f^i_j = \omega^i_j$ and calculating

$$(d_\nabla e_j)(X) = \nabla_X e_j = \sum_i \omega^i_j(X) \cdot e_i = (\omega^i_j \otimes e_i)(X) \ ,$$

proves the claim. □

The connection forms do not satisfy the product rule. Hence, they are not enough to represent the connection. The next lemma shows, that symbolically $\nabla = d + \omega^i_j$ holds, where it is understood, that d does not act on the frame.

Lemma E.5.4.
On U_α, for the frame $\{e_\beta\}$ it holds that:

$$d_\nabla(f^j e_j) = df^j \otimes e_j + f^j \omega^i_j \otimes e_i \ .$$

Proof E.5.5.
Using theorem E.4.6 and the linearity of d_∇, it follows that:

$$d_\nabla(f^j e_j) = df^j \otimes e_j + f^j d_\nabla e_j = df^j \otimes e_j + f^j \omega^i_j \otimes e_i \ .$$

□

The connection forms ω^i_j can be understood as components of an endomorphism valued differential form $\omega_\alpha \in \Omega^1(U_\alpha, \mathrm{End}(\mathbb{R}^n))$. This can be realized with the isomorphy of $\mathrm{End}(\mathbb{R}^n) = \mathbb{R}^n \otimes (\mathbb{R}^n)^*$. Let $\{\vartheta^\beta\}$ be the dual frame of $\{e_\beta\}$, then on U_α, the form ω_α can be written as:

$$\omega_\alpha = \omega^i_j \otimes \left(e_i \otimes \vartheta^j\right) \in \Omega^1(U_\alpha, \mathrm{End}(\mathbb{R}^n)) \ .$$

To avoid confusion between the tensor products, we define

$$\mathrm{End}(\mathbb{R}^n) \ni L^j_i = e_i \otimes \vartheta^j \qquad \rightsquigarrow \qquad \omega_\alpha = \omega^i_j \otimes L^j_i \ .$$

Going on with the definitions, we set $\mathbb{d} = d \otimes \mathrm{Id}_{\mathrm{End}(\mathbb{R}^n)}$, we find the following corollary:

Corollary E.5.6.
Locally, i.e. on every local trivialization set U_α for every local frame $\{e_\beta\}$, a connection d_∇ can be written as

$$d_\nabla = \mathbb{d} + \omega_\alpha \ .$$

Proof E.5.7.
With

$$\omega_\alpha e_j = \omega^i_\ell \otimes \left(e_i \otimes \vartheta^\ell\right)(e_j) = \delta_{\ell j} \cdot \omega^i_\ell \otimes e_i = \omega^i_j \otimes e_i \ ,$$

the claim follows directly form the calculation in proof E.5.5. □

Remark E.5.8.
In the following, we may identify $\mathbb{d}$ with d, understanding how d acts on vector valued forms by $\mathbb{d}$ and thus is well defined.

Definition E.5.9.
The **curvature differential forms** $F^i_j \in \Omega^2(U_\alpha)$ of the curvature vector form

F^{∇} on the local trivialization U_α for the local frame $\{e_\beta\}$ are differential 2-forms defined by

$$F^{\nabla}(X,Y)e_j = F^i_j(X,Y)e_i \ .$$

That F^i_j are proper 2-forms can be proven similarly as in proof E.5.3, using the properties of the curvature form F^{∇}. Since $F^{\nabla} \in \Omega^2(M, \mathrm{End}(E))$ already, we know, that $F^{\nabla}_\alpha \in \Omega^2(U_\alpha, \mathrm{End}(\mathbb{R}^n))$ exists on the local trivialization. It remains to connect F^{∇}_α to the forms F^i_j. On the local trivialization set U_α, for the local frame $\{e_\beta\}$ with dual frame $\{\vartheta^\beta\}$, we can write:

$$F^{\nabla}_\alpha = F^i_j \otimes L^j_i \ , \quad \text{where} \quad L^j_i = e_i \otimes \vartheta^j \ .$$

