Riemannian Geometry

Fall Semester 2016
Version 0.98

2017-03-06

Michael Kunzinger
Roland Steinbauer

University of Vienna
Faculty of Mathematics
Oskar-Morgenstern-Platz 1
A-1090 Wien
AUSTRIA

Preface

These lecture notes are based on the lecture course "Differentialgeometrie 2" taught by M.K. in the fall semesters of 2008 and 2012. The material has been slightly reorganised to serve as a script for the course "Riemannian geometry" by R.S. in the fall term 2016. It can be considered as a continuation of the lecture notes "Differential Geometry 1" of M.K. [10] and we will extensively refer to these notes.

Basically this is a standard introductory course on Riemannian geometry which is strongly influenced by the textbook "Semi-Riemannian Geometry (With Applications to Relativity)" by Barrett O'Neill [12]. The necessary prerequisites are a good knowledge of basic differential geometry and analysis on manifolds as is traditionally taught in a 3–4 hours course.

M.K., R.S., September 2016

Note added by R.S: This version (numbered 0.98) has still to be considered pre-final for the following two reasosn: some of the figures are missing or need improvement and some of the introductory and motivating texts need some editing. The definitions, results and proofs, however, are in a reasonably final form.

R.S., March 2017

Contents

Chapter 1

Semi-Riemannian Manifolds

In classical/elementary differential geometry of hypersurfaces in $\mathbb{R}^n$ and, in particular, of surfaces in $\mathbb{R}^3$ one finds that all intrinsic properties of the surface ultimately depend on the scalar product induced on the tangent spaces by the standard scalar product of the ambient Euclidean space. Our first goal is to generalise the respective notions of length, angle, curvature and the like to the setting of abstract manifolds. We will, however, allow for nondegenerate bilinear forms which are not necessarily positive definite to include central applications, in particular, general relativity. We will start with an account on such bilinear forms.

1.1 Scalar products

Contrary to basic linear algebra where one typically focusses on positive definite scalar products semi-Riemannian geometry uses the more general concept of nondegenerate bilinear forms. In this subsection we develop the necessary algebraic foundations.

1.1.1 Definition (Bilinear forms). *Let V be a finite dimensional vector space. A* bilinear form *on V is an $\mathbb{R}$-bilinear mapping $b: V \times V \to \mathbb{R}$. It is called* symmetric *if*

$$b(v, w) = b(w, v) \quad \textit{for all } v, w \in V. \tag{1.1.1}$$

A symmetric bilinear form is called

(i) Positive (negative) definite, *if $b(v, v) > 0$ (< 0) for all $0 \neq v \in V$,*

(ii) Positive (negative) semidefinite, *if $b(v, v) \geq 0$ (≤ 0) for all $v \in V$,*

(iii) nondegenerate, *if $b(v, w) = 0$ for all $w \in V$ implies $v = 0$.*

Finally we call b (semi)definite *if one of the alternatives in (i) (resp. (ii)) hold true. Otherwise we call b* indefinite.

In case b is definite it is semidefinite and nondegenerate and conversely if b is semidefinite and nondegenerate it is already definite. Indeed in the positive case suppose there is $0 \neq v \in V$ with $b(v,v) = 0$. Then for arbitrary $w \in V$ we find that

$$b(v+w, v+w) = \underbrace{b(v,v)}_{0} + 2b(v,w) + b(w,w) \geq 0 \tag{1.1.2}$$

$$b(v-w, v-w) = \underbrace{b(v,v)}_{0} - 2b(v,w) + b(w,w) \geq 0, \tag{1.1.3}$$

since b is positive semidefinite and so $2|b(v,w)| \leq b(w,w)$. But replacing w by λw with λ some positive number we obtain

$$2\ |b(v,w)| \leq \lambda\ b(w,w) \tag{1.1.4}$$

and since we may choose λ arbitrarily small we have $b(v,w) = 0$ for all w which by nondegeneracy implies $v = 0$, a contradiction.

If b is a symmetric bilinear form on V and if W is a subspace of V then clearly the restriction $b|_W$ (defined as $b|_{W\times W}$) of b to W is again a symmetric bilinear form. Obviously if b is (semi)definite then so is $b|_W$.

1.1.2 Definition (Index). *We define the* index r *of a symmetric bilinear form b on V by*

$$r := \max\{\dim W |\ W \textit{ subspace of } V \textit{ with } b|_W \textit{ negative definite}\}. \tag{1.1.5}$$

By definition we have $0 \leq r \leq \dim V$ and $r = 0$ iff b is positive definite.

Given a symmetric bilinear form b we call the function

$$q:\ V \to \mathbb{R}, \quad q(v) = b(v,v) \tag{1.1.6}$$

the *quadratic form associated with b*. Frequently it is more convenient to work with q than with b. Recall that by polarisation $b(v,w) = 1/2(q(v+w) - q(v) - q(w))$ we can recover b from q and so all the information of b is also encoded in q.

Let $\mathcal{B} = \{e_1, \dots, e_n\}$ be a basis of V, then

$$(b_{ij}) := (b(e_i, e_j))_{i,j=1}^n \tag{1.1.7}$$

is called the *matrix of b* with respect to $\mathcal{B}$. It is clearly symmetric and entirely determines b since $b(\sum v_i\, e_i, \sum w_j\, e_j) = \sum b_{ij} v_i w_j$. Moreover nondegeneracy of b is characterised by its matrix (w.r.t. any basis):

1.1.3 Lemma. *A symmetric bilinear form is nondegenerate iff its matrix w.r.t. one (and hence any) basis is invertible.*

Proof. Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis of V. Given $v \in V$ we have $b(v,w) = 0$ for all w iff $0 = b(v, e_i) = b(\sum_j v_j e_j, e_i) = \sum_j b_{ij} v_j$ for all $1 \leq i \leq n$. So b is degenerate iff there are $(v_1, \ldots, v_n) \neq (0, \ldots, 0)$ with $\sum_j b_{ij} v_j = 0$ for all i. But this means that the kernel of (b_{ij}) is non trivial and (b_{ij}) is singular. □

We now introduce a terminology slightly at odds with linear algebra standards but which reflects our interest in non positive definite symmetric bilinear forms.

1.1.4 Definition (scalar product, inner product). *A* scalar product g *on a vector space* V *is a nondegenerate symmetric bilinear form. An* inner product *is a positive definite scalar product.*

1.1.5 Example.

(i) *The* example of an inner product is the *standard scalar product* of Euclidean space $\mathbb{R}^n$: $v \cdot w = \sum_i v_i w_i$.

(ii) The most simple example of a vector space with indefinite scalar product is *two-dimensional Minkowski space* $\mathbb{R}^2$ with

$$g = \eta : \ \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \qquad g(v,w) = -v_1 w_1 + v_2 w_2. \tag{1.1.8}$$

Obviously g is bilinear and symmetric. To see that it is nondegenerate suppose that $g(v,w) = 0$ for all $w \in \mathbb{R}^2$. Setting $w = (1,0)$ and $w = (0,1)$ gives $v_1 = 0$ and $v_2 = 0$, respectively and so $v = 0$. Hence η is a scalar product but it is not an inner product since it is indefinite:

$$g\big((1,0),(1,0)\big) = -1 < 0, \quad \text{but} \quad g\big((0,1),(0,1)\big) = 1 > 0. \tag{1.1.9}$$

The corresponding quadratic form is $q(v) = -v_1^2 + v_2^2$.

In the following V will always be a (finite dimensional, real) vector space with a scalar product g in the sense of 1.1.4. A vector $0 \neq v \in V$ with $q(v) = 0$ will be called a *null vector*. Such vectors exists iff g is indefinite. Note that the zero vector 0 is *not* a null vector.

1.1.6 Example. We consider the lines $q = c$ and $q = -c$ ($c > 0$) in two-dimensional Minkowski space of Example 1.1.5(ii). They are either hyperbolas or straight lines in case $c = 0$, see Figure 1.1.

A pair of vectors $u, w \in V$ is called *orthogonal*, $u \perp w$, if $g(u,w) = 0$. Analogously we call subspaces U, W of V orthogonal, if $g(u,w) = 0$ for all $u \in U$ and all $w \in W$. **Warning:** In case of indefinite scalar products vectors that are orthogonal need not to be at right angles to one another as the following example shows.

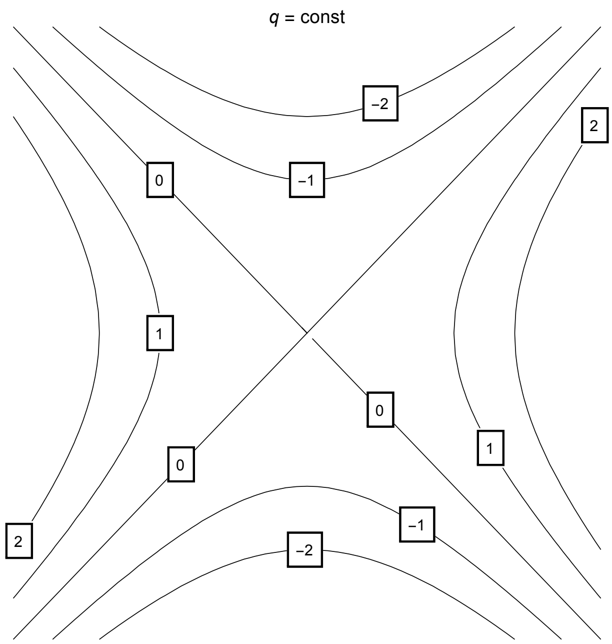

Figure 1.1: Contours of q in 2-dimensional Minkowski space

1.1.7 Example. The following pairs of vectors v, v' are orthogonal in two-dimensional Minkowski space, see Figure 1.2: $w = (1, 0)$ and $w' = (0, 1)$, $u = (1, b)$ and $u' = (b, 1)$ for some $b > 0$, $z = (1, 1) = z'$.

The example of the vectors z, z' above hints at the fact that null vectors are precisely those vectors that are orthogonal to themselves.
If W is a subspace of V let

$$W^\perp := \{v \in V : \ v \perp w \quad \text{for all } w \in W\}. \tag{1.1.10}$$

Clearly $W^\perp$ is a subspace of V which we call W *perp*. **Warning:** We cannot call $W^\perp$ the orthogonal complement of W since in general $W + W^\perp \neq V$, e.g. if $W = \text{span}(z)$ in Example 1.1.7 we even have $W^\perp = W$. However $W^\perp$ has two familiar properties.

1.1.8 Lemma. *Let W be a subspace of V. Then we have*

(i) $\dim W + \dim W^\perp = \dim V$,

(ii) $(W^\perp)^\perp = W$.

Proof.

(i) Let $\{e_1, \ldots, e_k\}$ be a basis of W which we extend to a basis $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ of V. Then we have

$$v \in W^\perp \Leftrightarrow g(v, e_i) = 0 \quad \text{for } 1 \leq i \leq k \ \Leftrightarrow \ \sum_{j=1}^{n} g_{ij} v_j = 0 \quad \text{for } 1 \leq i \leq k. \tag{1.1.11}$$

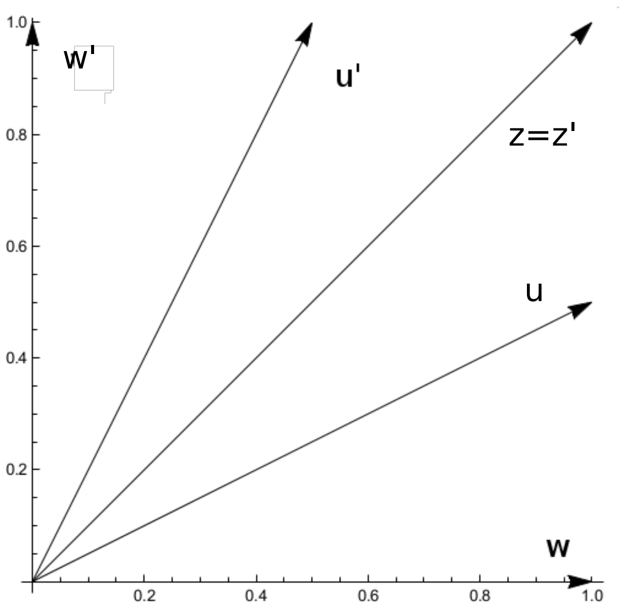

Figure 1.2: Pairs of orthogonal vectors in 2-dimensional Minkowski space

Now by Lemma 1.1.3 (g_{ij}) is invertible and hence the rows in the above linear system of equations are linearly independent and its space of solutions has dimension $n-k$. So $\dim W^\perp = n-k$.

(ii) Let $w \in W$, then $w \perp W^\perp$ and $w \in (W^\perp)^\perp$, which implies $W \subseteq (W^\perp)^\perp$. Moreover, by (i) we have $\dim W = \dim(W^\perp)^\perp = k$ and so $W = (W^\perp)^\perp$. □

A symmetric bilinear form g on V is nondegenerate, iff $V^\perp = \{0\}$. A subspace W of V is called *nondegenerate*, if $g|_W$ is nondegenerate. If g is an inner product, then any subspace W is again an inner product space, hence nondegenerate. If g is indefinite, however, there always exists degenerate subspaces, e.g. $W = \mathrm{span}(w)$ for any null vector w. Hence a subspace W of a vector space with scalar product in general is *not* a vector space with scalar product. Indeed W could be degenerate. We now give a simple characterisation of nondegeneracy for subspaces.

1.1.9 Lemma (Nondegenerate subspaces). *A subspace W of a vector space V with scalar product is nondegenerate iff*

$$V = W \oplus W^\perp. \tag{1.1.12}$$

Proof. By linear algebra we know that

$$\dim(W + W^\perp) + \dim(W \cap W^\perp) = \dim W + \dim W^\perp = \dim V. \tag{1.1.13}$$

So equation (1.1.12) holds iff

$$\dim(W \cap W^\perp) = 0 \;\Leftrightarrow\; \{0\} = W \cap W^\perp = \{w \in W : w \perp W\}$$

which is equivalent to the nondegeneracy of W. □

As a simple consequence we obtain by using 1.1.8(ii), i.e., $(W^\perp)^\perp = W$.

1.1.10 Corollary. *W is nondegenerate iff $W^\perp$ is nondegenerate.*

Our next objective is to deal with *orthonormal bases* in a vector space with scalar product. To begin with we define unit vectors. However, due to the fact that q may take negative values we have to be a little careful.

1.1.11 Definition (Norm). *We define the* norm *of a vector $v \in V$ by*

$$|v| := |g(v,v)|^{\frac{1}{2}}. \tag{1.1.14}$$

A vector $v \in V$ is called a unit vector *if $|v| = 1$, i.e., if $g(v,v) = \pm 1$. A family of pairwise orthogonal unit vectors is called* orthonormal.

Observe that an orthonormal system of $n = \dim V$ elements automatically is a basis. The existence of orthonormal bases (ONB) is guaranteed by the following statement.

1.1.12 Lemma (Existence of orthonormal bases). *Every vector space $V \neq \{0\}$ with scalar product possesses an orthonormal basis.*

Proof. There exists $v \neq 0$ with $g(v,v) \neq 0$, since otherwise by polarisation we would have $g(v,w) = 0$ for all pairs of vectors v, w, which implies that g is degenerate. Now $v/|v|$ is a unit vector and it suffices to show that any orthonormal system $\{e_1, \dots, e_k\}$ can be extended by one vector.
So let $W = \text{span}\{e_1, \dots, e_k\}$. Then by Lemma 1.1.3 W is nondegenerate and so is $W^\perp$ by corollary 1.1.10. Hence by the argument given above $W^\perp$ contains a unit vector e_{k+1} which extends $\{e_1, \dots, e_k\}$. □

The matrix of g w.r.t. any ONB is diagonal, more precisely

$$g(e_i, e_j) = \delta_{ij}\varepsilon_j, \quad \text{where } \varepsilon_j := g(e_j, e_j) = \pm 1. \tag{1.1.15}$$

In the following we will always order any ONB $\{e_i, \dots, e_n\}$ in such a way that in the so-called *signature* $(\varepsilon_1, \dots, \varepsilon_n)$ the negative signs come first. Next we give the representation of a vector w.r.t an ONB. Once again we have to be careful about the signs.

1.1.13 Lemma. *Let* $\{e_1, \dots, e_n\}$ *be an ONB for* V*. Then any* $v \in V$ *can be uniquely written as*

$$v = \sum_{i=1}^{n} \varepsilon_i \, g(v, e_i) \, e_i. \tag{1.1.16}$$

Proof. We have that

$$\langle v - \sum_i \varepsilon_i g(v, e_i) e_i, e_j \rangle = \langle v, e_j \rangle - \sum_i \varepsilon_i \, \langle v, e_i \rangle \, \underbrace{\langle e_i, e_j \rangle}_{\varepsilon_i \delta_{ij}} = 0 \tag{1.1.17}$$

for all j and so by nondegeneracy $v = \sum_i \varepsilon_i g(v, e_i) e_i$. Uniqueness now simply follows since $\{e_1, \dots, e_n\}$ is a basis. □

If a subspace W is nondegenerate we have by Lemma 1.1.9 that $V = W \oplus W^\perp$. Let now π be the orthogonal projection of V onto W. Since any ONB $\{e_1, \dots, e_k\}$ of W can be extended to an ONB of V (cf. the proof of 1.1.12) we have for any $v \in V$

$$\pi(v) = \sum_{j=1}^{k} \varepsilon_j \, g(v, e_j) \, e_j. \tag{1.1.18}$$

Next we give a more vivid description of the index r of g (see Definition 1.1.2), which we will also call the index of V and denote it by ind V

1.1.14 Proposition (Index and signature). *Let* $\{e_1, \dots, e_n\}$ *be any ONB of* V*. Then the index of* V *equals the number of negative signs in the signature* $(\varepsilon_1, \dots, \varepsilon_n)$*.*

Proof. Let exactly the first m of the ε_i be negative. In case g is definite we have $m = r = 0$ or $m = r = n = \dim V$ and we are done.
So suppose $0 < m < n$. Obviously g is negative definite on $S = \mathrm{span}\{e_1, \dots, e_m\}$ and so $m \leq r$.
To show the converse let W be a subspace with $g|_W$ negative definite and define

$$\pi : W \to S, \quad \pi(w) := -\sum_{i=1}^{m} g(w, e_i) e_i. \tag{1.1.19}$$

Then π is obviously linear and we will show below that it is injective. Then clearly $\dim W \leq \dim S$ and since W was arbitrary $r \leq \dim S = m$.
Finally π is injective since if $\pi(w) = 0$ then by Lemma 1.1.13 $w = \sum_{i=m+1}^{n} g(w, e_i) e_i$. Since $w \in W$ we also have $0 \geq g(w, w) = \sum_{i=m+1}^{n} g(w, e_i)^2$ which implies $g(w, e_j) = 0$ for all $j > m$. But then $w = 0$. □

The index of a nondegenerate subspace can now easily be related to the index of V.

1.1.15 Corollary. *Let W be a nondegenerate subspace of V. Then*

$$\text{ind}\, V = \text{ind}\, W + \text{ind}\, W^{\perp}. \tag{1.1.20}$$

Proof. Let $\{e_1, \ldots, e_k\}$ be an ONB of W and $\{e_{k+1}, \ldots, e_n\}$ be an ONB of $W^{\perp}$ such that $\{e_1, \ldots, e_n\}$ is an ONB of V, cf. the proof of 1.1.12. Now the assertion follows from Proposition 1.1.14. □

To end this section we will introduce *linear isometries.* Let (V_1, g_1) and (V_2, g_2) be vector spaces with scalar product.

1.1.16 Definition (Linear isometry). *A linear map $T : V_1 \to V_2$ is said to* preserve the scalar product *if*

$$g_2(Tv, Tw) = g_1(v, w). \tag{1.1.21}$$

A linear isometry *is a linear bijective map that preserves the scalar product.*

In case there is no danger of missunderstanding we will write equation (1.1.21) also as

$$\langle Tv, Tw\rangle = \langle v, w\rangle. \tag{1.1.22}$$

If equation (1.1.21) holds true then T automatically preserves the associated quadratic forms, i.e., $q_2(Tv) = q_1(v)$ for all $v \in V$. The converse assertions clearly holds by polarisation.
A map that preserves the scalar product is automatically *injective* since $Tv = 0$ by virtue of (1.1.21) implies $g_1(v, w) = 0$ for all w and so $v = 0$ by nondegeneracy. Hence a linear mapping is an isometry iff $\dim V_1 = \dim V_2$ and equation (1.1.21) holds. Moreover we have the following characterisation.

1.1.17 Proposition (Linear isometries). *Let (V_1, g_1) and (V_2, g_2) be vector spaces with scalar product. Then the following are equivalent:*

(i) $\dim V_1 = \dim V_2$ *and ind* $V_1 =$ *ind* V_2

(ii) There exists a linear isometry $T : V_1 \to V_2$

Proof. (i)⇒(ii): Choose ONBs $\{e_1, \ldots, e_n\}$ of V_1 and $\{e'_1, \ldots, e'_n\}$ of V_2. By Proposition 1.1.14 we may assume that $\langle e_i, e_i\rangle = \langle e'_i, e'_i\rangle$ for all i. Now we define a linear map T via $Te_i = e'_i$. Then clearly $\langle Te_i, Te_j\rangle = \langle e_i, e_j\rangle$ for all i, j and T is an isometry.
(ii)⇒(i): If T is an isometry then $\dim V_1 = \dim V_2$ and T maps any ONB of V_1 to an ONB of V_2. But then equation (1.1.21) and Proposition 1.1.14 imply that ind $V_1 =$ ind V_2. □

1.2 Semi-Riemannian metrics

In this section we start our program to transfer the setting of elementary differential geometry to abstract manifolds. The key element is to equip each tangent space with a scalar product that varies smoothly on the manifold. We start right out with the central definition.

1.2.1 Definition (Metric). *A semi-Riemannian* metric tensor *(or* metric*, for short) on a smooth manifold*[1] *M is a smooth, symmetric and nondegenerate* $(0,2)$*-tensor field g on M of constant index.*

In other words g smoothly assigns to each point $p \in M$ a symmetric nondegenerate bilinear form $g(p) \equiv g_p : T_pM \times T_pM \to \mathbb{R}$ such that the index r_p of g_p is the same for all p. We call this common value r_p the *index* r of the metric g. We clearly have $0 \leq r \leq n = \dim M$. In case $r = 0$ all g_p are inner products on T_pM and we call g a *Riemannian metric*, cf. [10, 3.1.14]. In case $r = 1$ and $n \geq 2$ we call g a *Lorentzian metric.*

1.2.2 Definition. *A* semi-Riemannian manifold (SRMF) *is a pair* (M, g)*, where g is a metric on M. In case g is Riemannian or Lorentzian we call* (M, g) *a* Riemannian manifold (RMF) *or* Lorentzian manifold (LMF)*, respectively.*

We will often sloppily call just M a (S)RMF or LMF and write $\langle\ ,\ \rangle$ instead of g and use the following convention

- $g_p(v, w) = \langle v, w\rangle \in \mathbb{R}$ for vectors $v, w \in T_pM$ and $p \in M$, and
- $g(X, Y) = \langle X, Y\rangle \in \mathcal{C}^\infty(M)$ for vector fields $X, Y \in \mathfrak{X}(M)$.

If (V, φ) is a chart of M with coordinates $\varphi = (x^1, \ldots, x^n)$ and natural basis vector fields $\partial_i \equiv \frac{\partial}{\partial x^i}$ we write

$$g_{ij} = \langle \partial_i, \partial_j\rangle \qquad (1 \leq i, j \leq n) \tag{1.2.1}$$

for the local components of g on V. Denoting the dual basis covector fields of ∂_i by dx^i we have

$$g|_U = g_{ij}\, dx^i \otimes dx^j, \tag{1.2.2}$$

where we have used the summation convention (see [10, p. 54]) which will be in effect from now on.
Since g_p is nondegenerate for all p the matrix $(g_{ij}(p))$ is invertible by Lemma 1.1.3 and we write $(g^{ij}(p))$ for its inverse. By the inversion formula for matrices the g^{ij} are smooth functions on V and by symmetry of g we have $g^{ij} = g^{ji}$ for all i and j.

[1] In accordance with [10] we assume all smooth manifolds to be second countable and Hausdorff. For background material on topological properties of manifolds see [10, Section 2.3].

1.2.3 Example.

(i) We consider $M = \mathbb{R}^n$. Clearly $T_pM \cong \mathbb{R}^n$ for all points p and the standard scalar product induces a Riemannian metric on $\mathbb{R}^n$ which we denote by

$$\langle v_p, w_p\rangle = v \cdot w = \sum_i v^i\, w^i. \tag{1.2.3}$$

We will always consider $\mathbb{R}^n$ equipped with this Riemannian metric.

(ii) Let $0 \leq r \leq n$. Then

$$\langle v_p, w_p\rangle = -\sum_{i=1}^{r} v^i\, w^i + \sum_{j=r+1}^{n} v^j\, w^j \tag{1.2.4}$$

defines a metric on $\mathbb{R}^n$ of index r. We will denote $\mathbb{R}^n$ with this metric tensor by $\mathbb{R}^n_r$. Clearly $\mathbb{R}^n_0$ is $\mathbb{R}^n$ in the sense of (i). For $n \geq 2$ the space $\mathbb{R}^n_1$ is called n-dimensional *Minkowski space*. In case $n = 4$ this is the simplest spacetime in the sense of Einstein's general relativity. In fact it is the flat spacetime of special relativity.

Setting $\varepsilon_i = -1$ for $1 \leq i \leq r$ and $\varepsilon_i = 1$ for $r+1 \leq i \leq n$ the metric of $\mathbb{R}^n_r$ takes the form

$$g = \varepsilon_i\; dx^i \otimes dx^i = \varepsilon_i\; e^i \otimes e^i. \tag{1.2.5}$$

As is clear from section 1.1 a nonvanishing index allows for the existence of null vectors. Here we further pursue this line of ideas.

1.2.4 Definition (Causal character). *Let M be a SRMF, $p \in M$. We call $v \in T_pM$*

(i) spacelike *if $\langle v, v\rangle > 0$ or if $v = 0$,*

(ii) null *if $\langle v, v\rangle = 0$ and $v \neq 0$,*

(iii) timelike *if $\langle v, v\rangle < 0$.*

The above notions define the so-called *causal character* of v. The set of null vectors in T_pM is called the *null cone* at p respectively *light cone* at p in the Lorentzian case. In this case we also refer to null vectors as *lightlike* and call a vector *causal* if it is either timelike or lightlike.

1.2.5 Example. Let v be a vector in 2-dimensional Minkowski space $\mathbb{R}^2_1$. Then v is null iff $v_1^2 = v_2^2$, i.e., iff $v_1 = \pm v_2$, see also figure 1.3.

The above terminology is of course motivated by physics and relativity in particular. Setting the speed of light $c = 1$ then a flash of light emitted at the origin of Minkowski space travels along the light cone. A point is timelike if a signal with speed $v < 1$ can reach it and it is spacelike if it needs superluminal speed to reach it.

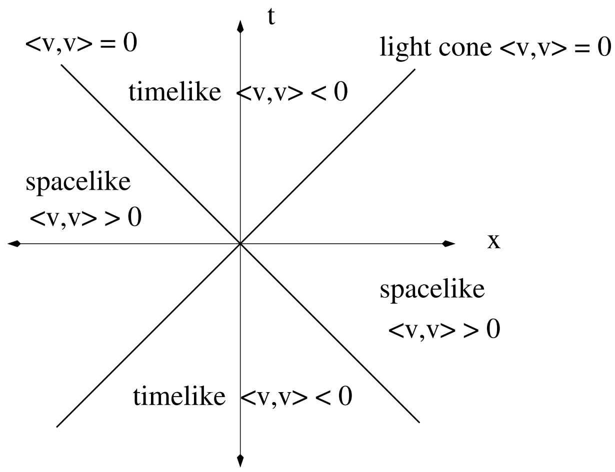

Figure 1.3: The lightcone in 2-dimensional Minkowski space

Let q be the quadratic form associated with g, i.e., $q(v) = \langle v, v\rangle$ for all $v \in T_pM$. Then by polarisation q determines the metric but it is *not* a tensor field since for $X \in \mathfrak{X}(M)$ and $f \in \mathcal{C}^\infty(M)$ we clearly have $q(fX) = f^2 q(X)$ (cf. [10, 2.6.19]). It is nevertheless frequently used in 'classical terminology' where it is called the *line element* and denoted by ds^2. Locally one writes

$$q = ds^2 = g_{ij}\, dx^i\, dx^j, \tag{1.2.6}$$

where juxtaposition of differentials means multiplication in each tangent space, that is

$$q(X) = g_{ij}\, dx^i(X)\, dx^j(X) = g_{ij} X^i X^j \tag{1.2.7}$$

for a vector field locally given by $X = X^i\partial_i$. Finally we write for the *norm* of a tangent vector

$$\|v\| := |q(v)|^{1/2} = |\langle v, v\rangle|^{1/2}. \tag{1.2.8}$$

1.2.6 Remark. The origin of the somewhat strange notation ds^2 is the following heuristic consideration: Consider two 'neighbouring points' p and p' with coordinates $(x^1, \dots, x^n)$ and $(x^1 + \triangle x^1, \dots, x^n + \triangle x^n)$. Then the 'tangent vector' $\triangle p = \sum \triangle x^i \partial_i$ at p points approximately to p' and so the 'distance' $\triangle s$ from p to p' should approximately be given by

$$\triangle s^2 = \|\triangle p\|^2 = \langle \triangle p, \triangle p\rangle = \sum_{i,j} g_{ij}(p) \triangle x^i \triangle x^j. \tag{1.2.9}$$

Let N be a submanifold of a RMF M with embedding $j : N \hookrightarrow M$. Then the pull back j^*g of the metric g to the submanifold N is given by (see [10, 2.7.24])

$$(j^*g)(p)(v, w) = g(j(p))(T_pjv, T_pjw) = g(p)(v, w), \tag{1.2.10}$$

where in the final equality we have identified $T_pj(T_pN)$ with T_pN, see [10, 3.4.11].

Hence $j^*g(p)$ is just the restriction of g_p to the subspace T_pN of T_pM. Since g is Riemannian this restriction is positive definite and so j^*g turns N into a RMF.
However, if M is only a SRMF then the $(0,2)$-tensor field j^*g on N need not be a metric. Indeed (cf. section 1.1) j^*g is a metric and hence (N, j^*g) a SRMF iff every T_pN is nondegenerate in T_pM and the index of T_pN is the same for all $p \in N$. Of course this index can be different from the index of g.
These considerations lead to the following definition.

1.2.7 Definition (Semi-Riemannian submanifold). *A submanifold N of a SRMF M is called a* semi-Riemannian submanifold (SRSMF) *if j^*g is a metric on N.*

If N is Riemannian or Lorentzian these terms replace semi-Riemannian in the above definition. However, note that while every SRSMF of a RMF is again Riemannian, a LMF can have Lorentzian as well as Riemannian submanifolds.

Finally we turn to isometries, i.e., diffeomorphisms that preserve the metric.

1.2.8 Definition. *Let (M, g_M) and (N, g_N) be SRMF and $\phi : M \to N$ be a diffeomorphism. We call ϕ an* isometry *if $\phi^* g_N = g_M$.*

Recall that the defining property of an isometry by [10, 2.7.24] in some more detail reads

$$\langle T_p\phi(v), T_p\phi(w)\rangle = g_N(\phi(p))\big(T_p\phi(v), T_p\phi(w)\big) = g_M(p)(v,w) = \langle v, w\rangle \qquad (1.2.11)$$

for all $v, w \in T_pM$ and all $p \in M$. Since ϕ is a diffeomorphism, every tangent map $T_p\phi : T_pM \to T_{\phi(p)}N$ is a linear isometry (cf. 1.1.16). Also note that $\phi^*g_N = g_M$ is equivalent to $\phi^*q_N = q_M$ since g is always uniquely determined by q. If there is an isometry between the SRMFs M and N we call them *isometric*.

1.2.9 Remark.

(i) Clearly id_M is an isometry. Moreover the inverse of an isometry is an isometry again and if ϕ_1 and ϕ_2 are isometries so is $\phi_1 \circ \phi_2$. Hence the isometries of M form a group, called the *isometry group* of M.

(ii) Given two vector spaces V and W with scalar product and a local isometry $\phi : V \to W$. If we consider V and W as SRMF, then ϕ is also an isometry of SRMF.

(iii) If V is a vector space with scalar product and $\dim V = n$, ind $V = r$ then V as a SRMF is isometric to $\mathbb{R}^n_r$. Just choose an ONB $\{e_1, \dots, e_n\}$ of V. Then the coordinate mapping $V \ni v = \sum v^i e_i \mapsto (v^1, \dots, v^n) \in \mathbb{R}^n$ is a linear isometry and (ii) proves the claim.

