Geometry and Physics

Catherine Meusburger, Karl-Hermann Neeb

30.7.2012

Contents

1 Manifolds, vector fields, curves and flows

1.1 Smooth manifolds and smooth maps

Definition 1.1. [Charts, Atlas] Let M be a topological space.

(a) A pair (φ, U), consisting of an open subset $U \subseteq M$ and a homeomorphism $\varphi\colon U \to \varphi(U) \subseteq \mathbb{R}^n$ of U onto an open subset of $\mathbb{R}^n$ is called an *n-dimensional chart of M*.

(b) Two n-dimensional charts (φ, U) and (ψ, V) of M are said to be *C^k-compatible* ($k \in \mathbb{N} \cup \{\infty\}$) if $U \cap V = \emptyset$ or if $U \cap V \neq \emptyset$ and the homeomorphism

$$\psi \circ \varphi^{-1}|_{\varphi(U\cap V)}\colon \varphi(U \cap V) \subset \mathbb{R}^n \to \psi(U \cap V) \subset \mathbb{R}^n$$

is a C^k-diffeomorphism.

(c) An *n-dimensional C^k-atlas of M* is a family $\mathcal{A} := (\varphi_i, U_i)_{i\in I}$ of n-dimensional charts of M with the following properties:

(A1) $\bigcup_{i\in I} U_i = M$, i.e. $(U_i)_{i\in I}$ is an open covering of M.

(A2) All charts (φ_i, U_i), $i \in I$, are pairwise C^k-compatible: for all $i, j \in I$ the homeomorphisms

$$\varphi_{ji} := \varphi_j \circ \varphi_i^{-1}|_{\varphi_i(U_{ij})}\colon \varphi_i(U_{ij}) \to \varphi_j(U_{ij}) \qquad U_{ij} := U_i \cap U_j$$

are C^k-maps.

(d) A chart (φ, U) is called *compatible* with a C^k-atlas $(\varphi_i, U_i)_{i\in I}$ if it is C^k-compatible with all charts of the atlas $\mathcal{A}$. A C^k-atlas $\mathcal{A}$ is called *maximal* if it contains all charts compatible with it. A maximal C^k-atlas is also called a *C^k-differentiable structure on M*. For $k = \infty$ we also call it a *smooth structure*.

Remark 1.2.

1. Every atlas $\mathcal{A}$ is contained in a unique maximal atlas: We simply add all charts compatible with $\mathcal{A}$, and thus obtain a maximal atlas. This atlas is unique (Exercise 1.2). This implies that every C^k-atlas $\mathcal{A}$ defines a unique C^k-differentiable structure on M.
2. A given topological space M may carry different differentiable structures. Examples are the exotic differentiable structures on $\mathbb{R}^4$ (the only $\mathbb{R}^n$ carrying exotic differentiable structures) and the 7-sphere $\mathbb{S}^7$.

Definition 1.3. [C^k-manifold] An *n-dimensional C^k-manifold* is a pair $(M, \mathcal{A})$ consisting of a Hausdorff space M and a maximal n-dimensional C^k-atlas $\mathcal{A}$ for M. For $k = \infty$ we call it a *smooth manifold*.

Example 1.4. [Open subsets of $\mathbb{R}^n$] Let $U \subseteq \mathbb{R}^n$ be an open subset. Then U is a Hausdorff space with respect to the induced topology. The inclusion map $\varphi\colon U \to \mathbb{R}^n$ defines a chart (φ, U) which already defines a smooth atlas of U, turning U into an n-dimensional smooth manifold.

Example 1.5. [Products of manifolds] Let M and N be smooth manifolds of dimensions d, resp., k and

$$M \times N = \{(m, n)\colon m \in M, n \in N\}$$

the product set, which we endow with the product topology.

We show that $M \times N$ carries a natural structure of a smooth $(d+k)$-dimensional manifold. Let $\mathcal{A} = (\varphi_i, U_i)_{i\in I}$ be an atlas of M and $\mathcal{B} = (\psi_j, V_j)_{j\in J}$ an atlas of N. Then the product sets $W_{ij} := U_i \times V_j$ are open in $M \times N$ and the maps

$$\gamma_{ij} := \varphi_i \times \psi_j \colon U_i \times V_j \to \mathbb{R}^d \times \mathbb{R}^k \cong \mathbb{R}^{d+k}, \quad (x, y) \mapsto (\varphi_i(x), \psi_j(y))$$

are homeomorphisms onto open subsets of $\mathbb{R}^{d+k}$. On $\gamma_{i'j'}(W_{ij} \cap W_{i'j'})$ we have

$$\gamma_{ij} \circ \gamma_{i'j'}^{-1} = (\varphi_i \circ \varphi_{i'}^{-1}) \times (\psi_j \circ \psi_{j'}^{-1}),$$

which is a smooth map. Therefore $(\varphi_{ij}, W_{ij})_{(i,j)\in I\times J}$ is a smooth atlas on $M \times N$.

Example 1.6. [Submanifolds of $\mathbb{R}^n$] Smooth *k-dimensional submanifolds of* $\mathbb{R}^n$ are often defined as follows:

A smooth k-dimensional submanifold of $\mathbb{R}^n$ is a subset $M \subset \mathbb{R}^n$ such that for every point $p \in M$ there is an open neighborhood $U_p \subset \mathbb{R}^n$ and a smooth function $f_p : U_p \to \mathbb{R}^{n-k}$ such that $M \cap U_p = f_p^{-1}(0)$ and $\mathrm{rank}(\mathtt{d}f_p(x)) = n - k$ for all $x \in U_p \cap M$.

Every smooth k-dimensional submanifold $M \subset \mathbb{R}^n$ has a natural structure as a k-dimensional manifold. Firstly, M has a natural structure of a Hausdorff space as a subset of the Hausdorff space $\mathbb{R}^n$. Moreover, by the Implicit Function Theorem there exist for each $p \in M$ an open neighbourhood $V_p \subset U_p \subset \mathbb{R}^n$ of p, an open neighbourhood $W_p \subset \mathbb{R}^n$ of 0 and a smooth diffeomorphism $\varphi_p : V_p \to W_p$ such that

$$\varphi_p(U_p \cap M) = (\mathbb{R}^k \times \{0\}) \cap W_p$$

If we identify $\mathbb{R}^k \times \{0\}$ with $\mathbb{R}^k$ then $(\varphi_p|_{V_p\cap M}, V_p \cap M)$ is a chart for M and for $V_p \cap V_q \cap M \neq \emptyset$,

$$\begin{aligned}\varphi_q \circ \varphi_p^{-1}|_{\varphi_p(V_p\cap V_q\cap M)} &= (\varphi_q|_{V_q\cap M}) \circ (\varphi_p|_{V_p\cap M})^{-1}|_{\varphi_p(V_p\cap V_q\cap M)}\\ \varphi_p(V_p \cap V_q \cap M) &\to \varphi_q(V_p \cap V_q \cap M)\end{aligned}$$

is a smooth map onto an open subset of $\mathbb{R}^d$. We thus obtain a smooth atlas of M.

Many manifolds that play an important role in physics are submanifolds of $\mathbb{R}^n$. A particularly simple example are quadrics, which cover already many relevant examples.

Example 1.7. [quadrics] A *quadric* Q in $\mathbb{R}^n$ is the set of zeros of a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(x) = \langle x, Ax + b\rangle + c.$$

where $A \in \mathrm{Mat}(n, \mathbb{R})$ is a symmetric matrix, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ and $\langle\, ,\, \rangle$ denotes the Euclidean scalar product on $\mathbb{R}^n$. The gradient of the function f is given by

$$\mathrm{grad}\, f(x) = 2Ax + b,$$

and its zeros are precisely the solutions of the linear equation $2Ax + b = 0$. If none of its solutions lies on the affine plane $\langle x, b\rangle + 2c = 0$, we have

$$f(x) = \tfrac{1}{2}\langle x, 2Ax + b + b\rangle + c = \tfrac{1}{2}\langle x, b\rangle + c \neq 0 \quad \forall x \in (\mathrm{grad} f)^{-1}(0)$$

and the quadric $Q = f^{-1}(0)$ is an $(n-1)$-dimensional smooth manifold. In particular, this is the case for all quadrics with $b = 0$, $c \neq 0$ and $A \in \mathrm{GL}_n(\mathbb{R})$ and for all quadrics with $A = 0$ and $b \neq 0$. This includes the following cases, all of which are $(n-1)$-dimensional submanifolds of $\mathbb{R}^n$:

- $A = 0$, $b \neq 0$: *affine hyperspaces* of $\mathbb{R}^n$

$$M = \{x \in \mathbb{R}^n : \langle x, b\rangle + c = 0\} \qquad \text{with } b \neq 0.$$

- $A = \mathbf{1}_n$, $b = 0$ and $c = -1$: the $(n-1)$-*sphere*

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \langle x, x\rangle = 1\}$$

- $A = \mathrm{diag}(1, -1, \ldots, -1)$, $b = 0$ and $c = -1$: $(n-1)$-*dimensional hyperbolic space*

$$\mathbb{H}^{n-1} = \Big\{x \in \mathbb{R}^n : x_n^2 - \sum_{i=1}^{n-1} x_i^2 = 1\Big\}.$$

- $A = \mathrm{diag}(1, -1, \ldots, -1)$, $b = 0$, $c = 1$: $(n-1)$-dimensional *de Sitter space*

$$\mathrm{dS}_{n-1} = \Big\{x \in \mathbb{R}^n : x_n^2 - \sum_{i=1}^{n-1} x_i^2 = -1\Big\}.$$

- $A = \mathrm{diag}(\underbrace{1, \ldots, 1}_{p\times}, \underbrace{-1, \ldots, -1}_{(q+1)\times})$, $b = 0$, $c = 1$: (p,q)-*anti de Sitter space*

$$\mathrm{AdS}_{(p,q)} = \Big\{x \in \mathbb{R}^n : \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{n} = -1\Big\}, \qquad p + q = n - 1.$$

With the results of Example 1.6, it is easy to check that a given subset of $\mathbb{R}^n$ is a submanifold and to determine its dimension. However, for many purposes this is not enough since one needs an explicit description of M in terms of coordinates. Although the definition of the smooth structure on M is based on a maximal smooth atlas, in practice it is advisable to describe a manifold with as few charts as possible. If an n-dimensional manifold cannot be realised as an open subset of $\mathbb{R}^n$, it is clear that one needs at least two charts, and in many examples, this is already sufficient.

Example 1.8. [The n-dimensional sphere] We consider the unit sphere

$$\mathbb{S}^n := \{(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \colon x_0^2 + x_1^2 + \ldots + x_n^2 = 1\}$$

in $\mathbb{R}^n$, endowed with the subspace topology, turning it into a compact space.

(a) To specify a smooth manifold structure on $\mathbb{S}^n$, we consider the open subsets

$$U_i^\varepsilon := \{x \in \mathbb{S}^n \colon \varepsilon x_i > 0\}, \quad i = 0, \ldots, n, \quad \varepsilon \in \{\pm 1\}.$$

These $2(n+1)$ subsets form a covering of $\mathbb{S}^n$. We have homeomorphisms

$$\varphi_i^\varepsilon \colon U_i^\varepsilon \to B := \{x \in \mathbb{R}^n \colon \|x\|_2 < 1\}$$

onto the open unit ball in $\mathbb{R}^n$, given by

$$\varphi_i^\varepsilon(x) = (x_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

and with continuous inverse map

$$(y_1, \ldots, y_n) \mapsto \Big(y_1, \ldots, y_i, \varepsilon\sqrt{1 - \|y\|_2^2}, y_{i+1}, \ldots, y_n\Big).$$

This leads to charts $(\varphi_i^\varepsilon, U_i^\varepsilon)$ of $\mathbb{S}^n$.

It is easy to see that these charts are pairwise compatible. We have $\varphi_i^\varepsilon \circ (\varphi_i^{\varepsilon'})^{-1} = \mathrm{id}_B$, and for $i < j$, we have

$$\varphi_i^\varepsilon \circ (\varphi_j^{\varepsilon'})^{-1}(y) = \Big(y_1, \ldots, y_i, y_{i+2}, \ldots, y_j, \varepsilon'\sqrt{1 - \|y\|_2^2}, y_{j+1}, \ldots, y_n\Big),$$

which is a smooth map

$$\varphi_j^{\varepsilon'}(U_i^\varepsilon \cap U_j^{\varepsilon'}) \to \varphi_i^\varepsilon(U_i^\varepsilon \cap U_j^{\varepsilon'}).$$

(b) There is another atlas of $\mathbb{S}^n$ consisting only of two charts, where the maps are slightly more complicated.

We call the unit vector $e_0 := (1, 0, \ldots, 0)$ the *north pole* of the sphere and $-e_0$ the *south pole*. We then have the corresponding *stereographic projection maps*

$$\varphi_+ \colon U_+ := \mathbb{S}^n \setminus \{e_0\} \to \mathbb{R}^n, \quad (y_0, y) \mapsto \frac{1}{1 - y_0} y$$

and

$$\varphi_- \colon U_- := \mathbb{S}^n \setminus \{-e_0\} \to \mathbb{R}^n, \quad (y_0, y) \mapsto \frac{1}{1 + y_0} y.$$

Both maps are bijective with inverse maps

$$\varphi_\pm^{-1}(x) = \Big(\pm \frac{\|x\|_2^2 - 1}{\|x\|_2^2 + 1}, \frac{2x}{1 + \|x\|_2^2}\Big)$$

(Exercise 1.4). This implies that (φ_+, U_+) and (φ_-, U_-) are charts of $\mathbb{S}^n$. That both are smoothly compatible, hence a smooth atlas, follows from

$$(\varphi_+ \circ \varphi_-^{-1})(x) = (\varphi_- \circ \varphi_+^{-1})(x) = \frac{x}{\|x\|^2}, \quad x \in \mathbb{R}^n \setminus \{0\},$$

which is the inversion at the unit sphere.

Example 1.9. [Counterexample: the Double Cone] Consider for $n \geq 2$ the double cone

$$M = \{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} : \ x_0^2 - (x_1^2 + \ldots + x_n^2) = 0\}.$$

As a subset of the Hausdorff space $\mathbb{R}^{n+1}$, M is a Hausdorff space. However, M is not a smooth manifold. To show this, note that $M \setminus \{0\}$ is an n-dimensional manifold by Example 1.6:

$$M = f^{-1}(0) \qquad \text{with} \quad f : \mathbb{R}^{n+1} \to \mathbb{R}, \quad f(x) = x_0^2 - (x_1^2 + \ldots + x_n^2)$$

and

$$\text{grad} f(x_0, \ldots, x_n) = 2(x_0, -x_1, \ldots, -x_n) \neq 0 \qquad \forall (x_0, \ldots, x_n) \neq (0, \ldots, 0).$$

Suppose now that there exists a chart (φ, U) with $0 \in U$, $U \subset M$ open, and a homeomorphism $\varphi : U \to \varphi(U) \subset \mathbb{R}^2$. Then $U \setminus \{0\}$ cannot be connected, because it contains at least one point $x \in M$ with $x_0 > 0$ and one point $y \in M$ with $y_0 < 0$. After applying a translation and restricting φ to a suitable subset, we can suppose $\varphi(0) = 0$ and

$$\varphi(U) = B_\epsilon(0) = \{x \in \mathbb{R}^n : \; ||x||_2 \leq \epsilon\}$$

with $\epsilon > 0$. This implies that $U \setminus \{0\} = \varphi^{-1}(B_\epsilon(0) \setminus \{0\})$ is connected, since φ is a homeomorphism and $B_\epsilon(0) \setminus \{0\}$ is connected for $n \geq 2$. Contradiction.

Although a large class of manifolds can be realised as submanifolds of $\mathbb{R}^n$ for some $n \in \mathbb{N}$, viewing a manifold as a subset of $\mathbb{R}^n$ is inadequate for several reasons. Firstly, it often leads to a very complicated description of the manifolds. Secondly, it is contrary to the sprit of differential geometry, in which the central structures are the charts and transitions between them, not the embedding into $\mathbb{R}^n$. This is also reflected in its applications in physics. In general relativity, a spacetime is described by a manifold, and this description is crucial for the interpretation of the theory. Embedding this manifold into $\mathbb{R}^n$ would correspond to introducing an absolute time and space outside of the spacetime manifold and hence to a Newtonian viewpoint.

For this reason, it natural to ask if one can construct a manifold "from scratch" instead of defining it as a certain subset of $\mathbb{R}^n$ or, more generally, of a Hausdorff topological space. This is possible and is called the *gluing construction* of manifolds. The idea is to start from certain open subsets $V_i \subset \mathbb{R}^n$ and C^k-homeomorphisms $\varphi_{ji} : V_{ij} \to V_{ji}$ between certain open subsets $V_{ij} \subset V_i$, which identify (glue together) the subsets V_{ij} and V_{ji}.

Definition 1.10. [Gluing data] A set of *gluing data* on $\mathbb{R}^n$ is a countable triple

$$((V_i)_{i \in I}, (V_{i,j})_{i,j \in I}, (\varphi_{ji})_{(i,j) \in K})$$

consisting of

- pairwise disjoint, non-empty open subsets $V_i \subset \mathbb{R}^n$, the *parametrisation domains*
- open subsets $V_{ij} \subset V_i$ such that $V_{ii} = V_i$ for all $i \in I$ and $V_{ij} = \emptyset$ if and only if $V_{ji} = \emptyset$ for all $i, j \in I$. The non-empty sets V_{ij} with $j \neq i$ are called *gluing domains*.
- C^k-diffeomorphisms $\varphi_{ji} : V_{ij} \to V_{ji}$ for $(i,j) \in K := \{(i,j) \in I \times I : V_{ij} \neq \emptyset\}$, the *gluing functions*

which satisfy the following conditions:

(G1) *cocycle condition*: for all $i, j, k \in I$ with $V_{ji} \cap V_{jk} \neq \emptyset$:

$$\varphi_{ji}^{-1}(V_{ji} \cap V_{jk}) \subset V_{ik} \qquad \text{and} \qquad \varphi_{ki}|_{\varphi_{ji}^{-1}(V_{ji}\cap V_{jk})} = \varphi_{kj} \circ \varphi_{ji}|_{\varphi_{ji}^{-1}(V_{ji}\cap V_{jk})}$$

(G2) *Hausdorff condition*: for all pairs of points $x \in (\partial V_{ij}) \cap V_i$, $y \in (\partial V_{ji}) \cap V_j$ with $(i, j) \in K$, $i \neq j$ there exist open neighbourhoods V_x of x and V_y of y with

$$(V_y \cap V_{ji}) \cap \varphi_{ji}(V_x \cap V_{ij}) = \emptyset.$$

Remark 1.11. (a) The cocycle condition implies $\varphi_{ii} = \mathrm{id}_{V_i} : V_i \to V_i$ for all $i \in I$ and $\varphi_{ij} = \varphi_{ji}^{-1}$ for all $(i, j) \in K$.

(b) A set of gluing data defines an equivalence relation on $V = \coprod_{i\in I} V_i$:

$$x \sim y \quad \Leftrightarrow \quad \exists (i, j) \in K \text{ with } x \in V_{ij}, y = \varphi_{ji}(x).$$

Reflexivity: For $x \in V_j$ the relation $\varphi_{jj}(x) = x$ implies $x \sim x$.
Symmetry: If $x \sim y$ and $(i, j) \in K$ with $y = \varphi_{ji}(x)$, then $\varphi_{ij} = \varphi_{ji}^{-1}$ implies $x = \varphi_{ij}(y)$ and hence $y \sim x$.
Transitivity: If $x \sim y$ and $y \sim z$ then there exist index pairs $(i, j), (j, k) \in K$ such that $x \in V_{ij}$, $y = \varphi_{ji}(x) \in V_{ji}$ and $y \in V_{jk}$, $z = \varphi_{kj}(y) \in V_{kj}$. As $V_{ji} \cap V_{jk} \neq \emptyset$, the cocyle condition implies $x \in \varphi_{ji}^{-1}(V_{ij} \cap V_{jk}) \subset V_{ik}$ and $\varphi_{ki}(x) = \varphi_{kj} \circ \varphi_{ji}(x) = \varphi_{kj}(y) = z$, so that $x \sim z$.