The choice of F^i_j together with L^j_i leads to:

$$F^{\nabla}_\alpha e_j = F^i_\ell \otimes L^\ell_i e_j = \delta_{j\ell} F^i_\ell \otimes e_i = F^i_j \otimes e_j \quad \Rightarrow \quad F^{\nabla}_\alpha(X,Y)e_j = F^i_j(X,Y)e_i \ .$$

Thus, F^i_j are indeed the coefficients of the curvature form F^{∇}_α on the local trivialization, for the local frame $\{e_\beta\}$. This result can also be taken as proof for the F^i_j being proper differential 2-forms.

Theorem E.5.10.
The curvature differential forms are related to the connection differential forms by the ***structure equation****:*

$$d\omega^i_j = -\omega^i_k \wedge \omega^k_j + F^i_j \ .$$

Proof E.5.11.
With the invariant definition of the exterior derivative, we calculate:

$$d\omega^i_j(X,Y) = X(\omega^i_j(Y)) - Y(\omega^i_j(X)) - \omega^i_j([X,Y]) \ .$$

By definition we have:

$$\omega^i_k \wedge \omega^k_j(X,Y) = \omega^i_k(X)\omega^k_j(Y) - \omega^i_k(Y)\omega^k_j(X) \ .$$

Last, a direct calculation, using $\nabla_X e_j = \omega^i_j(X)e_i$ etc., shows that

$$\begin{aligned} F^i_j(X,Y)e_i = F^{\nabla}(X,Y)e_j &= \Big(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}\Big)\, e_j \\ &= \Big(X(\omega^i_j(Y)) + \omega^i_k(X)\omega^k_j(Y) - Y(\omega^i_j(X)) \\ &\qquad -\omega^i_k(Y)\omega^k_j(X) - \omega^i_j([X,Y])\Big)\, e_i \\ &= \Big(d\omega^i_j(X,Y) + \omega^i_k \wedge \omega^k_j(X,Y)\Big)\, e_i \ . \end{aligned}$$

That is: $d\omega^i_j(X,Y) + \omega^i_k \wedge \omega^k_j(X,Y) = F^i_j(X,Y)$, which proves the claim. □

Theorem E.5.12.
Let $(U_\alpha, \varphi_\alpha)$ be a local trivialization. For the endomorphism valued forms on the local trivialization, ***structure equation*** *takes the form*

$$F_\alpha^\nabla = d\omega_\alpha + \omega \wedge_\circ \omega \ .$$

The wedge product $\wedge_\circ$ is defined w.r.t. the natural composition $\circ\colon End(\mathbb{R}^n) \oplus End(\mathbb{R}^n) \to End(\mathbb{R}^n)$ of endomorphisms.

Proof E.5.13.
Choosing a local frame $\{e_\beta\}$ together with dual frame $\{\vartheta^\beta\}$ on U_α, allows to write $F_\alpha^\nabla = F_j^i \otimes L_i^j$ and $\omega_\alpha = \omega_j^i \otimes L_i^j$. As remarked above, we understand d to be $\mathbb{d}$, i.e. $d\omega_\alpha = (d\omega_j^i) \otimes L_i^j$. With $Q(L_i^k, L_\ell^j) = L_i^k \circ L_\ell^j = \delta_{k\ell} L_i^j$ and corollary E.3.7 (in a slightly more general form) we find:

$$\begin{aligned}\omega_\alpha \wedge_\circ \omega_\alpha &= (\omega_k^i \otimes L_i^k) \wedge_\circ (\omega_j^\ell \otimes L_\ell^j) = \omega_k^i \wedge \omega_j^\ell \otimes L_i^k \circ L_\ell^j = \omega_k^i \wedge \omega_j^\ell \otimes \delta_{k\ell} L_i^j \\ &= \omega_k^i \wedge \omega_j^k \otimes L_i^j \ .\end{aligned}$$