1.3 The Levi-Civita Connection

The aim of this chapter is to define on SRMFs a 'directional derivative' of a vector field (or more generally a tensor field) in the direction of another vector field . This will be done by generalising the *covariant derivative* on hypersurfaces of $\mathbb{R}^n$, see [10, Section 3.2] to general SRMFs. Recall that for a hypersurface M in $\mathbb{R}^n$ and two vector fields X, $Y \in \mathfrak{X}(M)$ the *directional derivative* $D_X Y$ of Y in direction of X is given by ([10, (3.2.3), (3.2.4)])

$$D_X Y(p) = (D_{X_p} Y)(p) = (X_p(Y^1), \dots, X_p(Y^n)), \tag{1.3.1}$$

where Y^i $(1 \leq i \leq n)$ are the components of Y. Although X and Y are supposed to be tangential to M the directional derivative $D_X Y$ need *not* be tangential. To obtain an intrinsic notion one defines on an oriented hypersurface the *covariant derivative* $\nabla_X Y$ of Y in direction of X by the tangential projection of the directional derivative, i.e., ([10, 3.2.2])

$$\nabla_X Y = (D_X Y)^{\tan} = D_X Y - \langle D_X Y, \nu \rangle \, \nu, \tag{1.3.2}$$

where ν is the Gauss map ([10, 3.1.3]) i.e., the unit normal vector field of M such that for all p in the hypersurface $\det(\nu_p, e^1, \dots, e^{n-1}) > 0$ for all positively oriented bases $\{e^1, \dots, e^{n-1}\}$ of $T_p M$.
This construction clearly uses the structure of the ambient Euclidean space, which in case of a general SRMF is no longer available. Hence we will rather follow a different route and define the covariant derivative as an operation that maps a pair of vector fields to another vector field and has a list of characterising properties. Of course these properties are nothing else but the corresponding properties of the covariant derivative on hypersurfaces, that is we turn the analog of [10, 3.2.4] into a definition.

1.3.1 Definition (Connection). *A* (linear) connection *on a $\mathcal{C}^\infty$-manifold M is a map*

$$\nabla : \ \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \qquad (X, Y) \mapsto \nabla_X Y \tag{1.3.3}$$

such that the following properties hold

(∇1) $\nabla_X Y$ *is $\mathcal{C}^\infty(M)$-linear in X*
(i.e., $\nabla_{X_1 + f X_2} Y = \nabla_{X_1} Y + f \nabla_{X_2} Y \quad \forall f \in \mathcal{C}^\infty(M), X_1, X_2 \in \mathfrak{X}(M)$),

(∇2) $\nabla_X Y$ *is $\mathbb{R}$-linear in Y*
(i.e., $\nabla_X (Y_1 + a Y_2) = \nabla_X Y_1 + a \nabla_X Y_2 \quad \forall a \in \mathbb{R}, Y_1, Y_2 \in \mathfrak{X}(M)$),

(∇3) $\nabla_X (f Y) = X(f) Y + f \nabla_X Y$ *for all $f \in \mathcal{C}^\infty(M)$ (Leibniz rule).*

We call $\nabla_X Y$ the covariant derivative *of Y in direction X w.r.t. the connection ∇.*

1.3.2 Remark (Properties of ∇).

(i) Property ($\nabla 1$) implies that for fixed Y the map $X \mapsto \nabla_X Y$ is a tensor field. This fact needs some explanation. First recall that by [10, 2.6.19] tensor fields are precisely $\mathcal{C}^\infty(M)$-multilinear maps that take one forms and vector fields to smooth functions, more precisely $\mathcal{T}^r_s(M) \cong L^{r+s}_{\mathcal{C}^\infty(M)}(\Omega^1(M) \times \cdots \times \mathfrak{X}(M), \mathcal{C}^\infty(M))$. Now for $Y \in \mathfrak{X}(M)$ fixed, $A = X \mapsto \nabla_X Y$ is a $\mathcal{C}^\infty(M)$-multininear map $A : \mathfrak{X}(M) \to \mathfrak{X}(M)$ which naturally is identified with the mapping

$$\bar{A} : \Omega^1(M) \times \mathfrak{X}(M) \to \mathcal{C}^\infty(M), \qquad \bar{A}(\omega, X) = \omega(A(X)) \tag{1.3.4}$$

which is $\mathcal{C}^\infty(M)$-multilinear by ($\nabla 1$), hence a $(1, 1)$ tensor field.

Hence we can speak of $(\nabla_X Y)(p)$ for any p in M and moreover given $v \in T_pM$ we can define $\nabla_v Y := (\nabla_X Y)(p)$, where X is any vector field with $X_p = v$.

(ii) On the other hand the mapping $Y \to \nabla_X Y$ for fixed X is *not* a tensor field since ($\nabla 2$) merely demands $\mathbb{R}$-linearity.

In the following our aim is to show that on any SRMF there is exactly one connection which is compatible with the metric. However, we need a supplementary statement, which is of substantial interest of its own. In any vector space V with scalar product g we have an identification of vectors in V with covectors in V^* via

$$V \ni v \mapsto v^\flat \in V^* \quad \text{where} \quad v^\flat(w) := \langle v, w \rangle \quad (w \in V). \tag{1.3.5}$$

Indeed this mapping is injective by nondegeneracy of g and hence an isomorphism. We will now show that this construction extends to SRMFs providing a identification of vector fields and one forms.

1.3.3 Theorem (Musical isomorphism). *Let M be a SRMF. For any $X \in \mathfrak{X}(M)$ define $X^\flat \in \Omega^1(M)$ via*

$$X^\flat(Y) := \langle X, Y \rangle \qquad \forall Y \in \mathfrak{X}(M). \tag{1.3.6}$$

Then the mapping $X \mapsto X^\flat$ is a $\mathcal{C}^\infty(M)$-linear isomorphism from $\mathfrak{X}(M)$ to $\Omega^1(M)$.

Proof. First $X^\flat : \mathfrak{X}(M) \to \mathcal{C}^\infty(M)$ is obviously $\mathcal{C}^\infty(M)$-linear, hence in $\Omega^1(M)$, cf. [10, 2.6.19]. Also the mapping $\phi : X \mapsto X^\flat$ is $\mathcal{C}^\infty(M)$-linear and we show that it is an isomorphism.

ϕ is injective: Let $\phi(X) = 0$, i.e., $\langle X, Y \rangle = 0$ for all $Y \in \mathfrak{X}(M)$ implying $\langle X_p, Y_p \rangle = 0$ for all $Y \in \mathfrak{X}(M)$, $p \in M$. Now let $v \in T_pM$ and choose a vector field $Y \in \mathfrak{X}(M)$ with $Y_p = v$. But then by nondegeneracy of $g(p)$ we obtain

$$\langle X_p, v \rangle = 0 \quad \Rightarrow \quad X_p = 0, \tag{1.3.7}$$

and since p was arbitrary we infer X=0.

ϕ is surjective: Let $\omega \in \Omega^1(M)$. We will construct $X \in \mathfrak{X}(M)$ such that $\phi(X) = \omega$. We do so in three steps.

(1) The local case: Let $(\varphi = (x^1, \ldots, x^n), U)$ be a chart and $\omega|_U = \omega_i dx^i$. We set $X|_U := g^{ij}\omega_i \frac{\partial}{\partial x^j} \in \mathfrak{X}(U)$. Since (g^{ij}) is the inverse matrix of (g_{ij}) we have

$$\langle X|_U, \frac{\partial}{\partial x^k}\rangle = g^{ij}\,\omega_i\,\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\rangle = \omega_i\,g^{ij}\,g_{jk} = \omega_i\,\delta^i_{\ k} = \omega_k = \omega|_U(\frac{\partial}{\partial x^k}), \tag{1.3.8}$$

and by $\mathcal{C}^\infty(M)$-linearity we obtain $X^\flat|_U = \omega|_U$.

(2) The change of charts works: We show that for any chart $(\psi = (y^1, \ldots, y^n), V)$ with $U \cap V \neq \emptyset$ we have $X_U|_{U\cap V} = X_V|_{U\cap V}$. More precisely with $\omega|_V = \bar\omega_j dy^j$ and $g|_V = \bar g_{ij} dy^i \otimes dy^j$ we show that $g^{ij}\omega_i \frac{\partial}{\partial x^j} = \bar g^{ij}\bar\omega_i \frac{\partial}{\partial y^j}$.

To begin with recall that $dx^j = \frac{\partial x^j}{\partial y^i} dy^i$ ([10, 2.7.27(ii)]) and so

$$\omega|_{U\cap V} = \omega_j dx^j = \omega_j \frac{\partial x^j}{\partial y^i} dy^i = \bar\omega_i dy^i, \quad \text{implying} \quad \bar\omega_i = \omega_m \frac{\partial x^m}{\partial y^i}.$$

Moreover by [10, 2.4.15] we have $\frac{\partial}{\partial y^i} = \frac{\partial x^k}{\partial y^i}\frac{\partial}{\partial x^k}$ which gives

$$\bar g_{ij} = g\Big(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\Big) = g\Big(\frac{\partial x^k}{\partial y^i}\frac{\partial}{\partial x^k}, \frac{\partial x^l}{\partial y^j}\frac{\partial}{\partial x^l}\Big) = \frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}\, g\Big(\frac{\partial}{\partial y^k}, \frac{\partial}{\partial y^l}\Big) = \frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}\, g_{kl},$$

and so by setting $A = (a_{ki}) = (\frac{\partial x^k}{\partial y^i})$ we obtain

$$(\bar g_{ij}) = A^t (g_{ij}) A \quad \text{hence} \quad (\bar g^{ij}) = A^{-1} g^{ij} (A^{-1})^t \quad \text{and so} \quad \bar g^{ij} = \frac{\partial y^i}{\partial x^k}\, g^{kl}\, \frac{\partial y^j}{\partial x^l}.$$

Finally we obtain

$$\bar g^{ij}\,\bar\omega_i\,\frac{\partial}{\partial y^j} = \frac{\partial y^i}{\partial x^k}\,g^{kl}\,\frac{\partial y^j}{\partial x^l}\,\omega_m\,\frac{\partial x^m}{\partial y^i}\,\frac{\partial x^n}{\partial y^j}\,\frac{\partial}{\partial x^n} = g^{kl}\,\delta^m_{\ k}\,\omega_m\,\delta^n_{\ l}\,\frac{\partial}{\partial x^n} = g^{mn}\,\omega_m\,\frac{\partial}{\partial x^n}.$$

(3) Globalisation: By (2) $X(p) := X|_U(p)$ (where U is any chart neighbourhood of p) defines a vector field on M . Now choose a cover $\mathcal{U} = \{U_i |\, i \in I\}$ of M by chart neighbourhoods and a subordinate partition of unity $(\chi_i)_i$ such that $\mathrm{supp}(\chi_i) \subseteq U_i$ (cf. [10, 2.3.10]). For any $Y \in \mathfrak{X}(M)$ we then have

$$\begin{aligned} \langle X, Y\rangle &= \langle X, \sum_i \chi_i Y\rangle = \sum_i \langle X, \chi_i Y\rangle = \sum_i \langle X|_{U_i}, \chi_i, Y\rangle \\ &= \sum_i \omega|_{U_i}(\chi_i Y) = \sum_i \omega(\chi_i Y) = \omega(\sum_i \chi_i Y) = \omega(Y), \end{aligned} \tag{1.3.9}$$

and we are done. □

Hence in semi-Riemannian geometry we can always identify vectors and vector fields with covectors and one forms, respectively: X and $\phi(X) = X^\flat$ contain the same information and are called *metrically equivalent*. One also writes $\omega^\sharp = \phi^{-1}(\omega)$ and this notation is the source of the name 'musical isomorphism'. Especially in the physics literature this isomorphism is often encoded in the notation. If $X = X^i\partial_i$ is a (local) vector field then one denotes the metrically equivalent one form by $X^\flat = X_i dx^i$ and we clearly have $X_i = g_{ij}X^j$ and $X^i = g^{ij}X_j$. One also calls these operations the raising and lowering of indices. The musical isomorphism naturally extends to higher order tensors as we shall see in section 3.2, below.

The next result is crucial for all the following. It is sometimes called the *fundamental Lemma of semi-Riemannian geometry.*

1.3.4 Theorem (Levi Civita connection). *Let (M, g) be a SRMF. Then there exists one and only one connection ∇ on M such that (besides the defining properties* $(\nabla 1)-(\nabla 3)$ *of 1.3.1) we have for all $X, Y, Z \in \mathfrak{X}(M)$*

(∇4) $[X, Y] = \nabla_X Y - \nabla_Y X$ *(torsion free condition)*

(∇5) $Z\langle X, Y\rangle = \langle\nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle$ *(metric property).*

The map ∇ is called the Levi-Civita connection *of (M, g) and it is uniquely determined by the so-called* Koszul-formula

$$\begin{aligned} 2\langle\nabla_X Y, Z\rangle =& X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle \\ &- \langle X, [Y, Z]\rangle + \langle Y, [Z, X]\rangle + \langle Z, [X, Y]\rangle. \end{aligned} \tag{1.3.10}$$

Proof. *Uniqueness:* If ∇ is a connection with the additional properties (∇4), (∇5) then the Koszul-formula (1.3.10) holds: Indeed denoting the right hand side of (1.3.10) by $F(X, Y, Z)$ we find

$$\begin{aligned} F(X, Y, Z) =& \langle\nabla_X Y, Z\rangle + \cancel{\langle Y, \nabla_X Z\rangle} + \cancel{\langle\nabla_Y Z, X\rangle} + \cancel{\langle Z, \nabla_Y X\rangle} - \cancel{\langle\nabla_Z X, Y\rangle} - \cancel{\langle X, \nabla_Z Y\rangle} \\ &- \cancel{\langle X, \nabla_Y Z\rangle} + \cancel{\langle X, \nabla_Z Y\rangle} + \cancel{\langle Y, \nabla_Z X\rangle} - \cancel{\langle Y, \nabla_X Z\rangle} + \langle Z, \nabla_X Y\rangle - \cancel{\langle Z, \nabla_Y X\rangle} \\ =& 2\langle\nabla_X Y, Z\rangle. \end{aligned}$$

Now by injectivity of ϕ in Theorem 1.3.3, $\nabla_X Y$ is uniquely determined.

Existence: For fixed X, Y the mapping $Z \mapsto F(X, Y, Z)$ is $\mathcal{C}^\infty(M)$-linear as follows by a straight forward calculation using [10, 2.5.15(iv)]. Hence $Z \mapsto F(X, Y, Z) \in \Omega^1(M)$ and by 1.3.3 there is a (uniquely defined) vector field which we call $\nabla_X Y$ such that $2\langle\nabla_X Y, Z\rangle = F(X, Y, Z)$ for all $Z \in \mathfrak{X}(M)$. Now $\nabla_X Y$ by definition obeys the Koszul-formula and it remains to show that the properties (∇1)–(∇5) hold.

(∇1) $\nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y$ follows from the fact that $F(X_1+X_2, Y, Z) = F(X_1, Y, Z) + F(X_2, Y, Z)$. Now let $f \in \mathcal{C}^\infty(M)$ then we have by [10, 2.5.15(iv)]

$$2\langle\nabla_{fX} Y - f\nabla_X Y, Z\rangle = F(X, fY, Z) - fF(X, Y, Z) = \ldots = 0, \tag{1.3.11}$$

where we have left the straight forward calculation to the reader. Hence by another appeal to Theorem 1.3.3 we have $\nabla_{fX} Y = f\nabla_X Y$.

(∇2) follows since obviously $Y \mapsto F(X,Y,Z)$ is $\mathbb{R}$-linear.

(∇3) Again by [10, 2.5.15(iv)] we find

$$\begin{aligned} 2\langle \nabla_X fY, Z\rangle &= F(X, fY, Z) \\ &= X(f)\langle Y,Z\rangle - \cancel{Z(f)\langle X,Y\rangle} + \cancel{Z(f)\langle X,Y\rangle} + X(f)\langle Z,Y\rangle + fF(X,Y,Z) \\ &= 2\,\langle X(f)Y + f\nabla_X Y, Z\rangle, \end{aligned} \tag{1.3.12}$$

and the claim again follows by 1.3.3.

(∇4) We calculate

$$\begin{aligned} 2\langle \nabla_X Y - \nabla_Y X, Z\rangle &= F(X,Y,Z) - F(Y,X,Z) \\ &= \ldots = \langle Z, [X,Y]\rangle - \langle Z, [Y,X]\rangle = 2\langle [X,Y], Z\rangle \end{aligned} \tag{1.3.13}$$

and onother appeal to 1.3.3 gives the statement.

(∇5) We calculate

$$2\Big(\langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle\Big) = F(Z,X,Y) + F(Z,Y,X) = \ldots = 2Z\big(\langle X,Y\rangle\big). \qquad \square$$

1.3.5 Remark. In the case of M being an oriented hypersurface of $\mathbb{R}^n$ the covariant derivative is given by (1.3.2). By [10, 3.2.4, 3.2.5] ∇ satisfies (∇1)–(∇5) and hence is the Levi-Civita connection of M (with the induced metric).

Next we make sure that ∇ is local in both slots, a result of utter importance.

1.3.6 Lemma (Localisation of ∇). *Let $U \subseteq M$ be open and let $X, Y, X_1, X_2, Y_1, Y_2 \in \mathfrak{X}(M)$. Then we have*

(i) If $X_1|_U = X_2|_U$ then $\big(\nabla_{X_1} Y\big)\big|_U = \big(\nabla_{X_2} Y\big)\big|_U$, and

(ii) If $Y_1|_U = Y_2|_U$ then $\big(\nabla_X Y_1\big)\big|_U = \big(\nabla_X Y_2\big)\big|_U$.

Proof.

(i) By remark 1.3.2(i): $X \mapsto \nabla_X Y$ is a tensor field hence we even have that $X_1|_p = X_2|_p$ at any point $p \in M$ implying $(\nabla_{X_1} Y)|_p = (\nabla_{X_2} Y)|_p$.

(ii) It suffices to show that $Y|_U = 0$ implies $(\nabla_X Y)|_U = 0$. So let $p \in U$ and $\chi \in \mathcal{C}^\infty(M)$ with $\mathrm{supp}(\chi) \subseteq U$ and $\chi \equiv 1$ on a neighbourhood of p. By (∇3) we then have

$$0 = (\nabla_X \underbrace{\chi Y}_{=0})|_p = \underbrace{X(\chi)}_{=0}|_p Y_p + \underbrace{\chi(p)}_{=1}(\nabla_X Y)|_p \text{ and so } (\nabla_X Y)|_U = 0. \tag{1.3.14}$$

$\square$

1.3.7 Remark. Lemma 1.3.6 allows us to restrict ∇ to $\mathfrak{X}(U) \times \mathfrak{X}(U)$: Let $X, Y \in \mathfrak{X}(U)$ and $V \subseteq \bar{V} \subseteq U$ (cf. [10, 2.3.12]) and extend X, Y by vector fields $\tilde{X}, \tilde{Y} \in \mathfrak{X}(M)$ such that $\tilde{X}|_V = X|_V$ and $\tilde{Y}|_V = Y|_V$. (This can be easily done using a partition of unity subordinate to the cover $U, M \setminus \bar{V}$, cf. [10, 2.3.14].) Now we may set $(\nabla_X Y)|_V := (\nabla_{\tilde{X}} \tilde{Y})|_V$ since by 1.3.6 this definition is independent of the choice of the extensions $\tilde{X}$, $\tilde{Y}$. Moreover we may write U as the union of such V's and so $\nabla_X Y$ is a well-defined element of $\mathfrak{X}(U)$. In particular, this allows to insert the local basis vector fields ∂_i into ∇, which will be extensively used in the following.

1.3.8 Definition (Christoffel symbols). *Let $(\varphi = (x^1, \ldots, x^n), U)$ be a chart of the SRMF M. The* Christoffel symbols *(of the second kind) with respect to φ are the $\mathcal{C}^\infty$-functions $\Gamma^i_{jk} : U \to \mathbb{R}$ defined by*

$$\nabla_{\partial_i} \partial_j =: \Gamma^k_{ij} \partial_k \qquad (1 \leq i, j \leq n). \tag{1.3.15}$$

Since $[\partial_i, \partial_j] = 0$, property ($\nabla$4) immediately gives the symmetry of the Christoffel symbols in the lower pair of indices, i.e., $\Gamma^k_{ij} = \Gamma^k_{ji}$. Observe that Γ is not a tensor and so the Christoffel symbols do not exhibit the usual transformation behaviour of a tensor field under the change of charts. The next statement, in particular, shows how to calculate the Christoffel symbols from the metric.

1.3.9 Proposition (Christoffel symbols explicitly). *Let $(\varphi = (x^1, \ldots, x^n), U)$ be a chart of the SRMF (M, g) and let $Z = Z^i \partial_i \in \mathfrak{X}(U)$. Then we have*

(i) $\Gamma^k_{ij} =: g^{km} \Gamma_{mij} = \dfrac{1}{2} g^{km} \left(\dfrac{\partial g_{jm}}{\partial x^i} + \dfrac{\partial g_{im}}{\partial x^j} - \dfrac{\partial g_{ij}}{\partial x^m} \right)$,

(ii) $\nabla_{\partial_i} Z^j \partial_j = \left(\dfrac{\partial Z^k}{\partial x^i} + \Gamma^k_{ij} Z^j \right) \partial_k$.

The $\mathcal{C}^\infty(M)$-functions Γ_{kij} are sometimes called the Christoffel symbols of the first kind.
Proof.

(i) Set $X = \partial_i$, $Y = \partial_j$ and $Z = \partial_m$ in the Koszul formula (1.3.10). Since all Lie-brackets vanish we obtain

$$2\langle \nabla_{\partial_i} \partial_j, \partial_m \rangle = \partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}, \tag{1.3.16}$$

which upon multiplying with g^{km} gives the result.

(ii) follows immediately from (∇3) and 1.3.6. □

1.3.10 Lemma (The connection of flat space). *For* $X, Y \in \mathfrak{X}(\mathbb{R}^n_r)$ *with* $Y = (Y^1, \ldots, Y^n) = Y^i \partial_i$ *let*

$$\nabla_X Y = X(Y^i)\partial_i. \tag{1.3.17}$$

Then ∇ *is the Levi-Civita connection on* $\mathbb{R}^n_r$ *and in natural coordinates (i.e., using* id *as a global chart) we have*

(i) $g_{ij} = \delta_{ij}\varepsilon_j$ *(with* $\varepsilon_j = -1$ *for* $1 \leq j \leq r$ *and* $\varepsilon_j = +1$ *for* $r < j \leq n$*),*

(ii) $\Gamma^i_{jk} = 0$ *for all* $1 \leq i, j, k \leq n$.

Proof. Recall that in the terminology of [10, Sec. 3.2] we have $\nabla_X Y = D_X Y = p \mapsto DY(p)\, X_p$ which coincides with (1.3.17). The validity of (∇1)–(∇5) has been checked in [10, 3.2.4,5] and hence ∇ is the Levi-Civita connection. Moreover we have

(i) $g_{ij} = \langle \partial_i, \partial_j \rangle = \langle e_i, e_j \rangle = \varepsilon_i \delta_{ij}$, and

(ii) $\Gamma^i_{jk} = 0$ by (i) and 1.3.9(i). □

Next we consider vector fields with vanishing covariant derivatives.

1.3.11 Definition (Parallel vector field). *A vector field* X *on a SRMF* M *is called* parallel *if* $\nabla_Y X = 0$ *for all* $Y \in \mathfrak{X}(M)$.

1.3.12 Example. The coordinate vector fields in $\mathbb{R}^n_r$ are parallel: Let $Y = Y^j \partial_j$ then by 1.3.10(ii) $\nabla_Y \partial_i = Y^j \nabla_{\partial_j} \partial_i = 0$. More generally on $\mathbb{R}^n_r$ the constant vector fields are precisely the parallel ones, since

$$\nabla_Y X = 0\ \forall Y \ \Leftrightarrow\ DX(p)Y(p) = 0\ \forall Y\, \forall p \ \Leftrightarrow\ DX = 0 \ \Leftrightarrow\ X = \text{const}. \tag{1.3.18}$$

In light of this example the notion of a parallel vector field generalises the notion of a constant vector field. We now present an explicit example.

1.3.13 Example (Cylindrical coordinates on $\mathbb{R}^3$). Let (r, φ, z) be cylindrical coordinates on $\mathbb{R}^3$, i.e., $(x, y, z) = (r\cos\varphi, r\sin\varphi, z)$, see figure 1.4. This clearly is a chart on $\mathbb{R}^3 \setminus \{x \geq 0,\ y = 0\}$. Its inverse $(r, \varphi, z) \mapsto (r\cos\varphi, r\sin\varphi, z)$ is a parametrisation, hence we have (cf. [10, below 2.4.11] or directly [10, 2.4.15])

$$\partial_r = \cos\varphi\partial_x + \sin\varphi\partial_y,$$
$$\partial_\varphi = rX \quad \text{with } X = -\sin\varphi\partial_x + \cos\varphi\partial_y,$$
$$\partial_z = \partial_z.$$

Setting $y^1 = r$, $y^2 = \varphi$, $y^3 = z$ we obtain

$$g_{11} = \langle\partial_r, \partial_r\rangle = 1,$$
$$g_{22} = \langle\partial_\varphi, \partial_\varphi\rangle = r^2(\cos^2\varphi + \sin^2\varphi) = r^2,$$
$$g_{33} = \langle\partial_z, \partial_z\rangle = 1,$$
$$g_{ij} = 0 \quad \text{for all } i \neq j.$$

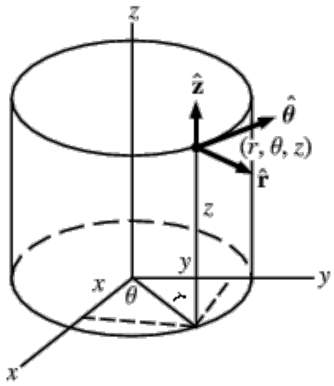

Figure 1.4: Cylindrical coordinates
♣ fix notation ♣

So we have

$$(g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.19}$$

and hence

$$g = g_{ij}dy^i \otimes dy^j = dr \otimes dr + r^2 d\varphi \otimes d\varphi + dz \otimes dz =: dr^2 + r^2 d\varphi^2 + dz^2.$$

By (1.3.19) $\{\partial_r, \partial_\varphi, \partial_z\}$ is orthogonal and hence (r, φ, z) is an orthogonal coordinate system. For the Christoffel symbols we find (by 1.3.9(i))

$$\Gamma^1_{22} = \frac{1}{2}g^{1l}\Gamma_{l22} = \frac{1}{2}g^{11}\Gamma_{122} = \frac{1}{2}\,1\,(\underbrace{g_{12,2}}_{=0} + \underbrace{g_{21,2}}_{=0} - g_{22,1}) = \frac{-1}{2}\,2r = -r,$$
$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{2}g^{2l}\Gamma_{l21} = \frac{1}{2}g^{22}\Gamma_{221} = \frac{1}{2}\frac{1}{r^2}(g_{22,1} + \underbrace{g_{12,2}}_{=0} - \underbrace{g_{21,2}}_{=0}) = \frac{1}{2r^2}\,2r = \frac{1}{r},$$

and all other $\Gamma^i_{jk} = 0$. Hence we have $\nabla_{\partial_i}\partial_j = 0$ for all i, j with the exception of

$$\nabla_{\partial_\varphi}\partial_\varphi = -r\partial_r\,, \quad \text{and } \nabla_{\partial_\varphi}\partial_r = \nabla_{\partial_r}\partial_\varphi = \frac{1}{r}\partial_\varphi = X\,.$$

By figure 1.4 we see that ∂_r and ∂_φ are parallel if one moves in the z-direction. We hence expect that $\nabla_{\partial_z}\partial_\varphi = 0 = \nabla_{\partial_z}\partial_r$ which also results from our calculations. Moreover ∂_z is parallel since it is a coordinate vector field in the natural basis of $\mathbb{R}^3$, cf. 1.3.12.

Our next aim is to extend the covariant derivative to tensor fields of general rank. We will start with a slight detour introducing the notion of a *tensor derivation* and its basic properties and then use this machinery to extend the covariant derivative to the space $\mathcal{T}^r_s(M)$ of tensor fields of rank (r, s).

Interlude: Tensor derivations

In this brief interlude we introduce some basic operations on tensor fields which will be essential in the following. We recall (for more information on tensor fields see e.g. [10, Sec. 2.6]) that a tensor field $A \in \mathcal{T}^r_s(M) = \Gamma(M, T^r_s(M))$ is a (smooth) section of the (r,s)-tensor bundle $T^r_s(M)$ of M. That is to say that for any point $p \in M$, the value of the tensor field $A(p)$ is a multilinear map

$$A(p): \quad \underbrace{T_pM^* \times \cdots \times T_pM^*}_{r \text{ times}} \times \underbrace{T_pM \times \cdots \times T_pM}_{s \text{ times}} \to \mathbb{R}. \tag{1.3.20}$$

Locally in a chart $(\psi = (x^1, \ldots, x^n), V)$ we have

$$A|_V = A^{i_1 \ldots i_r}_{j_1 \ldots j_s}\, \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \tag{1.3.21}$$

where the coefficient functions are given for $q \in V$ by

$$A^{i_1 \ldots i_r}_{j_1 \ldots j_s}(q) = A(q)(dx^{i_1}|_q, \ldots, dx^{i_r}|_q, \partial_{j_1}|_q, \ldots, \partial_{j_s}|_q). \tag{1.3.22}$$

The space $\mathcal{T}^r_s(M)$ of sections can be identified with the space

$$L^{r+s}_{\mathcal{C}^\infty(M)}(\underbrace{\Omega^1(M) \times \cdots \times \Omega^1(M)}_{r\text{-times}} \times \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{s\text{-times}}, \mathcal{C}^\infty(M)) \tag{1.3.23}$$

of $\mathcal{C}^\infty(M)$-multilinear maps from one-forms and vector fields to smooth functions. Recall also the special cases $\mathcal{T}^0_0(M) = \mathcal{C}^\infty(M)$, $\mathcal{T}^1_0(M) = \mathfrak{X}(M)$, and $\mathcal{T}^0_1(M) = \Omega^1(M)$.

Additionally we will also frequently deal with the following situation, which generalises the one of 1.3.2(i): If $A : \mathfrak{X}(M)^s \to \mathfrak{X}(M)$ is a $\mathcal{C}^\infty(M)$-multilinear mapping then we define

$$\begin{aligned} \bar{A} : \Omega^1(M) \times \mathfrak{X}(M)^s &\to \mathcal{C}^\infty(M) \\ \bar{A}(\omega, X_1, \ldots, X_s) &:= \omega(A(X_1, \ldots, X_s)). \end{aligned} \tag{1.3.24}$$

Clearly $\bar{A}$ is $\mathcal{C}^\infty(M)$-multilinear and hence a $(1,s)$-tensor field and we will frequently and tacitly identify $\bar{A}$ and A.

We start by introducing a basic operation on tensor fields that shrinks their rank from (r,s) to $(r-1, s-1)$. The general definition is based on the following special case.

1.3.14 Lemma ($(1,1)$-contraction). *There is a unique $\mathcal{C}^\infty(M)$-linear map $\mathcal{C} : \mathcal{T}^1_1(M) \to \mathcal{C}^\infty(M)$ called the $(1,1)$-contraction such that*

$$\mathcal{C}(X \otimes \omega) = \omega(X) \qquad \textit{for all } X \in \mathfrak{X}(M) \textit{ and } \omega \in \Omega^1(M). \tag{1.3.25}$$

Proof. Since $\mathcal{C}$ is to be $\mathcal{C}^\infty(M)$-linear it is a pointwise operation, cf. [10, 2.6.19] and we start by giving a local definition. For the natural basis fields of a chart $(\varphi = (x^1, \ldots, x^n), V)$

we necessarily have $\mathcal{C}(\partial_j \otimes dx^i) = dx^i(\partial_j) = \delta^i_j$ and so for $\mathcal{T}^1_1(M) \ni A = \sum A^i_j \partial_i \otimes dx^j$ we are forced to define

$$\mathcal{C}(A) = \sum_i A^i_i = \sum_i A(dx^i, \partial_j). \tag{1.3.26}$$

It remains to show that the definition is independent of the chosen chart. Let $(\psi = (y^1, \dots, y^n), V)$ be another chart then we have using [10, 2.7.27(iii)] as well as the summation convention

$$A(dy^m, \partial_m) = A\left(\frac{\partial y^m}{\partial x^i}\, dx^i, \frac{\partial x^j}{\partial y^m} \partial x_j\right) = \underbrace{\frac{\partial y^m}{\partial x^i} \frac{\partial x^j}{\partial y^m}}_{\delta^j_i}\, A(dx^i, \partial x_j) = A(dx^i, \partial_i). \tag{1.3.27}$$

□

To define the *contraction for general rank tensors* let $A \in \mathcal{T}^r_s(M)$, fix $1 \leq i \leq r$, $1 \leq j \leq s$ and let $\omega^1, \dots, \omega^{r-1} \in \Omega^1(M)$ and $X_1, \dots, X_{s-1} \in \mathfrak{X}(M)$. Then the map

$$\Omega(M) \times \mathfrak{X}(M) \ni (\omega, X) \mapsto A(\omega^1, \dots, \underset{i}{\omega}, \dots, \omega^{r-1}, X_1, \dots, \underset{j}{X}, \dots, X_{s-1}) \tag{1.3.28}$$

is a $(1,1)$-tensor. We now apply the $(1,1)$-contraction $\mathcal{C}$ of 1.3.14 to (1.3.28) to obtain a $\mathcal{C}^\infty(M)$-function denoted by

$$(\mathcal{C}^i_j A)(\omega^1, \dots, \omega^{r-1}, X_1, \dots, X_{s-1}). \tag{1.3.29}$$

Obviously $\mathcal{C}^i_j A$ is $\mathcal{C}^\infty(M)$-linear in all its slots, hence it is a tensor field in $\mathcal{T}^{r-1}_{s-1}(M)$ which we call the (i,j)*-contraction of* A. We illustrate this concept by the following examples.