Proposition 1.12. *For every set of gluing data, the quotient $M = \left(\coprod_{i\in I} V_i\right) / \sim$ is an n-dimensional C^k-manifold.*

Proof. We consider the maps $t_i = p \circ \iota_i : V_i \to M$, $x \mapsto [x]$, where $\iota_i : V_i \to \coprod_{i\in I} V_i$ are the inclusion maps and $p : \coprod_{i\in I} V_i \to M$ the projection on the equivalence classes. We equip M with the finest topology such that all maps t_i are continuous, i.e. a subset $O \subseteq M$ is open if and only if, for every i, the inverse image $t_i^{-1}(O)$ is open. If $W \subseteq V_i$ is open, then $t_j^{-1} t_i(W) = \varphi_{ji}(W \cap V_{ij})$ is open in V_j for every $j \in I$, and therefore $t_i(W)$ open in M. As t_i is obviously injective, it is an homeomorphism onto an open subset of M.

For all $i \in I$, we thus obtain a chart $\varphi_i := t_i^{-1} : t_i(V_i) \to V_i$ for M, $M = \bigcup_{i\in I} t_i(V_i)$ and, for all $(i, j) \in K$,

$$\varphi_j \circ \varphi_i^{-1} = \varphi_{ji} : V_{ij} \to V_{ji},$$

is a C^k-function by definition. This shows that $(\varphi_i, t_i(V_i))_{i\in I}$ is an n-dimensional C^k-Atlas for M, and by adding all maps compatible with this atlas, we obtain a C^k-diffferentiable structure on M (see Remark 1.2).

It remains to prove that M is Hausdorff. For this, we first show that

$$t_i(V_i) \cap t_j(V_j) = \begin{cases} t_i(V_{ij}) = t_j(V_{ji}) & (i, j) \in K \\ \emptyset & (i, j) \notin K. \end{cases}$$

In fact, $t_i(x) \in t_i(V_i)$ is contained in $t_j(V_j)$ for some $i \neq j$ if and only if there exists a $y \in V_j$ with $t_i(x) = t_j(y)$, i.e. $(i, j) \in K$, $x \in V_{ij}$ and $\varphi_{ji}(x) = y$.

Consider now two distinct points in $p, q \in M$ with $p \neq q$. Then there exist $i, j \in I$ and $x_i \in V_i$, $y_j \in V_j$ with $p = t_i(x)$, $q = t_j(y)$. If $V_{ij} = \emptyset$, then $t_i(V_i)$ and $t_j(V_j)$ are disjoint open sets in M with $p \in t_i(V_i)$ and $q \in t_j(V_j)$. Otherwise, there are three cases:

(a) $q \in t_i(V_i)$ or $p \in t_j(V_j)$: Assume that $q \in t_i(V_i)$. Then there exist disjoint open neighbourhoods $V_x \subset V_i$ of x and $V_y \subset V_i$ of $\varphi_{ij}(y)$. Their images $t_i(V_x), t_i(V_y)$ contain, respectively, p and q and are open and disjoint since t_i is a homeomorphism. The case $p \in t_j(V_j)$ is similar.

This leaves the cases where $i \neq j$, $y \notin V_{ji}$ and $x \notin V_{ij}$.

(b) $y \notin \overline{V_{ji}}$ or $x \notin \overline{V_{ij}}$: Again, it suffices to deal with the first case. Then $t_i(V_i)$ and $t_j(V_j \setminus \overline{V_{ji}})$ are disjoint open subsets of M, containing p, resp., q.

(c) $x \in \partial V_{ij}$ and $y \in \partial V_{ji}$. Then the Hausdorff condition implies that there exist open neighbourhoods V_x of x and V_y of y with $(V_y \cap V_{ji}) \cap \varphi_{ji}(V_x \cap V_{ij}) = \emptyset$. This implies that the images $t_i(V_x)$ and $t_j(V_y)$ are disjoint open subsets of M with $p \in t_i(V_x)$ and $q \in t_j(V_y)$. $\square$

Remark 1.13. The Hausdorff condition is necessary to ensure that $M = \coprod_{i \in I} V_i/\sim$ is a Hausdorff space. Consider $\mathbb{R}^2$ with the parametrisation and gluing domains

$$V_1 =]-3, -1[\times]0, 1[, \quad V_2 =]1, 3[\times]0, 1[, \quad V_{12} =]-3, -2[\times]0, 1[, \quad V_{21} =]1, 2[\times]0, 1[$$

and the gluing function $\varphi_{21} : V_{12} \to V_{12}$, $\varphi_{21}(x_1, x_2) = (x_1 + 4, x_2)$. Then the sets V_1, V_2, V_{12}, V_{21} are open, φ_{12} is a diffeomorphism and the cocycle condition is satisfied trivially, since there are only two gluing domains, V_{12} and V_{21}.

However, $M = V_1 \coprod V_2/\sim$ is not Hausdorff. For all $x_2 \in]0, 1[$, the points $t_1((-2, x_2))$, $t_2((2, x_2)) \in V_1 \coprod V_2/\sim$ are distinct since $(-2, x_2) \in V_1 \setminus V_{12}$ and $(2, x_2) \in V_2 \setminus V_{21}$. However, they cannot be separated by disjoint open subsets of M , since every open neighbourhood $U_{-2} \subset M$ of $t_1((-2, x_2))$ contains a point $t_1((-2-\epsilon, x_2))$ and and every open neighbourhood $U_2 \subset M$ of $t_2((2, x_2))$ a point $t_2((2-\epsilon, x_2))$. As $t_1(-2-\epsilon, x_2) = t_2(\varphi_{21}(-2-\epsilon, x_2)) = t_1(2-\epsilon, x_2)$, it follows that $U_2 \cap U_{-2} \neq \emptyset$.

Example 1.14. [Cylinder and Möbius strip by gluing] We consider $\mathbb{R}^2$ with the parametrisation and gluing domains

$$\begin{aligned} V_1 &=]-4, -1[\times]0, 1[, & V_{12} &= (]-2, -1[\times]0, 1[) \cup (]-4, -3[\times]0, 1[) \subset V_1, \\ V_2 &=]1, 4[\times]0, 1[, & V_{21} &= (]1, 2[\times]0, 1[) \cup (]3, 4[\times]0, 1[) \subset V_2 \end{aligned}$$

and the gluing function $\varphi_{21} : V_{12} \to V_{21}$

$$\varphi_{21}(x_1, x_2) = \begin{cases} (x_1 + 3, x_2) & (x_1, x_2) \in]-2, -1[\times]0, 1[\\ (7 + x_1, x_2) & (x_1, x_2) \in]-4, -3[\times]0, 1[\end{cases}$$

The cocycle condition is again satisfied trivially, and to show that this defines a set of gluing data, the only condition to be checked is the Hausdorff condition. We have

$$\partial V_{12} \cap V_1 = (\{-3\}\times]0, 1[)\dot{\cup}(\{-2\}\times]0, 1[) \qquad \partial V_{21} \cap V_2 = (\{2\}\times]0, 1[)\dot{\cup}(\{3\}\times]0, 1[).$$

If $x \in \partial V_{12} \cap V_1$, $y \in \partial V_{21} \cap V_2$ and $0 < \epsilon < \frac{1}{2}$ then

$$\begin{aligned} &\varphi_{21}(V_{12} \cap B_\epsilon(x)) = B_\epsilon(\varphi_{21}(x)) \cap V_{21} \subset (]1, 1+\epsilon[\times]0, 1[) \cup (]4-\epsilon, 4[\times]0, 1[) \\ &V_{21} \cap B_\epsilon(y) \subset (]2-\epsilon, 2+\epsilon[\times]0, 1[) \cup (]3-\epsilon, 3+\epsilon[\times]0, 1[). \end{aligned}$$

This implies $\varphi_{21}(B_\epsilon(x)\cap V_{12})\cap(B_\epsilon(y)\cap V_{21})=\emptyset$. We can thus take $V_x=B_\epsilon(x)$, $V_y=B_\epsilon(y)$. Consequently, the Hausdorff condition is satisfied and $V_1\coprod V_2/\sim$ is a smooth manifold. One can show that $V_1\times V_2$ is homeomorphic to a cylinder $\mathbb{S}^1\times]0,1[$.

If we take the same parametrisation and gluing domains V_1, V_2, V_{12}, V_{21} but modify the gluing function to

$$\varphi_{21}(x_1,x_2)=\begin{cases}(x_1+3,x_2) & (x_1,x_2)\in]-2,-1[\times]0,1[\\ (7+x_1,1-x_2) & (x_1,x_2)\in]-4,-3[\times]0,1[,\end{cases}$$

it is easy to see that the Hausdorff condition is again satisfied. The resulting smooth manifold $V_1\coprod V_2/\sim$ is a Möbius strip.

Remark 1.15. Note that this construction is a refinement of the gluing of topological spaces in topology, where the openness of the subsets V_{ij} is not required. The difference is that the above gluing construction defines an n-dimensional C^k-manifold and not not just a topological space. To ensure this, additional conditions on the gluing data are necessary which are absent in the gluing of topological spaces.

Definition 1.16. [Smooth Maps, Diffeomorphisms] Let M and N be differentiable manifolds.

(a) We call a continuous map $f\colon M\to N$ *smooth in* $p\in M$ if there exist charts (φ,U) of M with $p\in U$ and (ψ,V) of N with $f(p)\in V$ auch that the map

$$\psi\circ f\circ\varphi^{-1}\colon\varphi(f^{-1}(V))\to\psi(V),\quad \varphi(x)\mapsto\psi(f(x)) \tag{1}$$

is smooth in a neighborhood of $\varphi(p)$. We call a continuous map $f\colon M\to N$ *smooth* if it is smooth in each point of M and write $C^\infty(M,N)$ for the set of smooth maps $f\colon M\to N$. If $N=\mathbb{R}$ we set $C^\infty(M):=C^\infty(M,\mathbb{R})$.

(b) A smooth map $f\colon M\to N$ is called a *smooth isomorphism* or a *diffeomorphism* if there exists a smooth map $g\colon N\to M$ with $g\circ f=\mathrm{id}_M$ and $f\circ g=\mathrm{id}_N$. We write $\mathrm{Diff}(M,N)$ for the set of diffeomorphisms of M to N and $\mathrm{Diff}(M):=\mathrm{Diff}(M,M)$. Two manifolds M and N are called *diffeomorphic* if there exists a diffeomorphism $f:M\to N$.

Remark 1.17. (a) The identity map $\mathrm{id}_M:M\to M$, $p\mapsto p$ is smooth, since for any chart (φ,U) of M the map $\varphi\circ\varphi^{-1}=\mathrm{id}_{\varphi(U)}$ is smooth.

(b) If $f\colon M\to N$ and $g\colon N\to Q$ are continuous maps with f smooth in $p\in M$ and g smooth in $f(p)$, then the composition $g\circ f$ is smooth in p. For charts (φ,U), (ψ,V), resp., (η,W) of M, N, resp., Q, with $p\in U$, $f(p)\in V$ and $g\circ f(p)\in W$ we have

$$\eta\circ(g\circ f)\circ\varphi^{-1}=(\eta\circ g\circ\psi^{-1})\circ(\psi\circ f\circ\varphi^{-1}),$$

on its natural domain, which contains a neighborhood of $\varphi(p)$.

(c) It follows from (a) and (b) that "diffeomorphic" is an equivalence relation on the class of smooth n-dimensional manifolds.

(d) It follows from (b) that, if $f\colon M\to N$ is smooth in p, then for any two charts (χ,W) of M with $p\in W$ and (ξ,Z) of N with $f(p)\in Z$, the map

$$\xi\circ f\circ\chi^{-1}\colon\chi(f^{-1}(Z))\to\xi(Z)$$

is smooth. Smoothness does not depend on the choice of charts.

(e) If U is an open subset of $\mathbb{R}^n$, then a map $f\colon U \to M$ to a smooth m-dimensional manifold M is smooth if and only if for each chart (φ, V) of M the map

$$\varphi \circ f\colon f^{-1}(V) \to \mathbb{R}^n$$

is smooth. Smoothness of maps $f\colon M \to \mathbb{R}^n$ can be checked more easily. Since the identity is a chart of $\mathbb{R}^n$, the smoothness condition simply means that for each chart (φ, U) of M the map

$$f \circ \varphi^{-1}\colon \varphi(f^{-1}(V) \cap U) \to \mathbb{R}^n$$

is smooth.

(f) Any chart (φ, U) of a smooth n-dimensional manifold M defines a diffeomorphism $U \to \varphi(U) \subseteq \mathbb{R}^n$, when U is endowed with the canonical manifold structure as an open subset of M. In fact, by definition, we may use (φ, U) as an atlas of U. Then the smoothness of φ is equivalent to the smoothness of the map $\varphi \circ \varphi^{-1} = \mathrm{id}_{\varphi(U)}$, which is trivial. Likewise, the smoothness of $\varphi^{-1}\colon \varphi(U) \to U$ is equivalent to the smoothness of $\varphi \circ \varphi^{-1} = \mathrm{id}_{\varphi(U)}$.

Example 1.18. If M and N are differentiable manifolds and $M \times N$ their product, then the following maps are smooth:

(a) the projection maps $p_M\colon M \times N \to M$ and $p_N\colon M \times N \to N$.

(b) for $x \in M$, $y \in N$, the embeddings

$$i_x\colon N \to M \times N, \quad y \mapsto (x, y) \qquad i^y\colon M \to M \times N, \quad x \mapsto (x, y)$$

(c) the diagonal embedding $\Delta_M\colon M \to M \times M, x \mapsto (x, x)$.

Definition 1.19. [Smooth Curve, Piecewise Smooth Curve]

(a) If $I \subseteq \mathbb{R}$ is an open interval, then a smooth map $\gamma\colon I \to M$ is called a *smooth curve.*

(b) For a not necessarily open interval $I \subseteq \mathbb{R}$, a map $\gamma\colon I \to \mathbb{R}^n$ is called smooth if all derivatives $\gamma^{(k)}$ exist in all points of I and define continuous functions $I \to \mathbb{R}^n$. Based on this generalization of smoothness for curves, a curve $\gamma\colon I \to M$ is said to be smooth, if for each chart (φ, U) of M the curves $\varphi \circ \gamma\colon \gamma^{-1}(U) \to \mathbb{R}^n$ are smooth.

(c) A curve $\gamma\colon [a, b] \to M$ is called *piecewise smooth* if γ is continuous and there exists a subdivision $x_0 = a < x_1 < \ldots, < x_N = b$ such that $\gamma|_{[x_i, x_{i+1}]}$ is smooth for $i = 0, \ldots N - 1$.

Exercises for Section 1.1

Exercise 1.1. Let $M := \mathbb{R}$, endowed with its standard topology. Show that C^k-compatibility of 1-dimensional charts is not an equivalence relation.

Exercise 1.2. Show that each n-dimensional C^k-atlas is contained in a unique maximal one.

Exercise 1.3. Let If M_i, $i = 1, \ldots, n$, be smooth manifolds of dimension d_i. Show that the product space $M := M_1 \times \ldots \times M_n$ carries the structure of a $(d_1 + \ldots + d_n)$-dimensional manifold.

Exercise 1.4. (a) Verify the details in Example 1.8, where we describe an atlas of $\mathbb{S}^n$ by stereographic projections.

(b) Show that the two atlasses of $\mathbb{S}^n$ constructed in Example 1.8 and the atlas obtained from the realization of $\mathbb{S}^n$ as a quadric in $\mathbb{R}^{n+1}$ define the same differentiable structure.

Exercise 1.5. Determine a smooth atlas for the n-dimensional hyperboloid

$$\mathbb{H}^n = \{x \in \mathbb{R}^{n+1} : x_{n+1}^2 - (x_1^2 + x_2^2 + \ldots + x_n^2) = 1\}$$

with as few maps as possible. What conclusions do you draw from this about the relation between the manifolds $\mathbb{H}^n$ and $\mathbb{R}^n$?

Exercise 1.6. Let $0 < r < R$. Determine a set of gluing data on $\mathbb{R}^2$ for the torus

$$T = \{(R\cos t + r\sin\varphi, R\sin t + r\sin\varphi, r\cos\varphi) : \ t, \varphi \in [0, 2\pi]\} \subset \mathbb{R}^3.$$

Exercise 1.7. Show that the set $A := C^\infty(M, \mathbb{R})$ of smooth real-valued functions on M is a real algebra. If $g \in A$ is nonzero and $U := g^{-1}(\mathbb{R}^\times)$, then $\frac{1}{g} \in C^\infty(U, \mathbb{R})$.

Exercise 1.8. Let $f_1 \colon M_1 \to N_1$ and $f_2 \colon M_2 \to N_2$ be smooth maps. Show that the map

$$f_1 \times f_2 \colon M_1 \times M_2 \to N_1 \times N_2, \quad (x, y) \mapsto (f_1(x), f_2(y))$$

is smooth.

Exercise 1.9. Let $f_1 \colon M \to N_1$ and $f_2 \colon M \to N_2$ be smooth maps. Show that the map

$$(f_1, f_2) \colon M \to N_1 \times N_2, \quad x \mapsto (f_1(x), f_2(x))$$

is smooth.

Exercise 1.10. Let N be an open subset of the smooth manifold M. Show that if $\mathcal{A} = (\varphi_i, U_i)_{i\in I}$ is a smooth atlas of M, $V_i := U_i \cap N$ and $\psi_i := \varphi_i|_{V_i}$, then $\mathcal{B} := (\psi_i, V_i)_{i\in I}$ is a smooth atlas of N.

Exercise 1.11. Let $V_1, \ldots, V_k$ and V be finite-dimensional real vector space and

$$\beta \colon V_1 \times \ldots \times V_k \to V$$

be a k-linear map. Show that β is smooth with

$$\mathtt{d}\beta(x_1, \ldots, x_k)(h_1, \ldots, h_k) = \sum_{j=1}^{k} \beta(x_1, \ldots, x_{j-1}, h_j, x_{j+1}, \ldots, x_k).$$

Exercise 1.12. Let M be a compact smooth manifold containing at least two points. Then each atlas of M contains at least two charts. In particular the atlas of $\mathbb{S}^n$ obtained from stereographic projections is minimal.

Exercise 1.13. Let X and Y be topological spaces and $q \colon X \to Y$ a quotient map, i.e. q is surjective and $O \subseteq Y$ is open if and only if $q^{-1}(O)$ is open in X. Show that a map $f \colon Y \to Z$ (Z a topological space) is continuous if and only if the map $f \circ q \colon X \to Z$ is continuous.

Exercise 1.14. Show that a smooth function $f \colon \mathbb{R} \to \mathbb{R}$ is a diffeomorphism if and only if either

(1) $f' > 0$ and $\lim_{x\to\pm\infty} f(x) = \pm\infty$.

(2) $f' < 0$ and $\lim_{x\to\pm\infty} f(x) = \mp\infty$.

1.2 Tangent Vectors and Tangent Maps

Definition 1.20. [tangent vector, tangent bundle] Let M be a smooth manifold and $p \in M$.

(a) A *tangent vector* to M in p is an equivalence class of smooth curves $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ under the equivalence relation

$$\gamma_1 \sim \gamma_2 \quad \Leftrightarrow \quad \text{there exists a chart } (U, \varphi) \text{ with } p \in U \text{ and } (\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0).$$

(b) The *tangent space* $T_p(M)$ is the set of all tangent vectors in p. The disjoint union of all tangent spaces on M

$$T(M) := \coprod_{p \in M} T_p(M)$$

is called the *tangent bundle of* M. We write $\pi_{TM} \colon TM \to M$ for the projection, mapping $T_p(M)$ to $\{p\}$.