The rest follows from theorem E.5.10. □

Index

A

absolute time . . . 1
active transformation . . . 7

B

base space . . . 58
$\vec{\beta}$-factor . . . 14
Bianchi identity
  -algebraic Bianchi identity . . . 87
  -differential Bianchi identity . . . 87
boost transformation . . . 6
bundle isomorphism . . . 61
bundle map . . . 61

C

Cartan-formula . . . 70
Christoffel symbols . . . 58, 63
cocycle condition . . . 60
complexification . . . 45
connection . . . 61
  -compatible with metric . . . 74
  -connection differential forms . . . 110
  -definition with vector forms . . . 106
  -induced tensor connection . . . 64
  -linear connection . . . 63
  -symmetric linear connection . . . 75
contraction of tensor . . . 67
contravariant . . . 14, 47
covariant . . . 14, 47
covariant curve derivative . . . 79
covariant derivative . . . 64
covariant divergence . . . 97
covariant Hessian . . . 30
curvature
  -curvature differential forms . . . 111
  -curvature form . . . 109

D

density
  -mathematical . . . 95
  -physics . . . 96
  -pseudo density . . . 96
dual bundle . . . 99

E

Einstein field equations . . . 23
  -lineraization . . . 29
Einstein tensor . . . 26
Einstein-Hilbert action . . . 25
Einstein-Hilbert-Matter action . . . 26
energy-momentum tensor . . . 23
Energy-momentum-relation . . . 18
equivalence principal . . . 24
Euler-Lagrange-equations . . . 55
  -covariant . . . 98
event . . . 1
exterior covariant derivative . . . 107

F

fiber . . . 58
flat isomorphism . . . 49
flat manifold . . . 86
flatness condition . . . 85
four acceleration . . . 15
four force . . . 17
four momentum . . . 17
four position . . . 14
four velocity . . . 15
Fréchet-derivative . . . 52
Fréchet-differentiable . . . 52
functional . . . 51
functional derivative . . . 52
  - in physics . . . 56
fundamental lemma of Riemannian geometry . . . 76

G

Gâteaux-derivative . . . 51
Gâteaux-differentiable . . . 51
Galilean spacetime . . . 1
Galileo group . . . 2
γ-factor . . . 7
Gauss' law for gravity . . . 3
geodesic . . . 80
  -Riemannian geodesic . . . 82
geodesic equation . . . 80

H

harmonic coordinates 30
Homomorphism bundle 100

I

inertial system 5
instantaneous rest frame 16
integral curve 68
interior product 104
intrinsic curvature 85
isometry 77

J

Jacobi's formula 93

K

Killing equation 90
Killing field 89

L

Lagrange-density 53
-matter Lagrange-density 26
Laplace-Beltrami operator 30
Levi-Civita-connection 76
-naturality 77
Lie-derivative 69
light cone 9
light like 12
local flow 69
local trivialization 58
Lorentz contraction 11
Lorentz group 6
Lorentz transformations 6

M

mass density 3
metric determinant 92
Minkowski diagram 9
Minkowski metric 5
Minkowski space 5

N

Newton's shell theorem 4
Newtonian gravitational field 2
normal neighborhood 83
null vector 23

P

parallel along curve 80
parallel translate 80
parallel transport 80
passive transformation 7
Poincaré group 6
principle of relativity 5
proper length 10
proper mass 17
proper time 10, 18
proper time differential 20
proper time interval 20
pseudo Riemannian metric 59
pullback 104
pullback bundle 100

R

radiation coordinates 33
rapidity 8
relativistic mass 18
relativistic momentum 18
rest frame 10
retarded time 35
Ricci tensor 88
Riemann curvature tensor 85
-covariant Riemann tensor
.... 86
Riemannian curvature endomorphism
85
Riemannian metric 59
Riemannian normal coordinates
.... 83