1.3.15 Examples (Contraction).

(i) Let $A \in \mathcal{T}^2_3(M)$ then $\mathcal{C}^1_3 A \in \mathcal{T}^1_2(M)$ is given by

$$\mathcal{C}^1_3 A(\omega, X, Y) = \mathcal{C}\big(A(\,.\,, \omega, X, Y, .)\big) \tag{1.3.30}$$

which locally takes the form

$$(\mathcal{C}^1_3 A)^k_{ij} = (\mathcal{C}^1_3(A)(dx^k, \partial_i, \partial_j) = \mathcal{C}\big(A(\,.\,, dx^k, \partial_i, \partial_j, .)\big) = A(dx^m, dx^k, \partial_i, \partial_j, \partial_m) = A^{mk}_{ijm},$$

where of course we again have applied the summation convention.

(ii) More generally the components of $\mathcal{C}^k_l A$ of $A \in \mathcal{T}^r_s(M)$ in local coordinates take the form $A^{i_1 \dots \overset{k}{m} \dots i_r}_{j_1 \dots \underset{l}{m} \dots j_s}$.

Now we may define the notion of a tensor derivation announced above as map on tensor fields that satisfies a product rule and commutes with contractions.

1.3.16 Definition (Tensor derivation). *A* tensor derivation $\mathcal{D}$ *on a smooth manifold* M *is a family of* $\mathbb{R}$*-linear maps*

$$\mathcal{D} = \mathcal{D}^r_s : \ \mathcal{T}^r_s(M) \to \mathcal{T}^r_s(M) \quad (r, s \geq 0) \tag{1.3.31}$$

such that for any pair A, B *of tensor fields we have*

(i) $\mathcal{D}(A \otimes B) = \mathcal{D}A \otimes B + A \otimes \mathcal{D}B$

(ii) $\mathcal{D}(\mathcal{C}A) = \mathcal{C}(\mathcal{D}A))$ *for any contraction* $\mathcal{C}$*.*

The product rule in the special case $f \in \mathcal{C}^\infty(M) = \mathcal{T}^0_0(M)$ and $A \in \mathcal{T}^r_s(M)$ takes the form

$$\mathcal{D}(f \otimes A) = \mathcal{D}(fA) = (\mathcal{D}f)A + f\mathcal{D}A. \tag{1.3.32}$$

Moreover for $r = 0 = s$ the tensor derivation $\mathcal{D}^0_0$ is a derivation on $\mathcal{C}^\infty(M)$ (cf. [10, 2.5.12]) and so by [10, 2.5.13] there exists a unique vector field $X \in \mathfrak{X}(M)$ such that

$$\mathcal{D}f = X(f) \quad \text{for all } f \in \mathcal{C}^\infty(M). \tag{1.3.33}$$

Despite the fact that tensor derivations are *not* $\mathcal{C}^\infty(M)$-linear and hence not pointwise defined[2] (cf. [10, 2.6.19]) they are local operators in the following sense.

1.3.17 Proposition (Tensor derivations are local). *Let* $\mathcal{D}$ *be a tensor derivation on* M *and let* $U \subseteq M$ *be open. Then there exists a unique tensor derivation* $\mathcal{D}_U$ *on* U*, called the* restriction of $\mathcal{D}$ to U *statisfying*

$$\mathcal{D}_U(A|_U) = (\mathcal{D}A)|_U \tag{1.3.34}$$

for all tensor fields A *on* M*.*

Proof. Let $B \in \mathcal{T}^r_s(U)$ and $p \in U$. Choose a cut-off function $\chi \in \mathcal{C}^\infty_0(U)$ with $\chi \equiv 1$ in a neighbourhood of p. Then $\chi B \in \mathcal{T}^r_s(M)$ and we define

$$(\mathcal{D}_U B)(p) := \mathcal{D}(\chi B)(p). \tag{1.3.35}$$

We have to check that this definition is valid and leads to the asserted properties.

(1) The definition is independent of χ: choose another cut-off function $\tilde{\chi}$ at p and set $f = \chi - \tilde{\chi}$. Then choosing a function $\varphi \in \mathcal{C}^\infty_0(U)$ with $\varphi \equiv 1$ on $\mathrm{supp}(f)$ we obtain

$$\mathcal{D}(fB)(p) = \mathcal{D}(f\varphi B)(p) = \mathcal{D}(f)|_p(\varphi B)(p) + \underbrace{f(p)}_{=0}\mathcal{D}(\varphi B)(p) = 0, \tag{1.3.36}$$

since we have with a vector field X as in (1.3.33) that $\mathcal{D}f(p) = X(f)(p) = 0$ by the fact that $f \equiv 0$ near the point p.

[2] Recall from analysis that taking a derivative of a function is *not* a pointwise operation: It depends on the values of the function in a neighbourhood.

(2) $\mathcal{D}_U B \in \mathcal{T}^r_s(U)$ since for all $V \subseteq U$ open we have $\mathcal{D}_U B|V = \mathcal{D}(\chi B)|_V$ by definition if $\chi \equiv 1$ on V. Now observe that $\chi B \in \mathcal{T}^r_s(M)$.

(3) Clearly $\mathcal{D}_U$ is a tensor derivation on U since $\mathcal{D}$ is is a tensor derivation on M.

(4) $\mathcal{D}_U$ has the restriction property (1.3.34) since if $B \in \mathcal{T}^r_s(M)$ we find for all $p \in U$ that $\mathcal{D}_U(B|_U)(p) = \mathcal{D}(\chi B|_U)(p) = \mathcal{D}(\chi B)(p)$ and $\mathcal{D}(\chi B)(p) = \mathcal{D}(B)(p)$ since $\mathcal{D}((1-\chi)B)(p) = 0$ by the same argument as used in (1.3.36).

(5) Finally $\mathcal{D}_U$ is uniquely determined: Let $\tilde{\mathcal{D}}_u$ be another tensor derivation that satisfies (1.3.34) then for $B \in \mathcal{T}^r_s(U)$ we again have $\tilde{\mathcal{D}}_u((1-\chi)B)(p) = 0$ and so by (4)

$$\tilde{\mathcal{D}}_U(B)(p) = \tilde{\mathcal{D}}_U(\chi B)(p) = \mathcal{D}(\chi B)(p) = \mathcal{D}_U(B)(p)$$

for all $p \in U$. □

We next state and prove a product rule for tensor derivations.

1.3.18 Proposition (Product rule). *Let $\mathcal{D}$ be a tensor derivation on M. Then we have for $A \in \mathcal{T}^r_s(M)$, $\omega^1, \dots, \omega^r \in \Omega(M)$, and $X_1, \dots, X_s \in \mathfrak{X}(M)$*

$$\begin{aligned}\mathcal{D}\Big(A(\omega^1,\dots,\omega^r,X_1,\dots,X_s)\Big) =& (\mathcal{D}A)(\omega^1,\dots,\omega^r,X_1,\dots,X_s)\\ &+\sum_{i=1}^r A(\omega^1,\dots,\mathcal{D}\omega^i,\dots,\omega^r,X_1,\dots,X_s)\\ &+\sum_{j=1}^s A(\omega^1,\dots,\omega^r,X_1,\dots,\mathcal{D}X_j,\dots,X_s).\end{aligned} \tag{1.3.37}$$

Proof. We only show the case $r = 1 = s$ since the general case follows in complete analogy. We have $A(\omega, X) = \bar{\mathcal{C}}(A \otimes \omega \otimes X)$ where $\bar{\mathcal{C}}$ is a composition of two contractions. Indeed in local coordinates $A \otimes \omega \otimes X$ has components $A^i_j \omega_k X^l$ and $A(\omega, X) = A(\omega_i dx^i, X^j \partial_j) = \omega_i X^j A(dx^i, \partial_j) = A^i_j \omega_i X^j$ and the claim follows from 1.3.15(ii).
By 1.3.16(i)–(ii) we hence have

$$\begin{aligned}\mathcal{D}\big(A(\omega,X)\big) &= \mathcal{D}\big(\bar{\mathcal{C}}(A\otimes\omega\otimes X)\big) = \bar{\mathcal{C}}\mathcal{D}(A\otimes\omega\otimes X)\\ &= \bar{\mathcal{C}}(\mathcal{D}A\otimes\omega\otimes X) + \bar{\mathcal{C}}(A\otimes\mathcal{D}\omega\otimes X) + \bar{\mathcal{C}}(A\otimes\omega\otimes\mathcal{D}X)\\ &= \mathcal{D}A(\omega,X) + A(\mathcal{D}\omega,X) + A(\omega,\mathcal{D}X).\end{aligned} \tag{1.3.38}$$

□

The product rule (1.3.37) can obviously be solved for the term involving $\mathcal{D}A$ resulting in a formula for the tensor derivation of a general tensor field A in terms of $\mathcal{D}$ only acting on functions, vector fields, and one-forms. Moreover for a one form ω we have by (1.3.37)

$$(\mathcal{D}\omega)(X) = \mathcal{D}(\omega(X)) - \omega(\mathcal{D}X) \tag{1.3.39}$$

and so the action of a tensor derivation is determined by its action on functions and vector fields alone, a fact which we state as follows.

1.3.19 Corollary. *If two tensor derivations $\mathcal{D}_1$ and $\mathcal{D}_2$ agree on functions $\mathcal{C}^\infty(M)$ and on vector fields $\mathfrak{X}(M)$ then they agree on all tensor fields, i.e., $\mathcal{D}_1 = \mathcal{D}_2$.*

More importantly a tensor derivation can be constructed from its action on just functions and vector fields in the following sense.

1.3.20 Theorem (Constructing tensor derivations). *Given a vector field $V \in \mathfrak{X}(M)$ and an $\mathbb{R}$-linear mapping $\delta : \mathfrak{X}(M) \to \mathfrak{X}(M)$ obeying the product rule*

$$\delta(fX) = V(f)X + f\delta(X) \quad \text{for all } f \in \mathcal{C}^\infty(M),\, \mathfrak{X} \in \mathfrak{X}(M). \tag{1.3.40}$$

Then there exists a unique tesor derivation $\mathcal{D}$ on M such that $\mathcal{D}^0_0 = V : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ and $\mathcal{D}^1_0 = \delta : \mathfrak{X}(M) \to \mathfrak{X}(M)$.

Proof. Uniqueness is a consequence of 1.3.19 and we are left with constructing $\mathcal{D}$ using the product rule.
To begin with, by (1.3.39) we necessarily have for any one-form ω

$$(\mathcal{D}\omega)(X) \equiv (\mathcal{D}^0_1\omega)(X) = V(\omega(X)) - \omega(\delta(X)), \tag{1.3.41}$$

which obviously is $\mathbb{R}$-linear. Moreover, $\mathcal{D}\omega$ is $\mathcal{C}^\infty(M)$-linear hence a one-form since

$$\begin{aligned}\mathcal{D}\omega(fX) &= V(\omega(fX)) - \omega(\delta(fX)) = V(f\omega(X)) - \omega(V(f)X) - \omega(f\delta(X))\\ &= fV(\omega(X)) + \cancel{V(f)\omega(X)} - \cancel{V(f)\omega(X)} - f\omega(\delta(X))\\ &= f\big(V(\omega(X)) - \omega(\delta(X)\big) = f\mathcal{D}\omega(X).\end{aligned} \tag{1.3.42}$$

Similarly for higher ranks $r + s \geq 2$ we have to define $\mathcal{D}^r_s$ by the product rule (1.3.37): Again it is easy to see that $\mathcal{D}^r_s$ is $\mathbb{R}$-linear and that $\mathcal{D}^r_s A$ is $\mathcal{C}^\infty(M)$-multilinear hence in $\mathcal{T}^r_s(M)$.
We now have to verify (i), (ii) of Definition 1.3.16. We only show $\mathcal{D}(A \otimes B) = \mathcal{D}A \otimes B + B \otimes \mathcal{D}A$ in case $A, B \in \mathcal{T}^1_1(M)$, the general case being completely analogous:

$$\begin{aligned}\big(\mathcal{D}(A\otimes B)\big)(\omega^1,\omega^2,X_1,X_2) &= V(A(\omega^1,X_1)\cdot B(\omega^2,X_2))\\ &\quad - \Big(A(\mathcal{D}\omega^1,X_1)B(\omega^2,X_2) + A(\omega^1,X_1)B(\mathcal{D}\omega^2,X_2)\Big)\\ &\quad - \Big(A(\omega^1,\mathcal{D}X_1)B(\omega^2,X_2) + A(\omega^1,X_1)B(\omega^2,\mathcal{D}X_2)\Big)\\ &= \Big(V\big(A(\omega^1,X_1)\big) - A(\mathcal{D}\omega^1,X_1) - A(\omega^1,\mathcal{D}X_1)\Big)B(\omega^2,X_2)\\ &\quad + A(\omega^1,X_1)\Big(V\big(B(\omega^2,X_2)\big) - B(\mathcal{D}\omega^2,X_2) - B(\omega^2,\mathcal{D}X_2)\Big)\\ &= (\mathcal{D}A\otimes B + A\otimes\mathcal{D}B)(\omega^1,\omega^2,X_1,X_2).\end{aligned}$$

Finally, we show that $\mathcal{D}$ commutes with contractions. We start by considering $\mathcal{C} : \mathcal{T}^1_1(M) \to \mathcal{C}^\infty(M)$. Let $A = X \otimes \omega \in \mathcal{T}^1_1(M)$, then we have by (1.3.41)

$$\mathcal{D}(\mathcal{C}(X \otimes \omega)) = \mathcal{D}(\omega(X)) = V(\omega(X)) = \omega(\delta(X)) + \mathcal{D}(\omega)(X), \tag{1.3.43}$$

which agrees with

$$\mathcal{C}(\mathcal{D}(X \otimes \omega)) = \mathcal{C}(\mathcal{D}X \otimes \omega + X \otimes \mathcal{D}\omega) = \omega(\mathcal{D}X) + (\mathcal{D}\omega)(X). \tag{1.3.44}$$

Obviously the same holds true for (finite) sums of terms of the form $\omega^i \otimes X_i$. Since $\mathcal{D}$ is local (Proposition 1.3.17) and $\mathcal{C}$ is even pointwise it suffices to prove the statement in local coordinates. But there each $(1,1)$-tensor is a sum as mentioned above. The extension to the general case is now straight forward. We only explicitly check it for $A \in \mathcal{T}^1_2(M)$:

$$\begin{aligned}
\big(\mathcal{D}^0_1(\mathcal{C}^1_2 A)\big)(X) &= \mathcal{D}^0_0\big((\mathcal{C}^1_2 A)(X)\big) - (\mathcal{C}^1_2 A)(\mathcal{D}^1_0 X) = \mathcal{D}^0_0\big(\mathcal{C}(A(\,.\,,X,.))\big) - \mathcal{C}\big(A(\,.\,,\mathcal{D}X,.)\big)\\
&= \mathcal{C}\Big(\mathcal{D}^1_1\big(A(\,.\,,X,.)\big) - A(\,.\,,\mathcal{D}X,.)\Big)\\
&= \mathcal{C}\Big((\omega,Y) \mapsto \mathcal{D}\big(A(\omega,X,Y)\big) - A(\mathcal{D}\omega,X,Y) - A(\omega,X,\mathcal{D}Y) - A(\omega,\mathcal{D}X,Y)\Big)\\
&= \mathcal{C}\big((\omega,Y) \mapsto (\mathcal{D}A)(\omega,X,Y)\big) = \big(\mathcal{C}^1_2(\mathcal{D}A)\big)(X).
\end{aligned}$$

□

As a first important example of a tensor derivation we discuss the Lie derivative.

1.3.21 Example (Lie derivative on $\mathcal{T}^r_s$). Let $X \in \mathfrak{X}(M)$. Then we define the tensor derivative L_X, called the *Lie derivative* with respect to X by setting

$$\begin{aligned}
L_X(f) &= X(f) \quad \text{for all } f \in \mathcal{C}^\infty(M), \text{ and}\\
L_X(Y) &= [X,Y] \quad \text{for all vector fields } Y \in \mathfrak{X}(M).
\end{aligned}$$

Indeed this definition generalises the Lie derivative or Lie bracket of vector fields to general tensors in $\mathcal{T}^r_s(M)$ since by Theorem 1.3.20 we only have to check that $\delta(Y) = L_X(Y) = [X,Y]$ satisfies the product rule (1.3.40). But this follows immediately form the corresponding property of the Lie bracket, see [10, 2.5.15(iv)].

Finally we return to the Levi-Civita covariant derivative on a SRMF (M,g), cf. 1.3.4. We want to extend it from vector fields to arbitrary tensor fields using Theorem 1.3.20. A brief glance at the assumptions of the latter theorem reveals that the defining properties $(\nabla 2)$ and $(\nabla 3)$ are all we need. So the following definition is valid.

1.3.22 Definition (Covariant derivative for tensors). *Let M be a SRMF and $X \in \mathfrak{X}(M)$. The* (Levi-Civita) covariant derivative ∇_X *is the uniquely determined tensor derivation on M such that* *(i)* $\nabla_X f = X(f)$ *for all* $f \in \mathcal{C}^\infty(M)$*, and*

(ii) $\nabla_X Y$ *is the Levi-Civita covariant derivative of* Y *w.r.t.* X *as given by 1.3.4.*

The covariant derivative w.r.t. a vector field X is a generalisation of the directional derivative. Similar to the case of multi-dimensional calculus in $\mathbb{R}^n$ we may collect all such directional derivatives into one differential. To do so we need to take one more technical step.

1.3.23 Lemma. *Let* $A \in \mathcal{T}^r_s(M)$*, then the mapping*

$$\mathfrak{X}(M) \ni X \mapsto \nabla_X A \in \mathcal{T}^r_s(M)$$

is $\mathcal{C}^\infty(M)$*-linear.*

Proof. We have to show that for X_1, $X_2 \in \mathfrak{X}(M)$ and $f \in \mathcal{C}^\infty(M)$ we have

$$\nabla_{X_1 + fX_2} A = \nabla_{X_1} A + f \nabla_{X_2} A \quad \text{for all } A \in \mathcal{T}^r_s(M). \tag{1.3.45}$$

However, by 1.3.20 we only have to show this for $A \in \mathcal{T}^0_0(M) = \mathcal{C}^\infty(M)$ and $A \in \mathcal{T}^1_0(M) = \mathfrak{X}(M)$. But for $A \in \mathcal{C}^\infty(M)$ equation (1.3.45) holds by definition and for $A \in \mathfrak{X}(M)$ this is just property (∇1). □

1.3.24 Definition (Covariant differential). *For* $A \in \mathcal{T}^r_s(M)$ *we define the* covariant differential $\nabla A \in \mathcal{T}^r_{s+1}(M)$ *of* A *as*

$$\nabla A(\omega^1, \dots, \omega^r, X_1, \dots, X_s, X) := (\nabla_X A)(\omega^1, \dots, \omega^r, X_1, \dots, X_s) \tag{1.3.46}$$

for all $\omega^1, \dots, \omega^r \in \Omega^1(M)$ *and* $X_1, \dots, X_s \in \mathfrak{X}(M)$*.*

1.3.25 Remark.

(i) In case $r = 0 = s$ the covariant differential is just the exterior derivative since for $f \in \mathcal{C}^\infty(M)$ and $X \in \mathfrak{X}(M)$ we have

$$(\nabla f)(X) = \nabla_X f = X(f) = df(X). \tag{1.3.47}$$

(ii) ∇A is a 'collection' all the covariant derivatives $\nabla_X A$ into one object. The fact that the covariant rank is raised by one, i.e., that $\nabla A \in \mathcal{T}^r_{s+1}(M)$ for $A \in \mathcal{T}^r_s(M)$ is the source of the name *covariant* derivative/differential.

(iii) In complete analogy with vector fields (cf. Definition 1.3.11) we call $A \in \mathcal{T}^r_s(M)$ *parallel* if $\nabla_X A = 0$ for all $X \in \mathfrak{X}(M)$ which we can now simply write as $\nabla A = 0$.

(iv) The metric condition ($\nabla 5$) just says that g itself is parallel since by the product rule 1.3.18 we have for all X, Y, $Z \in \mathfrak{X}(M)$

$$(\nabla_Z g)(X,Y) = \nabla_Z(g(X,Y)) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) = 0 \tag{1.3.48}$$

where the vanishing is due to ($\nabla 5$).

(v) If in a local chart the tensor field $A \in \mathcal{T}^r_s(M)$ has components $A^{i_1\dots i_r}_{j_1\dots j_s}$ the components of its covariant differential $\nabla A \in \mathcal{T}^r_{s+1}(M)$ are denoted by $A^{i_1\dots i_r}_{j_1\dots j_s;k}$ and take the form

$$A^{i_1\dots i_r}_{j_1\dots j_s;k} = \frac{\partial A^{i_1\dots i_r}_{j_1\dots j_s}}{\partial x^k} + \sum_{l=1}^{r} \Gamma^{i_l}_{km} A^{i_1\dots m\dots i_r}_{j_1\dots\dots\dots j_s} - \sum_{l=1}^{s} \Gamma^{m}_{kj_l} A^{i_1\dots\dots\dots i_r}_{j_1\dots m\dots j_s}\,. \tag{1.3.49}$$

Our next topic is the notion of a covariant derivative of vector fields which are not defined on all of M but just, say on (the image of) a curve. Of course then we can only expect to be able to define a derivative of the vector field in the direction of the curve. Intuitively such a notion corresponds to the rate of change of the vector field as we go along the curve. We begin by making precise the notion of such vector fields but do not restrict ourselves to the case of curves.

1.3.26 Definition (Vector field along a mapping). *Let N, M be smooth manifolds and let $f \in \mathcal{C}^\infty(N, M)$. A* vector field along f *is a smooth mapping*

$$Z:\ N \to TM \quad \text{such that } \pi \circ Z = f, \tag{1.3.50}$$

where $\pi : TM \to M$ is the vector bundle projection. We denote the $\mathcal{C}^\infty(N)$-module of all vector fields along f by $\mathfrak{X}(f)$.

The definition hence says that $Z(p) \in T_{f(p)}M$ for all points $p \in N$. In the special case of $N = I \subseteq \mathbb{R}$ a real interval and $f = c : I \to M$ a $\mathcal{C}^\infty$-curve we call $\mathfrak{X}(c)$ the space of *vector fields along the curve c*. In particular, in this case $t \mapsto \dot c(t) \equiv c'(t)$ is an element of $\mathfrak{X}(c)$. More precisely we have (cf. [10, below (2.5.3)]) $c'(t) = T_t c(1) = T_t c(\frac{\partial}{\partial t}|_t) \in T_{c(t)}M$. Also recall for later use that for any $f \in \mathcal{C}^\infty(M)$ we have $c'(t)(f) = T_t c(\frac{d}{dt}|_t)(f) = \frac{d}{dt}|_t(f \circ c)$ and consequently in coordinates $\varphi = (x^1, \dots, x^n)$ the local expression of the velocity vector takes the form $c'(t) = c'(t)(x^i)\partial_i|_{c(t)} = \frac{d}{dt}|_t(x^i \circ c)\partial_i|_{c(t)}$. (For more details see e.g. [12, 1.17 and below].)
In case M is a SRMF we may use the Levi-Civita covariant derivative to define the derivative Z' of $Z \in \mathfrak{X}(c)$ along the curve c.

1.3.27 Proposition (Induced covariant derivative). *Let $c : I \to M$ be a smooth curve into the SRMF M. Then there exists a unique mapping $\mathfrak{X}(c) \to \mathfrak{X}(c)$*

$$Z \mapsto Z' \equiv \frac{\nabla Z}{dt} \tag{1.3.51}$$

called the induced covariant derivative *such that* *(i)* $(Z_1 + \lambda Z_2)' = Z_1' + \lambda Z_2' \qquad (\lambda \in \mathbb{R})$,

(ii) $(hZ)' = \dfrac{dh}{dt}\, Z + h\, Z' \qquad (h \in \mathcal{C}^\infty(I, \mathbb{R}))$,

(iii) $(X \circ c)'(t) = \nabla_{c'(t)} X \qquad (t \in I,\ X \in \mathfrak{X}(M))$.

In addition we have

(iv) $\frac{d}{dt} \langle Z_1, Z_2 \rangle = \langle Z_1', Z_2 \rangle + \langle Z_1, Z_2' \rangle$.

Observe that in 1.3.27(iii) we have $X \circ c \in \mathfrak{X}(c)$ and since by 1.3.2(i) we have $\nabla_{c'(t)} X \in T_{c(t)} M$ also the right hand side makes sense.

Proof. (Local) uniqueness: Let $Z \mapsto Z'$ be a mapping as above that satisfies (i)–(iii) and let $(\phi = (x^1, \dots, x^n), U)$ be a chart and let $c : I \to U$ be a smooth curve. For $Z \in \mathfrak{X}(c)$ we then have

$$Z(t) = \sum_i Z(t)(x^i) \partial_i|_{c(t)} =: \sum_i Z^i(t) \partial_i|_{c(t)} \equiv Z^i(t) \partial_i|_{c(t)}. \tag{1.3.52}$$

By (i)–(iii) we then obtain

$$Z'(t) = \frac{dZ^i}{dt}\, \partial_i|_{c(t)} + Z^i(t)(\partial_i \circ c)' = \frac{dZ^i}{dt}\, \partial_i|_{c(t)} + Z^i(t) \nabla_{c'(t)} \partial_i. \tag{1.3.53}$$

So Z' is completely determined by the Levi-Civita connection ∇ and hence unique.

Existence: For any $J \subseteq I$ such that $c(J)$ is contained in a chart domain U we define the mapping $Z \mapsto Z'$ by equation (1.3.53). Then properties (i)–(iii) hold. Indeed (i) is obvious. Property (ii) follows from the following straight forward calculation

$$(hZ)'(t) = \frac{d(hZ^i)}{dt}\, \partial_i|_{c(t)} + h(t) Z^i(t) \nabla_{c'(t)} \partial_i = h(t) Z'(t) + \frac{dh}{dt} Z^i(t) \partial_i|_{c(t)}.$$

Finally to prove (iii) let $X \in \mathfrak{X}(M)$, $X|_U = X^i \partial_i$. Then (as explained prior to the proposition) $\frac{d}{dt}(X^i \circ c) = c'(t)(X^i)$ and so by (∇_3)

$$\begin{aligned} \nabla_{c'(t)} X = \nabla_{c'(t)}(X^i \partial_i) &= c'(t)(X^i) \partial_i|_{c(t)} + X^i(c(t)) \nabla_{c'(t)} \partial_i \\ &= \frac{d(X^i \circ c)}{dt}\, \partial_i|_{c(t)} + X^i \circ c(t) \nabla_{c'(t)} \partial_i = (X \circ c)'(t). \end{aligned} \tag{1.3.54}$$

Now suppose J_1, J_2 are two subintervals of I with corresponding maps $F_i : Z \mapsto Z'$. Then on $J_1 \cap J_2$ both F_i satisfy properties (i)–(iii) hence coincide by the uniqueness argument from above and we obtain a well-defined map on the whole of I.

Finally we obtain (iv) since we have using the chart (φ, U)

$$\langle Z_1', Z_2 \rangle + \langle Z_1, Z_2' \rangle = \frac{dZ_1^i}{dt}\, Z_2^j \langle \partial_i, \partial_j \rangle + Z_1^i Z_2^j \langle \nabla_{c'} \partial_i, \partial_j \rangle + Z_1^i \frac{dZ_2^j}{dt} \langle \partial_i, \partial_j \rangle + Z_1^i Z_2^j \langle \partial_i, \nabla_{c'} \partial_j \rangle$$

and on the other hand

$$\frac{d}{dt}\langle Z_1, Z_2\rangle = \frac{d}{dt}\big(Z_1^i Z_2^j \langle\partial_i, \partial_j\rangle\big) = \frac{dZ_1^i}{dt} Z_2^j\langle\partial_i, \partial_j\rangle + Z_1^i \frac{dZ_2^j}{dt}\langle\partial_i, \partial_j\rangle + Z_1^i Z_2^j \frac{d}{dt}\langle\partial_i, \partial_j\rangle.$$

The result now follows from from differentiating $\langle\partial_i, \partial_j\rangle$ (actually $\langle\partial_i, \partial_j\rangle \circ c$):

$$\frac{d}{dt}\langle\partial_i, \partial_j\rangle = c'(t)\langle\partial_i, \partial_j\rangle = \langle\nabla_{c'}\partial_i, \partial_j\rangle + \langle\partial_i, \nabla_{c'}\partial_j\rangle,$$

where we have used (∇5). □

We now write Z' in terms of the Christoffel symbols. In a chart $(\varphi = (x^1, \dots x^n), V)$ we have

$$\nabla_{c'(t)}\partial_i = \nabla_{\frac{d(x^j\circ c)}{dt}\partial_j}\partial_i = \frac{d(x^j \circ c)}{dt}\nabla_{\partial_j}\partial_i = \frac{d(x^j \circ c)}{dt}\Gamma^k_{ij}\partial_k \tag{1.3.55}$$

and hence

$$Z'(t) = \left(\frac{dZ^k}{dt}(t) + \Gamma^k_{ij}(c(t))\,\frac{d(x^j \circ c)}{dt}(t)\, Z^i(t)\right)\partial_k|_{c(t)}. \tag{1.3.56}$$

In the special case that $Z = c'$ we call $Z' = c''$ the *acceleration* of c. Also we call a vector field $Z \in \mathfrak{X}(c)$ *parallel* if $Z' = 0$. The above formula (1.3.56) shows that this condition actually is expressed by a system of linear ODEs of first order implying the following result.

1.3.28 Proposition (Parallel vector fields). *Let $c : I \to M$ be a smooth curve into a SRMF M. Let $a \in I$ and $z \in T_{c(a)}M$. Then there exists an unique parallel vector field $Z \in \mathfrak{X}(c)$ with $Z(a) = z$.*

Proof. As noted above locally Z obeys (1.3.56), which is a system of linear first order ODEs. Given an initial condition such an equation possesses a unique solution defined on the whole interval where the coefficient functions are given. Hence the claim follows from covering $c(I)$ by chart neighbourhoods. □

This result gives rise to the following notion.

1.3.29 Definition (Parallel transport). *Let $c : I \to M$ be a smooth curve into a SRMF M. Let $a, b \in I$ and write $c(a) = p$ and $c(b) = q$. For $z \in T_pM$ let Z_z be as in 1.3.28 with $Z_z(a) = z$. Then we call the mapping*

$$P = P_a^b(c) : T_pM \ni z \mapsto Z_z(b) \in T_qM \tag{1.3.57}$$

the parallel transport *(or* parallel translation*) along c from p to q.*

Finally we have the following crucial property of parallel translation.

1.3.30 Proposition. *Parallel transport is a linear isometry.*

Proof. Let $z, y \in T_pM$ with parallel vector fields Z_z, Z_y. Since then also $Z_y + Z_y$ and λZ_z (for $\lambda \in \mathbb{R}$) are parallel we have $P(z+y) = Z_z(b) + Z_y(b) = P(z) + P(y)$ and $P(\lambda z) = \lambda Z_z(b) = \lambda P(z)$ so that P is linear.
Let $P(z) = 0$ then by uniqueness $Z_z = 0$ hence also $z = 0$. So P is injective hence bijective. Finally $\langle Z_z, Z_y \rangle$ is constant along c since

$$\frac{d}{dt} \langle Z_z, Z_y \rangle = \langle Z_z', Z_y \rangle + \langle Z_z, Z_y' \rangle = 0 \tag{1.3.58}$$

and so

$$\langle P(z), P(y) \rangle = \langle Z_z(b), Z_y(b) \rangle = \langle Z_z(a), Z_y(a) \rangle = \langle z, y \rangle. \tag{1.3.59}$$

□

Chapter 2

Geodesics

In Euclidean space the shortest path between to arbitrary points is uniquely given by the straight line connecting these two points. That is, straight lines have two decisive properties: they are parallel, i.e., their velocity vector is parallely transported and they globally minimise length. Already in spherical geometry matters become more involved. The curves which possess a parallel velocity vector are the great circles (e.g. the meridians). Intuitively they are the 'straightest possible' lines on the sphere in the sense that their curvature equals the curvature of the sphere and so is the smallest possible. Also they minimize length but only locally. Indeed a great circle starting in a point p is initially minimising but stops to be minimising after it passes through the antipodal point $-p$. Also between antipodal points (e.g. the north and the south pole) there are infinitely many great circles which all have the same length.