Remark 1.21. (a) If $\gamma_1, \gamma_2 : (-\epsilon, \epsilon) \to M$ with $\gamma_1(0) = \gamma_2(0) = p$ are equivalent curves, then we have $(\psi \circ \gamma_1)'(0) = (\psi \circ \gamma_2)'(0)$ for all charts (V, ψ) with $p \in V$, since

$$(\psi \circ \gamma_i)'(0) = \mathtt{d}_{(\varphi(p))}(\psi \circ \varphi^{-1})(\varphi \circ \gamma_i)'(0).$$

(b) If $U \subseteq \mathbb{R}^n$ is an open subset and $p \in U$, then each smooth curve $\gamma \colon I \to U$ with $\gamma(0) = p$ is equivalent to the curve $\eta_v(t) := p + tv$ for $v = \gamma'(0)$. Hence each equivalence class contains exactly one curve η_v. We may therefore think of a tangent vector in $p \in U$ as a vector $v \in \mathbb{R}^n$ attached to the point p, and the map

$$\mathbb{R}^n \to T_p(U), \quad v \mapsto [\eta_v]$$

is a bijection. In this sense, we identify all tangent spaces $T_p(U)$ with $\mathbb{R}^n$, so that we obtain a bijection

$$T(U) \cong U \times \mathbb{R}^n.$$

As an open subset of the product space $T(\mathbb{R}^n) \cong \mathbb{R}^{2n}$, the tangent bundle $T(U)$ inherits a natural manifold structure.

(c) If V is a vector space, then we identify $T(V)$, as in (b), in a natural way with $V \times V$. Accordingly we have

$$T_p(f)(v) = (f(p), \mathtt{d}f(p)v),$$

for a map $\mathtt{d}f \colon T(M) \to V$ with $\mathtt{d}f(p) := \mathtt{d}f|_{T_p(M)}$.

(d) For each $p \in M$ and any chart (φ, U) with $p \in U$, the map

$$T_p(\varphi) \colon T_p(M) \to \mathbb{R}^n, \quad [\gamma] \mapsto (\varphi \circ \gamma)'(0)$$

is well-defined and injective by the definition of the equivalence relation. Moreover, the curve

$$\gamma(t) := \varphi^{-1}(\varphi(p) + tv),$$

which is smooth and defined on some neighborhood of 0, satisfies $(\varphi \circ \gamma)'(0) = v$. Hence $T_p(\varphi)$ is a bijection.

Lemma 1.22. *Let M be an m-dimensional manifold. Then there is a unique vector space structure on $T_p(M)$ such that for each chart (φ, U) of M with $p \in U$ the map*

$$T_p(\varphi)\colon T_p(M) \to \mathbb{R}^m, \quad [\gamma] \mapsto (\varphi \circ \gamma)'(0)$$

is a linear isomorphism. If N is an n-dimensional manifold and $f : M \to N$ a smooth map, then for all $p \in M$

$$T_p(f) : T_pM \to T_{f(p)}N, \quad [\gamma] \mapsto [f \circ \gamma]$$

defines a linear map between T_pM and $T_{f(p)}N$. The collection of all these maps defines a map

$$T(f)\colon T(M) \to T(N) \quad \textit{with} \quad T_p(f) = T(f)|_{T_p(M)}, p \in M.$$

It is called the tangent map of f

Proof. The bijectiion $T_p(\varphi)$ from Remark 1.21(d) defines a vector space structure on $T_p(M)$ by

$$v + w := T_p(\varphi)^{-1}(T_p(\varphi)v + T_p(\varphi)w) \quad \text{and} \quad \lambda v := T_p(\varphi)^{-1}(\lambda T_p(\varphi)v)$$

for $\lambda \in \mathbb{R}$, $v, w \in T_p(M)$. It remains to show that this vector space structure does not depend on the choice of the chart. For any other chart (ψ, V) with $p \in V$ we have

$$T_p(\psi) = \mathtt{d}_{\varphi(p)}(\psi \circ \varphi^{-1}) \circ T_p(\varphi).$$

As $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ is a diffeomorphism, $A_p := \mathtt{d}_{\varphi(p)}(\psi \circ \varphi^{-1})$ is a linear automorphism of $\mathbb{R}^n$, so that

$$\begin{aligned}
T_p(\psi)^{-1}(T_p(\psi)v + T_p\psi w) &= T_p\varphi^{-1} \circ A_p^{-1}(A_p \circ T_p(\varphi)v + A_p \circ T_p(\varphi)v) \\
&= T_p(\varphi)^{-1}(T_p(\varphi)v + T_p(\varphi)w) \\
T_p(\psi)^{-1}(\lambda T_p(\psi)v) &= T_p\varphi^{-1} \circ A_p^{-1}(\lambda A_p \circ T_p(\varphi)v) = T_p(\varphi)^{-1}(\lambda T_p(\varphi)v).
\end{aligned}$$

This shows that the vector space structure does not depend on the choice of the map and is well-defined.

Consider now a smooth map $f : M \to N$ between smooth manifolds M, N. We need to show that $T_p(f)$ is well defined and linear. For any chart (φ, U) of N with $f(p) \in U$ and any chart (ψ, V) of M with $p \in V$, we have

$$\begin{aligned}
T_{f(p)}(\varphi)[f \circ \gamma] &= (\varphi \circ f \circ \gamma)'(0) = \mathtt{d}_{\psi(p)}(\varphi \circ f \circ \psi^{-1})(\psi \circ \gamma)'(0) \\
&= \mathtt{d}_{\psi(p)}(\varphi \circ f \circ \psi^{-1})T_p(\psi)[\gamma].
\end{aligned}$$

This relation shows that $T_p(f)$ does not depend on the choice of the representative γ and it is linear, since the maps $T_{f(p)}(\varphi)$, $T_p(\psi)$ and $\mathtt{d}_{\psi(p)}(\varphi \circ f \circ \psi^{-1})$ are linear. $\square$

Example 1.23. [Open subsets] (a) For an open subset $U \subseteq \mathbb{R}^n$ and $p \in U$, the vector space structure on $T_p(U) = \{p\} \times \mathbb{R}^n$ is simply given by

$$(p, v) + (p, w) := (p, v + w) \quad \text{and} \quad \lambda(p, v) := (p, \lambda v)$$

for $v, w \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.

(b) If $f\colon U \to V$ is a smooth map between open subsets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$, $p \in U$, and $\eta_v(t) = p + tv$, then the tangent map satisfies

$$T(f)(p,v) = [f \circ \eta_v] = (f \circ \eta_v)'(0) = (f(p), \mathtt{d}f(p)\eta_v'(0)) = (f(p), \mathtt{d}f(p)v).$$

The main difference to the map $\mathtt{d}f$ is the book keeping; here we keep track of what happens to the point p and the tangent vector v. We may also write

$$T(f) = (f \circ \pi_{TU}, \mathtt{d}f)\colon TU \cong U \times \mathbb{R}^n \to TV \cong V \times \mathbb{R}^n,$$

where $\pi_{TU}\colon TU \to U, (p,v) \mapsto p$, is the projection map.

(c) If (φ, U) is a chart of M and $p \in U$, then we identify $T(\varphi(U))$ with $\varphi(U) \times \mathbb{R}^n$ and obtain for $[\gamma] \in T_p(M)$:

$$T(\varphi)([\gamma]) = (\varphi(p), [\varphi \circ \gamma]) = (\varphi(p), (\varphi \circ \gamma)'(0)),$$

which is consistent with our previously introduced notation $T_p(\varphi)$.

Example 1.24. [Submanifolds of $\mathbb{R}^n$] Let $M \subset \mathbb{R}^n$ be a smooth k-dimensional submanifold and $p \in M$. Then there exists an open neighbourhood $U_p \subset \mathbb{R}^n$ of p and a smooth function $f_p : U_p \to \mathbb{R}^{n-k}$ such that $M \cap U_p = f_p^{-1}(\{0\})$ and $\operatorname{rank}(\mathtt{d}_p f_p) = n - k$. For every smooth curve $\gamma : (-\epsilon, \epsilon) \to M \cap U_p$ with $\gamma(0) = p$ we have $f_p \circ \gamma(t) = 0$ for all t and hence

$$0 = (f_p \circ \gamma)'(0) = \mathtt{d}_p f_p(\gamma'(0)) \quad \Rightarrow \quad \gamma'(0) \in \ker(\mathtt{d}_p f_p).$$

We can therefore identify the tangent space $T_p(M)$ with $\ker(d_p f_p)$. If (φ_p, V_p) is a chart as in Example 1.6 with $\varphi_p(V_p \cap M) = (\mathbb{R}^k \times \{0\}) \cap W_p$ then $\mathtt{d}_p\varphi_p : \mathbb{R}^k \to \ker(\mathtt{d}_p f_p)$ is a vector space isomorphism and

$$T_p(\varphi_p)[\gamma] = (\varphi_p \circ \gamma)'(0) = \mathtt{d}_p\varphi_p(\gamma'(0)).$$

Lemma 1.25. (Chain rule for tangent maps) *For smooth maps $f\colon M \to N$ and $g\colon N \to L$, the tangent maps satisfy*

$$T(g \circ f) = T(g) \circ T(f).$$

Proof. We recall from Remark 1.17 that $g \circ f\colon M \to L$ is a smooth map, so that $T(g \circ f)$ is defined. For $p \in M$ and $[\gamma] \in T_p(M)$, we further have

$$T_p(g \circ f)[\gamma] = [g \circ f \circ \gamma] = T_{f(p)}(g)[f \circ \gamma] = T_{f(p)}(g)T_p(f)[\gamma].$$

Since p was arbitrary, this implies the lemma. □

So far we only considered the tangent bundle $T(M)$ of a smooth manifold M as a set, but this set also carries a natural topology and a smooth manifold structure.

Definition 1.26. [Manifold structure on $T(M)$] Let M be a smooth manifold. First we introduce a topology on $T(M)$. For each chart (φ, U) of M, we have a tangent map

$$T(\varphi)\colon T(U) \to T(\varphi(U)) \cong \varphi(U) \times \mathbb{R}^n,$$

where we consider $T(U) = \bigcup_{p\in U} T_p(M)$ as a subset of $T(M)$. We define a topology on $T(M)$ by declaring a subset $O \subseteq T(M)$ to be open if for each chart (φ, U) of M, the set $T(\varphi)(O \cap T(U))$ is an open subset of $T(\varphi(U))$. It is easy to see that this defines indeed a Hausdorff topology on $T(M)$ for which all the subsets $T(U)$ are open and the maps $T(\varphi)$ are homeomorphisms onto open subsets of $\mathbb{R}^{2n}$ (Exercise 1.16).

Since for two charts (φ, U), (ψ, V) of M, the map

$$T(\varphi \circ \psi^{-1}) = T(\varphi) \circ T(\psi)^{-1} \colon T(\psi(V)) \to T(\varphi(U))$$

is smooth, for each atlas $\mathcal{A}$ of M, the collection $(T(\varphi), T(U))_{(\varphi,U)\in\mathcal{A}}$ is a smooth atlas of $T(M)$. We thus obtain on $T(M)$ the structure of a smooth manifold.

Lemma 1.27. *If $f\colon M \to N$ is a smooth map, then its tangent map $T(f)$ is smooth.*

Proof. Let $p \in M$ and choose charts (φ, U) and (ψ, V) of M, resp., N with $p \in U$ and $f(p) \in V$. Then the map

$$T(\psi) \circ T(f) \circ T(\varphi)^{-1} = T(\psi \circ f \circ \varphi^{-1}) \colon T(\varphi(f^{-1}(V) \cap U)) \to T(V)$$

is smooth, and this implies that $T(f)$ is a smooth map. $\square$

Remark 1.28. For smooth manifolds $M_1, \ldots, M_n$, the projection maps

$$\pi_i \colon M_1 \times \cdots \times M_n \to M_i, \quad (p_1, \ldots, p_n) \mapsto p_i$$

induce a diffeomorphism

$$(T(\pi_1), \ldots, T(\pi_n)) \colon T(M_1 \times \cdots \times M_n) \to TM_1 \times \cdots \times TM_n$$

(Exercise 1.17).

Exercises for Section 1.2

Exercise 1.15. Show that for a submanifold $M \subset \mathbb{R}^n$ which is a quadric $M = f^{-1}(0)$ with $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle x, Ax + b\rangle + c$, the tangent space $T_p(M)$ is isomorphic to the orthogonal complement

$$(2Ap + b)^\perp = \{y \in \mathbb{R}^n : \langle y, 2Ap + b\rangle = 0\}.$$

Determine the tangent spaces $T_p\mathbb{S}^n$ and $T_p\mathbb{H}^n$ of the n-sphere and the n-dimensional hyperbolic space.

Exercise 1.16. Let M be a smooth manifold. We call a subset $O \subseteq T(M)$ open if for each chart (φ, U) of M, the set $T(\varphi)(O \cap T(U))$ is an open subset of $T(\varphi(U))$. Show that:

(1) This defines a topology on $T(M)$.

(2) All subsets $T(U)$ are open.

(3) The maps $T(\varphi)\colon TU \to T(\varphi(U)) \cong \varphi(U) \times \mathbb{R}^n$ are homeomorphisms onto open subsets of $\mathbb{R}^{2n} \cong T(\mathbb{R}^n)$.

(4) The projection $\pi_{TM}\colon T(M) \to M$ is continuous.

(5) $T(M)$ is Hausdorff.

Exercise 1.17. For smooth manifolds $M_1, \ldots, M_n$, the projection maps

$$\pi_i\colon M_1 \times \cdots \times M_n \to M_i, \quad (p_1, \ldots, p_n) \mapsto p_i$$

induce a diffeomorphism

$$(T(\pi_1), \ldots, T(\pi_n))\colon T(M_1 \times \cdots \times M_n) \to TM_1 \times \cdots \times TM_n.$$

Exercise 1.18. Let N and $M_1, \ldots, M_n$ be a smooth manifolds. Show that a map

$$f\colon N \to M_1 \times \cdots \times M_n$$

is smooth if and only if all its component functions $f_i\colon N \to M_i$ are smooth.

Exercise 1.19. Let $f\colon M \to N$ be a smooth map between manifolds, $\pi_{TM}\colon TM \to M$ the tangent bundle projection and $\sigma_M\colon M \to TM$ the zero section. Show that for each smooth map $f\colon M \to N$ we have

$$\pi_{TN} \circ Tf = f \circ \pi_{TM} \quad \text{ and } \quad \sigma_N \circ f = Tf \circ \sigma_M.$$

Exercise 1.20. [Inverse Function Theorem for manifolds] Let $f\colon M \to N$ be a smooth map and $p \in M$ such that $T_p(f)\colon T_p(M) \to T_{f(p)}(N)$ is a linear isomorphism. Show that there exists an open neighborhood U of p in M such that the restriction $f|_U\colon U \to f(U)$ is a diffeomorphism onto an open subset of N.

Exercise 1.21. Let $\mu\colon E \times F \to W$ be a bilinear map and M a smooth manifold. For $f \in C^\infty(M, E)$, $g \in C^\infty(M, F)$ and $p \in M$ set $h(p) := \mu\big(f(p), g(p)\big)$. Show that h is smooth with

$$T(h)v = \mu\big(T(f)v, g(p)\big) + \mu\big(f(p), T(g)v\big) \quad \text{for } v \in T_p(M).$$

1.3 Vector fields

Throughout this subsection M denotes an n-dimensional smooth manifold.

Definition 1.29. [Vector Field, Lie Derivative] Let M be a n-dimensional manifold and denote by $\pi_{TM}\colon TM \to M$ the canonical projection mapping $T_p(M)$ to p. A *(smooth) vector field* X on M is a smooth section of the tangent bundle TM, i.e. a smooth map $X\colon M \to TM$ with $\pi_{TM} \circ X = \mathrm{id}_M$. We denote by $\mathcal{V}(M)$ for the space of all vector fields on M.

If $f \in C^\infty(M, V)$ is a smooth function on M with values in some finite-dimensional vector space V and $X \in \mathcal{V}(M)$, then we obtain a smooth function on M via

$$\mathcal{L}_X f := \mathtt{d}f \circ X\colon M \to TM \to V.$$

We thus obtain for each $X \in \mathcal{V}(M)$ a linear operator $\mathcal{L}_X$ on $C^\infty(M, V)$. The function $\mathcal{L}_X f$ is also called the *Lie derivative* of f with respect to X.

Remark 1.30. (a) If U is an open subset of $\mathbb{R}^n$, then $TU = U \times \mathbb{R}^n$ with the bundle projection

$$\pi_{TU} \colon U \times \mathbb{R}^n \to U, \quad (x, v) \mapsto x.$$

Therefore each smooth vector field is of the form $X(x) = (x, \widetilde{X}(x))$ for some smooth function $\widetilde{X} \colon U \to \mathbb{R}^n$, and we may thus identify $\mathcal{V}(U)$ with the space $C^\infty(U, \mathbb{R}^n)$ of smooth $\mathbb{R}^n$-valued functions on U.

(b) The space $\mathcal{V}(M)$ carries a natural vector space structure given by

$$(X + Y)(p) := X(p) + Y(p), \quad (\lambda X)(p) := \lambda X(p).$$

More generally, we can multiply vector fields with smooth functions

$$(fX)(p) := f(p)X(p), \quad f \in C^\infty(M, \mathbb{R}), X \in \mathcal{V}(M).$$

Remark 1.31. [Time-dependent vector fields] In many physics applications such as classical mechanics or electrodynamics, one considers so-called time-dependent vector fields. From a mathematical viewpoint, these are simply smooth functions $X : I \times M \to TM$ with $X(t, p) \in T_p(M)$ for all $t \in I$, $p \in M$. Equivalently, one can consider time-dependent vector fields as smooth vector fields on the product manifold $I \times M$ with

$$X(t, p) \in \operatorname{Im}(T_p(i_t)) \subset T_{(t,p)}(M) \qquad \forall (t, p) \in I \times M,$$

where $i_t : M \to \{t\} \times M$, $p \mapsto (t, p)$ denotes the embedding of M into the product manifold $I \times M$.

Remark 1.32. [Basic Vector Fields] (a) Let (φ, U) be a chart of M and $\varphi_1, \ldots, \varphi_n \colon U \to \mathbb{R}$ the corresponding coordinate functions. Then we obtain on U vector fields b_j^φ, $j = 1, \ldots, n$, defined by

$$b_j^\varphi(p) := T_p(\varphi)^{-1} e_j$$

where $e_1, \ldots, e_n$ is the standard basis for $\mathbb{R}^n$. We call these vector fields the *φ-basic vector fields* on U. The expression basic vector field is doubly justified. On the one hand, $\big(b_1^\varphi(p), \ldots, b_n^\varphi(p)\big)$ is a basis for $T_p(M)$ for every $p \in U$. On the other hand, the definition shows that every $X \in \mathcal{V}(U)$ can be written uniquely as

$$X = \sum_{j=1}^{n} x_j \cdot b_j^\varphi \quad \text{with} \quad x_j \in C^\infty(U).$$

(b) For functions $f \in C^\infty(U)$, we denote by

$$\frac{\partial f}{\partial \varphi_j} := \mathcal{L}_{b_j^\varphi} f$$

its Lie derivatives with respect to the basic vector fields. This notation is justified by the following observation: The smooth curves

$$\gamma_i : (-\epsilon, \epsilon) \to M, \quad \gamma_i(t) = \varphi^{-1}(\varphi(p) + te_j)$$

satisfy $T_p(\varphi)(\gamma_j) = e_j$, which implies

$$\frac{\partial f}{\partial \varphi_j}(p) := \mathcal{L}_{b_j^\varphi} f(p) = \mathtt{d}f \circ T(\varphi)^{-1} e_j = \frac{d}{dt}\Big|_{t=0} (f \circ \varphi^{-1})(\varphi(p) + te_j) = \partial_j (f \circ \varphi^{-1})(\varphi(p)).$$

(c) The Lie derivatives of f with respect to the basic vector fields coincide with partial derivatives of the function $f \circ \varphi^{-1}|_U : \varphi(U) \to \mathbb{R}$. As $\mathtt{d}f : TM \to T\mathbb{R} \cong \mathbb{R} \times \mathbb{R}$ is linear on each tangent space, the Lie derivative of $f \in C^\infty(U)$ with respect to $X = \sum_{i=1}^n x_i b_i^\varphi \in \mathcal{V}(U)$ then takes the form

$$\mathcal{L}_X f = \sum_{i=1}^n x_i \frac{\partial f}{\partial \varphi_i}.$$

(d) If (φ, U), (ψ, V) are coodinate charts on M with $U \cap V \neq \emptyset$, then it is easy to show that the basic vector fields for φ and ψ are related on $U \cap V$ by

$$b_j^\psi(p) = \sum_{k=1}^n \frac{\partial \varphi_k}{\partial \psi_j}(p) b_k^\varphi(p), \qquad \frac{\partial f}{\partial \psi_j}(p) = \sum_{k=1}^n \frac{\partial \varphi_k}{\partial \psi_j}(p) \frac{\partial f}{\partial \varphi_k} \qquad \forall p \in U \cap V,$$

and we have the identity

$$\sum_{k=1}^n \frac{\partial \psi_j}{\partial \varphi_k}(p) \frac{\partial \varphi_k}{\partial \psi_i}(p) = \frac{\partial \psi_j}{\partial \psi_i}(p) = \delta_{ij}$$

(see Exercise 1.23).