S

scalar curvature 88
Schwarzschild metric 37
Schwarzschild radius 37
section 59
sharp isomorphism 49
space like 12
structure equation 112
structure group 60

T

tensor
-of type (r, s) 48
-pure tensor 44
-rank 48
-tensor product 40

-universal propertyt 40
tensor bundle 100
tidal force 24
time dilation 11
time like 11
Torsion tensor field 75
total space 58
trace reversed perturbation 31
transition function 60

V

Variation of a functional 51
vector bundle 58
vector bundle projection 58
vector bundle valued differential-k form 101
vector field along curve 78
-extendible 78
vector form 101

W

weak equivalence principal 24
wedge product 103
Whitney bundle 100
world line 14

Bibliography

[AS10] A. Altland and B.D. Simons. *Condensed Matter Field Theory*. Cambridge books online. Cambridge University Press, 2010.

[Bau14] H. Baum. *Eichfeldtheorie: Eine Einführung in die Differentialgeometrie auf Faserbündeln*. Masterclass. Springer Berlin Heidelberg, 2014.

[Car97] S. M. Carroll. *Lecture Notes on General Relativity*. Ed. by S. M. Carroll. 1997. URL: https://arxiv.org/pdf/gr-qc/9712019.pdf.

[Fli16] T. Fließbach. *Allgemeine Relativitätstheorie*. Springer Berlin Heidelberg, 2016.

[HE75] S.W. Hawking and G.F.R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1975.

[HO07] S. Holtz and B. Obst. *Tensorprodukt*. Ed. by S. Holtz and B. Obst. Vorlesungsskript. 2007. URL: https://www-irm.mathematik.hu-berlin.de/~teschke/commalg/Tensor.pdf.

[Jän05] K. Jänich. *Vektoranalysis*. Springer-Lehrbuch. Springer, 2005.

[KM13] I. Kolar and P.W. Michor. *Natural Operations in Differential Geometry*. Springer Berlin Heidelberg, 2013.

[Lee97] J.M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Graduate Texts in Mathematics. Springer New York, 1997.

[Mor01] S. Morita. *Geometry of Differential Forms*. Iwanami series in modern mathematics. American Mathematical Society, 2001.

[RS18] J. W. Robbin and D. A. Salamon. *Introduction ro differential topology*. Ed. by J. W. Robbin and D. A. Salamon. 2018. URL: https://people.math.ethz.ch/~salamon/PREPRINTS/difftop.pdf.

[RW05] M. Roczen and H. Wolter. “Lineare Algebra individuell, Kapitel 4”. In: *Lineare Algebra individuell*. 2005. URL: https://www.math.hu-berlin.de/~roczen/papers/tensor.pdf.

[SW12] R.K. Sachs and H.H. Wu. *General Relativity for Mathematicians*. Graduate Texts in Mathematics. Springer New York, 2012.

[Wer11] D. Werner. *Funktionalanalysis*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2011.

[Wik18] Wikipedia. *Vector-valued differential form*. Ed. by Wikipedia. 2018. URL: https://en.wikipedia.org/wiki/Vector-valued_differential_form.

[Win07] S. Winitzki. *Advanced General Relativity*. Ed. by S. Winitzki. 2007. URL: https://sites.google.com/site/winitzki/index/topics-in-general-relativity.

[Woo16] N. Woodhouse. *Spezielle Relativitätstheorie*. Masterclass. Springer Berlin Heidelberg, 2016.

[Zir98] M. Zirnbauer. “Elektrodynamik”. 1998. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/06Elektrodynamik_SS98.pdf.

[Zir15a] M. Zirnbauer. *Klassische Theoretische Physik I*. Ed. by M. Zirnbauer. 2015. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/KTP1_SS2015.pdf.

[Zir15b] M. Zirnbauer. *Klassische Theoretische Physik II*. Ed. by M. Zirnbauer. 2015–2016. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/KTP2_WS1516.pdf.