In this section we study geodesics on SRMFs that is curves with parallel velocity vector and their respective properties. We introduce the exponential map, a main tool of SR geometry, which maps straight lines through the origin in the tangent space T_pM to so-called radial geodesics of the manifold thorough p. It is in turn used to define normal neighbourhoods of a point which have the property that any other point can be reached by a unique radial geodesic. Also the exponential map allows to introduce normal coordinates, that is charts which are well adapted to the geometry of the manifold. We prove the Gauss lemma which states that the exponential map is a radial isometry, i.e., it preserves angles with radial directions. We introduce the length functional and then turn to the Riemannian case. We define the Riemannian distance d between two points as the infimum of the length of all curves connecting these two points; it defines a metric in the topological sense on M. Now the Gauss lemma can be seen to say that small distance spheres centered at a point are perpendicular to radial geodesics and that radial geodesics minimise length. We prove the Hopf-Rinow theorem which says that the Riemannian distance function encodes the topological and also the metric structure of the manifold. As a consequence in a complete Riemannian manifold every pair of points can be joined by a minimising geodesic. We finally remark the differences to the Lorentzian case which indeed is much more involved.

2.1 Geodesics and the exponential map

In this subsection we generalise the notion of a straight line in Euclidean space. As in [10, Sec. 3.3] we define a *geodesic* to be a curve c such that its tangent vector c' is parallel along c. Equivalently we have for the acceleration $c'' = 0$. Locally this condition translates into a system of nonlinear ODEs of second order. More precisely we have:

2.1.1 Proposition (Geodesic equation). *Let $(\varphi = (x^1, \dots, x^n), U)$ be a chart of the SRMF M and let $c : I \to U$ be a smooth curve. Then c is a geodesic iff the local coordinate expressions $x^k \circ c$ of c obey the* geodesic equation,

$$\frac{d^2(x^k \circ c)}{dt^2} + \Gamma^k_{ij} \circ c \, \frac{d(x^i \circ c)}{dt} \, \frac{d(x^j \circ c)}{dt} = 0 \qquad (1 \leq k \leq n). \tag{2.1.1}$$

Proof. The curve c is a geodesic iff $(c')' = 0$. We have $c'(t) = \frac{d(x^k \circ c)}{dt} \partial_k|_{c(t)}$ and inserting $\frac{d(x^k \circ c)}{dt}$ for Z^k in (1.3.56) we obtain (2.1.1). □

It is most common to abbreviate the local expressions $x^k \circ c$ of c by c^k. Using this notation the geodesic equation (2.1.1) takes the form

$$\frac{d^2 c^k}{dt^2} + \Gamma^k_{ij} \frac{dc^i}{dt} \frac{dc^j}{dt} = 0 \quad \text{or even shorter} \quad \ddot{c}^k + \Gamma^k_{ij} \dot{c}^i \dot{c}^j = 0 \qquad (1 \leq k \leq n). \tag{2.1.2}$$

Obviously the geodesic equation is a system of nonlinear ODEs of second order and so by basic ODE-theory we obtain the following result on existence and uniqueness of geodesics.

2.1.2 Lemma (Existence of geodesics). *Let $p \in (M, g)$ and let $v \in T_pM$. Then there exists a real interval I around 0 and a unique geodesic $c : I \to M$ with $c(0) = p$ and $c'(0) = v$.*

We call c as in 2.1.2 (irrespectively of the interval I) the geodesic starting a p with initial velocity v.

2.1.3 Examples (Geodesics of flat space). The geodesic equations in $\mathbb{R}^n_r$ are trivial, i.e., they take the form $\frac{d^2 c^k}{dt^2} = 0$ since all Christoffel symbols Γ^k_{ij} vanish. Hence the geodesics are the straight lines $c(t) = p + tv$.

2.1.4 Examples (Geodesic on the cylinder). Let $M \in \mathbb{R}^3$ be the cylinder of radius 1 and ψ the chart $(\cos\varphi, \sin\varphi, z) \mapsto (\varphi, z)$ $(\varphi \in (0, 2\pi))$. The natural basis of T_pM with $(p = (\cos\varphi, \sin\varphi))$ w.r.t. ψ is then given by (cf. [10, 2.4.11])

$$\partial_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}, \quad \partial_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.1.3}$$

Also we have $g_{11} = \langle \partial_\varphi, \partial_\varphi \rangle = 1$, $g_{12} = g_{21} = \langle \partial_\varphi, \partial_z \rangle = 0$, $g_{22} = \langle \partial_z, \partial_z \rangle = 1$ and so

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad (g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.1.4}$$

which immediately implies that $\Gamma^i_{\ jk} = 0$ for all i, j, k. Hence the geodesic equations for a curve $c(t) = (\cos\varphi(t), \sin\varphi(t), z(t))$ using the notation $c^1(t) = x^1 \circ c(t) = \varphi(t)$, $c^2(t) = x^2 \circ c(t) = z(t)$ take the form

$$\ddot{\varphi}(t) = 0, \qquad \ddot{z}(t) = 0, \tag{2.1.5}$$

which are readily solved to obtain $\varphi(t) = a_1 t + a_0$, $z(t) = b_1 t + b_0$. So we find

$$c(t) = \Big(\cos(a_1 t + a_0), \sin(a_1 t + a_0), b_1 t + b_0) \tag{2.1.6}$$

revealing that the geodesics of the cylinder are helices with initial point and speed given by the a_i, b_i, see figure 2.1, left. This also includes the extreme cases of circles of latitude $z = c$ ($b_1 = 0$, $b_0 = c$) and generators ($a_1 = 0$). Another way to see that these are the

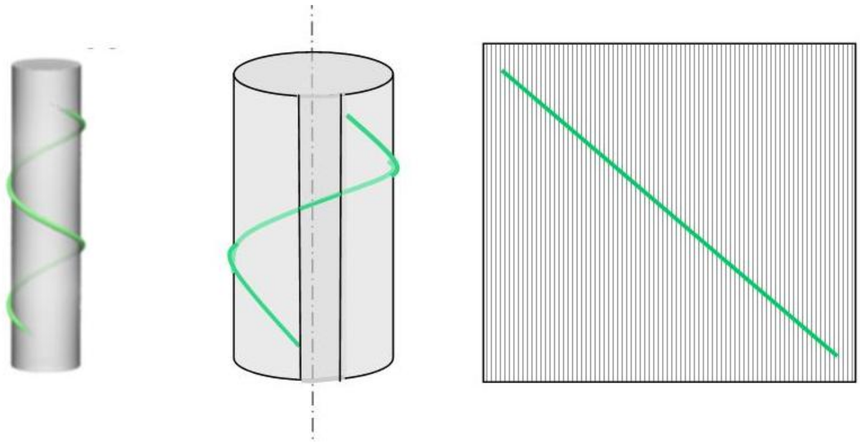

Figure 2.1: Geodesics on the cylinder

geodesics of the cylinder is by 'unwrapping' the cylinder to the plane, see Figure 2.1, right. This of course amounts to applying the chart ψ, which in this case is an isometry since $g_{ij} = \delta_{ij}$. Now the geodesics of $\mathbb{R}^2$ are straight lines which via ψ^{-1} are wrapped as helices onto the cylinder.

We now return to the general study of geodesics. Being solutions of a nonlinear ODE geodesics will in general not be defined for all values of their parameter. However, ODE theory provides us with unique maximal(ly extended) solutions. To begin with we have:

2.1.5 Lemma (Uniqueness of geodesics). *Let $c_1, c_2 : I \to M$ geodesics. If there is $a \in I$ such that $c_1(a) = c_2(a)$ and $\dot{c}_1(a) = \dot{c}_2(a)$ then we already have $c_1 = c_2$.*

Proof. Suppose under the hypothesis of the lemma there is $t_0 \in I$ such that $c_1(t_0) \neq c_2(t_0)$. Without loss of generality we may assume $t_0 > a$ and set $b = \inf\{t \in I : \ t > a$ and $c_1(t) \neq c_2(t)\}$. We now argue that $c_1(b) = c_2(b)$ and $\dot{c}_1(b) = \dot{c}_2(b)$. Indeed if $b = a$ this follows by assumption. Also in case $b > a$ we have that $c_1 = c_2$ on (a, b) and the claim follows by continuity.
Now since also $t \mapsto c_i(b + t)$ are geodesics $(i = 1, 2)$ Lemma 2.1.2 implies that c_1 and c_2 agree on a neighbourhood of b which contradicts the definition of b. □

2.1.6 Proposition (Maximal geodesics). *Let $v \in T_pM$. Then there exists a unique geodesic c_v such that*

(i) $c_v(0) = p$ *and* $c_v'(0) = v$

(ii) the domain of c_v *is maximal, that is if* $c : J \to M$, $0 \in J$ *is any geodesic with* $c(0) = p$ *and* $c'(0) = v$ *then* $J \subseteq I$ *and* $c_v|_J = c$.

Proof. Let $G := \{c : I_c \to M$ a geodesic: $0 \in I_c, \ c(0) = p, \ c'(0) = v\}$, then by 2.1.2 $G \neq \emptyset$ and by 2.1.5 we have for any pair $c_1, \ c_2 \in G$ that $c_1|_{I_{c_1} \cap I_{c_2}} = c_2|_{I_{c_1} \cap I_{c_2}}$. So the geodesics in G define a unique geodesic satisfying the assertions of the statement. □

2.1.7 Definition (Completeness). *We call a geodesic c_v as in 2.1.6* maximal*. If in a SRMF M all maximal geodesics are defined on the whole of $\mathbb{R}$ we call M* geodesically complete*.*

2.1.8 Examples (Complete manifolds).

(i) $\mathbb{R}^n_r$ is geodesically complete as well as the cylinder of 2.1.4.

(ii) $\mathbb{R}^n_r \setminus \{0\}$ is not geodesically complete since all geodesics of the form $t \mapsto tv$ are defined either on $\mathbb{R}^+$ or $\mathbb{R}^-$ only.

We next turn to the study of the causal character of geodesics. We begin with the following definition.

2.1.9 Definition (Causal character of curves). *A curve c into a SRMF M is called* spacelike, timelike *or* null *if for all t its velocity vector $c'(t)$ is spacelike, timelike or null, respectively. We call c* causal *if it is timelike or null. These properties of c are commonly referred to as its* causal character*.*

In general a curve need not to have a causal character, i.e., its velocity vector could change its causal character along the curve. However, geodesics do have a causal character: Indeed if c is a geodesic then by definition c' is parallel along c. But parallel transport by 1.3.28 is an isometry so that $\langle c'(t), c'(t)\rangle = \langle c'(t_0), c'(t_0)\rangle$ for all t. This fact can also be seen directly by the following simple calculation

$$\frac{d}{dt}\langle c'(t), c'(t)\rangle = 2\langle c''(t), c'(t)\rangle = 0. \tag{2.1.7}$$

Moreover we also clearly have that the speed of a geodesic is constant, i.e., $\|c'(t)\| = \|c'(t_0)\|$ for all t. The following technical result is significant.

2.1.10 Lemma (Geodesic parametrisation). *Let $c : I \to M$ be a nonconstant geodesic. A reparametrisation $c \circ h : J \to M$ of c is a geodesic iff h is of the form $h(t) = at + b$ with $a, b \in \mathbb{R}$.*

Proof. To begin with note that $(c\circ h)'(t) = c'(h(t))h'(t)$ since $(c\circ h)'(t) = T_t(c\circ h)(\frac{d}{dt}|_t) = T_{h(t)}c(T_t h(\frac{d}{dt}|_t)) = T_{h(t)}c(h'(t)\frac{d}{dt}|_t)) = h'(t)T_{h(t)}c(\frac{d}{dt}|_t) = h'(t)c'(h(t))$.
Now, if $Z \in \mathfrak{X}(c)$ then $Z\circ h \in \mathfrak{X}(c\circ h)$ and from (1.3.56) we have $(Z\circ h)'(t) = Z'(h(t))h'(t)$. Applying this to $Z = c'$, we obtain with 1.3.27(ii)

$$(c\circ h)''(t) = \Big(c'(h(t))h'(t)\Big)' = h''(t)c'(h(t)) + h'(t)^2\underbrace{c''(h(t))}_{=0}. \tag{2.1.8}$$

Now since c is nonconstant we have $c'(t) \neq 0$ and we obtain

$c\circ h$ is a geodesic $\Leftrightarrow$ $(c\circ h)'' = 0$ $\Leftrightarrow$ $h'' = 0$ $\Leftrightarrow$ $h(t) = at + b$. □

This result shows that the parametrisation of a geodesic has a geometric meaning. More generally a curve that has a reparametrisation as a geodesic is called a *pregeodesic*.

Next we turn to a deeper analysis of the geodesic equations as a system of second order ODEs. The first result is concerned with the dependence of a geodesic on its initial speed and is basically a consequence of smooth dependence of solutions of ODEs on the data.

2.1.11 Lemma (Dependence on the initial speed). *Let $v \in T_pM$ then there exists a neighbourhood $\mathcal{N}$ of v in TM and an interval I around 0 such that the mapping*

$$\mathcal{N} \times I \ni (w, s) \mapsto c_w(s) \in M \tag{2.1.9}$$

is smooth.

Proof. c_w is the solution of the second order ODE (2.1.1) which depends smoothly on s as well as on $c_w(0) =: p$ and $c_w'(0) = w$. This follows from ODE theory e.g. by rewriting (2.1.1) as a first order system as in [10, 2.5.16]. □

Our next aim is to make the rewriting of the geodesic equations as a first order system explicit. To this end let $(\psi = (x^1, \dots x^n), V)$ be a chart. Then $T\psi : TV \to \psi(V) \times \mathbb{R}^n$,

$T\psi = (x^1, \dots, x^n, y^1, \dots, y^n)$ is a chart of TM. Now $c : I \to V$ is a geodesic iff $t \mapsto (c^1(t), \dots, c^n(t), y^1(t), \dots, y^n(t))$ solves the following first order system

$$\begin{aligned} \frac{dc^k}{dt} &= y^k(t) \\ \frac{dy^k}{dt} &= -\Gamma^k_{ij}(x(t))\, y^i(t)\, y^j(t) \qquad (1 \le k \le n). \end{aligned} \tag{2.1.10}$$

Locally (2.1.10) is an ODE on TM since its right hand side is a vector field on TV, i.e., an element of $\mathfrak{X}(TV)$. The geodesics hence correspond to the flow lines of such a vector field. More precisely we have:

2.1.12 Theorem (Geodesic flow). *There exists a uniquely defined vector field $G \in \mathfrak{X}(TM)$, the so-called* geodesic field *or* geodesic spray *with the following property: The projection $\pi : TM \to M$ establishes a one-to-one correspondence between (maximal) integral curves of G and (maximal) geodesics of M.*

Proof. Given $v \in TM$ the mapping $s \mapsto c_v'(s)$ is a smooth curve in TM. Let $G_v := G(v)$ be the initial speed of this curve, i.e., $G_v := \frac{d}{ds}|_0(c_v'(s)) \in T_v(TM)$ (since $c_v'(0) = v$). Now by 2.1.11 $G \in \mathfrak{X}(TM)$. We now prove the following two statements:

(i) If c is a geodesic of M then c' is an integral curve of G.
Indeed let $\alpha(s) := c'(s)$ and for any fixed t set $w := c'(t)$ and $\beta(s) := c_w'(s)$. By 2.1.5 we have $c(t+s) = c_w(s)$ (since $c_w(0) = c(t)$ and $c_w'(0) = w = c'(t)$). Differentiation w.r.t. s yields $\alpha(s+t) = c_w'(s) = \beta(s)$. So we have in $T(TM)$ that $\alpha'(t+s) = \beta'(s)$. In particular, $\alpha'(t) = \beta'(0) = G_w = G_{\alpha(t)}$ and hence α is an integral curve of G.

(ii) If α is an integral curve of G then $\pi \circ \alpha$ is a geodesic of M.
Let $\alpha : I \to TM$ and $t \in I$. Since $\alpha(t) \in TM$ we have by (i) that the map $s \mapsto c_{\alpha(t)}'(s)$ an integral curve of G. For $s = 0$ both integral curves $c_{\alpha(t)}'(s)$ and $s \mapsto \alpha(s+t)$ attain the value $\alpha(t)$. Hence by [10, 2.5.17] they coincide on their entire domain. So we obtain for all s that

$$\alpha(s+t) = c_{\alpha(t)}'(s) \;\Rightarrow\; \pi(\alpha(s+t)) = c_{\alpha(t)}(s) \;\Rightarrow \pi \circ \alpha \quad \text{is a geodesis of } M.$$

Finally we have $\pi \circ c' = c$ and $(\pi\alpha)'(t) = \frac{d}{ds}|_0 \pi(\alpha(t+s)) = \frac{d}{ds}|_0 c_{\alpha(t)}(s) = c_{\alpha(t)}'(0) = \alpha(t)$ and so the maps $c \mapsto c'$ and $\alpha \mapsto \pi \circ \alpha$ are inverse to each other. Also G is unique since its integral curves are prescribed. □

The flow of the vector field G in the above Theorem is called *the geodesic flow* of M. Next we will introduce the *exponential* map which is one of *the* essential tools of semi-Riemannian geometry.

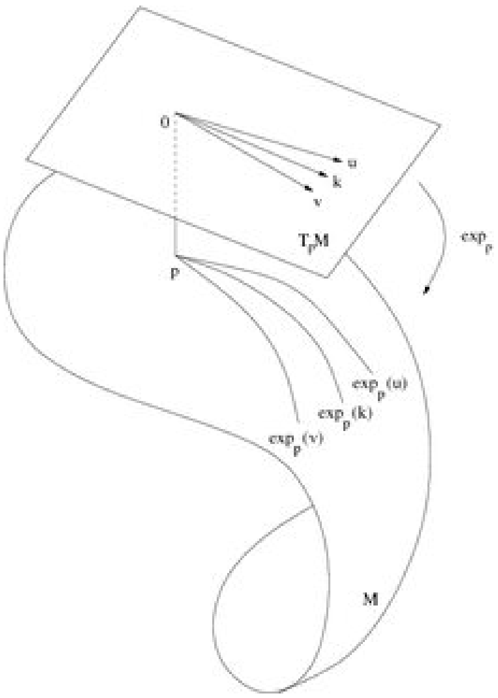

Figure 2.2: $\exp_p$ maps straight lines through $0 \in T_pM$ to geodesics through p.
♣ `Also indicate` U `and` $\tilde{U}$. ♣

2.1.13 Definition (Exponential map). *Let p be a point in the SRMF M and set $\mathcal{D}_p := \{v \in T_pM : c_v$ is at least defined on $[0,1]\}$. The* exponential map *of M at p is defined as*

$$\exp_p : \mathcal{D}_p \to M, \qquad \exp_p(v) := c_v(1). \tag{2.1.11}$$

Observe that $\mathcal{D}_p$ is the maximal domain of $\exp_p$. In case M is geodesically complete we have $\mathcal{D}_p = T_pM$ for all p.
Let now $v \in T_pM$ and fix $t \in \mathbb{R}$. The geodesic $s \mapsto c_v(ts)$ (cf. 2.1.10) has initial speed $tc_v'(0) = tv$ and so we have

$$c_{tv}(s) = c_v(ts) \tag{2.1.12}$$

for all t, s for which one and hence both sides of (2.1.12) are defined. This implies for the exponential map that

$$\exp_p(tv) = c_{tv}(1) = c_v(t) \tag{2.1.13}$$

which has the following immediate geometric interpretation: the exponential map $\exp_p$ maps straight lines $t \mapsto tv$ through the origin in T_pM to geodesics $c_v(t)$ through p in M, see Figure 2.2. This is actually done in a diffeomorphic way.

2.1.14 Theorem (Exponential map). *Let p be a point in a SRMF M. Then there exist neighbourhoods $\tilde{U}$ of 0 in T_pM and U of p in M such that the exponential map $\exp_p : \tilde{U} \to U$ is a diffeomorphism.*

Proof. The plan of the proof is to first show that $\exp_p$ is smooth on a suitable neighbourhood of $0 \in T_pM$ and then to apply the inverse function theorem.
To begin with we recall that by 2.1.11 the mapping $(w, s) \mapsto c_w(s)$ is smooth on $\mathcal{N} \times I$, where $\mathcal{N}$ is an open neighbourhood in TM and I is an interval around 0 which we assume w.l.o.g. to be $I = (-a, a)$. Now set $\mathcal{N}_p := \{\frac{v}{a'} : v \in \mathcal{N} \cap T_pM\}$, where a' is a fixed number with $a' > 1/a$. By (2.1.12) the map $(w, s) \mapsto c_w(s)$ is then defined on $\mathcal{N}_p \times J$ where $J \supseteq [0, 1]$. Indeed for $w = v/a' \in \mathcal{N}_p$ we have $c_w(s) = c_{v/a'}(s) = c_v(s/a')$ with $s/a' \in I = (-a, a)$ and so $s \in a'I \supseteq [0, 1]$. So in total $\exp_p : \mathcal{N}_p \to M$ is smooth.
We next show that $T_0(\exp_p) : T_0(T_pM) \to T_pM$ is a linear isomorphism. To this end let $v \in T_0(T_pM)$ which we may identify with T_pM (cf. [10, 2.4.10]). Set $\rho(t) = tv$. Then $v = \rho'(0) = T_0\rho(\frac{d}{dt}|_0)$ and so

$$T_0(\exp_p)(v) = T_0(\exp_p)(\rho'(0)) = T_0(\exp_p \circ\rho)(\frac{d}{dt}|_0) = T_0(c_v)(\frac{d}{dt}|_0) = c_v'(0) = v. \quad (2.1.14)$$

Hence $T_0 \exp_p = \mathrm{id}_{T_pM}$ and the claim now follows from the inverse function theorem, see e.g. [10, 1.1.5]. □

A subset S of a vector space is called *star shaped* (around 0) if $v \in S$ and $t \in [0, 1]$ implies that $tv \in S$. ♣ `Insert figure!` ♣ If U and $\tilde{U}$ are as in 2.1.14 and $\tilde{U}$ is star shaped then we call U a *normal neighbourhood* of p. In this case U is star shaped in the following sense.

2.1.15 Proposition (Radial geodesics). *Let U be a normal neighbourhood of p. Then for each $q \in U$ there exists a unique geodesic $\sigma : [0, 1] \to U$ from p to q called the* radial geodesic *from p to q. Moreover we have $\sigma'(0) = \exp_p^{-1}(q) \in \tilde{U}$.*

Proof. By assumption $\tilde{U} \subseteq T_pM$ is star shaped and $\exp_p : \tilde{U} \to U$ is a diffeomorphisms. Let $q \in U$ and set $v := \exp_p^{-1}(q) \in \tilde{U}$. Since $\tilde{U}$ is star shaped we have that $\rho(t) = tv \in \tilde{U}$ for $t \in [0, 1]$. Hence $\sigma = \exp_p \circ\rho$ which by (2.1.13) is a geodesic and connects p with q is contained in U. Moreover we have by (2.1.14)

$$\sigma'(0) = T_0(\exp_p)(\rho'(0)) = T_0 \exp_p(v) = v = \exp_p^{-1}(q). \quad (2.1.15)$$

♣ `Insert figure of radial geodesic` ♣ Let now $\tau : [0, 1] \to U$ be an arbitrary geodesic in U connecting p with q. Set $w := \tau'(0)$. Then the geodesics τ and $t \mapsto \exp_p(tw)$ both have the same velocity vector at p hence coincide by 2.1.5.
We show that $w \in \tilde{U}$. Suppose to the contrary that $w \notin \tilde{U}$ and set $\tilde{t} := \sup\{t \in [0, 1] : tw \in \tilde{U}\}$. Then $\tilde{t}w \in \partial\tilde{U}$. Now $\tau([0, 1]) \subseteq U$ is compact and so is $(\exp_p |\tilde{U})^{-1}(\tau([0, 1]))$ in $\tilde{U}$ and hence it has a positive distance to $\partial\tilde{U}$ and hence to $\tilde{t}w$. Now by the definition of the supremum there is $t_0 < \tilde{t}$ arbitrarily close to $\tilde{t}$ such that $\tilde{U} \ni t_0w \notin (\exp_p |\tilde{U})^{-1}(\tau([0, 1]))$.

But then $\exp_p(t_0 w) \notin \tau([0,1])$, which contradicts the fact that $\tau = t \mapsto \exp_p(tw)$ and so $w \in \tilde{U}$[1].
Finally we have $\exp_p(w) = \tau(1) = q = \exp_p(v)$ and by injectivity of the exponential map $v = w$. But this implies that $\sigma = \tau$. □

By a *broken geodesic* we mean a piecewise smooth curve whose smooth parts are geodesics. In $\mathbb{R}^n_r$ these are just the polygons. We now can prove the following criterion for connectedness.

2.1.16 Lemma (Connectedness via broken geodesics). *A SRMF is connected iff every pair of points may be joined by a broken geodesic.*

Proof. The condition is obviously sufficient. Let now M be connected and choose $p \in M$. Set $C := \{q \in M :\ q$ can be connected to p by a broken geodesic$\}$. Let now $q \in M$ and U be a normal neighbourhood of q. If $q \in C$ then by 2.1.15 $U \subseteq C$ hence C is open. Also if $q \notin C$ then $U \subseteq M \setminus C$ and so C is also closed. Hence $C = M$. □

Normal neighbourhoods are particularly useful in constructing special coordinate systems, called *Riemannian normal coordinates (RNC)* which are of great importance for explicit calculations. Let $p \in M$ with $U = \exp_p(\tilde{U})$ a normal neighbourhood of p and let $\mathcal{B} = \{e_1, \dots, e_n\}$ be an orthonormal basis of T_pM. The Riemannian normal coordinate system $(\varphi = (x^1, \dots, x^n), U)$ around p defined by $\mathcal{B}$ assigns to any point $q \in U$ the coordinates of $\exp_p^{-1}(q) \in \tilde{U} \subseteq T_pM$ w.r.t. $\mathcal{B}$, i.e.,

$$\exp_p^{-1}(q) = x^i(q) e_i. \tag{2.1.16}$$

If $\mathcal{B}' = \{f^1, \dots, f^n\}$ is the dual basis of $\mathcal{B}$ then we have

$$x^i \circ \exp_p = f^i \quad \text{on } \tilde{U}. \tag{2.1.17}$$

Indeed set $q = \exp_p(w)$ in (2.1.16) then $w = x^i(\exp_p(w)) e_i$.
The most important properties of RNCs around p are that in this coordinates the metric at p is precisely the flat metric and also at p all Christoffel symbols vanish. It is essential to point out that these properties only hold at the point in which the coordinates are centered and in general fail already arbitrarily near to p. However, tensor fields are defined pointwise and so in many situations it is very beneficial to check certain tensorial identities in (the center point of) RNCs. More precisely we have:

2.1.17 Proposition (Normal coordinates). *Let $x^1, \dots, x^n$ be RNC around p. Then we have for all i, j, k*

[1] Observe that this argument even shows that the entire segment $\{tw | t \in [0,1]\}$ lies in $\tilde{U}$. Just replace w by some $\tilde{w} = sw$ for $s \in [0,1]$. This fact will be used in ♣ `ref!` ♣, below.

(i) $g_{ij}(p) = \delta_{ij}\varepsilon_j$, *and*

(ii) $\Gamma^i_{jk}(p) = 0$.

Proof. We first show that $\partial_i|_p = e_i$. Let $v \in \tilde{U} \subseteq T_pM$, $v = a^i e_i$. By (2.1.13), (2.1.17) we have

$$x^i(c_v(t)) = x^i(\exp_p(tv)) = f^i(tv) = ta^i \tag{2.1.18}$$

and so by (2.1.16), $\exp_p^{-1} \circ c_v(t) = (a^1 t, \dots, a^n t)$. Hence we have $T_p \exp_p^{-1}(c_v'(0)) = (a^1, \dots a^n)$ which gives $v = c_v'(0) = a^i \partial_i|_p$. Since $v = a^i e_i$ was arbitrary we obtain $\partial_i|_p = e_i$ which immediately gives (i).
Now since $x^i \circ c_v(t) = ta^i$ the geodesic equation for c_v reduces to $\Gamma^k_{ij}(c_v(t))a^i a^j = 0$ for all k. Inserting $t = 0$ we have $\Gamma^k_{ij}(p)a^j a^i = 0$ for all $a = (a^1, \dots a^n) \in \mathbb{R}^n$. So for fixed k the quadratic form $a \mapsto \Gamma^k_{ij}(p)a^i a^j$ vanishes and by polarisation we find $(a, b) \mapsto \Gamma^k_{ij}(p)a^i b^j = 0$ and so $\Gamma^k_{ij}(p) = 0$, hence (ii) holds. □

2.1.18 Examples (Exponential map of $\mathbb{R}^n_r$). Let $v \in T_p(\mathbb{R}^n_r)$, then the geodesic c_v starting at p is just $t \mapsto p + tv$. Hence we have $\exp_p : v \mapsto c_v(1) = p + v$. This is a *global* diffeomorphism and even an isometry.

Our next goal is to prove an essential result that goes by the name of Gauss lemma and states that the exponential map is a 'radial isometry'. This means that the orthogonality to radial directions is preserved. We first need some preparations.

2.1.19 Definition (Two-parameter mappings). *Let $\mathcal{D} \subseteq \mathbb{R}^2$ be open and such that vertical and horizontal straight lines intersect $\mathcal{D}$ in intervals. A* two-parameter mapping *on $\mathcal{D}$ is a smooth map $f : \mathcal{D} \to M$.*

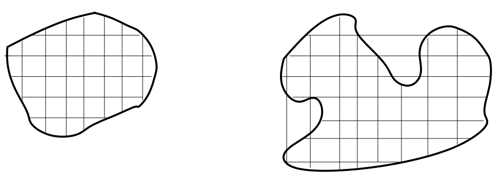

Figure 2.3: Sets $\mathcal{D}$ that (fail to) have the property of 2.1.19

For examples of sets that obey respectively lack the above property see Figure 2.3. Two-parameter maps are also often called *singular surfaces* since there is no condition on the rank of f.
Denoting the coordinates in $\mathbb{R}^2$ by (t, s) then a two-parameter map f defines two families of smooth curves, the t-parameter curves $s = s_0 : \; t \mapsto f(t, s_0)$ and the s-parameter

curves $t = t_0 : \; s \mapsto f(t_0, s)$. By definition all such curves are defined on intervals. The corresponding partial derivatives

$$f_t(t,s) := T_{(t,s)} f(\partial_t) \quad \text{and} \quad f_s(t,s) := T_{(t,s)} f(\partial_s) \tag{2.1.19}$$

are then vector fields along f in the sense of Definition 1.3.26. Observe that $f_t(t_0, s_0)$ is the velocity of the t-parameter curve $s = s_0$ at t_0 and analogous for $f_s(t_0, s_0)$.
If the image of f is contained in a chart $((x^1, \dots, x^n), V)$ of M then we denote the coordinate functions $x^i \circ f$ of f by f^i $(1 \le i \le n)$. We then have $f^i = x^i \circ f : \mathcal{D} \to \mathbb{R}$ and by [10, 2.4.14]

$$f_t = \frac{\partial f^i}{\partial t}\, \partial_i \quad \text{and} \quad f_s = \frac{\partial f^i}{\partial s}\, \partial_i. \tag{2.1.20}$$

Now let M be a SRMF and $Z \in \mathfrak{X}(f)$. Then $t \mapsto Z(t, s_0)$ and $s \mapsto Z(t_0, s)$ are vector fields along the t- and s-parameter curves $t \mapsto f(t, s_0)$ and $s \mapsto f(t_0, s)$, respectively. We denote the corresponding induced covariant derivatives by

$$Z_t = \frac{\nabla Z}{\nabla t} = \nabla_{\partial_t} Z \quad \text{and} \quad Z_s = \frac{\nabla Z}{\nabla s} = \nabla_{\partial_s} Z, \tag{2.1.21}$$

respectively. By (**??**) we have

$$\nabla_{\partial_t} Z(t,s) = Z_t(t,s) = \left(\frac{\partial Z^k}{\partial t}(t,s) + \Gamma^k_{ij}(f(t,s))\, Z^i(t,s)\, \frac{\partial f^j}{\partial t}(t,s) \right) \partial_k \tag{2.1.22}$$

and analogously for Z_s. In particular, for $Z = f_t$ we call $Z_t = f_{tt}$ the acceleration of the t-parameter curve and analogously for f_{ss}. We now note the following essential fact.

2.1.20 Lemma (Mixed second derivatives of 2-parameter maps commute). *Let M be a SRMF and $f : \mathcal{D} \to M$ a 2-parameter map. Then we have $\nabla_{\partial_t}(\partial_s f) = \nabla_{\partial_s}(\partial_t f)$ or for short $f_{ts} = f_{st}$.*

Proof. By (2.1.22) we have

$$f_{ts} = \left(\frac{\partial^2 f^k}{\partial t\, \partial s} + \Gamma^k_{ij} \circ f\, \frac{\partial f^i}{\partial t}\, \frac{\partial f^j}{\partial s} \right) \partial_k, \tag{2.1.23}$$

which by the symmetry of the Christoffel symbols is symmetric in i and j. □

As a final preparation we consider $x \in T_pM$. Since T_pM is a finite dimensional vector space we may identify $T_x(T_pM)$ with T_pM itself, cf. [10, 2.4.10]. Hence if $v_x \in T_x(T_pM)$ we will view v_x also as an element of T_pM. We call v_x *radial* if is a multiple of x.
Now we finally may state and prove the following result.