Lemma 1.33. (Properties of the Lie derivative) *The Lie derivative is a derivation:*

(a) *It is* local: $\mathcal{L}_X f(p) = \mathcal{L}_X(f|_V)(p)$ *for any open subset* $V \subset U$ *with* $p \in U$.

(b) *It is linear in* f*:*

$$\mathcal{L}_X(f+g) = \mathcal{L}_X f + \mathcal{L}_X g \qquad \mathcal{L}(\lambda f) = \lambda \mathcal{L}_X f \qquad \forall f, g \in C^\infty(U), \lambda \in \mathbb{R}$$

(c) *It satisfies the Leibnitz identity:*

$$\mathcal{L}_X(f \cdot g) = g \cdot \mathcal{L}_X f + f \cdot \mathcal{L}_X g \qquad \forall f, g \in C^\infty(U), X \in \mathcal{V}(U).$$

Proof. For any point $p \in M$ we can choose a chart (φ, U) with $p \in U$. The assertions then follow directly from the formulas in Remark 1.32 and the corresponding properties of the partial derivatives of functions $g : \mathbb{R}^n \to \mathbb{R}$. $\square$

Definition 1.34. (Lie bracket) Let (φ, U) be a chart of M and $X = \sum_{i=1}^n x_i b_i^\varphi$, $Y = \sum_{i=1}^n y_i b_i^\varphi$ smooth vector fields on U. The *Lie bracket* of X and Y is the smooth vector field $[X, Y] \in \mathcal{V}(U)$ defined by

$$[X, Y](p) = \sum_{j=1}^n \left(\sum_{i=1}^n x_i(p) \frac{\partial y_j}{\partial \varphi_i}(p) - y_i(p) \frac{\partial x_j}{\partial \varphi_i}(p) \right) b_j^\varphi(p) \qquad \forall p \in U.$$

Remark 1.35. In local coordinates, we find for the composition of two Lie derivative operators $\mathcal{L}_X$ and $\mathcal{L}_Y$ corresponding to the vector fields $X = \sum_i x_i b_i^\varphi$ and $Y = \sum_i y_i b_i^\varphi$ (cf. Remark 1.32):

$$\mathcal{L}_X \mathcal{L}_Y f = \sum_{i=1}^n x_i \frac{\partial}{\partial \varphi_i} \left(\sum_{j=1}^n y_i \frac{\partial f}{\partial \varphi_i} \right) = \sum_{i,j=1}^n x_i \frac{\partial y_j}{\partial \varphi_i} \frac{\partial f}{\partial \varphi_j} + x_i y_j \frac{\partial^2 f}{\partial \varphi_i \partial \varphi_j}.$$

This is not of the form $\mathcal{L}_Z f$ for any vector field Z because it contains second derivatives of f. However, the Schwarz Lemma implies that the term containing the second derivatives does not change if we exchange X and Y. This leads to the relation

$$\mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f = \sum_{i,j=1}^n \left(x_i \frac{\partial y_j}{\partial \varphi_i} - y_i \frac{\partial x_j}{\partial \varphi_i} \right) \frac{\partial f}{\partial \varphi_j} = \mathcal{L}_{[X,Y]} f. \tag{2}$$

Clearly, this relation determines the Lie bracket $[X, Y]$ uniquely because any vector field $Z \in \mathcal{V}(U)$ is determined by its Lie derivative on $C^\infty(U)$.

Lemma 1.36. (The Lie algebra structure on $\mathcal{V}(M)$)

(a) *The Lie bracket does not depend on the choice of the chart and defines a bilinear, antisymmetric map* $[\ ,\ ] : \mathcal{V}(M) \times \mathcal{V}(M) \to \mathcal{V}(M)$ *that satisfies the Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \qquad \forall X, Y, Z \in \mathcal{V}(M).$$

This equips $\mathcal{V}(M)$ *with the structure of a Lie algebra (cf. Definition 2.7).*

(b) *For all* $f \in C^\infty(M)$, $X, Y \in \mathcal{V}(M)$, *the Lie bracket satifies:*

$$[X, f \cdot Y] = f \cdot [X, Y] + \mathcal{L}_X f \cdot Y.$$

(c) *For all* $f \in C^\infty(M)$, $X, Y \in \mathcal{V}(M)$:

$$\mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f = \mathcal{L}_{[X,Y]} f$$

Proof. We first show that the Lie bracket does not depend on the choice of the chart. Once this is established, all other identities can then be derived by direct calculations using the formulas from Remark 1.32 and Definition 1.34 for the φ-basic vector fields for a chart (φ, U).

(a) Let (φ, U) and (ψ, V) be two charts on M with $U \cap V \neq \emptyset$ and $X, Y \in \mathcal{V}(U \cap V)$. Then we can uniquely express the vector fields X, Y in terms of the φ- and ψ-basic vector fields as

$$X = \sum_{i=1}^n x_i b_i^\varphi = \sum_{j=1}^n \widetilde{x}_j b_j^\psi, \qquad Y = \sum_{i=1}^n y_i b_i^\varphi = \sum_{j=1}^n \widetilde{y}_j b_j^\psi.$$

For a smooth function $f \in C^\infty(U \cap V)$ the function

$$\mathcal{L}_{[X,Y]} f = \mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f$$

can be calculated in both coordinate systems, which leads to the corresponding formulas for the Lie bracket (Remark 1.35):

$$[X,Y] = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_i \frac{\partial y_j}{\partial \varphi_i} - y_i \frac{\partial x_j}{\partial \varphi_i} \right) b_j^{\varphi} = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \widetilde{x}_i \frac{\partial \widetilde{y}_j}{\partial \psi_i} - \widetilde{y}_i \frac{\partial \widetilde{x}_j}{\partial \psi_i} \right) b_j^{\psi}.$$

This shows that the Lie bracket is independent of the choice of the chart. We can thus cover M by charts and define $[X,Y]$ by the formula in Definition 1.34 on the domain of each chart. As the resulting brackets agree on the overlap $U \cap V$ of any two charts (φ, U) and (ψ, V), this yields a map

$$[\cdot,\cdot] : \mathcal{V}(M) \times \mathcal{V}(M) \to \mathcal{V}(M).$$

The bilinearity and antisymmetry follow directly from the formula in Definition 1.34. For the Jacobi idenity, it is sufficient to show that it holds on the domain of each chart (φ, U). This can be verified by the following calculation:

$$\begin{aligned} \mathcal{L}_{[X,[Y,Z]]} f &= \mathcal{L}_X(\mathcal{L}_Y \mathcal{L}_Z - \mathcal{L}_Z \mathcal{L}_Y) f - (\mathcal{L}_Y \mathcal{L}_Z - \mathcal{L}_Z \mathcal{L}_Y)\mathcal{L}_X f \\ &= (\mathcal{L}_X \mathcal{L}_Y \mathcal{L}_Z - \mathcal{L}_X \mathcal{L}_Z \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_Z \mathcal{L}_X + \mathcal{L}_Z \mathcal{L}_Y \mathcal{L}_X) f. \end{aligned}$$

A short calculation shows that the cyclic sum over X, Y, Z of this expression vanishes, which shows that, for $J := [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]]$, we have $\mathcal{L}_J f = 0$ for every f, and hence that $J = 0$.

(b) It is again sufficient to show that this holds in the domain of each chart (φ, U). We calculate:

$$\begin{aligned} [X, f \cdot Y] &= \sum_{i,j=1}^{n} \left(x_i \frac{\partial (f y_j)}{\partial \varphi_i} - f y_i \frac{\partial x_j}{\partial \varphi_i} \right) b_j^{\varphi} \\ &= f \cdot \left(\sum_{i,j=1}^{n} \left(x_i \frac{\partial y_j}{\partial \varphi_i} - y_i \frac{\partial x_j}{\partial \varphi_i} \right) b_j^{\varphi} \right) + \left(\sum_{i=1}^{n} x_i \frac{\partial f}{\partial \varphi_i} \right) \left(\sum_{j=1}^{n} y_j b_j^{\varphi} \right) \\ &= f \cdot [X,Y] + \mathcal{L}_X f \cdot Y. \end{aligned}$$

(c) We know from (a) that the Lie bracket $[X,Y]$ for $X, Y \in \mathcal{V}(M)$ is a well-defined global vector field. That this vector field satisfies (c) in any local chart follows from Remark 1.35. □

Example 1.37. For open subsets $U \subset \mathbb{R}^n$, the space $\mathcal{V}(U)$ can be identified with $C^\infty(U, \mathbb{R}^n)$, and the Lie bracket of two vector fields $X = \sum_{i=1}^n x_i e_i$, $Y = \sum_{j=1}^n y_j e_j$ is given by

$$[X,Y](p) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_i(p) \partial_i y_j(p) - y_i(p) \partial_i x_j(p) \right) e_j = \mathtt{d}Y(p)X(p) - \mathtt{d}X(p)Y(p).$$

For the Lie derivative of a function $f \in C^\infty(U)$ with respect to X, we obtain

$$\mathcal{L}_X f(p) = \sum_{i=1}^{n} x_i(p) \partial_i f(p).$$

It remains to investigate how Lie derivatives of functions and the Lie brackets of vector fields behave under smooth maps $\varphi : M \to N$. For each vector field $X \in \mathcal{V}(M)$, we obtain a smooth map $T\varphi \circ X : M \to TN$. On the other hand, every vector field $Y \in \mathcal{V}(N)$ yields a smooth map $Y \circ \varphi : M \to TN$. This allows one to relate and compare vector fields on M and on N.

Definition 1.38. [φ-related vector fields] If $\varphi\colon M \to N$ is a smooth map, then we call two vector fields $X \in \mathcal{V}(M)$ and $Y \in \mathcal{V}(N)$ *φ-related* if

$$Y \circ \varphi = T\varphi \circ X\colon M \to TN. \tag{3}$$

Example 1.39. For every chart (φ, U) on M, the φ-basic vector fields on U are φ-related to the constant vector fields $e_j \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$, $e_j(p) = e_j$ since we have:

$$T_p(\varphi)b_j^\varphi(p) = T_p(\varphi)T_p(\varphi)^{-1}e_j(\varphi(p)) = e_j(\varphi(p)) = e_j \qquad \forall p \in U.$$

Example 1.40. We consider $M = \mathbb{S}^2$ and the vector field $X : \mathbb{S}^2 \to T\mathbb{S}^2$ defined by

$$X(p) = e_3 \times p \quad \forall p \in \mathbb{S}^2,$$

where $\times$ denotes the vector product in $\mathbb{R}^3$. The rotations around the x_3-axis

$$R : \mathbb{R}^3 \to \mathbb{R}^3, \quad R(x_1, x_2, x_3) = (\cos\alpha x_1 + \sin\alpha x_2, \cos\alpha x_2 - \sin\alpha x_1, x_3)$$

map $\mathbb{S}^2$ to itself and hence induce smooth maps $\varphi = R|_{\mathbb{S}^2} : \mathbb{S}^2 \to \mathbb{S}^2$. By Example 1.24, we have

$$T_p(\varphi) = R|_{p^\perp} : p^\perp \to \varphi(p)^\perp \qquad \forall p \in \mathbb{S}^2,$$

which implies

$$X(\varphi(p)) = e_3 \times \varphi(p) = Re_3 \times Rp = R(e_3 \times p) = T_p(\varphi)(X(p)).$$

The vector field X is φ-related to itself.

Lemma 1.41. (Related Vector Field Lemma) *Let M and N be smooth manifolds, $\varphi\colon M \to N$ a smooth map, $Y, Y' \in \mathcal{V}(N)$ and $X, X' \in \mathcal{V}(M)$. If X is φ-related to Y and X' is φ-related to Y', then the Lie bracket $[X, X']$ is φ-related to $[Y, Y']$, and*

$$\mathcal{L}_X \circ \varphi^* = \varphi^* \circ \mathcal{L}_Y, \tag{4}$$

where $\varphi^*\colon C^\infty(N) \to C^\infty(M)$, $f \mapsto f \circ \varphi$ is the pullback map.

Proof. The relation (4) follows for $f \in C^\infty(N)$ from the Chain Rule:

$$\mathcal{L}_X\varphi^* f = \mathtt{d}(f \circ \varphi)X = \mathtt{d}f \circ T(\varphi) \circ X = \mathtt{d}f \circ Y \circ \varphi = \varphi^*\mathcal{L}_Y f.$$

If, conversely,

$$\mathcal{L}_X\varphi^* f = \varphi^*\mathcal{L}_Y f$$

holds for all smooth functions defined on open subsets $U \subseteq N$, we can apply this relation to coordinate functions to obtain

$$T(\varphi) \circ X = Y \circ \varphi,$$

i.e., that X is φ-related to Y.

We further obtain

$$\begin{aligned}\mathcal{L}_{[X,X']}\varphi^* f &= (\mathcal{L}_X\mathcal{L}_{X'} - \mathcal{L}_{X'}\mathcal{L}_X)\varphi^* f = \mathcal{L}_X\varphi^*\mathcal{L}_{Y'} f - \mathcal{L}_{X'}\varphi^*\mathcal{L}_Y f \\ &= \varphi^*(\mathcal{L}_Y\mathcal{L}_{Y'} - \mathcal{L}_{Y'}\mathcal{L}_Y) f = \varphi^*\mathcal{L}_{[Y,Y']} f\end{aligned}$$

for smooth functions defined on subsets $U \subseteq N$, and by the preceding argument, this implies that $[X, X']$ is φ-related to $[Y, Y']$. $\square$

Exercises for Section 1.3

Exercise 1.22. Consider the n-sphere with the charts from Example 1.8 and determine the associated φ-basic vector fields.

Exercise 1.23. Let M be an n-dimensional manifold and (φ, U), (ψ, V) coodinate charts on M with $U \cap V \neq \emptyset$. Prove that the φ- and ψ-basic vector fields on $U \cap V$ are related by

$$b_j^\psi(p) = \sum_{k=1}^n \frac{\partial\varphi_k}{\partial\psi_j}(p) b_k^\varphi(p) \qquad \forall p \in U \cap V$$

and that the Lie derivatives with respect to the basic vector fields satisfy the relations

$$\frac{\partial f}{\partial\psi_j}(p) = \sum_{k=1}^n \frac{\partial\varphi_k}{\partial\psi_j}(p)\frac{\partial f}{\partial\varphi_k} \qquad \sum_{k=1}^n \frac{\partial\psi_j}{\partial\varphi_k}(p)\frac{\partial\varphi_k}{\partial\psi_i}(p) = \frac{\partial\psi_j}{\partial\psi_i}(p) = \delta_{ij}$$

Exercise 1.24. Let M be a smooth manifold, $X, Y \in \mathcal{V}(M)$ and $f, g \in C^\infty(M, \mathbb{R})$. Show that

(1) $\mathcal{L}_X(f \cdot g) = \mathcal{L}_X(f) \cdot g + f \cdot \mathcal{L}_X(g)$, i.e. the map $f \mapsto \mathcal{L}_X(f)$ is a derivation.

(2) $\mathcal{L}_{fX}(g) = f \cdot \mathcal{L}_X(g)$.

Exercise 1.25. Let $\mathcal{A}$ be a $\mathbb{K}$-algebra (not necessarily associative). Show that

(i) $\operatorname{der}(\mathcal{A}) := \{D \in \operatorname{End}(\mathcal{A})\colon (\forall a, b \in \mathcal{A})\, D(ab) = Da \cdot b + a \cdot Db\}$ is a Lie subalgebra of $\mathfrak{gl}(\mathcal{A}) = \operatorname{End}(\mathcal{A})_L$, i.e., closed under the commutator bracket $[D_1, D_2] := D_1D_2 - D_2D_1$.

(ii) If, in addition, $\mathcal{A}$ is commutative, then for $D \in \operatorname{der}(\mathcal{A})$ and $a \in \mathcal{A}$, the map $aD\colon \mathcal{A} \to \mathcal{A}, x \mapsto aDx$ also is a derivation.

Exercise 1.26. Let U be an open subset of $\mathbb{R}^{2n}$ and $\mathcal{P} = C^\infty(U, \mathbb{R})$ be the space of smooth functions on U and write $q_1, \ldots, q_m, p_1, \ldots, p_m$ for the coordinates with respect to a basis. Then $\mathcal{P}$ is a Lie algebra with respect to the *Poisson bracket*

$$\{f, g\} := \sum_{i=1}^n \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}.$$

Exercise 1.27. To each $A \in \mathfrak{gl}_n(\mathbb{R})$, we associate the linear vector field $X_A(x) := Ax$ on $\mathbb{R}^n$ Show that, for $A, B \in M_n(\mathbb{R})$, we have $X_{[A,B]} = -[X_A, X_B]$.

1.4 Integral Curves and Local Flows

Throughout this subsection M denotes an n-dimensional manifold.

Definition 1.42. Let $X \in \mathcal{V}(M)$ and $I \subseteq \mathbb{R}$ an open interval containing 0. A differentiable map $\gamma\colon I \to M$ is called an *integral curve* of X if

$$\gamma'(t) = X(\gamma(t)) \qquad \text{for each} \quad t \in I.$$

Note that the preceding equation implies that γ' is continuous and further that if γ is C^k, then γ' is also C^k. Therefore integral curves of smooth vector fields are automatically smooth.

If $J \supseteq I$ is an interval containing I, then an integral curve $\eta\colon J \to M$ is called an *extension* of γ if $\eta|_I = \gamma$. An integral curve γ is said to be *maximal* if it has no proper extension.

Remark 1.43. Integral curves can be defined analogously for time-dependent vector fields. By Remark 1.31, a time dependent vector field on M is a smooth function $X : I \times M \to TM$. A curve $\gamma : I \to M$ is called an *integral curve* of X if it satisfies

$$\gamma'(t) = X(t, \gamma(t)) \qquad \forall t \in I.$$

Remark 1.44. (a) If $U \subseteq \mathbb{R}^n$ is an open subset of $\mathbb{R}^n$, then we write a vector field $X \in \mathcal{V}(U)$ as $X(x) = (x, F(x))$, where $F\colon U \to \mathbb{R}^n$ is a smooth function. A curve $\gamma\colon I \to U$ is an integral curve of X if and only if it satisfies the ordinary differential equation

$$\gamma'(t) = F(\gamma(t)) \quad \text{for all} \quad t \in I.$$

(b) If (φ, U) is a chart of the manifold M and $X \in \mathcal{V}(M)$, then a curve $\gamma\colon I \to M$ is an integral curve of X if and only if the curve $\eta := \varphi \circ \gamma$ is an integral curve of the vector field $X_\varphi := T(\varphi) \circ X \circ \varphi^{-1} \in \mathcal{V}(\varphi(U))$ because

$$X_\varphi(\eta(t)) = T_{\gamma(t)}(\varphi)X(\gamma(t)) \quad \text{and} \quad \eta'(t) = T_{\gamma(t)}(\varphi)\gamma'(t).$$

Example 1.45. We consider the vector field $X : \mathbb{S}^2 \to T\mathbb{S}^2$, $X(p) = e_3 \times p$ from Example 1.40. Then for all $p \in \mathbb{S}^2$, the curve

$$\gamma_p : I \to \mathbb{R}^3, \quad \gamma(t) = \cos(t)(p - \langle p, e_3\rangle e_3) + \sin(t) e_3 \times p + \langle p, e_3\rangle e_3$$

defines an integral curve of X since we have:

$$\langle \gamma_p(t), \gamma_p(t)\rangle = (\langle p, p\rangle - \langle p, e_3\rangle^2)\cos^2(t) + (\langle p, p\rangle - \langle p, e_3\rangle^2)\sin^2(t) + \langle p, e_3\rangle^2 = 1$$

and

$$\gamma_p'(t) = -\sin(t)(p - \langle p, e_3\rangle e_3) + \cos(t) e_3 \times p = e_3 \times \gamma_p(t) = X(\gamma_p(t)).$$

Definition 1.46. Let $a < b \in [-\infty, \infty]$. For a continuous curve $\gamma\colon]a, b[\to M$ we say that

$$\lim_{t \to b} \gamma(t) = \infty$$

if for each compact subset $K \subseteq M$ there exists a $c < b$ with $\gamma(t) \notin K$ for $t > c$. Similarly, we define

$$\lim_{t \to a} \gamma(t) = \infty.$$

Theorem 1.47. (Existence and Uniqueness of Integral Curves) *Let $X \in \mathcal{V}(M)$ and $p \in M$. Then there exists a unique maximal integral curve $\gamma_p \colon I_p \to M$ with $\gamma_p(0) = p$. If $a := \inf I_p > -\infty$, then $\lim_{t\to a} \gamma_p(t) = \infty$ and if $b := \sup I_p < \infty$, then $\lim_{t\to b} \gamma_p(t) = \infty$.*

Proof. We have seen in Remark 1.44 that in local charts, integral curves are solutions of an ordinary differential equation with a smooth right hand side. We now reduce the proof to the Local Existence- and Uniqueness Theorem for ODE's.

Uniqueness: Let $\gamma, \eta \colon I \to M$ be two integral curves of X with $\gamma(0) = \eta(0) = p$. The continuity of the curves implies that

$$0 \in J := \{t \in I \colon \gamma(t) = \eta(t)\}$$

is a closed subset of I. In view of the Local Uniqueness Theorem for ODE's, for each $t_0 \in J$ there exists an $\varepsilon > 0$ with $[t_0, t_0 + \varepsilon] \subseteq J$, and likewise $[t_0 - \varepsilon, t_0] \subseteq J$. Therefore J is also open. Now the connectedness of I implies $I = J$, so that $\gamma = \eta$.

Existence: The Local Existence Theorem implies the existence of some integral curve $\gamma \colon I \to M$ on some open interval containing 0. For any other integral curve $\eta \colon J \to M$, the intersection $I \cap J$ is an interval containing 0, so that the uniqueness assertion implies that $\eta = \gamma$ on $I \cap J$.

Let $I_p \subseteq \mathbb{R}$ be the union of all open intervals I_j containing 0 on which there exists an integral curve $\gamma_j \colon I_j \to M$ of X with $\gamma_j(0) = p$. Then the preceding argument shows that

$$\gamma(t) := \gamma_j(t) \quad \text{for} \quad t \in I_j$$

defines an integral curve of X on I_p, which is maximal by definition. The uniqueness of the maximal integral curve also follows from its definition.

Limit condition: Suppose that $b := \sup I_p < \infty$. If $\lim_{t\to b} \gamma(t) = \infty$ does not hold, then there exists a compact subset $K \subseteq M$ and a sequence $t_m \in I_p$ with $t_m \to b$ and $\gamma(t_m) \in K$. As K can be covered with finitely many closed subsets homeomorphic to a closed subset of a ball in $\mathbb{R}^n$, after passing to a suitable subsequence, we may w.l.o.g. assume that K itself is homeomorphic to a compact subset of $\mathbb{R}^n$. Then a subsequence of $(\gamma(t_m))_{m\in\mathbb{N}}$ converges, and we may replace the original sequence by this subsequence, hence assume that $q := \lim_{m\to\infty} \gamma(t_m)$ exists.

The Local Existence Theorem for ODE's implies the existence of a compact neighborhood $V \subseteq M$ of q and $\varepsilon > 0$ such that the initial value problem

$$\eta(0) = x, \quad \eta' = X \circ \eta$$

has a solution on $[-\varepsilon, \varepsilon]$ for each $x \in V$. Pick $m \in \mathbb{N}$ with $t_m > b - \varepsilon$ and $\gamma(t_m) \in V$. Further let $\eta \colon [-\varepsilon, \varepsilon] \to M$ be an integral curve with $\eta(0) = \gamma(t_m)$. Then

$$\gamma(t) := \eta(t - t_m) \quad \text{for} \quad t \in [t_m - \varepsilon, t_m + \varepsilon],$$

defines an extension of γ to the interval $I_p \cup]t_m, t_m + \varepsilon[$ strictly containing $]a, b[$, hence contradicting the maximality of I_p. This proves that $\lim_{t\to b} \gamma(t) = \infty$. Replacing X by $-X$, we also obtain $\lim_{t\to a} \gamma(t) = \infty$. $\square$

Example 1.48. (a) On $M = \mathbb{R}$ we consider the vector field X given by the function $F(s) = 1 + s^2$, i.e. $X(s) = (s, 1 + s^2)$. The corresponding ODE is

$$\gamma'(s) = X(\gamma(s)) = 1 + \gamma(s)^2.$$

For $\gamma(0) = 0$ the function $\gamma(s) := \tan(s)$ on $I :=]-\frac{\pi}{2}, \frac{\pi}{2}[$ is the unique maximal solution because

$$\lim_{t \to \frac{\pi}{2}} \tan(t) = \infty \quad \text{and} \quad \lim_{t \to -\frac{\pi}{2}} \tan(t) = -\infty.$$

(b) Let $M :=]-1, 1[$ and $X(s) = (s, 1)$, so that the corresponding ODE is $\gamma'(s) = 1$. Then the unique maximal solution is

$$\gamma(s) = s, \quad I =]-1, 1[.$$

Note that we also have in this case

$$\lim_{s \to \pm 1} \gamma(s) = \infty$$

if we consider γ as a curve in the noncompact manifold M.

For $M = \mathbb{R}$ the same vector field has the maximal integral curve

$$\gamma(s) = s, \quad I = \mathbb{R}.$$

(c) For $M = \mathbb{R}$ and $X(s) = (s, -s)$, the differential equation is $\gamma'(t) = -\gamma(t)$, so that we obtain the maximal integral curves $\gamma(t) = \gamma_0 e^{-t}$. For $\gamma_0 = 0$ this curve is constant, and for $\gamma_0 \neq 0$ we have $\lim_{t\to\infty} \gamma(t) = 0$, hence $\lim_{t\to\infty} \gamma(t) \neq \infty$. This shows that maximal integral curves do not always leave every compact subset of M if they are defined on an interval that is unbounded from above.

The preceding example shows in particular that the global existence of integral curves can also be destroyed by deleting parts of the manifold M, i.e., by considering $M' := M \setminus K$ for some closed subset $K \subseteq M$.

Definition 1.49. A vector field $X \in \mathcal{V}(M)$ is said to be *complete* if all its maximal integral curves are defined on all of $\mathbb{R}$.

Corollary 1.50. *All vector fields on a compact manifold M are complete.*

Definition 1.51. Let M be a smooth manifold. A *local flow on M* is a smooth map

$$\Phi \colon U \to M,$$

where $U \subseteq \mathbb{R} \times M$ is an open subset containing $\{0\} \times M$, such that for each $x \in M$ the intersection $I_x := U \cap (\mathbb{R} \times \{x\})$ is an interval containing 0 and

$$\Phi(0, x) = x \quad \text{and} \quad \Phi\big(t, \Phi(s, x)\big) = \Phi(t + s, x)$$

hold for all t, s, x for which both sides are defined. The maps

$$\alpha_x \colon I_x \to M, \qquad t \mapsto \Phi(t, x)$$

are called the *flow lines*. The flow Φ is said to be *global* if $U = \mathbb{R} \times M$.

Lemma 1.52. *If $\Phi\colon U \to M$ is a local flow, then*

$$X^\Phi(x) := \frac{d}{dt}\Big|_{t=0} \Phi(t,x) = \alpha_x'(0)$$

defines a smooth vector field.

It is called the *velocity field* or the *infinitesimal generator* of the local flow Φ.

Lemma 1.53. *If $\Phi\colon U \to M$ is a local flow on M, then the flow lines are integral curves of the vector field X^Φ. In particular, the local flow Φ is uniquely determined by the vector field X^Φ.*

Proof. Let $\alpha_x\colon I_x \to M$ be a flow line and $s \in I_x$. For sufficiently small $t \in \mathbb{R}$ we then have

$$\alpha_x(s+t) = \Phi(s+t,x) = \Phi\big(t,\Phi(s,x)\big) = \Phi\big(t,\alpha_x(s)\big),$$

so that taking derivatives in $t = 0$ leads to $\alpha_x'(s) = X^\Phi(\alpha_x(s))$.

That Φ is uniquely determined by the vector field X^Φ follows from the uniqueness of integral curves (Theorem 1.47). □

Example 1.54. We consider $M = \mathbb{S}^2$ and $\Phi : \mathbb{R}\times\mathbb{S}^2 \to \mathbb{S}^2$ and the flow associated with a rotation around the x_3-axis:

$$\Phi(t,x) = R(t)x \qquad \text{with} \quad R = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then Φ is a flow because $\Phi(0,x) = x$ and $\Phi(t,\Phi(s,x)) = R(t)R(s)x = R(t+s)x$ for all $x \in \mathbb{S}^2$ and $t,s \in \mathbb{R}$. The associated velocity field $X_\varphi : \mathbb{S}^2 \to T\mathbb{S}^2$ is given by

$$X_\varphi(x) = \frac{d}{dt}\Big|_{t=0}\Phi(t,x) = R'(t)x = x_1e_2 - x_2e_1 = e_3 \times x.$$

Its flow lines are the curves $\gamma_x : \mathbb{R} \to \mathbb{S}^2$

$$\begin{aligned}\gamma_x(t) =& \varphi(t,x) = R(t)x = \cos t(x_1e_1 + x_2e_2) + \sin t(x_1e_2 - x_2e_1) + x_3e_3 \\ =& \cos t(x - \langle x,e_3\rangle e_3) + \sin(t)e_3 \times x + \langle x,e_3\rangle e_3.\end{aligned}$$

These are precisely the integral curves of the vector field X from Example 1.45.

As every flow determines a unique vector field, its velocity field, it is natural to ask if all vector fields on a manifold M arise as velocity fields of flows on M. That this is indeed the case is shown by the following theorem.

Theorem 1.55. *Each smooth vector field X is the velocity field of a unique local flow defined by*

$$\mathcal{D}_X := \bigcup_{x\in M} I_x \times \{x\} \quad \text{and} \quad \Phi(t,x) := \gamma_x(t) \quad \text{for} \quad (t,x) \in \mathcal{D}_X,$$

where $\gamma_x\colon I_x \to M$ is the unique maximal integral curve through $x \in M$.

Proof. If (s,x), $\big(t,\Phi(s,x)\big)$ and $(s+t,x)\in\mathcal{D}_X$, the relation

$$\Phi(s+t,x)=\Phi\big(t,\Phi(s,x)\big)\quad\text{ and }\quad I_{\Phi(s,x)}=I_{\gamma_x(s)}=I_x-s$$

follow from the fact that both curves

$$t\mapsto\Phi(t+s,x)=\gamma_x(t+s)\quad\text{ and }\quad t\mapsto\Phi\big(t,\Phi(s,x)\big)=\gamma_{\Phi(s,x)}(t)$$

are integral curves of X with the initial value $\Phi(s,x)$, hence coincide.

We claim that all maps

$$\Phi_t\colon M_t:=\{x\in M\colon (t,x)\in\mathcal{D}_X\}\to M,\quad x\mapsto\Phi(t,x)$$

are injective. In fact, if $p:=\Phi_t(x)=\Phi_t(y)$, then $\gamma_x(t)=\gamma_y(t)$, and on $[0,t]$ the curves $s\mapsto\gamma_x(t-s),\gamma_y(t-s)$ are integral curves of $-X$, starting in p. Hence the Uniqueness Theorem 1.47 implies that they coincide in $s=t$, which mans that $x=\gamma_x(0)=\gamma_y(0)=y$. From this argument it further follows that $\Phi_t(M_t)=M_{-t}$ and $\Phi_t^{-1}=\Phi_{-t}$.

It remains to show that $\mathcal{D}_X$ is open and Φ smooth. The local Existence Theorem provides for each $x\in M$ an open neighborhood U_x diffeomorphic to a cube and some $\varepsilon_x>0$, as well as a smooth map

$$\varphi_x\colon]-\varepsilon_x,\varepsilon_x[\,\times U_x\to M,\quad \varphi_x(t,y)=\gamma_y(t)=\Phi(t,y).$$

Hence $]-\varepsilon_x,\varepsilon_x[\,\times U_x\subseteq\mathcal{D}_X$, and the restriction of Φ to this set is smooth. Therefore Φ is smooth on a neighborhood of $\{0\}\times M$ in $\mathcal{D}_X$.

Now let J_x be the set¡ of all $t\in[0,\infty[$, for which $\mathcal{D}_X$ contains a neighborhood of $[0,t]\times\{x\}$ on which Φ is smooth. The interval J_x is open in $\mathbb{R}^+:=[0,\infty[$ by definition. We claim that $J_x=I_x\cap\mathbb{R}^+$. This entails that $\mathcal{D}_X$ is open because the same argument applies to $I_x\cap\,]-\infty,0]$.

We assume the contrary and find a minimal $\tau\in I_x\cap\mathbb{R}^+\setminus J_x$, because this interval is closed. Put $p:=\Phi(\tau,x)$ and pick a product set $I\times W\subseteq\mathcal{D}_X$, where W is an open neighborhood of p and $I=\,]-2\varepsilon,2\varepsilon[$ a 0-neighborhood, such that $2\varepsilon<\tau$ and $\Phi: I\times W\to M$ is smooth. By assumption, there exists an open neighborhood V of x such that Φ is smooth on $[0,\tau-\varepsilon]\times V\subseteq\mathcal{D}_X$. Then $\Phi_{\tau-\varepsilon}$ is smooth on V and

$$V':=\Phi_{\tau-\varepsilon}^{-1}\big(\Phi_\varepsilon^{-1}(W)\big)\cap V$$

is a neighborhood of x. Further,

$$V'=\Phi_{\tau-\varepsilon}^{-1}\big(\Phi_\varepsilon^{-1}(W)\big)\cap V=\Phi_\tau^{-1}(W)\cap V,$$

and Φ is smooth on $]\tau-2\varepsilon,\tau+2\varepsilon[\times V'$, because it is a composition of smooth maps:

$$]\tau-2\varepsilon,\tau+2\varepsilon[\times V'\to M,\quad (t,y)\mapsto\Phi\big(t-\tau,\Phi(\varepsilon,\Phi(\tau-\varepsilon,y))\big).$$

We thus arrive at the contradiction $\tau\in J_x$.

This completes the proof of the openness of $\mathcal{D}_X$ and the smoothness of Φ. The uniqueness of the flow follows from the uniqueness of the integral curves. $\square$

Remark 1.56. Let $X \in \mathcal{V}(M)$ be a complete vector field. If

$$\Phi^X : \mathbb{R} \times M \to M$$

is the corresponding global flow, then the maps $\Phi_t^X : x \mapsto \Phi^X(t, x)$ satisfy

(A1) $\Phi_0^X = \mathrm{id}_M$.

(A2) $\Phi_{t+s}^X = \Phi_t^X \circ \Phi_s^X$ for $t, s \in \mathbb{R}$.

It follows in particular that $\Phi_t^X \in \mathrm{Diff}(M)$ with $(\Phi_t^X)^{-1} = \Phi_{-t}^X$, so that we obtain a group homomorphism

$$\gamma_X : \mathbb{R} \to \mathrm{Diff}(M), \quad t \mapsto \Phi_t^X.$$

With respect to the terminology introduced below, (A1) and (A2) mean that Φ^X defines a *smooth action of* $\mathbb{R}$ *on* M. As Φ^X is determined by the vector field X, we call X the infinitesimal generator of this action. In this sense the smooth $\mathbb{R}$-actions on a manifold M are in one-to-one correspondence with the complete vector fields on M.

Remark 1.57. Let $\Phi^X : \mathcal{D}_X \to M$ be the maximal local flow of a vector field X on M. Let $M_t = \{x \in M : (t, x) \in \mathcal{D}_X\}$, and observe that this is an open subset of M. We have already seen in the proof of Theorem 1.55 above, that all the smooth maps $\Phi_t^X : M_t \to M$ are injective with $\Phi_t^X(M_t) = M_{-t}$ and $(\Phi_t^X)^{-1} = \Phi_{-t}^X$ on the image. It follows in particular, that $\Phi_t^X(M_t) = M_{-t}$ is open, and that

$$\Phi_t^X : M_t \to M_{-t}$$

is a diffeomorphism whose inverse is Φ_{-t}^X.

Proposition 1.58. (Smooth Dependence Theorem) *Let M and Λ be smooth manifolds and $\Psi : \Lambda \to \mathcal{V}(M)$ be a map for which the map*

$$\Lambda \times M \to T(M), \quad (\lambda, p) \mapsto \Psi_\lambda(p)$$

is smooth (the vector field Ψ_λ depends smoothly on the parameter λ). Then the subset

$$\mathcal{D} := \{(t, \lambda, p) \in \mathbb{R} \times \Lambda \times M : (t, p) \in \mathcal{D}_{\Phi_\lambda}\}$$

of $\mathbb{R} \times \Lambda \times M$ is open and the map $\mathcal{D} \to M, (t, \lambda, p) \mapsto \Phi^{\Psi_\lambda}(t, p)$ is smooth.

Proof. The parameters do not cause any additional problems, as can be seen by the following trick: On the product manifold $\Lambda \times M$ we consider the smooth vector field Y, given by

$$Y(\lambda, p) := (0_\lambda, \Psi_\lambda(p)) \in T_\lambda(\Lambda) \times T_p(M) \cong T_{(\lambda,p)}(\Lambda \times M).$$

Then the integral curves of Y are of the form $\gamma(t) = (\lambda, \gamma_p(t))$, where γ_p is an integral curve of the smooth vector field Ψ_λ on M. Therefore the assertion is an immediate consequence on the smoothness of the flow of Y on $\Lambda \times M$ (Theorem 1.55). $\square$

We take a closer look at the interaction of local flows and vector fields. It will turn out that this leads to a new concept of a directional derivative which works for general tensor fields. Let $X \in \mathcal{V}(M)$ and $\Phi^X : \mathcal{D}_X \to M$ its maximal local flow. For $f \in C^\infty(M)$ and $t \in \mathbb{R}$ we set

$$(\Phi_t^X)^* f := f \circ \Phi_t^X \in C^\infty(M_t).$$

Then we find

$$\lim_{t\to 0} \frac{1}{t}((\Phi_t^X)^* f - f) = \mathtt{d}f(X) = \mathcal{L}_X f \in C^\infty(M).$$

For a second vector field $Y \in \mathcal{V}(M)$, we define a smooth vector field on the open subset $M_{-t} \subseteq M$ by

$$(\Phi_t^X)_* Y := T(\Phi_t^X) \circ Y \circ \Phi_{-t}^X = T(\Phi_t^X) \circ Y \circ (\Phi_t^X)^{-1}$$

(cf. Remark 1.57) and define the *Lie derivative* by

$$\mathcal{L}_X Y := \lim_{t\to 0} \frac{1}{t}((\Phi_{-t}^X)_* Y - Y) = \frac{d}{dt}\Big|_{t=0} (\Phi_{-t}^X)_* Y,$$

which is defined on all of M since for each $p \in M$ the vector $((\Phi_t^X)_* Y)(p)$ is defined for sufficiently small t and depends smoothly on t.