2.1.21 Theorem (Gauss lemma). *Let M be a SRMF and let $p \in M$, $0 \neq x \in \mathcal{D}_p \subseteq T_pM$. Then for any v_x, $w_x \in T_x(T_pM)$ with v_x radial we have*

$$\langle (T_x \exp_p)(v_x), (T_x \exp_p)(w_x) \rangle = \langle v_x, w_x \rangle. \tag{2.1.24}$$

Proof. Since v_x is radial and (2.1.24) is linear we may suppose w.l.o.g. that $v_x = x$ and to simplify notations we choose to denote this vector by v. Also we write w instead of w_x. Let now

$$f(t, s) := \exp_p(t(v + sw)). \tag{2.1.25}$$

The mapping $(t, s) \mapsto t(v + sw)$ is continuous and maps $[0, 1] \times \{0\}$ into the set $\mathcal{D}_p$, the domain of $\exp_p$ in T_pM. By 2.2.3, below $\mathcal{D}_p$ is open and so there is $\varepsilon, \delta > 0$ such that f is defined on the square $\mathcal{D} := \{(t, s) :\ -\delta < t < 1 + \delta,\ -\varepsilon < s < \varepsilon\}$ (see figure 2.4) and so f is a 2-parameter mapping.

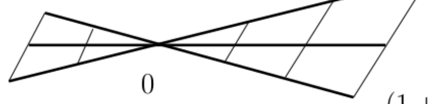

Figure 2.4: The shape of $h(\mathcal{D})$ with $h(t, s) = t(v + sw))$.♣ `correct figure` ♣

We now have $f_t(1, 0) = T_v \exp_p(v)$ and $f_s(1, 0) = T_v \exp_p(w)$ and so we have to show that $\langle f_t(1, 0), f_s(1, 0) \rangle = \langle v, w \rangle$.
The curves $t \mapsto f(t, s)$ are geodesics with initial speed $v + sw$. Hence $f_{tt} = 0$ and so $\langle f_t, f_t \rangle = \text{const} = \langle v + sw, v + sw \rangle$ since $f_t(0, s) = T_0 \exp_p(v + sw) = v + sw$.
Moreover by 2.1.20 $f_{st} = f_{ts}$ and so we have

$$\begin{aligned} \frac{\partial}{\partial t} \langle f_s, f_t \rangle &= \langle f_{st}, f_t \rangle + \langle f_s, \underbrace{f_{tt}}_{=0} \rangle = \langle f_{ts}, f_t \rangle \\ &= \frac{1}{2} \frac{\partial}{\partial s} \langle f_t, f_t \rangle = \frac{1}{2} \frac{\partial}{\partial s} \langle v + sw, v + sw \rangle = \langle v, w \rangle + s \langle w, w \rangle \end{aligned} \tag{2.1.26}$$

which implies

$$\left(\frac{\partial}{\partial t} \langle f_s, f_t \rangle \right)(t, 0) = \langle v, w \rangle \quad \text{for all } t. \tag{2.1.27}$$

Now $f(0, s) = \exp_p(0) = p$ for all s and so $f_s(0, 0) = 0$ which gives $\langle f_s, f_t \rangle(0, 0) = 0$. Now integrating (2.1.27) yields $\langle f_s, f_t \rangle(t, 0) = t\langle v, w \rangle$. Finally setting $t = 1$ we obtain $\langle f_t(1, 0), f_s(1, 0) \rangle = \langle v, w \rangle$. □

2.2 Geodesic convexity

Geodesically convex sets, which sometimes are also called totally normal, are normal neighbourhoods of all of their points. In this section we are going to prove existence of such sets around each point in a SRMF. The arguments will rest on global properties of the exponential map. To formulate these we need to introduce product manifolds as a preparation. We skip the obvious proof of the following lemma.

2.2.1 Definition & Lemma (Product manifold). *Let M^m and N^n be smooth manifolds of dimension m and n respectively. Let $(p,q) \in M \times N$ and let $(\varphi = (x^1, \dots, x^m), U)$ and $(\psi = (y^1, \dots, y^n), V)$ be charts of M around p and of N around q, respectively. Then we call $(\varphi \times \psi, U \times V)$ a* product chart *of $M \times N$ around (p,q). The family of all product charts defines a $\mathcal{C}^\infty$-structure on $M \times N$ and we call the resulting smooth manifold the* product manifold *of M with N. The dimension of $M \times N$ is $m + n$.*

2.2.2 Remark (Properties of product manifolds). Let M^m and N^n be $\mathcal{C}^\infty$-manifolds.

(i) The natural manifold topology of $M \times N$ is precisely the product topology of M and N since the $\varphi \times \psi$ are homeomorphisms by definition.

(ii) The projections $\pi_1 : M \times N \to M$, $\pi_1(p,q) = p$ and $\pi_2 : M \times N \to N$, $\pi_2(p,q) = q$ are smooth since $\varphi \circ \pi_1 \circ (\varphi \times \psi)^{-1} = \mathrm{pr}_1 : \varphi(U) \times \psi(V) \to \varphi(U)$ and similarly for π_2. From this local representation we even have that π_1, π_2 are surjective submersions. By ♣ `properly cite Cor. 1.1.23 in submanifold part of ANAMF` ♣ if follows that for $(p,q) \in M \times N$ the sets $M \times \{q\} = \pi_2^{-1}(\{q\})$ and $\{p\} \times N = \pi_1^{-1}(\{p\})$ are closed submanifolds of $M \times N$ of dimension m and n, respectively. Moreover the bijection $\pi_1|_{M\times\{q\}} : M \times \{q\} \to M$ is a diffeomorphism by ♣ `properly cite Thm. 1.1.5 in submanifold part of ANAMF` ♣ and analogously for $\pi_2|_{\{p\}\times N}$.

(iii) A mapping $f : P \to M \times N$ from a smooth manifold P into the product is smooth iff $\pi_1 \circ f$ and $\pi_2 \circ f$ are smooth. This fact is easily seen from the chart representations.

(iv) By (ii) we may identify $T_{(p,q)}(M \times \{q\})$ with T_pM and likewise for $T_{(p,q)}(\{p\} \times N)$ and T_qN. We now show that (using this identification)

$$T_{(p,q)}(M \times N) = T_pM \oplus T_qN. \tag{2.2.1}$$

Since $\dim T_pM = m$ and $\dim T_qN = n$ it suffices to show that $T_{(p,q)}(M \times \{q\}) \cap T_{(p,q)}(\{p\} \times N) = 0$. Suppose v is in this intersection. Now since $\pi_1|_{\{p\}\times N} \equiv p$ we have $T_{(p,q)}\pi_1|_{T_{(p,q)}(\{p\}\times N)} = 0$ and so $T_{(p,q)}\pi_1(v) = 0$. But by (ii) $T_{(p,q)}\pi_1|_{T_{(p,q)}(M\times\{q\})}$ is bijective and hence $v = 0$. In most cases one identifies $T_pM \oplus T_qn$ with $T_pM \times T_qN$ and hence writes $T_{(p,q)}(M \times N) \cong T_pM \times T_qN$.

(v) If (M, g_M) and (N, g_N) are SRMFs then $M \times N$ is again a SRMF with metric

$$g := \pi_1^*(g_M) + \pi_2^*(g_N), \quad \text{i.e., cf. [10, 2.7.24]}, \tag{2.2.2}$$
$$g_{(p,q)}(v, w) = g_M(p)\big(T_{(p,q)}\pi_1(v), T_{(p,q)}\pi_1(w)\big) + g_N(q)\big(T_{(p,q)}\pi_2(v), T_{(p,q)}\pi_2(w)\big).$$

Indeed g is obviously symmetric and it is nondegenerate since suppose $g(v, w) = 0$ for all $w \in T_{(p,q)}(M \times N)$. Now choosing $w \in T_{(p,q)}(M \times \{q\}) \cong T_pM$ we obtain by $T_{(p,q)}\pi_2(w) = 0$ that

$$g_M(p)\big(T_{(p,q)}\pi_1(v), \underbrace{T_{(p,q)}\pi_1(w)}_{(*)}\big) = 0. \tag{2.2.3}$$

Now since $T_{(p,q)}\pi_1$ is bijective the term $(*)$ attains all values in T_pM hence by nondegeneracy of g_M, $T_{(p,q)}\pi_1(v) = 0$ and analogously $T_{(p,q)}\pi_2(v) = 0$. Since $v = v_M + v_N$ where $v_M \in T_pM$ and $v_N \in T_qN$ we finally obtain $v = 0$.

The SRMF $(M \times N, g)$ is called the *semi-Riemannian product* of M with N.

Now we are in a position to collect together the maps $\exp_p : T_pM \supseteq \mathcal{D}_p \to M$ $(p \in M)$ to a single mapping. To begin with we assume M to be complete, i.e., we assume that $\exp_p$ is defined on all of T_pM for all $p \in M$. Let $\pi : TM \to M$ the projection and define the mapping

$$E : TM \to M \times M, \quad E(v) = (\pi(v), \exp_{\pi(v)} v), \tag{2.2.4}$$

that is for $v \in T_pM \subseteq TM$ we have $E(v) = (p, \exp_p(v))$. In case M is not complete the maximal domain of E is given by

$$\mathcal{D} := \{v \in TM : \ c_v \text{ exists at least on } [0,1]\}. \tag{2.2.5}$$

Now for each p we have that the maximal domain $\mathcal{D}_p$ of $\exp_p$ is $\mathcal{D}_p = \mathcal{D} \cap T_pM$. Observe that $\mathcal{D}$ also is the maximal domain of

$$\exp := \pi_2 \circ E : \ v_p \mapsto \exp_p(v_p) = c_{v_p}(1). \tag{2.2.6}$$

We now have the following result which we have already used in the proof of the Gauss lemma 2.1.21.

2.2.3 Proposition (The domain of exp). *The domain $\mathcal{D}$ of E is open in TM and the domain $\mathcal{D}_p$ of $\exp_p$ is open and star shaped around 0 in T_pM.*

Proof. Let $G \in \mathfrak{X}(TM)$ be the geodesic spray as in 2.1.12. Then as proven there the flow lines of G are the derivatives of geodesics, i.e., $Fl_t^G(v) = c_v'(t)$. The maximal domain $\tilde{\mathcal{D}}$ of Fl^G is open in $\mathbb{R} \times TM$ by [10, 2.5.17(iii)].
Clearly $\tilde{\mathcal{D}}$ also is the maximal domain of $\pi \circ \mathrm{Fl}^G : (t, v) \mapsto c_v(t)$. Now let $\Phi : TM \to \mathbb{R} \times TM$, $\Phi(v) = (1, v)$. Then Φ is smooth and we have $\mathcal{D} = \{v \in TM : \ \exists c_v(1)\} = \{v \in TM : (1, v) \in \tilde{\mathcal{D}}\} = \Phi^{-1}(\tilde{\mathcal{D}})$, hence open. But then also $\mathcal{D}_p = \mathcal{D} \cap T_pM$ is open in T_pM.

Let finally $v \in \mathcal{D}_p$ then c_v is defined on $[0,1]$ and by (2.1.12) we have $c_{tv}(1) = c_v(t)$. So $tv \in \mathcal{D}_p$ for all $t \in [0,1]$ and hence $\mathcal{D}_p$ is star shaped. □

We now introduce sets which generalise the notion of convex subsets of Euclidean space.

2.2.4 Definition (Convex sets). *An open subset $C \subseteq M$ of a SRMF is called* (geodesically) convex *if C is a normal neighbourhood of all of its points.*

By 2.1.15 for any pair p, q of points in a convex set C there hence exists a unique geodesic called $c_{pq} : [0,1] \to M$ which connects p to q and stays entirely in C. Observe that there also might be other geodesics in M that connect p and q. They, however, have to leave C, as is depicted in the case of the sphere in Figure 2.5.

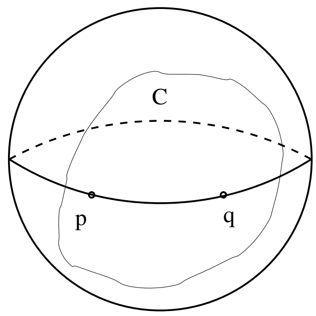

Figure 2.5: The unique geodesic connecting p and q within the convex set C is the shorter great circle arc between p and q.

To study further properties of the mapping E we need some further preparations.

2.2.5 Definition (Diagonal, null section). *Let M be a smooth manifold. The* diagonal *$\triangle_M \subseteq M \times M$ is defined by $\triangle_M := \{(p,p) : p \in M\}$. The* zero section *$TM_0$ of TM is defined as $TM_0 := \{0_p : p \in M\}$.*

Observe that in a chart (φ, U) around p we have that the map $f : U \times U \ni (p, p') \mapsto \varphi(p) - \varphi(p')$ is a submersion and we locally have that $\triangle_M = f^{-1}(0)$ and so $\triangle_M$ is a submanifold of $M \times M$, cf. ♣ `properly cite 1.1.22 of submf.pdf` ♣. Moreover the mapping $p \mapsto (p,p)$ is a diffeomorphism from M to $\triangle_M$. Finally by [10, 2.5.6] the zero section is precisely the base of the vector bundle $TM \to M$ and so it is also diffeomorphic to M. We now have:

2.2.6 Theorem (Properties of E). *The mapping E is a diffeomorphism from a neighbourhood of the zero section TM_0 of TM to a neighbourhood of the diagonal $\triangle_M \subseteq M \times M$.*

♣ `insert figure` ♣

Proof. We first show that every $x \in TM_0$ possesses a neighbourhood where E is a diffeomorphism. To begin with observe that we have $E(x) = (\pi(x), \exp_p(0)) = (p, p)$. We aim at applying the inverse function theorem (see eg. ♣ `properly cite 1.1.5 of submf.pdf` ♣) and so we have to show that $T_xE : T_x(TM) \to T_{E(x)}(M \times M)$ is bijective. But since $\dim TM = \dim(M \times M) = 2 \dim M$ it suffices to establishes injectivity at every point $x \in TM_0$. So let $x = 0_p \in TM_0$ and suppose $T_xE(v) = 0$ for $v \in T_x(TM)$. We have to show that $v = 0$.
Denoting by $\pi : TM \to M$ the bundle projection and by $\pi_1 : M \times M \to M$ the projection onto the first factor we have $\pi = \pi_1 \circ E$. So $T_x\pi(v) = T_{E(x)}\pi_1(T_xE(v)) = 0$ hence $v \in T_x(T_pM)$ since $\ker(T_x\pi) = T_x(T_pM)$ (cf. ♣ `properly cite 1.1.26 of submf.pdf` ♣). Now $E|_{T_pM} = \exp_p$ (identifying $\{p\} \times M$ with M) and so by (2.1.14)

$$0 = T_xE(v) = T_x(E|_{T_pM})(v) = T_{0_p} \exp_p(v) = v \tag{2.2.7}$$

as claimed.

In a second step we show that E is a diffeomorphism on a suitable domain. First we observe that for U in a basis of neighbourhoods of p and $\varepsilon > 0$ the sets of the form

$$W_{U,\varepsilon} := (T\psi)^{-1}\big(\psi(U) \times B_\varepsilon(0)\big) \tag{2.2.8}$$

form a basis of neighbourhoods of $v \in TM_0$ with ψ a chart of M around $\pi(v) = p$. Now cover TM_0 by such neighbourhoods W_{U_i,ε_i} such that $E|_{W_{U_i,\varepsilon_i}}$ is a diffeomorphism for all i. Then $W = \cup_i W_{U_i,\varepsilon_i}$ is a neighbourhood as claimed. Indeed E is a local diffeomorphism on W and we only have to show that it is also injective. To this end let $w_1, w_2 \in W$, $w_i \in W_{U_i,\varepsilon_i}$ $(i = 1, 2)$ with $E(w_1) = E(w_2)$. But then $\pi(w_1) = \pi(w_2) =: p \in U_1 \cap U_2$ and supposing w.l.o.g. that $\varepsilon_1 \leq \varepsilon_2$ we have $w_1, w_2 \in W_{U_2,\varepsilon_2}$. But then $w_1 = w_2$ by the fact that $E|_{W_{U_2,\varepsilon_2}}$ is a diffeomorphism. □

Now we are ready to state and prove the main result of this section. ♣ `insert figure` ♣

2.2.7 Theorem (Existence of convex sets). *Every point p in a SRMF M possesses a basis of neighbourhoods consisting of convex sets.*

Proof. Let V be a normal neighbourhood of p with Riemannian normal coordinates $\psi = (x^1, \dots, x^n)$ and define on V the function $N(q) = \sum_i (x^i(q))^2$. Then the sets $V(\delta) := \{q \in V : N(q) < \delta\}$ are diffeomorphic via ψ to the open balls $B_{\sqrt{\delta}}(0)$ in $\mathbb{R}^n$, hence they form a basis of neighbourhoods of p in M.
By 2.2.6 choosing δ small enough the map E is a diffeomorphism of an open neighbourhood W of $0_p \in TM$ to $V(\delta) \times V(\delta)$. We may choose W in such a way that $[0, 1]W \subseteq W$ (e.g. by setting $W = T\varphi^{-1}(\varphi(U) \times B(0))$ with φ a chart of M and $B(0)$ is a suitable Euclidean ball).
Let now $B \in \mathcal{T}^0_2(V)$ be the symmetric tensor field with

$$B_{ij}(q) = \delta_{ij} - \sum_l \Gamma^l_{ij}(q)\, x^l(q). \tag{2.2.9}$$

B by 2.1.17 is positive definite at p and hence if we further shrink δ and W it is positive definite on $V(\delta)$. We now claim that $V(\delta)$ is a normal neighbourhood of each of its points. Let $q \in V(\delta)$ and set $W_q = W \cap T_qM$ then by construction $E|_{W_q}$ is a diffeomorphism onto $\{q\} \times V(\delta)$ and so is $\exp_q|_{W_q}$ onto $V(\delta)$.
It remains to show that W_q is star shaped. For $q \neq \tilde{q} \in V(\delta)$ we set $v := E^{-1}(q, \tilde{q}) = \exp_q^{-1}(\tilde{q}) \in W_q$. Then $\sigma : [0,1] \to M$, $\sigma(t) = c_v(t)$ is a geodesic joining q with $\tilde{q}$. If we can show that σ lies in $V(\delta)$ then the proof of 2.1.15 (see the footnote) shows that $tv \in W_q$ for all $t \in [0,1]$ and we are done since any $v \in W_q$ is of the form $E^{-1}(q, \tilde{q})$.
Hence we assume by contradiction that σ leaves $V(\delta)$. Then there is a $t \in (0,1)$ with $N(\sigma(t)) \geq \delta$. ♣ `insert figure` ♣ Since $N(q), N(\tilde{q}) < \delta$ there is $t_0 \in (0,1)$ such that $t \mapsto N \circ \sigma$ takes a maximum in t_0. Then for $\sigma^i = x^i \circ \sigma$ we find by the geodesic equation

$$\frac{d^2(N\circ\sigma)}{dt^2} = 2\sum_i \big((\frac{d\sigma^i}{dt})^2 + \sigma^i \frac{d^2(\sigma^i)}{dt^2}\big) = 2\sum_{jk} \big(\delta_{jk} - \sum_i \Gamma^i_{\ jk}\sigma^i\big)\frac{d\sigma^j}{dt}\frac{d\sigma^k}{dt} \qquad (2.2.10)$$

and so

$$\frac{d^2(N\circ\sigma)}{dt^2}(t_0) = 2B(\sigma'(t_0), \sigma'(t_0)) > 0, \qquad (2.2.11)$$

since $\sigma' \neq 0$. But this contradicts the fact that $N \circ \sigma(t_0)$ is maximal. □

Convex neighbourhoods are of great technical significance as we can see e.g. in the following statement.

2.2.8 Corollary (Extendability of geodesics). *Let $c : [0,b) \to M$ be a geodesic. Then c is continuously extendible (as a curve) to $[0,b]$ iff it is extendible to $[0,b]$ as a geodesic.*

Proof. The 'only if' part of the assertion is clear. Let now $\tilde{c} : [0,b] \to M$ be a continuous extension of c. By 2.2.7, $\tilde{c}(b)$ has a convex neighbourhood C. Let now $a \in [0,b)$ be such that $\tilde{c}([a,b]) \subseteq C$. Then C is also a normal neighbourhood of $c(a)$ and $c|_{[a,b)}$ is a radial geodesic by 2.1.15 and can hence be extended until it reaches ∂C or to $[a,\infty)$. But since $\tilde{c}(b) \in C$ (and hence $\notin \partial C$) c can be extended as a geodesic beyond b. □

Let now C be a convex set in a SRMF M and let $p, q \in C$. We denote by σ_{pq} the unique geodesic in C such that $\sigma_{pq}(0) = p$ and $\sigma_{pq}(1) = q$, cf. 2.1.15. Then we write $\overrightarrow{pq} := \sigma'_{pq}(0) = \exp_p^{-1}(q) \in T_pM$ for the *displacement vector* of p and q. We then have

2.2.9 Lemma (Displacement vector). *Let $C \subseteq M$ be convex. Then the mapping*

$$\Phi : C \times C \to TM, \quad (p,q) \mapsto \vec{pq} \qquad (2.2.12)$$

is smooth and a diffeomorphism onto its image $\Phi(C \times C)$ in TM.

Proof. Let $(p_0, q_0) \in C \times C$, then $T_{q_0} \exp_{p_0}$ is regular. Also we have $\Phi(p,q) = \exp_p^{-1}(q)$ is the solution to the equation $\exp_p(f(p,q)) = q$ and so Φ is smooth by the implicit function theorem (see e.g. [10, 2.1.2]).
Moreover $E(\Phi(p,q)) = (p,q)$ and E is invertible at $\Phi(p,q)$ for all $(p,q) \in C \times C$ since $T_{\Phi(p,q)}E = \begin{pmatrix} \mathrm{id} & 0 \\ * & T_q \exp_p \end{pmatrix}$ is nonsingular. So locally $\Phi = E^{-1}$ and so Φ is a local diffeomorphism and, in particular, $\Phi(C \times C)$ is open.
Finally Φ is injective and hence it is a diffeomorphism with inverse $E^{-1}|_{\Phi(C\times C)}$. □

In Euclidean space the intersection of convex sets is convex. This statement fails to hold on SRMFs already in simple situations as the following counterexample shows.

2.2.10 Examples (Convex sets on S^1). Let $M = S^1$ and $p \in S^1$. Then the exponential function is a diffeomorphism from $(-\pi, \pi) \subseteq T_pS^1$ to $S^1 \setminus \{\bar{p}\}$, where $\bar{p}$ is the antipodal point of p, see ♣ `insert figure` ♣. The sets $C = S^1 \setminus \{p\}$ and $C' = S^1 \setminus \{\bar{p}\}$ are convex but their intersection $C \cap C'$ is not even connected hence, in particular, not convex.

However, we have the following result in a special case.

2.2.11 Lemma (Intersection of convex sets). *Let $C_1, C_2 \subseteq M$ be convex and suppose C_1 and C_2 are contained in a convex set $D \subseteq M$. Then the intersection $C_1 \cap C_2$ is convex.*

Proof. Let $p \in C_1 \cap C_2$. We show that $C_1 \cap C_2$ is a normal neighbourhood of p. To begin with $\exp_p : \tilde{D} \to D$ is a diffeomorphism. Set $\tilde{C}_i := \exp_p^{-1}(C_i)$ $(i = 1, 2)$, then $\exp_p : \tilde{C}_1 \cap \tilde{C}_2 \to C_1 \cap C_2$ is a diffeomorphism. Hence it only remains to show that $\tilde{C}_1 \cap \tilde{C}_2$ is star shaped.
Indeed let $v \in \tilde{C}_1 \cap \tilde{C}_2$ then $q := \exp_p(v) \in C_1 \cap C_2$ and so $\sigma_{pq}(t) = \exp_p(tv)$ is the unique geodesic in D that joins p and q. By convexity of the C_i we have that $\sigma_{pq}(t) = \exp_p(tv) \in C_1 \cap C_2$ for all $t \in [0, 1]$ and so $tv \in \tilde{C}_1 \cap \tilde{C}_2$ and we are done. □

The above lemma allows to prove the existence of a convex refinement of any open cover of a SRMF M. To be more precise we call a covering $\mathcal{C}$ of M by open and convex sets a *convex covering* if all non trivial intersections $C \cap C'$ of sets in $\mathcal{C}$ are convex. Then one can show that for any given open covering $\mathcal{O}$ of M there exists a convex covering $\mathcal{C}$ such that any set of $\mathcal{C}$ is contained in some element of $\mathcal{O}$, see [12, Lemma 5.10].

2.3 Arc length and Riemannian distance

In this section we define the *arc length* of (piecewise smooth) curves. On *Riemannian* manifolds this in turn allows to define a notion of distance between two points p and q as the infimum of the arc length of all curves connecting p and q. We will show that this *Riemannian distance function* encodes the topology of the manifold.

To begin with we introduce some notions on curves. ♣ `maybe move prior to 2.1.16` ♣ If $I = [a, b]$ is a closed interval we call a continuous curve $\alpha : [a, b] \to M$ piecewise smooth if there is a finite partition $a = t_0 < t_1 < \cdots < t_k = b$ of the interval $[a, b]$ such that all the restrictions $\alpha|_{[t_i,t_{i+1}]}$ are smooth. Thus at each of the so-called *break points*[2] t_i the curve may well have two distinct velocity vectors $\alpha'(t_i^-)$ and $\alpha'(t_i^+)$, representing the left- resp. right-sided derivative at t_i. If I is an arbitrary interval we call $\alpha : I \to M$ *piecewise smooth* if for all $a < b$ in I the restriction $\alpha|_{[a,b]}$ is piecewise smooth (in the previous sense). In particular, the break points have no cluster point in I.

2.3.1 Definition (Arc length). *Let $\alpha : [a, b] \to M$ be a piecewise smooth curve into a SRMF M. We define the* arc length *(or* length*, for short) of α by*

$$L(\alpha) := \int_a^b \|\alpha'(s)\| \, ds. \tag{2.3.1}$$

We recall that $\|\alpha'(s)\| = |\langle\alpha'(s), \alpha'(s)\rangle|^{1/2}$ and in coordinates $\|\alpha'(s)\| = |g_{ij}(\alpha(s)) \frac{d(x^i\circ\alpha)}{ds}(s) \frac{d(x^j\circ\alpha)}{ds}(s)|^{1/2}$.
On Riemannian manifolds the arc length behaves much as in the Euclidean setting, cf. [10, Ch. 1]. However in the semi-Riemannian case with an indefinite metric there are new effects. For example null curves always have vanishing arc length.
A reparametrisation of a piecewise smooth curve α is a piecewise smooth function $h : [c, d] \to [a, b]$ such that $h(c) = a$ and $h(d) = b$ (orientation preserving) or $h(c) = b$ and $h(d) = a$ (orientation reversing). If $h'(t)$ does not change sign then h is monotone and precisely as in [10, 1.1.2, 1.1.3] we have.

2.3.2 Lemma (Parametrisations). *Let $\alpha : I \to M$ be a piecewise smooth curve. Then we have:*

(i) The arc length 2.3.1 of α is invariant under monotone reparametrisations.

(ii) If $\|\alpha'(t)\| > 0$ for all t then α possesses a strictly monotone reparametrisation h such that for $\beta := \alpha \circ h$ we have $\|\beta'(s)\| = 1$ for all s. Such a parametrisation is called parametrisation by arc length.

Let now $p \in M$ and let U be a normal neighbourhood of p. The function

$$r : U \to \mathbb{R}^+, \quad r(q) := \|\exp_p^{-1}(q)\| \tag{2.3.2}$$

[2] We follow the widespread custom to call both the t_i as well as the $\alpha(t_i)$ breakpoints.

is called the *radius function* of M at p. In Riemannian normal coordinates we have

$$r(q) = |-\sum_{i=1}^{r}(x^i)^2(q) + \sum_{j=r+1}^{n}(x^j)^2(q)|^{1/2}. \tag{2.3.3}$$

Hence r is smooth except at its zeroes, hence off p as well as off the local null cone at p. In a normal neighbourhood the radius is given exactly by the length of radial geodesics as we prove next.

2.3.3 Lemma (Radius in normal neighbourhoods). *Let r be the radius in a normal neighbourhood U of a point p in a SRMF M. If σ is the radial geodesic from p to some $q \in U$ the we have*

$$L(\sigma) = r(q). \tag{2.3.4}$$

Proof. Let $v = \sigma'(0)$ then by 2.1.15 $v = \exp_p^{-1}(q)$. Now since σ is a geodesic, $\|\sigma'\|$ is constant and we have

$$L(\sigma) = \int_0^1 \|\sigma'(s)\|\, ds = \int_0^1 \|v\|\, ds = \|v\| = \|\exp_p^{-1}(q)\| = r(q). \tag{2.3.5}$$

□

From now on until the end of this chapter we will *exclusively deal with Riemannian manifolds.* Indeed it is only in the Riemannian case that the topology of the manifold is completely encoded in the metric.

2.3.4 Proposition (Radial geodesics are locally minimal). *Let M be a* Riemannian *manifold and let U be a normal neighbourhood of a point $p \in M$. If $q \in U$, then the radial geodesic $\sigma : [0,1] \to M$ from p to q is the unique shortest piecewise smooth curve in U from p to q, where uniqueness holds up to monotone reparametrisations.*

Proof. Let $c : [0,1] \to U$ be a piecewise smooth curve from p to q and set $s(t) = r(c(t))$, with r the radius function of (2.3.2). Since $\exp_p$ is a diffeomorphism we may for $t \neq 0$ uniquely write c in the form

$$c(t) = \exp_p(s(t)v(t)) =: f(s(t), t), \tag{2.3.6}$$

where v is a curve in T_pM with $\|v(t)\| = 1$ for all t. (This amounts to using polar coordinates in T_pM.) Here $f(s,t) = \exp_p(sv(t))$ is a two-parameter map on a suitable domain and the function $s : (0,1] \to \mathbb{R}^+$ is piecewise smooth. (Indeed we may suppose w.l.o.g. that $s(t) \neq 0$, i.e., $c(t) \neq p$ for all $t \in (0,1]$ since otherwise we may define t_0 to be the last parameter value when $c(t_0) = p$ and replace c by $c|_{[t_0,1]}$.)

Now except for possibly finitely many values of t we have (cf. (2.1.19) and below)

$$\frac{dc(t)}{dt} = \frac{\partial f}{\partial s}\big(s(t),t\big)\, s'(t) + \frac{\partial f}{\partial t}\big(s(t),t\big). \tag{2.3.7}$$

From $f(s,t) = \exp_p(sv(t))$ we have

$$\frac{\partial f}{\partial s} = (T_{sv(t)} \exp_p)(v(t)) \quad \text{and} \quad \frac{\partial f}{\partial t} = (T_{sv(t)} \exp_p)(sv'(t)) \tag{2.3.8}$$

and by the Gauss lemma 2.1.21 we find (since $(T_{sv(t)} \exp_p)(v(t))$ is radial)

$$\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}\rangle(s,t) = s\langle v(t), v'(t)\rangle = s\frac{1}{2}\frac{\partial}{\partial t}\underbrace{\|v(t)\|^2}_{=1} = 0 \tag{2.3.9}$$

and so $\frac{\partial f}{\partial s} \perp \frac{\partial f}{\partial t}$. Similarly we obtain via the Gauss lemma that $\|\frac{\partial f}{\partial s}\|^2 = \langle v(t), v(t)\rangle = 1$ which using (2.3.7) implies $\|\frac{dc}{dt}\|^2 = |s'(t)|^2 + \|\frac{\partial f}{\partial t}\|^2 \geq s'(t)^2$. This now gives for all $\varepsilon > 0$ that

$$\int_\varepsilon^1 \|c'(t)\|\, dt \geq \int_\varepsilon^1 |s'(t)|\, dt \geq \int_\varepsilon^1 s'(t)\, dt = s(1) - s(\varepsilon) \tag{2.3.10}$$

and hence in the limit $\varepsilon \to 0$ we find $L(c) \geq s(1) = r(q) = L(\sigma)$, where the final equality is due to 2.3.3.