Theorem 1.59. *$\mathcal{L}_X Y = [X, Y]$ for $X, Y \in \mathcal{V}(M)$.*

Proof. Fix $p \in M$. It suffices to show that $\mathcal{L}_X Y$ and $[X, Y]$ coincide in p. We may therefore work in a local chart, hence assume that $M = U$ is an open subset of $\mathbb{R}^n$.

Identifying vector fields with smooth $\mathbb{R}^n$-valued functions, we then have

$$[X, Y](x) = \mathtt{d}Y(x)X(x) - \mathtt{d}X(x)Y(x), \quad x \in U.$$

On the other hand,

$$\begin{aligned}((\Phi_{-t}^X)_* Y)(x) &= T(\Phi_{-t}^X) \circ Y \circ \Phi_t^X(x) \\ &= \mathtt{d}(\Phi_{-t}^X)(\Phi_t^X(x))Y(\Phi_t^X(x)) = \big(\mathtt{d}(\Phi_t^X)(x)\big)^{-1} Y(\Phi_t^X(x)).\end{aligned}$$

To calculate the derivative of this expression with respect to t, we first observe that it does not matter if we first take derivatives with respect to t and then with respect to x or vice versa. This leads to

$$\frac{d}{dt}\Big|_{t=0} \mathtt{d}(\Phi_t^X)(x) = \mathtt{d}\Big(\frac{d}{dt}\Big|_{t=0} \Phi_t^X\Big)(x) = \mathtt{d}X(x).$$

Next we note that for any smooth curve $\alpha\colon [-\varepsilon, \varepsilon] \to \mathrm{GL}_n(\mathbb{R})$ with $\alpha(0) = \mathbf{1}$ we have

$$(\alpha^{-1})'(t) = -\alpha(t)^{-1}\alpha'(t)\alpha(t)^{-1},$$

and in particular $(\alpha^{-1})'(0) = -\alpha'(0)$. Combining all this, we obtain with the Product Rule

$$\mathcal{L}_X(Y)(x) = -\mathtt{d}X(x)Y(x) + \mathtt{d}Y(x)X(x) = [X, Y](x). \qquad \square$$

Corollary 1.60. *If $X, Y \in \mathcal{V}(M)$ are complete vector fields, then their global flows $\Phi^X, \Phi^Y : \mathbb{R} \to \mathrm{Diff}(M)$ commute if and only if X and Y commute, i.e. $[X, Y] = 0$.*

Proof. (1) Suppose first that Φ^X and Φ^Y commute, i.e.,

$$\Phi^X(t) \circ \Phi^Y(s) = \Phi^Y(s) \circ \Phi^X(t) \quad \text{for } t, s \in \mathbb{R}.$$

Let $p \in M$ and $\gamma_p(s) := \Phi^Y_s(p)$ be the Y-integral curve through p. We then have

$$\gamma_p(s) = \Phi^Y_s(p) = \Phi^X_t \circ \Phi^Y_s \circ \Phi^X_{-t}(p),$$

and passing to the derivative in $s = 0$ yields

$$Y(p) = \gamma_p'(0) = T(\Phi^X_t)Y(\Phi^X_{-t}(p)) = \big((\Phi^X_t)_*Y\big)(p).$$

Passing now to the derivative in $t = 0$, we arrive at $[X, Y] = \mathcal{L}_X(Y) = 0$.

(2) Now we assume $[X, Y] = 0$. First we show that $(\Phi^X_t)_*Y = Y$ holds for all $t \in \mathbb{R}$. For $t, s \in \mathbb{R}$ we have

$$(\Phi^X_{t+s})_*Y = (\Phi^X_t)_*(\Phi^X_s)_*Y,$$

so that

$$\frac{d}{dt}(\Phi^X_t)_*Y = -(\Phi^X_t)_*\mathcal{L}_X(Y) = 0$$

for each $t \in \mathbb{R}$. Since for each $p \in M$ the curve

$$\mathbb{R} \to T_p(M), \quad t \mapsto \big((\Phi^X_t)_*Y\big)(p)$$

is smooth, and its derivative vanishes, it is constant $Y(p)$. This shows that $(\Phi^X_t)_*Y = Y$ for each $t \in \mathbb{R}$.

For $\gamma(s) := \Phi^X_t\Phi^Y_s(p)$ we now have $\gamma(0) = \Phi^X_t(p)$ and

$$\gamma'(s) = T(\Phi^X_t) \circ Y(\Phi^Y_s(p)) = Y(\Phi^X_t\Phi^Y_s(p)) = Y(\gamma(s)),$$

so that γ is an integral curve of Y. We conclude that $\gamma(s) = \Phi^Y_s(\Phi^X_t(p))$, and this means that the flows of X and Y commute. $\square$

Exercises for Section 1.4

Exercise 1.28. Let $M := \mathbb{R}^n$. For a matrix $A \in M_n(\mathbb{R})$, we consider the linear vector field $X_A(x) := Ax$. Determine the maximal flow Φ^X of this vector field.

Exercise 1.29. Let M be a smooth manifold and $Y \in \mathcal{V}(M)$ a smooth vector field on M. Suppose that Y generates a local flow $\Phi^Y : \mathcal{D}_Y \to M$ which is defined on an entire box of the form $[-\varepsilon, \varepsilon] \times M \subseteq \mathcal{D}_Y$. Show that this implies the completeness of Y.

Exercise 1.30. Let $\varphi \colon M \to N$ be a smooth map and $X \in \mathcal{V}(M)$, $Y \in \mathcal{V}(N)$ be φ-related vector fields. Show that for any integral curve $\gamma \colon I \to M$ of X, the curve $\varphi \circ \gamma \colon I \to N$ is an integral curve of Y.

Exercise 1.31. Let $X \in \mathcal{V}(M)$ be a vector field and write $X^{\mathbb{R}} \in \mathcal{V}(\mathbb{R})$ for the vector field on $\mathbb{R}$, given by $X^{\mathbb{R}}(t) = (t, 1)$. Show that, for an open interval $I \subseteq \mathbb{R}$, a smooth curve $\gamma \colon I \to M$ is an integral curve of X if and only if $X^{\mathbb{R}}$ and X are γ-related.

Exercise 1.32. Let $X \in \mathcal{V}(M)_c$ be a complete vector field and $\varphi \in \mathrm{Diff}(M)$. Then $\varphi_* X$ is also complete and

$$\Phi_t^{\varphi_* X} = \varphi \circ \Phi_t^X \circ \varphi^{-1} \quad \text{for} \quad t \in \mathbb{R}.$$

Exercise 1.33. Let M be a smooth manifold, $\varphi \in \mathrm{Diff}(M)$ and $X \in \mathcal{V}(M)_c$ be a complete vector field. Show that the following are equivalent:

(1) φ commutes with the flow maps Φ_t^X.

(2) For each integral curve $\gamma\colon I \to M$ of X, the curve $\varphi \circ \gamma$ also is an integral curve of X.

(3) $X = \varphi_* X = T(\varphi) \circ X \circ \varphi^{-1}$, i.e., X is φ-invariant.

Exercise 1.34. Let $X, Y \in \mathcal{V}(M)$ be two commuting complete vector fields, i.e., $[X, Y] = 0$. Show that the vector field $X + Y$ is complete and that its flow is given by

$$\Phi_t^{X+Y} = \Phi_t^X \circ \Phi_t^Y \quad \text{for all} \quad t \in \mathbb{R}.$$

Exercise 1.35. Let V be a finite-dimensional vector space and $\mu_t(v) := tv$ for $t \in \mathbb{R}^\times$. Show that:

(1) A vector field $X \in \mathcal{V}(V)$ is linear if and only if $(\mu_t)_* X = X$ holds for all $t \in \mathbb{R}^\times$.

(2) A diffeomorphism $\varphi \in \mathrm{Diff}(V)$ is linear if and only if it commutes with all the maps μ_t, $t \in \mathbb{R}^\times$.

2 Lie Groups

Symmetries of physical systems are most naturally modelled by the mathematical concept of a group. If S is the state space of a physical system, then a symmetry is mostly considered as a bijection of this set preserving additional structure on S. As composition of symmetries is a symmetry and any symmetry should have an inverse symmetry, we are thus lead to certain groups G of bijections of the set S.

Groups can be studied on three levels:

- the discrete level: no additional structure on G.
- the topological level: topological groups; G is endowed with a topology.
- the differentiable level: Lie groups; G is endowed with a smooth manifold structure.

The first level only provides a reasonable context for groups arising as symmetry groups of discrete structures, such as crystals, which do not permit any continuous (in the sense of “continuum”) symmetry operations. Whenever continuous symmetries exist, such as rotations of a round sphere, it is natural to study symmetries $(g_t)_{t \in \mathbb{R}}$ depending on a real parameter, such that

$$g_0 = \mathrm{id} \quad \text{and} \quad g_t g_s = g_{t+s} \quad \text{for} \quad t, s \in \mathbb{R}.$$

We thus obtain continuous one-parameter groups of a topological group G. As topological groups can still be rather wild, one then refines the structure on G in such a way that differentiation of one-parameter groups becomes meaningful. This leads to the concept of an "infinitesimal generator" of a one-parameter group (the great idea of Sophus Lie[1]) which is closely related to vector fields as infinitesimal generators of local flows on manifolds. It turns out that the concept of a Lie group, i.e. a group endowed with a smooth manifold structure compatible with the group operations, provides precisely the additional structure for which the set $\mathbf{L}(G)$ of infinitesimal generators of one-parameter groups carries the nice algebraic structure of a Lie algebra and, in addition, the structure of the group near the identity is completely determined by its Lie algebra, resp., its one-parameter groups.

2.1 The concept of a Lie group

In the context of smooth manifolds, the natural class of groups are those endowed with a manifold structure compatible with the group structure.

Definition 2.1. A *Lie group* is a group G, endowed with the structure of a smooth manifold, such that the group operations

$$m_G\colon G \times G \to G, \quad (x,y) \mapsto xy \quad \text{and} \quad \iota_G\colon G \to G, \quad x \mapsto x^{-1}$$

are smooth.

In the following, G denotes a Lie group with

- multiplication map $m_G\colon G \times G \to G, (x,y) \mapsto xy$,
- inversion map $\iota_G\colon G \to G, x \mapsto x^{-1}$, and
- neutral element $\mathbf{1}$.

For $g \in G$ we write

- $\lambda_g\colon G \to G, x \mapsto gx$ for the left multiplication maps (left translations),
- $\rho_g\colon G \to G, x \mapsto xg$ for the right multiplication maps (right translations), and
- $c_g\colon G \to G, x \mapsto gxg^{-1}$ for the conjugation with g.

A *morphism of Lie groups* is a smooth homomorphism of Lie groups $\varphi\colon G_1 \to G_2$.

Remark 2.2. All maps λ_g, ρ_g and c_g are smooth. Moreover, they are bijective with $\lambda_{g^{-1}} = \lambda_g^{-1}$, $\rho_{g^{-1}} = \rho_g^{-1}$ and $c_{g^{-1}} = c_g^{-1}$, so that they are diffeomorphisms of G.

Example 2.3. The additive group $G := (\mathbb{R}^n, +)$ is a Lie group because the maps

$$\mathbb{R}^{2n} \to \mathbb{R}^n, \quad (x,y) \mapsto x + y \quad \text{and} \quad \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto -x$$

are smooth.

[1]The Norwegian mathematician Marius Sophus Lie (1842–1899) was the first to study differentiability properties of groups in a systematic way. In the 1890s Sophus Lie developed his theory of differentiable groups (called continuous groups at a time when the concept of a topological space was not yet developed) to study symmetries of differential equations.

Example 2.4. Let $G := \mathrm{GL}_n(\mathbb{K})$ be the group of invertible $(n \times n)$-matrices with entries in the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Since the determinant function

$$\det\colon M_n(\mathbb{K}) \to \mathbb{K}, \quad \det(a_{ij}) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

is continuous and $\mathbb{K}^\times := \mathbb{K} \setminus \{0\}$ is open in $\mathbb{K}$, the set $\mathrm{GL}_n(\mathbb{K}) = \det^{-1}(\mathbb{K}^\times)$ is open in $M_n(\mathbb{K})$ and thus carries a canonical manifold structure.

For the smoothness of the multiplication map, it suffices to observe that

$$(ab)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

is the (ik)-entry in the product matrix. Since all these entries are quadratic polynomials in the entries of a and b, the product is a smooth map.

For $g \in \mathrm{GL}_n(\mathbb{K})$ we define $b_{ij}(g) := \det(g_{mk})_{m \neq j, k \neq i}$. According to Cramer's Rule, the inverse of g is given by

$$(g^{-1})_{ij} = \frac{(-1)^{i+j}}{\det g} b_{ij}(g).$$

The smoothness of the inversion therefore follows from the smoothness of the determinant (which is a polynomial) and the polynomial functions b_{ij} defined on $M_n(\mathbb{K})$.

Example 2.5. (a) (The circle group) We have already seen how to endow the circle

$$\mathbb{S}^1 := \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 = 1\}$$

with a manifold structure. Identifying it with the unit circle

$$\mathbb{T} := \{z \in \mathbb{C} \colon |z| = 1\}$$

in $\mathbb{C}$, it also inherits a group structure, given by

$$(x, y) \cdot (x', y') := (xx' - yy', xy' + x'y) \quad \text{and} \quad (x, y)^{-1} = (x, -y).$$

With these explicit formulas, it is easy to verify that $\mathbb{T}$ is a Lie group (Exercise 2.1).

(b) (The n-dimensional torus) In view of (a), we have a natural manifold structure on the *n-dimensional torus* $\mathbb{T}^n := (\mathbb{S}^1)^n$. The corresponding direct product group structure

$$(t_1, \ldots, t_n)(s_1, \ldots, s_n) := (t_1 s_1, \ldots, t_n s_n)$$

turns $\mathbb{T}^n$ into a Lie group (Exercise 2.2).

Lemma 2.6. *Let G be a Lie group with multiplication $m_G \colon G \times G \to G$. Then its tangent map satisfies*

$$T_{(g,h)}(m_G)(v, w) = T_g(\rho_h)v + T_h(\lambda_g)w \quad \text{for} \quad v \in T_g(G), w \in T_h(G) \tag{5}$$

Proof. For $v \in T_g(G)$ and $w \in T_h(G)$, the linearity of $T_{(g,h)}(m_G)$ implies that

$$T_{(g,h)}(m_G)(v, w) = T_{(g,h)}(m_G)(v, 0) + T_{(g,h)}(m_G)(0, w) = T_g(\rho_h)v + T_h(\lambda_g)w. \qquad \square$$

In the following we shall use the simplified notation

$$g \cdot v := T(\lambda_g)v \quad \text{and} \quad v \cdot g := T(\rho_g)v \quad \text{for} \quad g \in G, v \in TG. \tag{6}$$

Then (5) turns into

$$T_{(g,h)}(m_G)(v,w) = g \cdot w + v \cdot h.$$

For differential curves $\alpha(t)$, $\beta(t)$ in G, this leads to the product rule

$$(\alpha\beta)'(t) = \alpha(t)\beta'(t) + \alpha'(t)\beta(t). \tag{7}$$

2.2 The Lie algebra of a Lie group

Lie groups are non-linear objects. We now introduce the Lie algebra $\mathbf{L}(G)$ of a Lie group as a "first order approximation", resp., a "linearization" of G.

We start with the introduction of the concept of a Lie algebra.

Definition 2.7. (a) Let $\mathbb{K}$ be a field and L a $\mathbb{K}$-vector space. A bilinear map $[\cdot,\cdot]\colon L \times L \to L$ is called a *Lie bracket* if

(L1) $[x,x] = 0$ for $x \in L$ and

(L2) $\big[x,[y,z]\big] = \big[[x,y],z\big] + \big[y,[x,z]\big]$ for $x,y,z \in L$ (Jacobi identity).[2]

Note that, provided (L1) holds, the Jacobi identity can also be expressed in a more symmetric fashion by

$$\big[x,[y,z]\big] + \big[y,[z,x]\big] + \big[z,[x,y]\big] = 0.$$

A *Lie algebra* [3] (over $\mathbb{K}$) is a $\mathbb{K}$-vector space L, endowed with a Lie bracket. A subspace $E \subseteq L$ of a Lie algebra is called a *subalgebra* if $[E,E] \subseteq E$. A *homomorphism* $\varphi\colon L_1 \to L_2$ of Lie algebras is a linear map with $\varphi([x,y]) = [\varphi(x),\varphi(y)]$ for $x,y \in L_1$. A Lie algebra L is said to be *abelian* if $[x,y] = 0$ holds for all $x,y \in L$.

Remark 2.8. If $b_1,\ldots,b_n \in L$ is a basis of the Lie algebra L, then all information on the bilinear Lie bracket is contained in the brackets

$$[b_i,b_j] = \sum_{k=1}^{n} c_{ij}^k b_k$$

which in turn is contained in the n^3 numbers c_{ij}^k called the *structure constants of* L. Skew-symmetry and Jacobi identity of the Lie bracket can be expressed in terms of the structure constants as

$$c_{ij}^k = -c_{ji}^k \quad \text{and} \quad \sum_\ell c_{ij}^\ell c_{\ell k}^m + c_{jk}^\ell c_{\ell i}^m + c_{ki}^\ell c_{\ell j}^m = 0.$$

[2]Carl Gustav Jacob Jacobi (1804–1851), mathematician in Berlin and Königsberg (Kaliningrad). He found his famous identity about 1830 in the context of Poisson brackets, which are related to Hamiltonian Mechanics and Symplectic Geometry.

[3]The notion of a Lie algebra was coined in the 1920s by Hermann Weyl.

Example 2.9. Each associative algebra $\mathcal{A}$ is a Lie algebra $\mathcal{A}_L$ with respect to the commutator bracket

$$[a,b] := ab - ba.$$

In particular, the matrix algebra $M_n(\mathbb{K})$ and the endomorphism algebra $\mathrm{End}(V)$ of a vector space are Lie algebras with respect to the commutator bracket.

In fact, (L1) is obvious. For (L2), we calculate

$$[a, bc] = abc - bca = (ab - ba)c + b(ac - ca) = [a,b]c + b[a,c],$$

and this implies

$$[a,[b,c]] = [a,b]c + b[a,c] - [a,c]b - c[a,b] = [[a,b],c] + [b,[a,c]].$$

Example 2.10. For every smooth manifold M, the space $\mathcal{V}(M)$ of smooth vector fields on M is a Lie algebra.

Let G be a Lie group. A vector field $X \in \mathcal{V}(G)$ is called *left invariant* if

$$X(gh) = g \cdot X(h) \quad \text{for} \quad g, h \in G.$$

We write $\mathcal{V}(G)^l$ for the linear space of left invariant vector fields in $\mathcal{V}(G)$. Clearly $\mathcal{V}(G)^l$ is a linear subspace of $\mathcal{V}(G)$.

Lemma 2.11. *The vector space $\mathcal{V}(G)^l$ of left invariant vector fields on G is a Lie subalgebra of* $(\mathcal{V}(G), [\cdot,\cdot])$.

Proof. Writing the left invariance as $X \circ \lambda_g = T(\lambda_g) \circ X$, we see that it means that X is left invariant if and only if it is λ_g-related to itself for every $g \in G$. Therefore the Related Vector Field Lemma implies that if X and Y are left invariant, their Lie bracket $[X,Y]$ is also λ_g-related to itself for each $g \in G$, hence left invariant. □

Definition 2.12. [The Lie algebra of G] Next we observe that the left invariance of a vector field X implies that for each $g \in G$ we have $X(g) = g \cdot X(\mathbf{1})$, so that X is completely determined by its value $X(\mathbf{1}) \in T_{\mathbf{1}}(G)$. Conversely, for each $x \in T_{\mathbf{1}}(G)$, we obtain a left invariant vector field $x_l \in \mathcal{V}(G)^l$ with $x_l(\mathbf{1}) = x$ by $x_l(g) := g \cdot x$. That this vector field is indeed left invariant follows from

$$x_l(gh) = gh \cdot x = T(\lambda_{hg})x = T(\lambda_h \circ \lambda_g)x = T(\lambda_h)T(\lambda_g)x = h \cdot x_l(g)$$

for all $h, g \in G$. Hence

$$T_{\mathbf{1}}(G) \to \mathcal{V}(G)^l, \quad x \mapsto x_l$$

is a linear bijection. We thus obtain a Lie bracket $[\cdot,\cdot]$ on $T_{\mathbf{1}}(G)$ by

$$[x,y] := [x_l, y_l](\mathbf{1}).$$

It satisfies

$$[x,y]_l = [x_l, y_l] \quad \text{for all} \quad x, y \in T_{\mathbf{1}}(G). \tag{8}$$

The Lie algebra

$$\mathbf{L}(G) := (T_{\mathbf{1}}(G), [\cdot,\cdot]) \cong \mathcal{V}(G)^l$$

is called *the Lie algebra of G*.

Remark 2.13. Let $d_1, \ldots, d_n$ be a basis of $T_{\mathbf{1}}(G)$ and $D_j \in \mathcal{V}(G)^l$ denote the corresponding left invariant vector fields. To determine the Lie bracket on $\mathbf{L}(G) = T_{\mathbf{1}}(G)$, one can proceed as follows. In a local chart (φ, U) of G with $\mathbf{1} \in U$ we identify the vector fields $\widetilde{D}_j := \varphi_* D_J$ with smooth functions

$$\widetilde{D}_j = \sum_\ell \widetilde{D}_j^\ell \cdot b_\ell^\varphi \colon \varphi(U) \to \mathbb{R}^n.$$

Then their Lie bracket is given in coordinate free notation by

$$[\widetilde{D}_j, \widetilde{D}_k] = \mathtt{d}\widetilde{D}_k \cdot \widetilde{D}_j - \mathtt{d}\widetilde{D}_j \cdot \widetilde{D}_k,$$

and in terms of the component functions by

$$[\widetilde{D}_j, \widetilde{D}_k]^\ell = \sum_\alpha \widetilde{D}_j^\alpha \frac{\partial \widetilde{D}_k^\ell}{\partial \varphi_\alpha} - \widetilde{D}_k^\alpha \frac{\partial \widetilde{D}_j^\ell}{\partial \varphi_\alpha}.$$

Proposition 2.14. (Functoriality of the Lie algebra) *If $\varphi \colon G \to H$ is a morphism of Lie groups, then the tangent map*

$$\mathbf{L}(\varphi) := T_{\mathbf{1}}(\varphi) \colon \mathbf{L}(G) \to \mathbf{L}(H)$$

is a homomorphism of Lie algebras.

Proof. Let $x, y \in \mathbf{L}(G)$ and x_l, y_l be the corresponding left invariant vector fields. Then $\varphi \circ \lambda_g = \lambda_{\varphi(g)} \circ \varphi$ for each $g \in G$ implies that

$$T(\varphi) \circ T(\lambda_g) = T(\lambda_{\varphi(g)}) \circ T(\varphi),$$

and applying this relation to $x, y \in T_{\mathbf{1}}(G)$, we get

$$T\varphi \circ x_l = \big(\mathbf{L}(\varphi)x\big)_l \circ \varphi \quad \text{and} \quad T\varphi \circ y_l = \big(\mathbf{L}(\varphi)y\big)_l \circ \varphi, \tag{9}$$

i.e. x_l is φ-related to $\big(\mathbf{L}(\varphi)x\big)_l$ and y_l is φ-related to $\big(\mathbf{L}(\varphi)y\big)_l$. Therefore the Related Vector Field Lemma implies that

$$T\varphi \circ [x_l, y_l] = [\big(\mathbf{L}(\varphi)x\big)_l, \big(\mathbf{L}(\varphi)y\big)_l] \circ \varphi.$$

Evaluating at $\mathbf{1}$, we obtain $\mathbf{L}(\varphi)[x, y] = [\mathbf{L}(\varphi)(x), \mathbf{L}(\varphi)(y)]$, showing that $\mathbf{L}(\varphi)$ is a homomorphism of Lie algebras. □

Example 2.15. For the Lie group $G = (\mathbb{R}^n, +)$ we write its tangent bundle as $T\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$ and, accordingly, we write smooth vector fields as functions $X \colon \mathbb{R}^n \to \mathbb{R}^n$. In this picture, the differential of the translation maps $\lambda_x(y) = x + y$ is the identity, so that $X \in \mathcal{V}(\mathbb{R}^n)$ is left invariant if and only if it is constant. For constant vector fields X, Y we have

$$[X, Y](p) = \mathtt{d}Y(p)X(p) - \mathtt{d}X(p)Y(p) = 0.$$

Therefore the Lie algebra $\mathbf{L}(\mathbb{R}^n)$ is abelian, i.e. all brackets vanish.

Example 2.16. Since the Lie group $G = \mathrm{GL}_n(\mathbb{R})$ is an open subset of $M_n(\mathbb{R})$, we identify its tangent bundle with the subset $T\,\mathrm{GL}_n(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R}) \times M_n(\mathbb{R})$ and smooth vector fields with functions $X\colon \mathrm{GL}_n(\mathbb{R}) \to M_n(\mathbb{R})$. Here the left multiplications $\lambda_g(h) = gh$ are restrictions of linear maps, to that a vector field $X \in \mathcal{V}(\mathrm{GL}_n(\mathbb{R}))$ is left invariant if and only if

$$X(gh) = T(\lambda_g)X(h) = gX(h) \quad \text{for} \quad g, h \in \mathrm{GL}_n(\mathbb{R}).$$

Therefore the left invariance of a vector field X is equivalent to the existence of some $A \in M_n(\mathbb{R})$ with $X(g) = X_A(g) := gA$. For these vector fields we have $\mathtt{d}X_A(g)C = CA$ for $B \in M_n(\mathbb{R})$, so that

$$[X_A, X_B](g) = \mathtt{d}X_B(g)X_A(g) - \mathtt{d}X_A(g)X_B(g) = g(AB - BA).$$

Therefore the Lie algebra $\mathbf{L}(\mathrm{GL}_n(\mathbb{R}))$ is the space $M_n(\mathbb{R}) \cong T_1(\mathrm{GL}_n(\mathbb{R}))$, endowed with the *commutator bracket*

$$[A, B] = AB - BA.$$

This Lie algebra is denoted $\mathfrak{gl}_n(\mathbb{R})$, to express that it is the Lie algebra of $\mathrm{GL}_n(\mathbb{R})$.

2.3 The exponential function of a Lie group

In this section, we introduce a key tool of Lie theory which is a bridge between the "nonlinear" Lie group G and the "linear" Lie algebra $\mathbf{L}(G)$: the exponential function $\exp_G\colon \mathbf{L}(G) \to G$. It is a natural generalization of the matrix exponential map, which is obtained for $G = \mathrm{GL}_n(\mathbb{R})$ and its Lie algebra $\mathbf{L}(G) = \mathfrak{gl}_n(\mathbb{R})$.

Definition 2.17. Let G be a Lie group. A smooth function $\exp\colon \mathbf{L}(G) \to G$ is called an *exponential function* if for every $x \in \mathbf{L}(G)$ the curve

$$\gamma_x(t) := \exp(tx)$$

is a one-parameter group, i.e.

$$\gamma_x(t + s) = \gamma_x(t)\gamma_x(s) \quad \text{for } s, t \in \mathbb{R}, \quad \text{and} \quad \gamma_x'(0) = x. \tag{10}$$

Passing to the derivative of this relation with respect to s in 0, we see that any smooth one-parameter group $\gamma\colon \mathbb{R} \to G$ with $\gamma'(0) = x$ is the unique solution of the initial value problem

$$\gamma(0) = \mathbf{1} \quad \text{and} \quad \dot\gamma(t) = \gamma(t) \cdot x = x_l(\gamma(t)).$$

Therefore we call x the *infinitesimal generator of* γ_x (cf. (7)).

Theorem 2.18. *Every Lie group has a uniquely determined exponential function.*

Proof. (Sketch) If $\gamma_x(t)$ is a smooth one-parameter group of G with $\gamma_x'(0) = x$, then

$$\Phi_t(g) := g\gamma_x(t)$$

defines a flow on G whose infinitesimal generator $X^\Phi := \frac{d}{dt}|_{t=0}\Phi_t \in \mathcal{V}(G)$ is a left invariant vector field with $X^\Phi(\mathbf{1}) = x$. In particular, $\gamma_x(t)$ is the unique integral curve through $\mathbf{1}$. To prove the existence of an exponential function, one therefore has to study the flows generated by left invariant vector fields.

To show the existence of an exponential function, one first shows that all left invariant vector fields x_l are complete and defines $\exp(x) := \gamma_x(1)$, where γ_x is the unique integral curve of x_l through the identity. Then one verifies that $\exp(tx) = \gamma_x(t)$ for $t \in \mathbb{R}$, and the smoothness of exp follows from the smooth dependence of integral curves from parameters (cf. Section 1). $\square$

Remark 2.19. Let $x \in \mathbf{L}(G)$ and $x_l(g) = g \cdot x$ denote the corresponding left invariant vector field. Then its flow has the form $\Phi_t^{x_l}(g) = g \exp(tx)$, so that the corresponding Lie derivative is given on smooth functions on G by

$$(\mathcal{L}_x f)(g) := \mathcal{L}_{x_l}(g) := \frac{d}{dt}\Big|_{t=0} f(g \exp tx).$$

Accordingly, the right invariant vector field $x_r(g) = x \cdot g$ generates the flow $\Phi_t^{x_r}(g) = \exp(tx)g$ and the corresponding Lie derivative is

$$(\mathcal{R}_x f)(g) := \mathcal{L}_{x_r}(g) := \frac{d}{dt}\Big|_{t=0} f((\exp tx)g).$$

Remark 2.20. (a) For a Lie group G, the exponential function $\exp_G \colon \mathbf{L}(G) \to G$ satisfies

$$T_0(\exp_G) = \mathrm{id}_{\mathbf{L}(G)}$$

because for each $x \in \mathbf{L}(G)$ we have

$$T_0(\exp_G)x = \frac{d}{dt}\Big|_{t=0} \exp_G(tx) = x.$$

Therefore the Inverse Function Theorem implies that $\exp_G$ is a local diffeomorphism in 0 in the sense that there exists an open 0-neighborhood $U \subseteq \mathbf{L}(G)$ such that $\exp_G|_U \colon U \to \exp_G(U)$ is a diffeomorphism onto an open subset of G.

If $b_1, \ldots, b_n$ is a basis of $\mathbf{L}(G)$, then we thus obtain the so-called *canonical coordinates of the first kind* on an identity neighborhood of G:

$$\Phi \colon \mathbb{R}^n \to G, \quad x \mapsto \exp_G(x_1 b_1 + \ldots + x_n b_n)$$

(b) Sometimes it is more convenient to use *canonical coordinates of the second kind*

$$\Psi \colon \mathbb{R}^n \to G, \quad x \mapsto \exp_G(x_1 b_1) \cdot \ldots \cdot \exp_G(x_n b_n)$$

That Ψ is a local diffeomorphism in 0 follows from $T_0(\Psi)(x) = \sum_{i=1}^n x_i b_i$, which in turn follows by repeated application of the product rule (7). which leads to . Hence the claim follows from the Inverse Function Theorem.

Example 2.21. For $G = \mathbb{R}^n$ the identity $\exp_{\mathbb{R}^n} = \mathrm{id}$ is an exponential function because each curve $\gamma_x(t) = tx$ is a smooth one-parameter group with $\gamma_x'(0) = x$.

Example 2.22. For $G := \mathrm{GL}_n(\mathbb{R})$, the left invariant vector field A_l corresponding to a matrix A is given by

$$A_l(g) = T_{\mathbf{1}}(\lambda_g)A = gA$$

because $\lambda_g(h) = gh$ extends to a linear endomorphism of $M_n(\mathbb{R})$. The unique solution $\gamma_A \colon \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$ of the initial value problem

$$\gamma(0) = \mathbf{1}, \quad \gamma'(t) = A_l(\gamma(t)) = \gamma(t)A$$

is the curve describing the fundamental system of the linear differential equation defined by the matrix A:

$$\gamma_A(t) = e^{tA} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k.$$

It follows that $\exp_G(A) = e^A$ is the matrix exponential function.

Example 2.23. We consider the 3-dimensional *Heisenberg group*

$$H_3 := \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{R} \right\}$$

which clearly is a 3-dimensional submanifold of the Lie group $\mathrm{GL}_3(\mathbb{R}) \subseteq M_3(\mathbb{R}) \cong \mathbb{R}^{3\times 3}$ from which it inherits a Lie group structure. With respect to the obvious (x, y, z)-coordinates, we can identify H_3 with $\mathbb{R}^3$, endowed with the multiplication

$$\mathbf{x} \sharp \mathbf{x}' = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \sharp \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} := \begin{pmatrix} x + x' \\ y + y' \\ z + z' + xy' \end{pmatrix}.$$

From the canonical basis e_1, e_2, e_3 of $\mathbb{R}^3$ we obtain a the following left invariant vector fields on $H_3 \cong (\mathbb{R}^3, \sharp)$:

$$P(\mathbf{x}) = T_1(\lambda_{(x,y,z)})e_1 = e_1, \quad Q(\mathbf{x}) = T_1(\lambda_{(x,y,z)})e_2 = \begin{pmatrix} 0 \\ 1 \\ x \end{pmatrix}, \quad Z(\mathbf{x}) = T_1(\lambda_{(x,y,z)})e_3 = e_3$$

with the Lie brackets

$$[P, Q] = Z, \quad [P, Z] = [Q, Z] = 0.$$

In the matrix picture, these vector fields correspond to the matrices

$$\widehat{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \widehat{Q} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \widehat{Z} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and the exponential function is given by

$$\exp(p\widehat{P} + q\widehat{Q} + z\widehat{Z}) = \exp \begin{pmatrix} 0 & p & z \\ 0 & 0 & q \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & p & z + \frac{pq}{2} \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix}$$

and defines a diffeomorphism $\mathbf{L}(H_3) \to H_3$. On the other hand,

$$\exp(p\widehat{P})\exp(q\widehat{Q})\exp(z\widehat{Z}) = \begin{pmatrix} 1 & p & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & p & z+pq \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix},$$

so that the corresponding normal coordinates of the second kind are slightly different.

Remark 2.24. If the two elements $x, y \in \mathbf{L}(G)$ commute, then the corresponding left invariant vector fields commute, and this implies that the corresponding flows Φ^{x_l} and Φ^{y_l} commute (Corollary 1.60). In particular,

$$\exp(tx)\exp(sy) = \Phi^{y_l}_s \Phi^{x_l}_t(\mathbf{1}) = \Phi^{x_l}_t \Phi^{y_l}_s(\mathbf{1}) = \exp(sy)\exp(tx), \quad s, t \in \mathbb{R}.$$

Therefore the Trotter Formula implies that

$$\exp(x+y) = \exp x \exp y.$$

For $G = \mathrm{GL}_n(\mathbb{R})$ and $x, y \in M_n(\mathbb{R})$ with $xy = yx$, the corresponding relation is an easy consequence of the binomial formula:

$$\begin{aligned}\exp(x+y) &= \sum_{k=0}^\infty \frac{(x+y)^k}{k!} = \sum_{k=0}^\infty \frac{1}{k!}\sum_{\ell=0}^k \binom{k}{\ell} x^\ell y^{k-\ell} \\ &= \sum_{k=0}^\infty \sum_{\ell=0}^k \frac{x^\ell}{\ell!}\frac{y^{k-\ell}}{(k-\ell)!} = \Big(\sum_{p=0}^\infty \frac{x^p}{p!}\Big)\Big(\sum_{\ell=0}^\infty \frac{y^\ell}{\ell!}\Big) = \exp(x)\exp(y).\end{aligned}$$

Remark 2.25. We have seen above that the one-parameter group $\gamma_x \colon \mathbb{R} \to G$ of a Lie group G with $\gamma_x'(0) = x$ is a solution of the ordinary differential equation

$$\dot\gamma(t) = T_{\mathbf{1}}(\lambda_{\gamma(t)})x = \gamma(t) \cdot x,$$

which formally looks like a linear differential equation. For $G = \mathrm{GL}_n(\mathbb{R})$ the $\cdot$ really stands for a matrix product (cf. Example 2.22).

More generally, one frequently considers ODEs on Lie groups of the form

$$\dot\gamma = \gamma \cdot \xi \quad \text{where} \quad \xi \in C^\infty(I, \mathbf{L}(G)),$$

where $I \subseteq \mathbb{R}$ is an interval containing 0. Using similar arguments as for the familiar linear time-dependent ODEs, one can show that, for any initial value γ_0, these equations have a unique solution γ.

For $G = \mathrm{GL}_n(\mathbb{R})$, these solutions can actually be constructed by Picard iteration. For $\xi \in C([0,T], M_n(\mathbb{R}))$ we want to solve the linear initial value problem

$$\gamma(0) = \mathbf{1}, \quad \gamma'(t) = \gamma(t)\xi(t), \quad 0 \le t \le T. \tag{11}$$

Picard iteration yields a sequence of continuous curves:

$$\gamma_0(t) := \mathbf{1}, \quad \gamma_{n+1}(t) := \mathbf{1} + \int_0^t \gamma_n(\tau)\xi(\tau)\, d\tau,$$

so that

$$\gamma_n(t) = \mathbf{1} + \sum_{k=1}^{n} \int_0^t \int_0^{\tau_n} \cdots \int_0^{\tau_2} \xi(\tau_1)\xi(\tau_2)\cdots\xi(\tau_n)\ d\tau_1\, d\tau_2 \cdots d\tau_n.$$

For

$$\beta_n(t) := \int_0^t \int_0^{\tau_n} \cdots \int_0^{\tau_2} \xi(\tau_1)\xi(\tau_2)\cdots\xi(\tau_n)\ d\tau_1\, d\tau_2 \cdots d\tau_n,$$

we obtain the estimate

$$\|\beta_n(t)\| \leq \|\xi\|_\infty^n \int_0^t \int_0^{\tau_n} \cdots \int_0^{\tau_2}\ d\tau_1\, d\tau_2 \cdots d\tau_n = \|\xi\|_\infty^n \frac{t^n}{n!},$$

so that the limit $\gamma := \lim_{n\to\infty} \gamma_n = \mathbf{1} + \sum_{k=1}^\infty \beta_k$ exists uniformly on $[0, T]$. This in turn implies that γ satisfies the integral equation

$$\gamma(t) = \mathbf{1} + \int_0^t \gamma(\tau)\xi(\tau)\, d\tau.$$

Hence γ is C^1 with $\dot\gamma = \gamma \cdot \xi$. In view of the above construction of the curve γ, it is called the *product integral of* ξ.

2.4 Linear Lie groups

The following theorem is an important result on subgroups of Lie groups. Here the exponential function turns out to be an important tool to relate subgroups and Lie subalgebras.