Let now $L(c) = L(\sigma)$ then in the estimate (2.3.10) we have to have equality everywhere which enforces $\frac{\partial f}{\partial t}(s(t),t) = 0$ for all t. But since $T_{sv(t)} \exp_p$ is bijective this implies $v' \equiv 0$ and hence v has to be constant.
Moreover we need to have $|s'(t)| = s'(t) > 0$ and hence $c(t) = \exp_p(s(t)v)$ is a monotone reparametrisation of $\sigma(t)$: Indeed by 2.1.15 we have $\sigma(t) = \exp_p(t \exp_p^{-1}(q))$ and moreover $\exp_p^{-1}(q) = \exp_p^{-1}(c(1)) = s(1)v$ and so $\sigma(t) = \exp_p(ts(1)v)$. Now with $h(t) := \frac{s(t)}{s(1)}$ we have $\sigma(h(t)) = \exp_p(\frac{s(t)}{s(1)}\, s(1)v) = \exp_p(s(t)v) = c(t)$. □

In $\mathbb{R}^n$ the distance between two points $d(p,q) = \|p-q\|$ is at the same time the length of the shortest curve between these two points, i.e., the straight line segment from p to q. But this ceases to be true even in the simple case of $\mathbb{R}^2 \setminus \{(0,0)\}$. Here there is no shortest path between the two points $p = (-1,0)$ and $q = (1,0)$. However, the infimum of the arc length of all paths connecting theses two points clearly equals the Euclidean distance 2 between p and q. ♣ `insert Figure` ♣ This idea works also on general Riemannian manifolds and we start with the following definition.

2.3.5 Definition (Riemannian distance). *Let M be a* Riemannian *manifold and let p, $q \in M$. We define the set of 'permissible' paths connecting p and q by*

$$\Omega(p,q) := \{\alpha : \alpha \text{ is a piecewise smooth curve from } p \text{ to } q\}. \tag{2.3.11}$$

The Riemannian distance $d(p, q)$ *between* p *and* q *is then defined to be*

$$d(p,q) := \inf_{\alpha \in \Omega(p,q)} L(\alpha). \tag{2.3.12}$$

Furthermore we define just as in $\mathbb{R}^n$ the *ε-neighbourhood* of a point p in a RMF M for all $\varepsilon > 0$ by

$$U_\varepsilon(p) := \{q \in M : \ d(p,q) < \varepsilon\}. \tag{2.3.13}$$

Now in ε-neighbourhoods the usual Euclidean behaviour of geodesics holds true, more precisely we have:

2.3.6 Proposition (ε-neighbourhoods). *Let* M *be a RMF and let* $p \in M$*. Then for* ε *sufficiently small we have:*

(i) $U_\varepsilon(p)$ *is a normal neighbourhood,*

(ii) The radial geodesic σ *from* p *to any* $q \in U_\varepsilon(p)$ *is the unique shortest piecewise smooth curve* in M *connecting* p *with* q*. In particular, we have* $L(\sigma) = r(q) = d(p,q)$*.*

Note the fact that the radial geodesic in 2.3.6(ii) is *globally* the shortest curve from p to q and moreover that it is smooth.

Proof. Let U be a normal neighbourhood of p and denote by $\tilde{U}$ the corresponding neighbourhood of $0 \in T_pM$. Then for ε sufficiently small $\tilde{U}$ contains the star shaped open set $\tilde{N} = \tilde{N}_\varepsilon(0) := \{v \in T_pM : \ \|v\| < \varepsilon\}$. Therefore $N := \exp_p(\tilde{N})$ is also a normal neighbourhood.

By 2.3.4 for any $q \in N$ the radial geodesic σ from p to q is the unique shortest piecewise smooth curve *in* N from p to q. Moreover by 2.3.3 we have $L(\sigma) = r(q)$. Since $\sigma'(0) = \exp_p^{-1}(q) \in \tilde{N}$ we have $r(q) = \|\exp_p^{-1}(q)\| < \varepsilon$.

To finish the proof is suffices to show the following assertion:

$$\text{Any piecewise } \mathcal{C}^\infty\text{-curve } \alpha \text{ in } M \text{ starting in } p \text{ and leaving } N \text{ satisfies } L(\alpha) \geq \varepsilon. \tag{2.3.14}$$

Indeed then σ is the unique shortest piecewise smooth curve *in* M connecting p and q. But this implies that $r(q) = L(\sigma) = d(p,q)$ and we have shown (ii) but also $d(p,q) < \varepsilon$. Moreover for $q \notin N$ by (2.3.14) we have $d(p,q) \geq \varepsilon$. So in total $N = U_\varepsilon(p)$ implying (i). So it only remains to prove (2.3.14).

To begin with let $0 < a < \varepsilon$. We then have $\tilde{N}_a(0) \subseteq \tilde{N}$ and hence $\exp_p(\tilde{N}_a(0)) \subseteq N$. Now since α leaves N is also has to leave $\exp_p(\tilde{N}_a(0))$. Let $t_0 = \sup\{t : \ \alpha(t) \in \exp_p(\tilde{N}_a(0))\}$. Then $\alpha|_{[0,t_0]}$ is a curve connecting p with a point $q \in \partial \exp_p(\tilde{N}_a(0))$ which stays entirely in the slightly bigger normal neighbourhood $\exp_p(\tilde{N}_{a+\delta}(0))$ of p where δ is chosen such that $a + \delta < \varepsilon$. Now by 2.3.3 and 2.3.4 we have $L(\alpha) \geq L(\alpha|_{[0,t_0]}) \geq r(q) = a$. Finally for $a \to \varepsilon$ we obtain $L(\alpha) \geq \varepsilon$. □

2.3.7 Remark (Normal ε-neighbourhoods). Observe that the above proof also shows that any normal neighbourhood around p contains an ε-neighbourhood $U_\varepsilon(p)$ which itself is normal. Moreover it is open since $U_\varepsilon(p) = N = \exp_p(\tilde{N}_\varepsilon(0))$.

2.3.8 Example (Cylinder). Let M be the cylinder of 2.1.4 and denote by L any vertical line in M. Hence if $p \in M \setminus L$ then $M \setminus L$ is a normal neighbourhood of p and so $M \setminus L$ is convex. By 2.1.15 the radial geodesic σ (cf. Figure 2.6) is the unique shortest piecewise smooth curve from p to q in $M \setminus L$. However, the curve τ is obviously a shorter curve from p to q in M. *But τ leaves $M \setminus L$.*

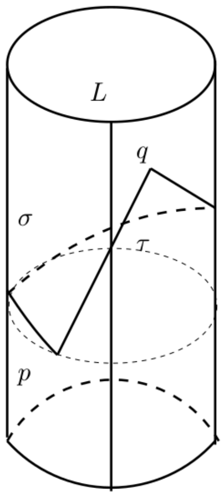

Figure 2.6: Convex neighbourhood on the cylinder. ♣ `improve figure` ♣

If p in M is arbitrary then the largest normal ε-neighbourhood of p is $U_\pi(p)$. This fact is most vividly seen from the (isometric hence equivalent) picture of the unwinded cylinder, see Figure 2.7. If $q \in U_\pi(p)$ then the radial geodesic σ form p to q is the unique shortest curve from p to q in all of M. If r is a point in $M \setminus U_\pi(p)$ there is still a shortest curve from p to r in M. If, however, r lies on the line L through the antipodal point p' of p then this curve is non-unique.

2.3.9 Theorem (Riemannian distance). *Let M be a connected RMF. Then the Riemannian distance function $d : M \times M \to \mathbb{R}$ is a metric (in the topological sense) on M. Furthermore the topology induced by d on M coincides with the manifold topology of M.*

Proof. First of all note that d is finite: let $p, q \in M$ then by connectedness there exists a piecewise C^1-curve α from connecting p and q (cf. 2.1.16) and so $d(p, q) \leq L(\alpha) < \infty$.
Now we show that d actually is a metric on M.

Positive definiteness: Clearly $d(p, q) > 0$ for all p, q. Now assume $d(p, q) = 0$. We have to show that $p = q$. Suppose to the contrary that $p \neq q$. Then by the Hausdorff property of M there exists a normal neighbourhood U of p not containing q. But then by 2.3.7 U contains a normal ε-neighbourhood $U_\varepsilon(p)$ and so by (2.3.14) $d(p, q) \geq \varepsilon > 0$.

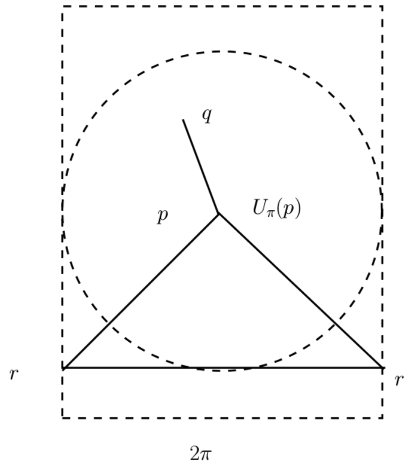

Figure 2.7: The largest normal ε-neighbourhood on the cylinder.♣ `improve figure` ♣

Symmetry: $d(p,q) = d(q,p)$ since $L(\alpha) = L(t \mapsto \alpha(-t))$.

Triangle inequality: Let $p, q, r \in M$. For $\varepsilon > 0$ let $\alpha \in \Omega(p,q)$, $\beta \in \Omega(q,r)$ such that $L(\alpha) < d(p,q) + \varepsilon$ and $L(\beta) < d(q,r) + \varepsilon$. Now define $\gamma = \alpha \cup \beta$. Then γ connects p with r and we have

$$d(p,r) \leq L(\gamma) = L(\alpha) + L(\beta) < d(p,q) + d(q,r) + 2\varepsilon \tag{2.3.15}$$

and since ε was arbitrary we conclude $d(p,r) \leq d(p,q) + d(q,r)$.

Finally by 2.2.7 the normal neighbourhoods provide a basis of neighbourhoods of p and by 2.3.7 every normal neighbourhood contains some $U_\varepsilon(p)$. Conversely by 2.3.7 every sufficiently small ε-neighbourhood of p is open and so d generates the manifold topology of M. □

We remark that there is also a proof which does not suppose the Riemannian metric to be smooth and the statement of the theorem remains true also for Riemannian metrics that are merely continuous.

2.3.10 Remark (Minimising curves). By the definition of the Riemannian distance function a piecewise smooth curve σ from p to q is a curve of minimal distance between these points if $L(\sigma) = d(p,q)$. In this case we call σ a *minimising curve.* In general there can be several such curves between a given pair of points; just consider the meridians running from the north pole to the south pole of the sphere.
Observe that every segment of a minimising curve form p to q is itself minimising (between its respective end points). Otherwise there would be a shorter curve between p and q. ♣ `insert figure` ♣.

We finally complete our picture concerning the relation between geodesics and minimising curves. By 2.1.15 radial geodesics are the unique minimising curves which connect the center of a normal neighbourhood U to any point $q \in U$ and *stay within* U. Moreover by making U smaller, more precisely by considering a normal ε-neighbourhood, radial geodesics become globally minimising curves, cf. 2.3.6. On the other hand geodesics that become too long, e.g. after leaving a normal neighbourhood of its starting point need no longer be minimising; just again think of the sphere. However, if a curve is minimising it has to be a geodesic and hence it is even smooth and not merely piecewise smooth, as the following statement says.

2.3.11 Corollary (Minimising curves are geodesics). *Let p and q be points in a RMF M. Let α be a minimising curve from p to q then α is (up to monotone reparametrisations) a geodesic from p to q.*

Proof. Let $\alpha : [0,1] \to M$ a minimising piecewise smooth curve between $\alpha(0) = p$ and $\alpha(1) = q$. We may now find a finite partition $I = \cup_{i=1}^{k} I_i$ such that each segment $\alpha_i := \alpha|_{I_i}$ is contained in a convex set. We may also suppose w.l.o.g. that every α_i is non constant since otherwise we can leave out the interval I_i. By 2.3.10 every α_i is minimising and hence by 2.3.4 a monotone reparametrisation of a geodesics. Hence by patching together these reparametrisations we obtain a possibly broken geodesic σ with break points at the end points of the I_i. We have $L(\sigma) = L(\alpha) = d(p,q)$. So the corollary follows from the following statement which shows that there are actually no break points.

If a geodesic segment c_1 ending at r and a geodesic segment c_2 starting at r combine to give a minimising curve segment c, then c is an (unbroken) geodesic. (2.3.16)

Observe that the intuitive idea behind this statement is that rounding off a corner of γ near r would make γ shorter ♣ `insert figure` ♣.
To formally prove (2.3.16) we once again choose a convex neighborhood U around r. Then the end part of c_1 and the starting part of c_2 combine to a minimising curve $\bar{c}$ in U. Since U is normal for each of its points and particular for some $r' \neq r$ on $\bar{c}$ it follows by 2.1.15 that $\bar{c}$ is a radial geodesic. So $\bar{c}$ and hence c has no break point at r. □

2.4 The Hopf-Rinow theorem

In this section we state and prove the main result on *complete Riemannian manifolds* which links the geodesics of the manifold to its structure as a metric space. The technical core of this result is contained in the following lemma.

2.4.1 Lemma (Globally defined exp_p.). *Let M be a connected RMF and let $p \in M$ such that the exponential map $\exp_p$ at p is defined on all of T_pM (i.e., $\mathcal{D}_p = T_pM$). Then for each $q \in M$ there is a minimising geodesic from p to q.*

Proof. Let $U_\varepsilon(p)$ be a normal ε-neighborhood of p, cf. 2.3.7. If $q \in U_\varepsilon(p)$ then the claim follows from 2.3.6(ii). So let $q \notin U_\varepsilon(p)$ and denote by r the radius function at p, see (2.3.2). Now for $\delta > 0$ sufficiently small (i.e., $\delta < \varepsilon$) the 'sphere' $S_\delta := \{m \in M : \ r(m) = \delta\} = \exp_p(\{v \in T_pM : \ \|v\| = \delta\})$ lies within $U_\varepsilon(p)$. Since S_δ is compact the continuous function $S_\delta \ni s \mapsto d(s,q)$ attains its minimum in some point $m \in S_\delta$. We now show that

$$d(p,m) + d(m,q) = d(p,q). \tag{2.4.1}$$

Clearly we have $\geq$ in (2.4.1) by the triangle inequality. Conversely let $\alpha : [0,b] \to M$ be any curve from p to q and let $a \in (0,b)$ be any value of the parameter such that $\alpha(a) \in S_\delta$. (Clearly α initially lies inside S_δ and then has to leave it, Figure ♣ `insert figure` ♣). Write $\alpha_1 = \alpha|_{[0,a]}$ and $\alpha_2 = \alpha|_{[a,b]}$. Then by 2.3.6(ii) and the definition of m we have

$$L(\alpha) = L(\alpha_1) + L(\alpha_2) \geq \delta + L(\alpha_2) \geq \delta + d(m,q). \tag{2.4.2}$$

Again appealing to 2.3.6(ii) this implies

$$d(p,q) \geq \delta + d(m,q) = d(p,m) + d(m,q) \tag{2.4.3}$$

and we have proven (2.4.1).

Recalling that by assumption $\mathcal{D}_p = T_pM$ let now $c : [0,\infty) \to M$ be the unit speed geodesic whose initial piece is the radial geodesic form p to m. We show that c is the asserted minimising geodesic from p to q. To begin with set $d := d(p,q)$ and

$$T := \{t \in [0,d] : \ t + d(c(t),q) = d\}. \tag{2.4.4}$$

It suffices to show that $d \in T$ since in this case we have $d(c(d),q) = 0$ hence $c(d) = q$. Moreover we then have $L(c|_{[0,d]}) = d = d(p,q)$ and so c is minimising.

Now to show that $d \in T$ we first observe that $c_{[0,t]}$ is minimising for any $t \in T$. Clearly we have $t = L(c|_{[0,t]}) \geq d(p,c(t))$. Conversely by definition of T it holds that $d \leq d(p,c(t)) + d(c(t),q) = d(p,c(t)) + d - t$ and so $t = L(c|_{[0,t]}) \leq d(p,c(t))$. Hence in total we have $L(c|_{[0,t]}) = d(p,c(t))$.
Now let $\tilde{t}$ such that $c(\tilde{t}) = m$. Then by (2.4.1) we have $d = d(p,q) = d(c(0),c(\tilde{t})) + d(c(\tilde{d}),q) = \tilde{t} + d(c(\tilde{d}),q)$ and so $\tilde{t} \in T$.

So T is non-empty, closed and contained in $[0, d]$, hence compact. Writing $t_0 := \max T \leq d$ it remains to show that $t_0 = d$.

We assume to the contrary that $t_0 < d$. Then let $U_{\varepsilon'}(c(t_0))$ be a normal ε-neighbourhood of $c(t_0)$ which does not contain q, see Figure ♣ `insert figure` ♣. The same argument as in the beginning of the proof shows that there exists a radial unit speed geodesic $\sigma : [0, \delta'] \to U_{\varepsilon'}(c(t_0))$ that joins $c(t_0)$ with some point $m' \in S_{\delta'}$ with $d(m', q)$ minimal on $S_{\delta'}$. As in (2.4.1) we obtain

$$d(c(t_0), m') + d(m', q) = d(c(t_0), q). \tag{2.4.5}$$

Now observing that $d(c(t_0), m') = L(\sigma|_{[0,\delta']}) = \delta'$ we obtain from the fact that $t_0 \in T$ and (2.4.5)

$$d = t_0 + d(c(t_0), q) = t_0 + \delta' + d(m', q). \tag{2.4.6}$$

Also $d = d(p, q) \leq d(p, m') + d(m', q)$ and so $d - d(m', q) = t_0 + \delta' \leq d(p, m')$. But this implies that the concatenation $\tilde{c}$ of $c|_{[0,t_0]}$ and σ is a curve joining p with m' which satisfies

$$d(p, m') \leq L(\tilde{c}) = t_0 + \delta' \leq d(p, m').$$

Hence $\tilde{c}$ is minimising and so by 2.3.11 an (unbroken) geodesic which implies that $\tilde{c} = c$. This immediately gives $m' = \sigma(\delta') = c(t_0 + \delta')$ and further by (2.4.6) we obtain

$$d - t_0 = \delta' + d(m', q) = \delta' + d(c(t_0 + \delta'), q).$$

But this means that $t_0 + \delta' \in T$ which contradicts the fact that $t_0 = \max T$. □

We may now head on to the main result of this section.

2.4.2 Theorem (Hopf-Rinow). *Let (M, g) be a connected RMF, then the following conditions are equivalent:*

(MC) The metric space (M, d) is complete (i.e., every Cauchy sequence converges).

(GC') There is $p \in M$ such that M is geodesically complete at p, i.e., $\exp_p$ is defined on all of T_pM.

(GC) (M, g) is geodesically complete.

(HB) M possesses the Heine-Borel property, i.e., every closed and bounded subset of M is compact.

The Hopf-Rinow theorem hence, in particular, guarantees that for connected Riemannian manifolds geodesic completeness coincides with completeness as a metric space. Therefore the term *complete Riemannian manifold* is unambiguous in the connected case and we will use it from now on. The theorem together with the previous Lemma 2.4.1 has the following immediate and major consequence.

2.4.3 Corollary (Geodesic connectedness). *In a connected complete Riemannian manifold any pair of points can be joined by a minimising geodesic.*

The converse of this result is obviously wrong; just consider the open unit disc in $\mathbb{R}^2$. From the point of view of semi-Riemannian geometry the striking fact of the corollary is that in a complete RMF two arbitrary points can be connected by a geodesic *at all.* This property called *geodesic connectedness* fails to hold in complete connected Lorentzian manifolds. The great benefit of the corollary, of course, is that it allows to use geodesic constructions globally.
We now proceed to the proof of the Hopf-Rinow theorem.

Proof of 2.4.2.
(MC)⇒(GC): Let $c : [0, b) \to M$ be a unit speed geodesic. We have to show that c can be extended beyond b as a geodesic. By 2.2.8 is suffices to show that c can be extended continuously (as a curve) to b. To this end let $(t_n)_n$ be a sequence in $[0, b)$ with $t_n \to b$. Then $d(c(t_n), c(t_m)) \leq |t_n - t_m|$ and so $(c(t_n))_n$ is a Cauchy sequence in M , which by (MC) is convergent to a point called $c(b)$. If $(t'_n)_n$ is another such sequence then since $d(c(t_n), c(t'_n)) \leq |t_n - t'_n|$ we find that $(c(t'_n))_n$ also converges to $c(b)$. Hence we have c extended continuously to $[0, b]$.

(GC)⇒(GC') is clear.

(GC')⇒(HB): Let $A \subseteq M$ be closed and bounded. For any $q \in A$, by 2.4.1 there is a minimising geodesic $\sigma_q : [0, 1] \to M$ from p to q. As in (2.3.5) we have $\|\sigma'_q(0)\| = L(\sigma_q) = d(p, q)$.
Now since A is bounded there is $R > 0$ such that $d(p, q) \leq R$ for all $q \in A$. So $\sigma'_q(0) \in \overline{B_R(0)} = \{v \in T_pM : \|v\| \leq R\}$ which clearly is compact. But then $A \subseteq \exp_p(\overline{B_R(0)})$ hence is contained in a compact set and thus is compact itself.

(HB)⇒(MC)[3]: The point set of every Cauchy sequence is bounded hence its closure is compact by (HB). So the sequence possesses a convergent subsequence and being Cauchy it is convergent itself. □

Next we prove that any smooth manifold M (which, recall our convention, is assumed to be Hausdorff and second countable) can be equipped with a Riemannian metric g. Indeed g can simply be constructed by glueing the Euclidean metrics in the charts of an atlas as is done in the proof below.
In fact much more is true. The Theorem of Numizo and Ozeki (see e.g. [8, 62.12]) guarantees that on every smooth manifold there exists a *complete* Riemannian metric.

2.4.4 Theorem (Existence of Riemannian metrics). *Let M be a smooth manifold then there exists a Riemanninan metric g on M.*

[3] Observe that this part of the proof is purely topological.

Proof. Cover M by charts $((x^1_\alpha, \ldots, x^n_\alpha), U_\alpha)$ and let $(\chi_\alpha)_\alpha$ be a partition of unity subordinate to this cover with $\mathrm{supp}(\chi_\alpha) \subseteq U_\alpha$, cf. [10, 2.3.14]. Then on U_α set

$$g_\alpha := \sum_i dx^i_\alpha \otimes dx^i_\alpha \quad \text{and finally define} \quad g := \sum_\alpha \chi_\alpha g_\alpha \quad \text{on } M. \tag{2.4.7}$$

Since a linear combination of positive definite scalar products with positive coefficients again is positive definite, g indeed is a Riemannian metric on M. □

This result has the following immediate topological consequence.

2.4.5 Corollary (Metrisability). *Every smooth manifold is metrisable.*

Proof. In case M is connected this is immediate from 2.4.4 and 2.3.9. In the general case we have $M = \mathring{\bigcup} M_i$ with all (countably many) M_i connected. On each M_i we have a metric d_i for which we may assume w.l.o.g. $d_i < 1$ (otherwise replace d_i by $\frac{d_i}{1+d_i}$). Then

$$d(x, y) := \begin{cases} d_i(x, y) & \text{if } x, y \in M_i \\ 1 & \text{else} \end{cases} \tag{2.4.8}$$

clearly is a metric on M. □

The construction used in the proof can also be employed to draw the following conclusion from general topology.

2.4.6 Corollary. *Every compact Riemannian manifold is complete.*

Proof. We first construct a metric on M from the distance functions in each of the (countably many) connected components as in (2.4.8). Then we just use the fact that any compact metric space is complete. In fact this follows from the proof of (HB)⇒(MC) in 2.4.2 by noting that the Heine-Borel property (HB) holds trivially in any compact metric space. □

To finish this section we remark that the Lorentzian situation is much more complicated, somewhat more precisely we have

2.4.7 Remark (Lorentzian analogs). On Lorentzian manifolds the above results are wrong in general! In some more detail we have the following.

(i) There does not exist a Lorentzian metric on any smooth manifold M. Observe that the above proof fails since linear combinations of nondegenerate scalar products need not be nondegenerate. In fact, there exist topological obstructions to the existence of Lorentzian metrics: M can be equipped with a Lorentzian metric iff there exists a nowhere vanishing vector field hence iff M is non compact or compact with Euler characteristic 0.

(ii) The Lorentzian analog to 2.3.4 is the following: Let U be a normal neighborhood of p. If there is a timelike curve form p to some point $q \in U$ then the radial geodesic from p to q is the *longest* curve from p to q.

(iii) There is no Lorentzian analog of the Hopf Rinow theorem 2.4.2. If a Lorentzian manifold is connected and geodesically complete it need not even be geodesically connected. For a counterexample see e.g. [12, p. 150].

Chapter 3

Curvature

In elementray differential geometry, that is in the geometry of 2-dimensional surfaces in $\mathbb{R}^3$, a notion of curvature was defined which is based on the idea on how the surface bends in the surrounding Euclidean space. More precisely (cf. [10, Sec. 3.1]) in this approach one to defines the Weingarten map as the derivative of the unit normal vector and the principal curvatures as its eigenvalues. It was Gauss who showed in his 'theorema egregium' that the product of the principal curvatures, i.e., the Gauss curvature is a quantity intrinsic to the surface (cf. [10, Thm. 3.1.18]). This lead Riemann to generalise the Gaussian curvature to manifolds with a positive definite metric, i.e., to the invention of Riemannian geometry. The semi-Riemannian case requires no significant changes and will be presented in this chapter.

More precisely we define the Riemannian curvature tensor in Section 3.1 as a measure for the non commutativity of second order covariant derivatives. After deriving its symmetries and local representation we discuss the geometric meaning of the Riemann tensor: On the one hand it measures the failure of a vector parallely transported along a closed curve to return to its initial value, on the other hand it is the obstruction to local flatness of the manifold.
In Section 3.2 we introduce the generalisation of the classical differential operators of multivariable calculus to SRMFs. Along the way we discuss the operations of type changing, i.e., the extension of the musical isomorphism to tensor fields of general rank and of metric contraction.
Finally in section 3.3 we introduce the Ricci and scalar curvature and write down the Einstein equations, that is the field equations of general relativity—Einstein's theory of space, time and gravity—which link the curvature of the Lorentzian spacetime manifold to its energy content.

3.1 The curvature tensor

To motivate the definition of the curvature tensor on a SRMF we consider the parallel transport of a vector along a curve. If we parallel transport a vector along a closed curve in the plane then upon returning to the starting point we end up with the same vector as we have started with. This, however, is not the case in the sphere.

If we parallel transport a vector say along a geodesic triangle on the sphere we do not end up with the vector we have started with, see Figure 3.1.

Now since parallel transport is defined via the covariant derivative, see page 30, the difference between the starting vector and the final vector can be expressed in terms of the *non commutativity* of covariant derivatives. This leads to the following formal definition, we will, however, return to the intuitive idea behind it at the end of this section.

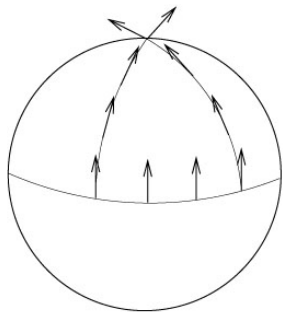

Figure 3.1: Parallel transport of a vector along a geodesic triangle on the sphere.

3.1.1 Definition & Lemma (Riemanninan curvature tensor). *Let (M, g) be a SRMF with Levi-Civita connection ∇. Then the mapping $R : \mathfrak{X}(M)^3 \to \mathfrak{X}(M)$ given by*

$$R_{XY}Z := \nabla_{[X,Y]}Z - [\nabla_X, \nabla_Y]Z \tag{3.1.1}$$

is a $(1, 3)$ tensor field called the Riemannian curvature tensor *of M.*

Proof. By (1.3.24) we only have to show that R is $\mathcal{C}^\infty(M)$-multilinear. So let $f \in \mathcal{C}^\infty(M)$. We then have by [10, 2.5.15(iv)] that $[X, fY] = X(f)Y + f[X, Y]$ and so

$$\begin{aligned} R_{X,fY}Z &= \nabla_{[X,fY]}Z - \nabla_X\nabla_{fY}Z + \nabla_{fY}\nabla_X Z \\ &= X(f)\nabla_Y Z + f\nabla_{[X,Y]}Z - \nabla_X(f\nabla_Y Z) + f\nabla_Y\nabla_X Z \\ &= X(f)\nabla_Y Z - X(f)\nabla_Y Z + fR_{XY}Z = fR_{XY}Z. \end{aligned} \tag{3.1.2}$$

Since by definition $R_{XY}Z = -R_{YX}Z$ we also find that $R_{fXY}Z = fR_{XY}Z$. Finally by an analogous calculation one finds that $R_{XY}fZ = fR_{X,Y}Z$. □

We will follow the widespread convention to also write $R(X, Y)Z$ for $R_{XY}Z$. Moreover we will also call R the Riemann tensor or curvature tensor for short. Since R is a tensor field one may also insert individual tangent vectors into its slots. In particular for $x, y \in T_pM$ the mapping

$$R_{xy} : T_pM \to T_pM, \qquad z \mapsto R_{xy}z \tag{3.1.3}$$

is called the *curvature operator*. We next study the symmetry properties of the curvature tensor.

3.1.2 Proposition (Symmetries of the Riemann tensor). *Let $x, y, z, v, w \in T_pM$ then for the curvature operator we have the following identities*

(i) $R_{xy} = -R_{yx}$ *(skew-symmetry)*

(ii) $\langle R_{xy}v, w\rangle = -\langle R_{xy}w, v\rangle$ *(skew-adjointness)*

(iii) $R_{xy}z + R_{yz}x + R_{zx}y = 0$ *(first Bianchi identity)*

(iv) $\langle R_{xy}v, w\rangle = \langle R_{vw}x, y\rangle$ *(symmetry by pairs)*

Proof. Since ∇_X and $[\,,\,]$ are local operations (see 1.3.6 and [10, 2.5.15(v)], respectively) it suffices to work on any neighbourhood of p. Moreover all identities are tensorial and we may extend $x, y, \dots$ in any convenient way to vector fields $X, Y, \dots$ on that neighbourhood. In the present case it is beneficial to do so in such a way that all Lie-brackets vanish which is achieved by taking the vector fields to have constant components w.r.t. a coordinate basis (Recall that $[\partial_i, \partial_j] = 0$, cf. [10, 2.5.15(vi)] ♣ `will be (vii) in the new version` ♣.) We then have

$$R_{XY}Z = \nabla_Y\nabla_X Z - \nabla_X\nabla_Y Z \tag{3.1.4}$$

and we go on proving the individual items of the proposition.

(i) follows directly form the anti symmetry of $[\,,\,]$. (Observe that (i) also is easy to see from the definition and we have already used it in the proof of 3.1.1.)

(ii) It suffices to show that $\langle R_{xy}v, v\rangle = 0$ since the assertion then follows by replacing v by $v + w$. Indeed then we have $0 = \langle R_{xy}(v + w), v + w\rangle = \langle R_{xy}v, w\rangle + \langle R_{xy}w, v\rangle$.

Now to prove the above statement we write using 1.3.4(∇5)

$$\begin{aligned}\langle R_{XY}V, V\rangle &= \langle \nabla_Y\nabla_X V, V\rangle - \langle \nabla_X\nabla_Y V, V\rangle \\ &= Y\langle \nabla_X V, V\rangle - \cancel{\langle \nabla_X V, \nabla_Y V\rangle} - X\langle \nabla_Y V, V\rangle + \cancel{\langle \nabla_Y V, \nabla_X V\rangle} \\ &= \frac{1}{2}YX\langle V, V\rangle - \frac{1}{2}XY\langle V, V\rangle = -\frac{1}{2}\underbrace{[X, Y]}_{=0}(\langle V, V\rangle) = 0.\end{aligned} \tag{3.1.5}$$

(iii) follows from the following more general reasoning. Let $F : \mathfrak{X}(M)^3 \to \mathfrak{X}(M)$ be an $\mathbb{R}$-multilinear map and define the mapping $S(F) : \mathfrak{X}(M)^3 \to \mathfrak{X}(M)$ as the sum of the cyclic permutations of F, i.e.,

$$S(F)(X, Y, Z) = F(X, Y, Z) + F(Y, Z, X) + F(Z, X, Y). \tag{3.1.6}$$

Then a cyclic permutation of X, Y, Z obviously leaves $S(F)(X, Y, Z)$ unchanged. Consequently we find using 1.3.4(∇4)

$$\begin{aligned}S(R)_{XY}Z &= S\nabla_Y\nabla_X Z - S\nabla_X\nabla_Y Z \\ &= S\nabla_X\nabla_Z Y - S\nabla_X\nabla_Y Z = -S\nabla_X[Y, Z] = 0.\end{aligned} \tag{3.1.7}$$

(iv) is a combinatorial exercise. By (iii) $\langle S(R)_{YV}X, W\rangle = 0$. Now summing over the four cyclic permutations of Y, V, X, W and writing out $S(R)$ one obtains 12 terms, 8 of which cancel in pairs by (i) and (ii) leaving

$$2\langle R_{XY}V, W\rangle + 2\langle R_{WV}X, Y\rangle = 0, \tag{3.1.8}$$

which by another appeal to (i) gives the asserted identity.

□

Next we derive a local formula for the Riemann tensor.

3.1.3 Lemma (Coordinate expression for R). *Let $(x^1, \dots, x^n)$ be local coordinates. Then we have $R_{\partial_k \partial_l}\partial_j = R^i_{jkl}\partial_i$, where*

$$R^i_{jkl} = \frac{\partial}{\partial x^l}\Gamma^i_{\ kj} - \frac{\partial}{\partial x^k}\Gamma^i_{\ lj} + \Gamma^i_{\ lm}\Gamma^m_{\ kj} - \Gamma^i_{\ km}\Gamma^m_{\ lj}. \tag{3.1.9}$$

Proof. Since $[\partial_i, \partial_j] = 0$ for all i, j we have by (3.1.4) that

$$R_{\partial_k\partial_l}\partial_j = \nabla_{\partial_l}\nabla_{\partial_k}\partial_j - \nabla_{\partial_k}\nabla_{\partial_l}\partial_j. \tag{3.1.10}$$

Now by 1.3.8 we have

$$\nabla_{\partial_l}(\nabla_{\partial_k}\partial_j) = \nabla_{\partial_l}(\Gamma^m_{\ kj}\partial_m) = \frac{\partial}{\partial x^l}\Gamma^m_{\ kj}\partial_m + \Gamma^m_{\ kj}\Gamma^r_{\ lm}\partial_r = \Big(\frac{\partial}{\partial x^l}\Gamma^i_{\ kj} + \Gamma^i_{\ lm}\Gamma^m_{\ kj}\Big)\,\partial_i. \tag{3.1.11}$$

Now exchanging k and l and subtracting the respective terms gives the assertion. □

In the remainder of this section we want to give two interpretations of the Riemann tensor to aid also an intuitive understanding of this pretty complicated geometric object.

(1) We show that the Riemann tensor is an obstruction to the manifold being locally flat, i.e., the manifold being covered by charts in which the metric is flat, see Theorem 3.1.7, below.

(2) We make precise the idea which we already discussed prior to Definition 3.1.1. We will establish that the curvature tensor is a measure for the failure of a vector to return to its starting value when parallely transported along a closed curve.

To arrive at (1) we need some preparations. Our proof of the result alluded to above depends on the construction of an especially simple coordinate system. Recall that the natural basis vector fields ∂_i in any coordinate system commute, i.e., $[\partial_i, \partial_j] = 0$ for all i, j (once again see e.g. [10, 2.5.15(vi)]). We now establish the converse of this result: any n-tuple of commuting and linearly independent local vector fields is the natural basis for some chart. We begin with the following characterisation of commuting flows resp. vector fields, which is of clear independent interest. ♣ `Insert references to or a short discussion of the Lie derivative of vector fields` ♣

3.1.4 Lemma (Commuting flows). *Let M be a smooth manifold and let $X, Y \in \mathfrak{X}(M)$. Then the following conditions are equivalent:*

(i) $[X, Y] = 0$*,*

(ii) $(\mathrm{Fl}_t^X)^* Y = Y$*, wherever the l.h.s. is defined,*

(iii) $\mathrm{Fl}_t^X \circ \mathrm{Fl}_s^Y = \mathrm{Fl}_s^Y \circ \mathrm{Fl}_t^X$*, wherever one (hence both) side(s) are defined.*

Proof. (i)⇒(ii): We have

$$\begin{aligned}\frac{d}{dt}(\mathrm{Fl}_t^X)^* Y &= \frac{d}{ds}\Big|_0 (\mathrm{Fl}_{t+s}^X)^* Y = \frac{\partial}{\partial s}\Big|_0 (T\mathrm{Fl}_{-(s+t)}^X \circ Y \circ \mathrm{Fl}_{s+t}^X) \\ &= \frac{\partial}{\partial s}\Big|_0 (T\mathrm{Fl}_{-t}^X \circ T\mathrm{Fl}_{-s}^X \circ Y \circ \mathrm{Fl}_s^X \circ \mathrm{Fl}_t^X) = T\mathrm{Fl}_{-t}^X \circ \underbrace{\Big(\frac{\partial}{\partial s}\Big|_0 (\mathrm{Fl}_s^X)^* Y\Big)}_{=L_X Y = [X,Y]} \circ \mathrm{Fl}_t^X \\ &= (\mathrm{Fl}_t^X)^* (L_X Y) = 0,\end{aligned} \tag{3.1.12}$$

and so $(\mathrm{Fl}_t^X)^* Y = (\mathrm{Fl}_0^X)^* Y = Y$.

(ii)⇒(i): follows directly from the definition of the Lie derivative.

(ii)⇒(iii): Observe that condition (ii) precisely says that $(\mathrm{Fl}_t^X)^* Y$ and Y are Fl_t^X-related. Recall that given a smooth function between two manifolds $f : M_1 \to M_2$ one says that $X_1 \in \mathfrak{X}(M_1)$ is f-related to $X_2 \in \mathfrak{X}(M_2)$, $X_1 \sim_f X_2$, if $Tf \circ X_1 = X_2 \circ f$.
Now we have in general that

$$X_1 \sim_f X_2 \;\Rightarrow\; f \circ \mathrm{Fl}_t^{X_1} = \mathrm{Fl}_t^{X_2} \circ f, \tag{3.1.13}$$

wherever one (hence both) side(s) are defined. Indeed we have

$$\frac{d}{dt}\Big|_0 (f \circ \mathrm{Fl}_t^{X_1}(p)) = T_{\mathrm{Fl}_t^{X_1}(p)} f \big(\frac{d}{dt}\big|_0 \mathrm{Fl}_t^{X_1}(p)\big) = T_{\mathrm{Fl}_t^{X_1}(p)} f \circ X \circ \mathrm{Fl}_t^{X_1}(p) = X_2(f(\mathrm{Fl}_t^{X_1}(p))),$$

and $f \circ \mathrm{Fl}_0^{X_1}(p) = f(p)$. So by the definition of the flow we obtain $f \circ \mathrm{Fl}_t^{X_1}(p) = \mathrm{Fl}_t^{X_2}(f(p))$.

Now we obtain for X, Y resp. its flows

$$\begin{aligned}\mathrm{Fl}_t^X \circ \mathrm{Fl}_s^Y = \mathrm{Fl}_s^Y \circ \mathrm{Fl}_t^X \;&\Leftrightarrow\; \mathrm{Fl}_s^Y = \mathrm{Fl}_{-t}^X \circ \mathrm{Fl}_s^Y \circ \mathrm{Fl}_t^X = \mathrm{Fl}_{-t}^X \circ \mathrm{Fl}_t^X \circ \mathrm{Fl}_s^{(\mathrm{Fl}_t^X)^* Y} = \mathrm{Fl}_s^{(\mathrm{Fl}_t^X)^* Y} \\ &\Leftrightarrow\; Y = (\mathrm{Fl}_t^X)^* Y\end{aligned}$$

□

3.1.5 Lemma (Coordinates adapted to given vector fields). *Let V be an open subset of a smooth manifold M. Given vector fields $X_1, \dots, X_n \in \mathfrak{X}(M)$ such that $[X_i, X_j] = 0$ for all i, j and $\{X_1(p), \dots, X_n(p)\}$ is a basis of T_pM for all $p \in V$ then around each p there is a chart $(\varphi = (x^1, \dots, x^n), U)$ with $U \subseteq V$ and*

$$X_i\big|_U = \frac{\partial}{\partial x^i} \qquad \text{for all } i = 1, \dots, n. \tag{3.1.14}$$

Proof. Fix some $p \in V$ and set $F : t = (t^1, \dots, t^n) \mapsto (\mathrm{Fl}^{X_1}_{t^1} \circ \dots \circ \mathrm{Fl}^{X_n}_{t^n})(p)$. Then F is smooth on an open neighbourhood W of $0 \in \mathbb{R}^n$ and we may assume that $F(W) \subseteq V$. Since $[X_i, X_j] = 0$ by 3.1.4 the flows of the X_i commute and we have

$$\frac{\partial}{\partial t^i} F(t) = \frac{\partial}{\partial t^i} \mathrm{Fl}^{X_i}_{t^i} \circ \mathrm{Fl}^{X_1}_{t^1} \circ \dots \circ \mathrm{Fl}^{X_n}_{t^n}(p) = X_i(F(t)). \tag{3.1.15}$$

Since the $X_i(F(t))$ are a basis it follows that F is a local diffeomorphism hence w.l.o.g. $F : W \to U \subseteq V$ is a diffeomorphism.
Now define $\varphi = F^{-1} : U \to W$. Then we have for $q = \varphi^{-1}(t)$ using (3.1.15)

$$\frac{\partial}{\partial x^i}\Big|_q = \big(T_q\varphi\big)^{-1}(e_i) = T_tF(e_i) = \frac{\partial}{\partial t^i} F(t) = X_i(q). \tag{3.1.16}$$

□

Next we establish an explicit expression for the commutator of the induced covariant derivative in terms of the curvature tensor, which obviously is of independent interest.

3.1.6 Proposition (Exchanging 2nd order derivatives.). *Let $f : \mathcal{D} \to M$ be a two-parameter map into a SRMF M and let $Z \in \mathfrak{X}(f)$. Then we have*

$$\left(\frac{\nabla}{dv}\frac{\nabla}{du} - \frac{\nabla}{dv}\frac{\nabla}{du} \right) Z \;=\; Z_{uv} - Z_{vu} \;=\; R(f_u, f_v)Z. \tag{3.1.17}$$

Proof. We work in a chart and write $Z = Z^k\partial_k$. Then by (1.3.56) we have $Z_u = Z^k_u\partial_k$ with $Z^k_u = \partial Z^k/\partial u + \Gamma^k_{lm}Z^l\partial f^m/\partial u$ and so

$$\begin{aligned} Z_{uv} = \Bigg(\frac{\partial^2 Z^k}{\partial v \partial u} + \frac{\partial \Gamma^k_{lm}}{\partial v} Z^l \frac{\partial f^m}{\partial u} + \Gamma^k_{lm} \frac{\partial Z^l}{\partial v}\frac{\partial f^m}{\partial u} + \Gamma^k_{lm} Z^l \frac{\partial^2 f^m}{\partial v \partial u} \\ + \Gamma^k_{ij}\frac{\partial Z^i}{\partial u}\frac{\partial f^j}{\partial v} + \Gamma^k_{ij}\Gamma^i_{lm} Z^l \frac{\partial f^m}{\partial u}\frac{\partial f^j}{\partial v} \Bigg) \partial_k \end{aligned} \tag{3.1.18}$$

and analogously for Z_{vu}. Upon inserting the symmetric terms cancel and we obtain

$$\begin{aligned} &Z_{uv} - Z_{vu} \\ &= \left(Z^l \left(\frac{\partial \Gamma^k_{lm}}{\partial v}\frac{\partial f^m}{\partial u} - \frac{\partial \Gamma^k_{lm}}{\partial u}\frac{\partial f^m}{\partial v} \right) + \Gamma^k_{ij}\Gamma^i_{lm} Z^l \left(\frac{\partial f^m}{\partial u}\frac{\partial f^j}{\partial v} - \frac{\partial f^m}{\partial v}\frac{\partial f^j}{\partial u} \right) \right) \partial k \end{aligned} \tag{3.1.19}$$

On the other hand we have

$$
\begin{aligned}
R(f_u,f_v)Z &= R\Big(\frac{\partial f^i}{\partial u}\partial_i, \frac{\partial f^j}{\partial v}\partial_j\Big)\big(Z^l\partial_l\big)\\
&= Z^l\,\frac{\partial f^i}{\partial u}\,\frac{\partial f^j}{\partial v}\,R(\partial_i,\partial_j)\partial_l = Z^l\,\frac{\partial f^i}{\partial u}\,\frac{\partial f^j}{\partial v}\;R^k_{lij}\partial_k\\
&= \left(Z^l\,\frac{\partial f^i}{\partial u}\,\frac{\partial f^j}{\partial v}\,\Big(\frac{\partial\Gamma^k_{il}}{\partial x^j} - \frac{\partial\Gamma^k_{jl}}{\partial x^i} + \Gamma^k_{jm}\Gamma^m_{il} - \Gamma^k_{im}\Gamma^m_{jl}\Big)\right)\partial_k \qquad (3.1.20)\\
&= Z^l\,\Big(\frac{\partial f^i}{\partial u}\,\underbrace{\frac{\partial\Gamma^k_{il}}{\partial x^j}\,\frac{\partial f^j}{\partial v}}_{\frac{\partial\Gamma^k_{il}}{\partial v}} - \frac{\partial f^j}{\partial v}\,\underbrace{\frac{\partial\Gamma^k_{jl}}{\partial x^i}\,\frac{\partial f^i}{\partial u}}_{\frac{\partial\Gamma^k_{jl}}{\partial u}}\Big)\,\partial_k + Z^l\,\Big(\,\Gamma^k_{jm}\Gamma^m_{il}\,\tfrac{\partial f^i}{\partial u}\,\tfrac{\partial f^j}{\partial v}\\
&\qquad\qquad - \Gamma^k_{im}\Gamma^m_{jl}\,\tfrac{\partial f^i}{\partial u}\,\tfrac{\partial f^j}{\partial v}\,\Big)\,\partial_k\\
&= Z^l\,\Big(\frac{\partial\Gamma^k_{ml}}{\partial v}\,\frac{\partial f^m}{\partial u} - \frac{\partial\Gamma^k_{ml}}{\partial u}\,\frac{\partial f^m}{\partial v}\Big)\,\partial_k + \Gamma^k_{jm}\Gamma^m_{il}Z^l\Big(\frac{\partial f^i}{\partial u}\,\frac{\partial f^j}{\partial v} - \frac{\partial f^i}{\partial v}\,\frac{\partial f^j}{\partial u}\Big)\,\partial_k,
\end{aligned}
$$

where for the last equality we have replaced the summation indices i and j by m in the first two terms. Now the r.h.s. of (3.1.20) equals the r.h.s. of (3.1.19) upon exchanging the summation indices i and m in the last term and we are done. □

With this we can now prove the following characterisation for the vanishing of the curvature tensor.

3.1.7 Theorem (Locally flat SRMF). *For any SRMF M the following are equivalent:*

(i) M is locally flat, *that is, for all points $p \in M$ there is a chart (U,φ) around p where the metric is flat, i.e., $(\varphi_* g)_{ij} = \varepsilon_j\delta_{ij}$ on $\varphi(U)$.*

(ii) The Riemann tensor vanishes.

Proof. (i)⇒(ii) follows simply from 3.1.3 and the fact that in the chart φ the Christoffel symbols all vanish.

(ii)⇒(i): The statement is local, so we may assume that $M = \mathbb{R}^n$ and $p = 0$. Let $e_1,\dots,e_n$ be an ONB at 0 and choose the e_i as coordinate axes for coordinates $x^1,\dots,x^n$ on $\mathbb{R}^n$. Then we have $\partial_i = e_i$ for $1 \le i \le n$.
Now for each i we first parallel transport e_i along the x^1-axis ($t \mapsto (t,0,\dots,0)$) and then from each point t_0 on the x^1-axis along the x^2-axis ($t \mapsto (t_0,t,0,\dots,0)$) and so on for the x^3, x^4,…-axes.
In this way we obtain vector fields $E_1,\dots,E_n \in \mathfrak{X}(\mathbb{R}^n)$ which are smooth since parallel transport is governed by an ODE whose solutions by ODE-theory depend smoothly on the initial data. Moreover since parallel transport preserves scalar products (1.3.28) the E_i form an ONB at every $q \in \mathbb{R}^n$.
Now for $1 \le k \le n$ set $M_k := \mathbb{R}^k \times \{0\} \subseteq \mathbb{R}^n$. By construction for all $1 \le j \le n$ we have that $E_j|_{M_k}$ is a vector field along the mapping $f_k : (t_1,\dots,t_k) \mapsto \sum_{i=1}^k t_i e_i = (t_1,\dots,t_k,0,\dots,0)$

and we may consider $\frac{\nabla}{\partial t_i}(E_j|_{M_k})$ for all $1 \leq i \leq k$, see 41 but now for all k parameters. Let now $j \in \{1, \dots, n\}$

We claim $\frac{\nabla}{dt_i}(E_j|_{M_k}) = 0$ for all $1 \leq i \leq k$.
We proceed by induction. For $k = 1$ the equality follows by definition of parallel transport.
$k \mapsto k+1$: Let $\frac{\nabla}{dt_i}(E_j|_{M_k}) = 0$ for all $1 \leq i \leq k$. By construction we also have $\frac{\nabla}{dt_{k+1}}(E_j|_{M_{k+1}}) = 0$. Now by 3.1.6 and our assumption that $R = 0$ we have

$$\frac{\nabla}{dt_{k+1}} \frac{\nabla}{dt_i} E_j(t_1, \dots, t_{k+1}, 0, \dots, 0) = \frac{\nabla}{dt_i} \frac{\nabla}{dt_{k+1}} E_j(t_1, \dots, t_{k+1}, 0, \dots, 0) = 0. \tag{3.1.21}$$

So $0 = \frac{\nabla}{dt_i} E_j(t_1, \dots, t_k, 0, \dots, 0)$ is parallely transported along the straight line $t_{k+1} \mapsto (t_1, \dots, t_k, t_{k+1}, 0, \dots, 0)$ hence vanishes on all of M_{k+1}.

Next we claim that $E_1, \dots, E_n \in \mathfrak{X}(\mathbb{R}^n)$ are parallel.
It suffices to show that $\nabla_{\partial_i} E_j = 0$ for all i, j. For fixed $x^1, \dots, x^n \in \mathbb{R}^n$ let $c_i : t \mapsto (x^1, \dots, x^{i-1}, t, x^{i+1}, \dots, x^n)$. Then we have by 1.3.27(iii)

$$\nabla_{\partial_i} E_j(x^1, \dots, x^n) = \nabla_{c_i'(x^i)} E_j = \frac{\nabla}{dt_i}(E_j \circ c)|_{t_i = x^i} = 0, \tag{3.1.22}$$

where for the last equality we have used the previous claim in case $k = n$.

Finally from the latter claim we have $[E_i, E_j] = \nabla_{E_i} E_j - \nabla_{E_i} E_j = 0$ for all i, j. So there are coordinates $((y^1, \dots, y^n), V)$ locally around p such that $E_j|_V = \partial_{y^i}$ for all $1 \leq j \leq n$. But in this coordinates we have

$$g_{ij} = g(\partial_{y^i}, \partial_{y^j}) = g(E_i, E_j) = \varepsilon_i \delta_{ij}. \tag{3.1.23}$$

□

Finally we come to item (2) of our list on page 65 which we make precise in the following remark. Here we closely follow [7, II.22 (p. 65)].

3.1.8 Remark (Riemann tensor and parallel transport along closed curves). Let M be a SRMF, $p \in M$, $Z \in T_pM$ and c be a curve in M with $c(0) = p$. Further let $(x^1, \dots, x^n)$ be coordinates around p and write as usual $c^i(t)$ for the coordinates of c w.r.t. this chart. Let now $Z(t) = Z^i(t)\partial_i|_{c(t)}$ be the vector field obtained from parallely transporting $Z = Z(0)$ along c. By (1.3.56) we then have

$$\frac{dZ^i(t)}{dt} + \Gamma^i_{kl}(x(c(t))) \frac{dc^k}{dt} Z^l(t) = 0. \tag{3.1.24}$$

Let now more generally $f : I \times J \to M$ a smooth two-parameter map with I and J intervals around 0. We denote by $x^i(u, v) = x^i \circ f(u, v)$ the local coordinates of f. For some fixed pair $(u, v) \in I \times J$ we define the 'corner points' (see Figure ♣ `insert figure` ♣)

$P = f(0,0)$, $Q = f(u,0)$, $R = f(u,v)$ and $S = f(0,v)$. Now we transport $Z = Z_P \in T_PM$ parallel along f first to Q then to R and S and then back to P. We will denote the values of this vector field Z on f at the 'corner points' by Z_Q, Z_R, Z_S and finally $\bar{Z}_P = Z(u,v)$. In general the resulting vector $\bar{Z}_P$ will not equal the staring value Z_P but depend smoothly on u and v. Indeed the solutions of the ODE (3.1.24) depend smoothly on the data Q, R and S as well as the right hand side which themselves depend smoothly on (u,v).
Now we expand the components of these vectors w.r.t. the coordinates x^i in a Taylor series which gives

$$Z^i_Q = Z^i_P + \Big(\frac{\partial Z^i}{\partial u}\Big)_P u + \frac{1}{2}\Big(\frac{\partial^2 Z^i}{\partial u^2}\Big)_P u^2 + \mathcal{O}(u^3), \tag{3.1.25}$$

$$Z^i_R = Z^i_Q + \Big(\frac{\partial Z^i}{\partial v}\Big)_Q v + \frac{1}{2}\Big(\frac{\partial^2 Z^i}{\partial v^2}\Big)_Q v^2 + \mathcal{O}(v^3), \tag{3.1.26}$$

$$Z^i_S = Z^i_R + \Big(\frac{\partial Z^i}{\partial u}\Big)_R u - \frac{1}{2}\Big(\frac{\partial^2 Z^i}{\partial u^2}\Big)_R u^2 + \mathcal{O}(u^3), \tag{3.1.27}$$

$$\bar{Z}^i_P = Z^i_S + \Big(\frac{\partial Z^i}{\partial v}\Big)_S v - \frac{1}{2}\Big(\frac{\partial^2 Z^i}{\partial v^2}\Big)_S v^2 + \mathcal{O}(v^3). \tag{3.1.28}$$

Inserting (3.1.25) into (3.1.26), (3.1.26) into (3.1.27), and (3.1.27) into (3.1.28) we obtain for the difference between the starting an the final vector

$$\begin{aligned} \Delta Z^i_P := \bar{Z}^i_P - Z^i_P = &\bigg(\Big(\frac{\partial Z^i}{\partial u}\Big)_P - \Big(\frac{\partial Z^i}{\partial u}\Big)_R\bigg)\, u \;+\; \bigg(\Big(\frac{\partial Z^i}{\partial v}\Big)_Q - \Big(\frac{\partial Z^i}{\partial v}\Big)_S\bigg)\, v \\ &+ \bigg(\Big(\frac{\partial^2 Z^i}{\partial u^2}\Big)_P - \Big(\frac{\partial^2 Z^i}{\partial u^2}\Big)_R\bigg)\, \frac{u^2}{2} + \bigg(\Big(\frac{\partial^2 Z^i}{\partial v^2}\Big)_Q - \Big(\frac{\partial^2 Z^i}{\partial v^2}\Big)_S\bigg)\, \frac{v^2}{2} + \dots \end{aligned} \tag{3.1.29}$$

Now we assume that the x^i are Riemannian normal coordinates at p which by 2.1.17(ii) leads to the vanishing of the first term in the Taylor expansion of the Christoffel symbols and we obtain

$$\Gamma^i{}_{jk}(x) = \sum_m \Big(\frac{\partial}{\partial x^m}\Gamma^i{}_{jk}\Big)_P x^m + \mathcal{O}(x^2). \tag{3.1.30}$$

The next and laborious step consist in calculating the coefficients $(\partial Z^i/\partial u)_P$, $(\partial^2 Z^i/\partial u^2)_P$ etc. from the ODE (3.1.24). We obtain

(i) $\displaystyle \frac{\partial Z^i}{\partial u}\bigg|_P = 0$

Indeed we have by parallel transport (cf. (3.1.24)) along the curve $u \mapsto x(u,0)$ that $\partial_u Z^i(u,0) = -\Gamma^i{}_{kl}(x(u,0))\,\partial_u Z^k(u,0)\,Z^l(u,0)$ and so $\partial_u Z^i(0,0) = -\Gamma^i{}_{kl}(P)\cdot\ldots = 0$, since we have assumed x^i to be Riemannian normal coordinates around P.

(ii) $\displaystyle \frac{\partial Z^i}{\partial v}\bigg|_Q = -\Big(\frac{\partial}{\partial x^m}\,\Gamma^i{}_{jk}\,\frac{\partial x^m}{\partial u}\,\frac{\partial x^j}{\partial v}\,Z^k\Big)_P u + \dots$

Parallel transport along $v \mapsto x(u,v)$ from Q to R gives $(\partial_u Z^i)_Q = -(\Gamma^i_{jk}\,\partial_v x^j\, Z^k)_Q$. Now by Taylor expansion at P (cf. (3.1.30)) we obtain $\Gamma^i_{jk}(Q) = \partial_m \Gamma^i_{jk}(P)\, x^m + \dots$ and also $x^m(Q) = 0 + \partial_u x^m(P)\, u + \partial_v x^m(P)\, 0 + \dots$, since $Q = (u,0)$. Similarly $\partial_v x^j(Q) = \partial_v x^j(P) + u \cdot \ldots + 0 \cdot \ldots + \dots$, and $Z^k(Q) = Z^k(P) + \partial_u Z^k(P)\, u + \partial_v Z^k(P)\, 0 + \dots$. Collecting terms together we finally arrive at the asserted result.

(iii) $$\frac{\partial Z^i}{\partial u}\Big|_R = -\Big(\frac{\partial}{\partial x^m}\,\Gamma^i_{jk}\,\frac{\partial x^m}{\partial u}\,\frac{\partial x^j}{\partial u}\,Z^k\Big)_P u - \Big(\frac{\partial}{\partial x^m}\,\Gamma^i_{jk}\,\frac{\partial x^m}{\partial v}\,\frac{\partial x^j}{\partial u}\,Z^k\Big)_P v + \dots$$

Now we use parallel transport from R to S along $s \mapsto x(u-s,v)$, $s \in [0,u]$ to obtain $\partial_s Z^i(R) = -\big(\Gamma^i_{jk}(x(u-s,v))\,\partial_s(x^j(u-s,v))\,Z^k\big)_R$. Next note that $\partial_s(Z^i(u-s,v)) = -\partial_u Z(u-s,v)$ and $\partial_s(x^j(u,s,v)) = -\partial_u x^j(u-s,v)$ and so we have $\partial_u Z^i(R) = -\big(\Gamma^i_{jk}(x(u-.,v))\,\partial_u x^j(u-.,v)\,Z^k\big)_R = -(\Gamma^i_{jk}\,\partial_u x^j\,Z^k)_R$. Again Taylor expansion now at P gives $\Gamma^i_{jk}(R) = \partial_m\Gamma^i_{jk}(P)x^m(R) + \dots = \partial_m\Gamma^i_{jk}(P)\partial_u x^m(P)\,u + \partial_v x^m(P)\,v + \dots$, and $\partial_j(R) = \partial_u x^j(P) + u \cdot \ldots + v \cdot \ldots$. Moreover by (3.1.25), (3.1.26) $Z^k(R) = Z^k(P) + \partial_u Z^k(P)u + \partial_v Z^k(Q)\,v + \dots$. Again collecting together the respective terms we obtain the asserted formula.

(iv) $$\frac{\partial Z^i}{\partial v}\Big|_S = -\Big(\frac{\partial}{\partial x^m}\,\Gamma^i_{jk}\,\frac{\partial x^m}{\partial v}\,\frac{\partial x^j}{\partial v}\,Z^k\Big)_P v + \dots$$

We parallely transport Z along $t \mapsto (0, v-t)$ from S to P to obtain $\partial_t\big(Z^i(0,v-t)\big) = -\big(\Gamma^i_{jk}(x(0,v-t))\,\partial_t(x^j(0,v-t))Z^k\big)_S$ which by analogous reasoning as in (iii) gives $\partial_v Z^i(S) = -(\Gamma^i_{jk}\,\partial_v\, x^j\, Z^k)_S$, where by Taylor expansion in P, $\Gamma^i_{jk}(S) = \partial_m\Gamma^i_{jk}(P)x^m(S) + \dots = \partial_m\Gamma^i_{jk}(P)\,(\partial_u x^m \cdot 0 + \partial_v x^m \cdot v)_P + \dots$ as well as $\partial_v x^j(S) = \partial_v x^j(P) + u\cdot\ldots + v\cdot\ldots + \dots$ and using (3.1.25)–(3.1.27) $Z^k(S) = Z^k(P) + \dots$ Once more collecting the terms gives the result.

(v) $$\frac{\partial^2 Z^i}{\partial u^2}\Big|_P = -\Big(\frac{\partial \Gamma^i_{jk}}{\partial x^m}\,\frac{\partial x^m}{\partial u}\,\frac{\partial x^j}{\partial u}\,Z^k\Big)_P + \dots$$

We again use parallel transport from P to Q which by (3.1.24) gives $\partial_u Z^i = -\Gamma^i_{jk}\partial_u x^j Z^k$ and hence $\partial^2 Z^i/\partial u^2 = -\partial_m\Gamma^i_{jk}\partial_u x^m \partial_u x^j Z^k - \Gamma^i_{jk}\partial^2 x^i/\partial u^2\, Z^k - \Gamma^i_{jk}\partial_u x^j \partial_u Z^k$ which gives without the need to use any expansion $(\partial^2 Z^i/\partial u^2)_p = (-\partial_m\Gamma^i_{jk}\partial_u x^m \partial_u x^j Z^k)_P$.

(vi) $$\frac{\partial^2 Z^i}{\partial u^2}\Big|_R = -\Big(\frac{\partial \Gamma^i_{jk}}{\partial x^m}\,\frac{\partial x^m}{\partial u}\,\frac{\partial x^j}{\partial u}\,Z^k\Big)_P + \dots$$

Here we use (3.1.24) along the curve $s \mapsto (u-s,v)$ to to obtain $\partial_s(Z^i(u-s,v)) = -\Gamma^i_{jk}(x(u-s,v))\partial_s(x^j(u-s,v))Z^k(x(u-s,v))$ and so $\partial_u Z^i(u-s,v) = -\Gamma^i_{jk}(x(u-s,v))\partial_u x^j(u-s,v)Z^k(x(u-s,v))$. But this gives $\partial^2 Z^i/\partial u^2(u-s,v) = -\partial_m\Gamma^i_{jk}(x(u-s,v))\partial_u x^m(u-s,v)Z^k(x(u-s,v)) - \Gamma^i_{jk}(x(u-s,v))\partial^2 x^j/\partial u^2(u-s,v)Z^k(x(u-s,v)) - \Gamma^i_{jk}(x(u-s,v))\partial_u x^j(u-s,v)\partial_s(Z^k(x(u-s,v)))$. Finally expansion at P gives $(\partial^2 Z^i/\partial u^2)_R = -(\partial_m\Gamma^i_{jk}\partial_u x^m\partial_u x^j Z^k)_P - (\Gamma^i_{jk}\partial^2 xj/\partial u^2 Z^k)_P - (\Gamma^i_{jk}\partial_u x^j \Gamma^k_{ml}\partial_u x^m Z^l)_p + \dots = -(\partial_m\Gamma^i_{jk}\partial_u x^m \partial_u x^j Z^k)_P$.

(vii) $\left.\dfrac{\partial^2 Z^i}{\partial v^2}\right|_Q = -\Big(\dfrac{\partial \Gamma^i{}_{jk}}{\partial x^m}\,\dfrac{\partial x^m}{\partial v}\,\dfrac{\partial x^j}{\partial v}\,Z^k\Big)_P + \dots$

From $\partial_v Z^i = -\Gamma^i{}_{jk}\partial_v x^j Z^k$ we obtain that $\partial^2 Z^i/\partial v^2 = -\partial_m\Gamma^i{}_{jk}\partial_v x^m \partial_v x^j Z^k - \Gamma^i{}_{jk}$ $\partial^2 x^j/\partial v^2 Z^k$ $-\Gamma^i{}_{jk}\partial_v x^j \partial_v Z^k$ and so as in (vi) we find $(\partial^2 Z^i/\partial v^2)_Q = -(\partial_m\Gamma^i{}_{jk}\partial_v x^m \partial_v x^j$ $Z^k)_P + \dots$.

(viii) $\left.\dfrac{\partial^2 Z^i}{\partial v^2}\right|_S = -\Big(\dfrac{\partial \Gamma^i{}_{jk}}{\partial x^m}\,\dfrac{\partial x^m}{\partial v}\,\dfrac{\partial x^j}{\partial v}\,Z^k\Big)_P + \dots$

As in (iv) and (vi) the minus signs coming from the inner derivatives compensate to give $\partial^2 Z^i/\partial v^2 = \partial_m\Gamma^i{}_{jk}\partial_v x^m\partial_v x^j Z^k - \Gamma^i{}_{jk}\partial^2 x^j/\partial v^2 Z^k - \Gamma^i{}_{jk}\partial_v x^j \partial_v Z^k$ and so gain as in (vi) $(\partial^2 Z^i/\partial v^2)_S = -(\partial_m\Gamma^i{}_{jk}\partial_v x^m\partial_v x^j Z^k)_P + \dots$.

Now we may plug (i)–(viii) into (3.1.29) to arrive at

$$\Delta Z^i_P = \left(\Big(\frac{\partial \Gamma^i{}_{jk}}{\partial x^m} - \frac{\partial \Gamma^i{}_{mk}}{\partial x^j}\Big)\,\frac{\partial x^j}{\partial u}\,\frac{\partial x^m}{\partial v}\,Z^k\right)_P uv + \dots \tag{3.1.31}$$

Further by 3.1.3 and 2.1.17(ii) we obtain

$$R^i_{kmj}\,|_P = \left(\frac{\partial \Gamma^i{}_{mk}}{\partial x^j} - \frac{\partial \Gamma^i{}_{jk}}{\partial x^m} + \Gamma^i{}_{jr}\Gamma^r{}_{mk} - \Gamma^i{}_{ms}\Gamma^s{}_{jk}\right)_P = -\left(\frac{\partial \Gamma^i{}_{jk}}{\partial x^m} - \frac{\partial \Gamma^i{}_{mk}}{\partial x^j}\right)_P \tag{3.1.32}$$

and so we finally obtain

$$\Delta Z^i_P = -\Big(R^i_{kmj}\,\frac{\partial x^j}{\partial u}\,\frac{\partial x^m}{\partial v}\,Z^k\Big)_P uv + \dots = -R^i\Big(\frac{\partial x}{\partial v}\,\frac{\partial x}{\partial u}\Big)(Z)\,|_P\,uv + \dots. \tag{3.1.33}$$

From this result we may now immediately draw the following conclusions:

(1) If $R = 0$ by 3.1.7 M is locally isometric to $\mathbb{R}^n_r$ and since parallel transport then is trivial we obtain $\Delta Z^i_P = 0$.

(2) In the general case ΔZ^i_P is of second order in (u, v) and depends on the curvature tensor R at P. Hence we may view R as an obstruction to the vanishing of ΔZ^i_P. In particular, (3.1.33) gives the following alternative characterisation of the Riemann tensor

$$R^i\Big(\frac{\partial x}{\partial v}\,\frac{\partial x}{\partial u}\Big)(Z)\,|_P = -\lim_{u,v\to 0}\,\frac{1}{uv}\,\Delta Z^i_P. \tag{3.1.34}$$

3.2 Some differential operators

The aim of this section is to introduce on a SRMF the generalisations of the classical differential operators of *gradient, divergence* and *Laplacian*. To deal with these in an appropriate way we need some preparations, namely we need to introduce the operations of *type-changing* of higher order tensor fields and *metric contraction*.

The former operation is nothing but the generalisation of the musical isomorphism of 1.3.3 of vector fields and one-forms to higher order tensors. To achieve this goal we proceed as follows: Given a tensor field $A \in \mathcal{T}^r_s(M)$ on a SRMF (M, g) we define for any $1 \leq a \leq r$, $1 \leq b \leq s$ the tensor field $\downarrow^a_b A \in \mathcal{T}^{r-1}_{s+1}(M)$ via

$$\begin{aligned}(\downarrow^a_b A)(\omega^1, \ldots, \omega^{r-1}, X_1 \ldots, X_{s+1})\\ := A(\omega^1, \ldots, \underset{\text{slot } a}{X^\flat_b}, \ldots, \omega^{r-1}, X_1 \ldots, X_{b-1}, X_{b+1}, \ldots, X_{s+1}),\end{aligned} \tag{3.2.1}$$

where $X^\flat_b$ is the metric equivalent one-form of the vector field X_b, cf. 1.3.3. So on the r.h.s. we extract the bth vector field and insert its metrically equivalent one-form in the ath slot among the one-forms. It is instructive to consider an example.

3.2.1 Example (Index lowering). Let A be a $(2, 2)$-tensor field on M, then $B :=\downarrow^1_2 A$ is the $(1, 3)$-tensor field given by

$$B(\omega, X, Y, Z) = A(Y^\flat, \omega, X, Z). \tag{3.2.2}$$

Now let $(x^1, \ldots, x^n)$ be local coordinates on M. Then first observe that $\partial^\flat_k = g_{km} dx^m$ since $\partial^\flat_k(V^m\partial_m) = \langle\partial_k, V^m\partial_m\rangle = g_{km}V^m$. And so we have

$$B^i_{jkl} = B(dx^i, \partial_j, \partial_k, \partial_l) = A(g_{km}dx^m, dx^i, \partial_j, \partial_l) = g_{km}A^{mi}_{jl}. \tag{3.2.3}$$

We now see that the operation $\downarrow^1_2$ changes the first upper index of A into the second lower index of B.

The operator $\downarrow^a_b\colon\ \mathcal{T}^r_s(M) \to \mathcal{T}^{r-1}_{s+1}(M)$ is classically also called the *lowering*[1] of the respective indices. It is obviously $\mathcal{C}^\infty$-linear and moreover it is an isomorphism with inverse $\uparrow^a_b\colon\ \mathcal{T}^{r'}_{s'}(M) \to \mathcal{T}^{r'+1}_{s'-1}(M)$ given by

$$\begin{aligned}(\uparrow^a_b A)(\omega^1, \ldots, \omega^{r'+1}, X_1, \ldots, X_{s'-1})\\ := A(\omega^1, \ldots, \omega^{a-1}, \omega^{a+1}, \ldots \omega^{r'+1}, X_1, \ldots, \underset{\text{slot } b}{(\omega^a)^\sharp}, \ldots, X_{s'-1}),\end{aligned} \tag{3.2.4}$$

where now $(\omega^b)^\sharp$ is the vector field metrically equivalent to the one-form ω^b. This operation extracts the the ath one-form and inserts its metrically equivalent vector field in the bth slot among the vector fields and is classically called the *raising* of the respective index. Generally we call all tensors which are derived from a given tensor by raising or lowering an index *metrically equivalent*.

We again look at an example where we also demonstrate that the lowering and raising of indices are inverse operations.

[1]This terminology reflects that classical differential geometry was exclusively written in coordinates.

3.2.2 Example (Index lowering). First observe that in local coordinates we have $(dx^i)^\sharp = g^{ij}\partial_j$ since $\langle g^{ij}\partial_j, V^k\partial_k\rangle = \delta^i_k V^k = V^i = dx^i(V^k\partial_k)$. Now for a $(1,3)$-tensor field B we have

$$(\uparrow^1_2 B)^{ij}_{kl} = (\uparrow^1_2 B)(dx^i, dx^j, \partial_k, \partial_l) = B(dx^j, \partial_k, g^{im}\partial_m, \partial_l) = g^{im}\, B^j_{kml}. \tag{3.2.5}$$

So as expected $\uparrow^1_2$ turns the second lower index into the first upper index using the inverse metric. Now to check that it is the inverse of $\downarrow^1_2$ we write using equations (3.2.3) and (3.2.5)

$$(\uparrow^1_2\downarrow^1_2 A)^{ij}_{kl} = g^{im}(\downarrow^1_2 A)^j_{kml} = g^{im}\, g_{mn}\, A^{nj}_{kl} = \delta^i_n\, A^{nj}_{kl} = A^{ij}_{kl}. \tag{3.2.6}$$

We give another example to emphasise how natural actually type changing is; in fact it often occurs in calculations without even being noticed.

3.2.3 Example (Type changing). As in (1.3.24) we consider a $(1,s)$-tensor field A given as a $\mathcal{C}^\infty$-multilinear map $A : \mathfrak{X}(M)^s \to \mathfrak{X}(M)$. Then we have (using $\bar{A}$ as in (1.3.24))

$$\begin{aligned}(\downarrow^1_1 \bar{A})(V, X_1, \dots, X_s) &= \bar{A}(V^\flat, X_1, \dots, X_s)\\ &= V^\flat\big(A(X_1, \dots, X_s)\big) = \langle V, A(X_1, \dots, X_s)\rangle.\end{aligned} \tag{3.2.7}$$

Finally we point at one peculiar issue in dealing with the coordinate expression for the Riemann tensor which arises due to historic reasons. Actually the coordinate version of differential geometry was developed long before the invariant approach and in harmonising these two this issue requires some care.

3.2.4 Remark (Coordinate expression of the Riemann tensor). We have written the coordinate expression of the curvature tensor in 3.1.3 (according to the classical pattern) as

$$R_{\partial_k\partial_l}(\partial_j) = R^i_{jkl}\partial_i, \text{ hence the order of arguments is } R_{XY}Z = R(Z, X, Y). \tag{3.2.8}$$

Indeed using the convention of (1.3.24) we obtain $R^i_{jkl} = \bar{R}(dx^i, \partial_j, \partial_k, \partial_l) = dx^i\big(R(\partial_j, \partial_k, \partial_l)\big)$. The components of the $(0,4)$-tensor $\downarrow^1_1 R$ are then given by

$$\begin{aligned}R_{ijkl} = (\downarrow^1_1 \bar{R})(\partial_i, \partial_j, \partial_k, \partial_l) &= \langle \partial_i, R(\partial_j, \partial_k, \partial_l)\rangle\\ &= \langle \partial_i, R_{\partial_k\partial_l}(\partial_j)\rangle = \langle \partial_i, R^m_{jkl}\partial_m\rangle = g_{im}R^m_{jkl},\end{aligned} \tag{3.2.9}$$

where we have used (3.2.7).

Next we turn to the operation of *metric contraction*. On smooth manifolds we have introduced the contraction $\mathcal{C}^i_j : \mathcal{T}^r_s(M) \to \mathcal{T}^{r-1}_{s-1}$ on page 22. There the ith contravariant index or slot is contracted with the jth covariant one, i.e., in coordinates $(\mathcal{C}^i_j A)^{i_1\dots i_{r-1}}_{j_1\dots j_{s-1}} = A^{i_1\dots \overset{i}{m}\dots i_r}_{j_1\dots \underset{j}{m}\dots j_s}$. On a SRMF we may use the metric to also contract two covariant or two contravariant slots by first raising respectively lowering the respective index, that is we combine the metric type changing with the contraction. More precisely let $1 \leq a < b \leq s$, then for arbitrary r we define $\mathcal{C}_{ab} : \mathcal{T}^r_s(M) \to \mathcal{T}^r_{s-2}(M)$ locally by

$$(\mathcal{C}_{ab}A)^{i_1\ldots\ldots i_r}_{j_1\ldots j_{s-2}} := g^{lm}\, A^{i_1\ldots\ldots\ldots\ldots i_r}_{j_1\ldots \underset{a}{l} \ldots \underset{b}{m} \ldots j_{s-2}}. \tag{3.2.10}$$

Analogously we define for $1 \leq a < b \leq r$ and for s arbitrary $\mathcal{C}^{ab} : \mathcal{T}^r_s(M) \to \mathcal{T}^{r-2}_s(M)$ locally by

$$(\mathcal{C}^{ab}A)^{i_1\ldots\ldots i_{r-2}}_{j_1\ldots\ldots j_s} := g_{lm}\, A^{i_1\ldots \overset{a}{l} \ldots \overset{b}{m} \ldots i_{s-2}}_{j_1\ldots\ldots\ldots\ldots j_r}. \tag{3.2.11}$$

We now have the following compatibility result.

3.2.5 Lemma (Metric contraction and ∇). *On a SRMF (M, g) the covariant derivative as well as the covariant differential commute with type changing and metric contraction.*

Proof. For the case of the covariant derivative and type changing it suffices to consider the case $\downarrow^a_1$ since the assertion then follows by permutation for $\downarrow^a_b$ and by the following argument for $\uparrow^a_b$: Let $B = \downarrow^a_b A$ then we have

$$\uparrow^a_b \nabla_V B = \uparrow^a_b \nabla_V(\downarrow^a_b A) = \uparrow^a_b \downarrow^a_b \nabla_V A = \nabla_V A = \nabla_V \uparrow^a_b B. \tag{3.2.12}$$

Now to consider $\downarrow^a_1 A$ first note that $\downarrow^a_1 A = \mathcal{C}^a_1(g \otimes A)$. Indeed in coordinates we have (cf. (3.2.3)) $\mathcal{C}^a_1(g \otimes A)^{i_1\ldots i_{r-1}}_{mj_1\ldots j_s} = g_{ml} A^{i_1\ldots l \ldots i_r}_{j_1\ldots\ldots j_s}$. Now by Definition 1.3.22, ∇_V is a tensor derivation which by Definition 1.3.16(ii) commutes with contractions and moreover satisfies the metric property ($\nabla 5$) (cf. 1.3.25(iv)) so we obtain

$$\nabla_V(\downarrow^a_1 A) = \nabla_V\big(\mathcal{C}^a_1(g \otimes A)\big) = \mathcal{C}^a_1 \nabla_V(g \otimes A) = \mathcal{C}^a_1(g \otimes \nabla_V A) = \downarrow^a_1 \nabla_V A. \tag{3.2.13}$$

By equations 3.2.10 resp. (3.2.11) metric contraction is just the composition of type changing and contraction, so ∇_V also commutes with this operation.

The analogous assertions for the covariant differential ∇ follow easily from those of ∇_V. We demonstrate this for type changing just in a special case which, however, makes clear how to proceed in the general case. Let $A \in \mathcal{T}^2_1(M)$, then $\nabla A \in \mathcal{T}^2_2(M)$, $\downarrow^1_1 \nabla A \in \mathcal{T}^1_3(M)$ and we have

$$\begin{aligned}(\downarrow^1_1 \nabla A)(\omega, X, Y, Z) &= \nabla A(X^\flat, \omega, Y, Z) = \nabla_Z A(X^\flat, \omega, Y) = (\downarrow^1_1 \nabla_Z A)(\omega, X, Y)\\ &= \big(\nabla_Z(\downarrow^1_1 A)\big)(\omega, X, Y) = \big(\nabla(\downarrow^1_1 A)\big)(\omega, X, Y, Z).\end{aligned} \tag{3.2.14}$$

Finally one easily verifies that ∇ commutes with tensor products and contractions and hence with metric contraction. □

Now we are finally in a position to introduce the above mentioned differential operators on SRMFs.

3.2.6 Definition (Gradient). *For a function $f \in \mathcal{C}^\infty(M)$ we define its* gradient grad(f) *(or* grad$f$*, for short) as the vector field metrically equivalent to $df \in \Omega^1(M)$, i.e.,*

$$\langle \mathrm{grad}(f), X\rangle = df(X) = X(f) \qquad \text{for all } X \in \mathfrak{X}(M). \tag{3.2.15}$$

We clearly see that while the differential df of a function is defined on any smooth manifold it needs a metric to define the gradient. In local coordinates we have $df = \partial_i f dx^i$ and so

$$\mathrm{grad} f = g^{ij}\, \frac{\partial f}{\partial x^i}\, \frac{\partial}{\partial x^j} = \uparrow_1^1 df \tag{3.2.16}$$

since $\langle g^{ij}\partial_i f \partial_j, v^k\partial_k\rangle = \partial_i f V^k \delta^i_k = \partial_i f V^i = \partial_i f dx^i(V^k\partial_k)$.
As a simple example we note that on flat space $\mathbb{R}^n_r$ we have grad$f = \varepsilon_i \partial_i f dx^i$ which on $\mathbb{R}^n$ reduces to the well known formula grad$f = \partial_i f \partial_i$.

3.2.7 Definition (Divergence). *For a tensor field A we call a* divergence *of A every contraction of the new covariant slot of ∇A with any of its original contravariant slots.*

We discuss some special cases. For a vector field $V \in \mathfrak{X}(M)$ the only possibility is div$V = \mathcal{C}(\nabla V)$ which in coordinates reads using 1.3.9(ii)

$$\mathrm{div} V = \mathcal{C}(\nabla V) = dx^i(\nabla_{\partial_i} V) = dx^i \left(\Big(\frac{\partial V^m}{\partial x^i} + \Gamma^m_{ik} V^k \Big) \partial_m \right) = \sum_i \Big(\frac{\partial V^i}{\partial x^i} + \Gamma^i_{ik} V^k \Big). \tag{3.2.17}$$

In the special case of flat space $\mathbb{R}^n_r$ we obtain the well-know formula from analysis div$V = \sum_i \partial_i V^i$.

3.2.8 Definition (Hessian). *The* Hesse tensor H^f, *or* Hessian *for short, of a function $f \in \mathcal{C}^\infty(M)$ is defined as the second covariant differential of f, i.e.,*

$$H^f = \nabla(\nabla f). \tag{3.2.18}$$

3.2.9 Lemma (Hessian explicitly). *The Hessian H^f of $f \in \mathcal{C}^\infty(M)$ is a symmetric $(0, 2)$-tensor field and we have*

$$H^f(X, Y) = XYf - (\nabla_X Y)f = \langle \nabla_X(\mathrm{grad} f), Y\rangle. \tag{3.2.19}$$

Proof. Since $\nabla f = df$ we have

$$\begin{aligned} H^f(X, Y) = \nabla(df)(X, Y) &= \big(\nabla_Y(df)\big)(X) \\ &= Y\big(df(X)\big) - df(\nabla_Y X) = Y(Xf) - (\nabla_Y X)(f). \end{aligned} \tag{3.2.20}$$

Symmetry now follows from the torsion free condition (∇4), i.e., by $XY - YX = [X, Y] = \nabla_X Y - \nabla_Y X$. Finally we have by the metric condition (∇5)

$$\langle \nabla_X(\mathrm{grad} f), Y\rangle = X\langle \mathrm{grad} f, Y\rangle - \langle \mathrm{grad} f, \nabla_X Y\rangle = XYf - \nabla_X Y(f) = H^f(X, Y).$$

□

In local coordinates we hence have for the Hessian

$$H^f_{ij} = H^f(\partial_i, \partial_j) = \partial_i\partial_j f - (\nabla_{\partial_i}\partial_j) f = \partial_i\partial_j f - \Gamma^k_{ij}\partial_k f. \tag{3.2.21}$$

3.2.10 Definition (Laplace). *The Laplace-Beltrami operator on a SRMF* (M, g) *is the mapping*

$$\triangle : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M), \qquad \triangle f = \mathrm{div}(\mathrm{grad} f). \tag{3.2.22}$$

More explicitly we have

$$\begin{aligned} \triangle f = \mathrm{div}(\mathrm{grad} f) &= \mathcal{C}\big(\nabla(\mathrm{grad} f)\big) = \mathcal{C}\big(\nabla(\uparrow^1_1 df)\big) \\ &= \mathcal{C}\big(\uparrow^1_1 \nabla(df)\big) = (\mathcal{C}\uparrow^1_1) H^f = \mathcal{C}_{12}(H^f), \end{aligned} \tag{3.2.23}$$

and we see that the Laplace-Beltrami operator is the metric contraction of the Hessian tensor. In local coordinates we hence obtain

$$\triangle f = g^{ij} H^f_{ij} = g^{ij}(\partial_i\partial_j f - \Gamma^k_{ij}\partial_k f), \tag{3.2.24}$$

which on flat $\mathbb{R}^n_r$ gives $\triangle f = \sum \varepsilon_i \partial^2 f/\partial x_i^2$. Obviously this gives the Laplace operator on $\mathbb{R}^n$ and the wave operator on Minkowski space $\mathbb{R}^n_1$. This is the reason why on Riemanninan and Lorentzian manifolds $\triangle$ often is called the Laplace and the wave operator, respectively.

3.3 The Einstein equations

In this section we introduce the famous Einstein equations, i.e., the fundamental equations of General Relativity (GR), Albert Einstein's eminent theory of space, time and gravitation, which is the currently best available physical description of our universe at large. Just a year ago and only briefly after its centennial GR has seen a spectacular success by the direct observation of gravitational waves emitted from a binary black hole merger.

The Einstein equations are the so called field equations of GR and link the geometry and, in particular, the curvature of the spacetime manifold to its energy-matter content. Here we collect the mathematical prerequisites for their formulation.

We start by introducing the *Ricci tensor* and the *curvature scalar*, two 'curvature quantities' derived from the Riemann tensor.

3.3.1 Definition (Ricci tensor). *Let* (M, g) *be a SRMF with Riemann tensor* R. *The* Ricci tensor *Ric is defined as the contraction* $\mathcal{C}^1_3 R \in \mathcal{T}^0_2(M)$.

The Ricci tensor's local coordinates are denoteby by R_{ij} and take the form (cf. page 22)

$$R_{ij} = R^m_{ijm}. \tag{3.3.1}$$

Moreover Ric is symmetric by pair symmetry of the Riemann tensor 3.1.2(iv) since using RNC centered at a point p we have there (cf 2.1.17(i)

$$\begin{aligned}\mathrm{Ric}(X,Y) = (\mathcal{C}^1_3 R)(X,Y) &= R(dx^i, X, Y, \partial_i)\\ &= dx^i\big(R(X,Y,\partial_i)\big) = \varepsilon_i\langle R_{Y\partial_i}X, \partial_i\rangle = \varepsilon_i\langle R_{X\partial_i}Y, \partial_i\rangle. \end{aligned}\tag{3.3.2}$$

Also we note the trace formula $\mathrm{Ric}(X,Y) = \mathrm{trace}(V \mapsto R_{XV}Y)$. A SRMF with Ric $= 0$ is called *Ricci flat*. Clearly any flat manifold $R = 0$ is also Ricci flat but the converse is not true as we shall discuss below and which is essential for GR.
We proceed introducing the curvature scalar or scalar curvature of a SRMF.

3.3.2 Definition (Scalar curvature). *The* scalar curvature S *of the SRMF* (M,g) *is defined as the contractions of the Ricci tensor,* $S = \mathcal{C}(Ric) \in \mathcal{C}^\infty(M)$.

Observe that since Ric $\in \mathcal{T}^0_2(M)$ the contraction $\mathcal{C}$ unambiguously stands for $\mathcal{C}_{11}$, cf. (3.2.10). In local coordinates we have

$$S = g^{ij}R_{ij} = g^{ij}R^m_{ijm} \tag{3.3.3}$$

For our further considerations we rely on the following property of the curvature tensor.

3.3.3 Proposition (Second Bianchi identity). *For* $x, y, z \in T_pM$ *we have*

$$(\nabla_z R)(x,y) + (\nabla_x R)(y,z) + (\nabla_y R)(z,x) = 0. \tag{3.3.4}$$

Proof. As in the proof of 3.1.2 we may extend x, y, z arbirarily to vector fields X, Y, Z on a neighbourhood U of p. In this case we choose these extensions in such a way that their coefficients are constant w.r.t. some normal coordinates at p. Then again all Lie brackets vanish on U and moreover by 2.1.17(ii) all Christoffel symbols vanish at p and hence also all covariant derivatives of X, Y, Z w.r.t. each other vanish at p.
Using the notation used in the proof of 3.1.2(iii) we have to show that $S\nabla_Z R(X,Y) = 0$. Now by the product rule 1.3.18 we have for arbitrary V at p

$$\big((\nabla_Z R)(X,Y)\big)(V) = \nabla_Z\big(R(X,Y)V\big) - R(\underbrace{\nabla_Z X}_{=0}, Y)V - R(X, \underbrace{\nabla_Z Y}_{=0})V - R(X,Y)(\nabla_Z V)$$

and so again at p

$$(\nabla_Z R)(X,Y) = [\nabla_Z, R(X,Y)] = [\nabla_Z, [\nabla_Y, \nabla_X], \tag{3.3.5}$$

where we have used (3.1.4). Since the Jacobi identity holds also for ∇_X (cf. [10, 2.5.15(iii)]), the sum over all cyclic permutations of $(\nabla_Z R)(X,Y)$ vanishes as claimed. □

3.3.4 Corollary (Divergence of Ricci). *We have* $dS = 2\mathrm{div}(Ric)$.

Proof. By 3.1.3 and 1.3.25(v) we have in coordinates

$$(\nabla R)(\partial_j, \partial_k, \partial_l, \partial_r) = (\nabla_{\partial_r} R)_{\partial_k \partial_l}(\partial_j) = R^i_{jkl;r}\partial_i, \tag{3.3.6}$$

which upon using 3.3.3 gives $R^i_{jkl;r} + R^i_{jlr;k} + R^i_{jrk;l} = 0$. Now interchanging r and k in the final term (which by 3.1.2(i) causes a sign change) and contracting i with r gives

$$0 = R^r_{jkl;r} + R^r_{jlr;k} - R^r_{jkr;l} = R^r_{jkl;r} + R_{jl;k} - R_{jk;l} \tag{3.3.7}$$

and so

$$g^{jk} R^r_{jkl;r} + g^{jk} R_{jl;k} - S_{;l} = 0. \tag{3.3.8}$$

Next we note that $R_{mjkl} = R_{jmlk}$. Indeed by 3.1.2(i),(ii) we find using (3.2.9) $R_{jmlk} = g_{rj} R^r_{mlk} = \langle \partial_j, \partial_r \rangle R^r_{mlk} = \langle \partial_j, R^r_{mlk}\partial_r \rangle = \langle R_{\partial_l \partial_k}(\partial_m), \partial_j \rangle = -\langle R_{\partial_k \partial_l}(\partial_m), \partial_j \rangle = \langle R_{\partial_k \partial_l}(\partial_j), \partial_m \rangle = R_{mjkl}$. Moreover we have $R^r_{jkl} = g^{rm} R_{mjkl}$ and so

$$g^{jk} R^r_{jkl;r} = g^{jk} g^{rm} R_{mjkl;r} = g^{jk} g^{rm} R_{jmlk;r} = g^{rm} R^k_{mlk;r} = g^{rm} R_{ml;j} = R^r_{l;r}. \tag{3.3.9}$$

So by (3.3.8) we find

$$R^r_{l;r} + R^k_{l;k} = 2R^r_{l;r} = S_{;l}. \tag{3.3.10}$$

Finally since Ric is symmetric (cf. (3.3.2)) we have $\mathcal{C}_{13}(\nabla \mathrm{Ric}) = \mathcal{C}_{23}(\nabla \mathrm{Ric}) = \mathrm{div}(\mathrm{Ric})$, which in coordinates reads $g^{rs} R_{sl;r} = R^r_{l;r}$. So (3.3.10) gives $2\mathrm{div}(\mathrm{Ric}) = \nabla S = ds$. □

We now very briefly discuss the basic principles of General Relativity. Naturally any discussion in the setting of this course has to be superficial and we refer e.g. to [13, Ch. 4] for a more appropriate account.

The stage of GR is *spacetime* which is the set of all events (t, x), labelled by a one-dimensional time coordinate and a three-dimensional space coordinate. Spacetime can be a model of e.g. the surroundings of a star, our solar system, or our universe as a whole. Mathematically spacetime is described by a 4-dimensional Lorentzian manifold (M, g), where the Lorentzian signature is chosen as to implement the causality structure already present in special relativity.

Now contrary to classical Newtonian physics gravity is *not* described as a force field on this manifold M but rather as the *curvature of spacetime*. This ground breaking idea which Einstein famously called his happiest thought, relies on taking the *principle of equivalence* to be the basic building block of the theory. Indeed, due to Galileo's principle of equivalence, all bodies fall the same in a gravitational field, so gravity can be thought of as being a 'property' of spacetime!

One can also argue why this 'property' has to be related to curvature. Generalising the Newtonian idea that bodies which move freely, i.e., without any force acting upon them move along straight paths, freely falling test bodies in GR should move along geodesics of spacetime. Now considering test bodies falling freely in a gravitational field of a point

mass in the Newtonian picture one sees that they undergo a relative acceleration, due to so-called tidal forces. Translated into the spacetime perspective this means that geodesics focus—and the quantity that focusses geodesics clearly is curvature.
Consequently the curvature of spacetime has to be related to physical forces, or better to the all the mass and energy it contains. (Mass is equivalent to energy by the famous equation $E = mc^2$.) Already in classical mechanics and electrodynamics the matter variables (forces, strain, stress, etc.) are described by a single object, the so-called energy momentum tensor T which is a symmetric $(0, 2)$-tensor field. Moreover T is divergence free and this property implements *energy conservation*, another basic principle in all of physics.
So specifically Einstein in 1915 was looking for the correct equation that relates T to the curvature of spacetime. In a time where Riemannian resp. Lorentzian geometry has by far not been developed to its present state he first tried several variants of the Ricci curvature in his attempt to describe the perihelion precession of the planet mercury. However, the Ricci tensor is not divergence free and so he finally introduced the following quantity.

3.3.5 Definition (Einstein tensor). *Let (M, g) be a Lorentzian manifold. We define the* Einstein tensor *as*

$$G := Ric - \frac{1}{2}\, S\, g. \tag{3.3.11}$$

The essential properties of G are now:

3.3.6 Lemma (Properties of G). *The Einstein tensor of a spacetime has the following properties:*

(i) G is a symmetric and divergence free $(0, 2)$-tensor field.

(ii) $Ric = G - \frac{1}{2}\mathcal{C}(G)g$.

Proof. (i) Symmetry of G follows immediately from symmetry of Ric and g. To calculate the divergence $\operatorname{div}(Sg) = \mathcal{C}_{13}(\nabla(Sg))$ we write

$$\begin{aligned}\nabla(Sg)(X, Y, Z) &= \nabla_Z(Sg)(X, Y) = \nabla_Z(Sg(X, y)) - Sg(\nabla_Z X, Y) - Sg(X, \nabla_Z Y)\\ &= (\nabla_Z S)\, g(X, Y) + S\nabla_Z(g(X, Y)) - Sg(\nabla_Z X, Y) - Sg(X, \nabla_Z Y).\end{aligned} \tag{3.3.12}$$

Moreover by $(\nabla 5)$ we have $0 = (\nabla_Z g)(X, Y) = \nabla_Z(g(X, Y)) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y)$ and so

$$\nabla(Sg)(X, Y, Z) = (\nabla_Z S)\, g(X, Y) = dS(Z)\, g(X, Y) = g \otimes dS(X, Y, Z). \tag{3.3.13}$$

Now for convenience proceeding in coordinates we find $(\nabla(Sg))_{ijk} = g_{ij}(dS)_k = g_{ij}\partial_k S$ and hence

$$\operatorname{div}(Sg)_j = \mathcal{C}_{13}(\nabla(Sg))_j = g^{ki} g_{ij} \frac{\partial S}{\partial x^k} = \frac{\partial S}{\partial x^j} = (dS)_j. \tag{3.3.14}$$

So we finally arrive at $\operatorname{div}(Sg) = dS$ and so by 3.3.4

$$\operatorname{div}(G) = \operatorname{div}(\mathrm{Ric} - \frac{1}{2}\,Sg) = \frac{1}{2}(dS - dS) = 0. \tag{3.3.15}$$

(ii) We have $\mathcal{C}(g) = g^{ij}g_{ij} = \delta^i_i = \dim M = 4$, hence by definitions 3.3.2 and 3.3.5 $\mathcal{C}(G) = \mathcal{C}(\mathrm{Ric}) - 1/2\,S\mathcal{C}(g) = S - 2S = -S$ and finally

$$\mathrm{Ric} = G + \frac{1}{2}Sg = G - \frac{1}{2}\,\mathcal{C}(G)g. \tag{3.3.16}$$

□

The significance of the previous results lie in the fact that (i) says that in the light of the above discussion G is a formally qualified candidate for the curvature quantity to be equated with T, while (ii) guarantees that it is also a sensible one, since it encodes the same information as Ric. So we finally arrive at:

The Einstein equations. If (M, g) is a spacetime with energy momentum tensor T then

$$G \;=\; \frac{8\pi N}{c^4}\,T. \tag{3.3.17}$$

Here $N = 6.67 \cdot 10^{-11} m^3/(kg \cdot s^2)$ is Newton's gravitational constant and $c = 2.99 \cdot 10^8 m/s$ is the speed of light in vacuum. Usually one sets $N/c^4 = 1$ which amounts to using so-called geometric units. In the very important special case of vacuum, i.e., in the absence of matter, the equations reduce to

$$\mathrm{Ric} \;=\; 0, \tag{3.3.18}$$

since taking the trace of (3.3.17) gives $\mathcal{C}(G) = S = 8\pi\mathcal{C}(T)$, which clearly vanishes for vacuum. So vacuum solutions to Einstein equations are Ricci flat but far from (locally) flat, i.e. $R = 0$, as is exemplified e.g. by the notorious *Schwarzschild metric* which is the (unique) spherically symmetric solution of (3.3.18) and provides the simplest model of a *black hole*.

Now in a sense General Relativity is the study of solutions of the Einstein equations. From the coordinate formulae one sees that they form a highly complicated system of (by symmetries of G) 10 coupled nonlinear (quasilinear, to be precise) partial differential equations for g. Although there are literally thousands of known exact solutions to (3.3.17) accompanied by a big wealth of deep results in Lorentzian geometry and also recently the global existence theory of Einstein's equations has made great advances it is still fair to say that one is far from reaching a comprehensive understanding of their full content. So General Relativity is a very active field of research today, combining many fascinating aspects of (Lorentzian) geometry and analysis.

Bibliography

[1] Abraham, R., Marsden, J.E., Foundations of Mechanics. Benjamin/Cummings, 1978.

[2] Bishop, R.L., Goldberg, S.I., Tensor Analysis on Manifolds. Dover, 1980.

[3] Boothby, W.M., An Introduction to Differentiable Manifolds and Riemannian Geometry. Academic Press, 1986.

[4] Brickel, F., Clark, R.S., Differentiable Manifolds. An Introduction. Van Nostrand, 1970.

[5] Čap, A., Differentialgeometrie 1 (Lecture Notes, `http://www.mat.univie.ac.at/∼cap/lectnotes.html`).

[6] do Carmo, M., Differential forms and applications. Springer, 1994.

[7] Eisenhart, L., Riemannian Geometry, 1949

[8] Kriegl, A., Differentialgeometrie (Lecture Notes, `http://www.mat.univie.ac.at/∼kriegl/Skripten/diffgeom.pdf`).

[9] Kühnel, W., Differentialgeometrie. Vieweg, 1999.

[10] Kunzinger, M. Differentialgeometrie (Lecture Notes, `www.mat.univie.ac.at/ mike/teaching/ss08/dg.pdf`).

[11] Michor, P.W., Topics in Differential Geometry, AMS, 2008.

[12] O'Neill, B., Semi-Riemannian Geometry. With Applications to Relativity. Academic Press, 1983.

[13] Wald, R., General Relativity, The University of Chicago Press, 1984.