Theorem 2.26. (von Neumann's Closed Subgroup Theorem) *Let H be a closed subgroup of the Lie group G. Then H is a submanifold of G and $m_H := m_G|_{H\times H}$ induces a Lie group structure on H such that the inclusion map $j_H\colon H \to G$ is a morphism of Lie groups for which $\mathbf{L}(j_H)\colon \mathbf{L}(H) \to \mathbf{L}(G)$ is an isomorphism of $\mathbf{L}(H)$ onto $\{x \in \mathbf{L}(G)\colon \exp(\mathbb{R}x) \subseteq H\}$.*

The preceding theorem shows in particular that very closed subgroup $G \subseteq \mathrm{GL}_n(\mathbb{K})$ is a Lie group with Lie algebra

$$\mathbf{L}(G) \cong \{x \in M_n(\mathbb{K})\colon\ \exp(\mathbb{R}x) \subseteq G\}.$$

These Lie groups are called *linear Lie groups.* Von Neumann's Theorem provides a direct way to calculate their Lie algebra $\mathbf{L}(G)$ as a Lie subalgebra of the Lie algebra $\mathfrak{gl}_n(\mathbb{K})$. Below we encounter various concrete examples of matrix groups that arise as automorphism groups of geometric structures on $\mathbb{R}^n$.

Lemma 2.27. *Let G be a Lie group and $H \subseteq G$ a subgroup which is a neighborhood of $\mathbf{1}$. Then H is open and closed, hence a Lie group, and $\mathbf{L}(H) = \mathbf{L}(G)$.*

Proof. Since the left multiplications λ_g are diffeomorphisms, the coset $gH = \lambda_g(H)$ is a neighborhood of g. For $g \in H$ the relation $gH = H$ thus shows that H is open. Then all cosets gH are open, and therefore $H = G \setminus \bigcup_{g\notin H} gH$ is closed.

For each $x \in \mathbf{L}(G)$, the one-parameter group $\gamma_x\colon \mathbb{R} \to G$ is continuous. Hence $\gamma_x^{-1}(H)$ is a non-empty open closed subset of the connected space $\mathbb{R}$, which implies that $\mathbb{R} = \gamma_x^{-1}(H)$, i.e. $\gamma_x(\mathbb{R}) \subseteq H$. This means that $\mathbf{L}(G) = \mathbf{L}(H)$. $\square$

To introduce some important classes of linear Lie groups, we fix some notation concerning matrices. We write a matrix $A = (a_{ij})_{i,j=1,\ldots,n}$ also as (a_{ij}) and define

$$A^\top := (a_{ji}), \quad \overline{A} := (\overline{a_{ij}}), \quad \text{and} \quad A^\dagger := \overline{A}^\top = (\overline{a_{ji}}).$$

Note that $A^\dagger = A^\top$ is equivalent to $\overline{A} = A$, which means that all entries of A are real.

Examples 2.28. (a) The subgroup

$$\mathrm{GL}_n(\mathbb{R})_+ := \{g \in \mathrm{GL}_n(\mathbb{R}) \colon \det g > 0\}$$

is the group of orientation preserving matrices. This is an open subgroup of $\mathrm{GL}_n(\mathbb{R})$ so that it has the same Lie algebra as $\mathrm{GL}_n(\mathbb{R})$ (Lemma 2.27).

(b) Since $\mathrm{vol}(gE) = |\det(g)| \,\mathrm{vol}(E)$ for a measurable subset $E \subseteq \mathbb{R}^n$,

$$\mathrm{VGL}_n(\mathbb{R}) := \{g \in \mathrm{GL}_n(\mathbb{R}) \colon |\det g| = 1\}$$

is the group of volume preserving matrices. That it is a subgroup follows from the multiplicativity of the determinant. From the relation

$$|\det(e^A)| = e^{\operatorname{tr} A}$$

it follows that $\exp(\mathbb{R}x) \subseteq \mathrm{VGL}_n(\mathbb{R})$ is equivalent to $\operatorname{tr} x = 0$, i.e.

$$\mathfrak{vgl}_n(\mathbb{R}) := \mathbf{L}(\mathrm{VGL}_n(\mathbb{R})) = \{x \in \mathfrak{gl}_n(\mathbb{R}) \colon \operatorname{tr} x = 0\}.$$

(c) The *special linear group*

$$\mathrm{SL}_n(\mathbb{R}) := \{g \in \mathrm{GL}_n(\mathbb{R}) \colon \det g = 1\} = \mathrm{GL}_n(\mathbb{R})_+ \cap \mathrm{VGL}_n(\mathbb{R})$$

is the group of those matrices preserving orientation and volume. Its Lie algebra is

$$\mathfrak{sl}_n(\mathbb{R}) = \mathbf{L}(\mathrm{SL}_n(\mathbb{R})) = \mathbf{L}(\mathrm{GL}_n(\mathbb{R})_+) \cap \mathbf{L}(\mathrm{VGL}_n(\mathbb{R})) = \mathfrak{vgl}_n(\mathbb{R}) = \{x \in \mathfrak{gl}_n(\mathbb{R}) \colon \operatorname{tr} x = 0.\}$$

Example 2.29. (Symmetry groups of bilinear forms)

(a) Any bilinear form β on $\mathbb{K}^n$ is of the form

$$\beta(x, y) = x^\top B y = \sum_{i,j=1}^n x_i b_{ij} y_j.$$

We say that a matrix $g \in \mathrm{GL}_n(\mathbb{K})$ preserves this form if

$$\beta(gx, gy) = \beta(x, y) \quad \text{for all} \quad x, y \in \mathbb{K}^n.$$

In view of $\beta(gx, gy) = x^\top g^\top B g y$, this is equivalent to the condition $g^\perp B g = B$, which leads us to the *general orthogonal groups*

$$\mathrm{O}_n(\mathbb{K}, B) := \{g \in \mathrm{GL}_n(\mathbb{K}) \colon g^\top B g = B\}.$$

Clearly, these are closed subgroups, because they are defined by the equation $g^\top Bg = B$. To determine their Lie algebra, we note that $e^{tx} \in \mathrm{O}_n(\mathbb{K}, B)$ for all $t \in \mathbb{R}$ leads to

$$B = e^{tx^\top} Be^{tx} = B + (x^\top B + Bx)t + \cdots ,$$

so that the derivative in 0 yields $x^\top B + Bx = 0$. If, conversely, this condition is satisfied, then the curve $\gamma(t) := e^{tx^\perp} Be^{tx}$ satisfies

$$\dot\gamma(t) = e^{tx^\top} x^\top Be^{tx} + e^{tx^\top} Bxe^{tx} = e^{tx^\top}(x^\top B + Bx)e^{tx} = 0,$$

to that γ is constant. As $\gamma(0) = B$, this means that $e^{tx} \in \mathrm{O}_n(\mathbb{K}, B)$ for every $t \in \mathbb{R}$. We thus arrive at

$$\mathfrak{o}_n(\mathbb{K}, B) := \mathbf{L}(\mathrm{O}_n(\mathbb{K}, B)) = \{x \in \mathfrak{gl}_n(\mathbb{K}) \colon x^\top B + Bx = 0\}.$$

(b) For $B = \mathbf{1}$ (the identity matrix), we obtain the *orthogonal group*

$$\mathrm{O}_n(\mathbb{K}) = \{g \in \mathrm{GL}_n(\mathbb{K}) \colon g^\top g = \mathbf{1}\} \quad \text{with} \quad \mathfrak{o}_n(\mathbb{K}) = \{x \in \mathfrak{gl}_n(\mathbb{K}) \colon x^\top + x = 0\}.$$

Intersecting with $\mathrm{SL}_n(\mathbb{K})$ leads to the *special orthogonal group*

$$\mathrm{SO}_n(\mathbb{K}) = \{g \in \mathrm{O}_n(\mathbb{K}) \colon \det g = 1\}$$

with Lie algebra

$$\mathfrak{so}_n(\mathbb{K}) = \mathfrak{o}_n(\mathbb{K}) = \{x \in \mathfrak{gl}_n(\mathbb{K}) \colon x^\top + x = 0\}.$$

Here we use that $x^\top = -x$ implies that $\operatorname{tr} x = \operatorname{tr} x^\top = -\operatorname{tr} x$, and therefore $\operatorname{tr} x = 0$.

(c) For $n = p + q$ and

$$I_{p,q} := \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix} \in M_{p+q}(\mathbb{R})$$

we obtain the *pseudo-orthogonal groups*

$$\mathrm{O}_{p,q}(\mathbb{R}) := \{g \in \mathrm{GL}_n(\mathbb{R}) \colon g^\top I_{p,q} g = I_{p,q}\},$$

where $\mathrm{O}_{n,0}(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})$. We write $\mathbb{R}^{p,q} := (\mathbb{R}^{p+q}, \beta_{p,q})$ for $\mathbb{R}^{p+q}$, endowed with the corresponding symmetric bilinear form

$$\beta(x, y) = x_1y_1 + \ldots + x_py_p - x_{p+1}y_{p+1} - \ldots - x_{p+q}y_{p+q}.$$

(d) For the skew-symmetric matrix $J := \begin{pmatrix} \mathbf{0} & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0} \end{pmatrix}$, the group

$$\mathrm{Sp}_{2n}(\mathbb{K}) := \{g \in \mathrm{GL}_{2n}(\mathbb{K}) \colon g^\top Jg = J\}$$

is called the *symplectic group*. The corresponding skew-symmetric bilinear form on $\mathbb{K}^{2n}$ is given by

$$\beta(x, y) = x^\top Jy = \sum_{i=1}^n x_iy_{n+i} - x_{n+i}y_i.$$

Example 2.30. On $\mathbb{C}^n$ one also considers hermitian forms, and the scalar product

$$\langle z, w\rangle := \sum_{j=1}^{n} z_j \overline{w_j}$$

is the most important one. Its symmetry group is the *unitary group*

$$\mathrm{U}_n(\mathbb{C}) = \{g \in \mathrm{GL}_n(\mathbb{C})\colon g^\dagger g = \mathbf{1}\} = \{g \in \mathrm{GL}_n(\mathbb{C})\colon (\forall z, w \in \mathbb{C}^n)\langle gz, gw\rangle = \langle z, w\rangle\}.$$

With similar calculations as for the real case, we obtain the Lie algebra

$$\mathfrak{u}_n(\mathbb{C}) := \mathbf{L}(\mathrm{U}_n(\mathbb{C})) = \{x \in \mathfrak{gl}_n(\mathbb{C})\colon x^\dagger + x = 0\}.$$

And for the *special unitary group*

$$\mathrm{SU}_n(\mathbb{C}) = \{g \in \mathrm{U}_n(\mathbb{C})\colon \det g = 1\} = \mathrm{U}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C})$$

we obtain

$$\mathfrak{su}_n(\mathbb{C}) := \mathbf{L}(\mathrm{SU}_n(\mathbb{C})) = \{x \in \mathfrak{gl}_n(\mathbb{C})\colon x^\dagger + x = 0, \operatorname{tr} x = 0\}.$$

Note that, although $\mathrm{SU}_n(\mathbb{C})$ and $\mathrm{U}_n(\mathbb{C})$ are groups of complex matrices, their Lie algebra is only a REAL vector space.

	Lie group G	Lie algebra $\mathfrak{g}$
general linear group	$\mathrm{GL}_n(\mathbb{K})$	$\mathfrak{gl}_n(\mathbb{K}) = M_n(\mathbb{K})$
volume preserving group	$\mathrm{VGL}_n(\mathbb{R}) :\ \lvert\det g\rvert = 1$	$\mathfrak{vgl}_n(\mathbb{R}) = \mathfrak{sl}_n(\mathbb{R}) :\ \operatorname{tr} x = 0$
special linear group	$\mathrm{SL}_n(\mathbb{K}) :\ \det g = 1$	$\mathfrak{sl}_n(\mathbb{K}) :\ \operatorname{tr} x = 0$
B -orthogonal group	$\mathrm{O}_n(\mathbb{K}, B) :\ g^\top B g = B$	$\mathfrak{o}_n(\mathbb{K}, B) :\ x^\top B + Bx = 0$
orthogonal group	$\mathrm{O}_n(\mathbb{K}) :\ g^\top g = \mathbf{1}$ $\mathrm{O}_n(\mathbb{R}) = \mathrm{O}(n)$	$\mathfrak{o}_n(\mathbb{K}) :\ x^\top + x = 0$
special orthogonal group	$\mathrm{SO}_n(\mathbb{K}) :\ g^\top g = \mathbf{1}, \det g = 1$	$\mathfrak{so}_n(\mathbb{K}) = \mathfrak{o}_n(\mathbb{K}) :\ x^\top + x = 0$
pseudo-orthogonal group	$\mathrm{O}_{p,q}(\mathbb{R}) = \mathrm{O}(p,q) :\ g^\top I_{p,q} g = I_{p,q}$	$\mathfrak{o}_{p,q}(\mathbb{R}) :\ x^\top I_{p,q} + I_{p,q} x = 0$
symplectic group	$\mathrm{Sp}_{2n}(\mathbb{R}) :\ g^\top J g = J$	$\mathfrak{sp}_{2n}(\mathbb{R}) :\ x^\top J + Jx = 0$
unitary group	$\mathrm{U}_n(\mathbb{C}) = \mathrm{U}(n) : g^\dagger g = \mathbf{1}$	$\mathfrak{u}_n(\mathbb{C}) : X^\dagger + X = 0$
special unitary group	$\mathrm{SU}_n(\mathbb{C}) = \mathrm{SU}(n) : g^\dagger g = \mathbf{1}, \det g = 1$	$\mathfrak{su}_n(\mathbb{C}) : X^\dagger + X = 0, \operatorname{tr} X = 0$

Example 2.31. Consider the group $\mathrm{SO}_3(\mathbb{R})$ of rotations of 3-space. Its Lie algebra is

$$\mathfrak{so}_3(\mathbb{R}) = \{X \in M_3(\mathbb{R})\colon X^\top = -X\}.$$

The exponential function of this group is closely related to rotations of $\mathbb{R}^3$. For the basis

$$J_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_3 := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

we have the commutator relations

$$[J_1, J_2] = J_3, \quad [J_2, J_3] = J_1 \quad \text{and} \quad [J_3, J_1] = J_2.$$

This can be written more compactly by using the *completely antisymmetric tensor* ε_{ijk} which is defined by

$$\varepsilon_{\sigma_1\sigma_2\sigma_3} = \prod_{i<j} \frac{\sigma_j - \sigma_i}{j - i}.$$

We then have

$$[J_i, J_j] = \sum_{k=1}^{3} \varepsilon_{ijk} J_k,$$

so that the structure constants of $\mathfrak{so}_3(\mathbb{R})$ with respect to the basis (J_1, J_2, J_3) given by ε_{ijk}.

The corresponding one-parameter groups are given by

$$e^{tJ_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}, \quad e^{tJ_2} = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}, \quad e^{tJ_3} = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so that e^{tJ_j} is a roation around the e_j-axis. To understand the geometry of e^{tX} for a general $X \in \mathfrak{so}_3(\mathbb{R})$, we recall from Exercise 2.9 the existence of a vector $x \in \mathbb{R}^3$ with $Xv = x \times v$ for $v \in \mathbb{R}^3$. Now $v_1 := \frac{x}{\|x\|}$ is a unit vector. Pick a unit vector $v_2 \perp v_1$ and put $v_3 := v_1 \times v_2 = Xv_2$. Then

$$Xv_1 = 0, \quad Xv_2 = \|x\| v_3 \quad \text{and} \quad Xv_3 = -\|x\| v_2.$$

This formula also shows that the operator norm of X on euclidean $\mathbb{R}^3$ equals $\|x\|$. With respect to the basis (v_1, v_2, v_3), the matrix of the linear map e^X is therefore given by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \|x\| & -\sin \|x\| \\ 0 & \sin \|x\| & \cos \|x\| \end{pmatrix}.$$

We conclude that e^{tX} is a one-parameter group of rotations around the axis $\mathbb{R}x$ where e^X rotates by the angle $\|x\|$. In particular, $e^X = \mathbf{1}$ for $\|x\| = 2\pi$.

As every element $g \in \mathrm{SO}_3(\mathbb{R})$ is a rotation (a consequence of the normal form of (3×3)-orthogonal matrices or the simple fact that 1 must be an eigenvalue of g), it follows in particular that the exponential function

$$\exp\colon \mathfrak{so}_3(\mathbb{R}) \to \mathrm{SO}_3(\mathbb{R})$$

is surjective. We actually find for each $g \in \mathrm{SO}_3(\mathbb{R})$ an $X \in \mathfrak{so}_3(\mathbb{R})$ with $\|X\| \leq \pi$ and $e^X = g$.

2.5 On the topology of matrix groups

In this subsection we take a brief look at the topological properties of matrix groups. Since compact groups behave much better than arbitrary topological groups, we first observe that real orthogonal and the complex unitary groups are compact. The compactness of a group has profound implications for its representation theory, which is mostly due to the existence of a biinvariant probability measure. In the theory of elementary particles the compactness of the corresponding symmetry group is responsible for the discreteness of the quantum numbers classifying these particles.

2.5.1 Compact matrix groups

Lemma 2.32. *The groups*

$$\mathrm{U}_n(\mathbb{C}), \quad \mathrm{SU}_n(\mathbb{C}), \quad \mathrm{O}_n(\mathbb{R}) \quad \textit{and} \quad \mathrm{SO}_n(\mathbb{R})$$

are compact.

Proof. Since all these groups are subsets of $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$, by the Heine–Borel Theorem we only have to show that they are closed and bounded.

Boundedness: In view of

$$\mathrm{SO}_n(\mathbb{R}) \subseteq \mathrm{O}_n(\mathbb{R}) \subseteq \mathrm{U}_n(\mathbb{C}) \quad \text{and} \quad \mathrm{SU}_n(\mathbb{C}) \subseteq \mathrm{U}_n(\mathbb{C}),$$

it suffices to see that $\mathrm{U}_n(\mathbb{C})$ is bounded. Let $g_1, \ldots, g_n$ denote the rows of the matrix $g \in M_n(\mathbb{C})$. Then $g^\dagger = g^{-1}$ is equivalent to $gg^\dagger = \mathbf{1}$, which means that $g_1, \ldots, g_n$ form an orthonormal basis for $\mathbb{C}^n$ with respect to the scalar product $\langle z, w\rangle = \sum_{j=1}^n z_j \overline{w_j}$ which induces the norm $\|z\| = \sqrt{\langle z, z\rangle}$. Therefore $g \in \mathrm{U}_n(\mathbb{C})$ implies $\|g_j\| = 1$ for each j, so that $\mathrm{U}_n(\mathbb{C})$ is bounded.

Closedness: The functions

$$f, h \colon M_n(\mathbb{K}) \to M_n(\mathbb{K}), \quad f(A) := AA^\dagger - \mathbf{1} \quad \text{and} \quad h(A) := AA^\top - \mathbf{1}$$

are continuous. Therefore the groups

$$\mathrm{U}_n(\mathbb{K}) := f^{-1}(\mathbf{0}) \quad \text{and} \quad \mathrm{O}_n(\mathbb{K}) := h^{-1}(\mathbf{0})$$

are closed. Likewise $\mathrm{SL}_n(\mathbb{K}) = \det^{-1}(\mathbf{1})$ is closed, and therefore the groups $\mathrm{SU}_n(\mathbb{C})$ and $\mathrm{SO}_n(\mathbb{R})$ are also closed because they are intersections of closed subsets. □

Proposition 2.33. (a) *The exponential function* $\exp\colon \mathfrak{u}_n(\mathbb{C}) \to \mathrm{U}_n(\mathbb{C})$ *is surjective. In particular,* $\mathrm{U}_n(\mathbb{C})$ *is arcwise connected.*

(b) *The group* $\mathrm{O}_n(\mathbb{R})$ *has the two arc components*

$$\mathrm{O}_n(\mathbb{R})_\pm := \{g \in \mathrm{O}_n(\mathbb{R}) \colon \det g = \pm 1\}$$

and the exponential function of $\mathrm{SO}_n(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})_+$ *is surjective.*

Proof. (a) First we consider $\mathrm{U}_n(\mathbb{C})$. To see that this group is arcwise connected, let $u \in \mathrm{U}_n(\mathbb{C})$. Then there exists an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of u. Let $\lambda_1, \ldots, \lambda_n$ denote the corresponding eigenvalues. Then the unitarity of u implies that $|\lambda_j| = 1$, and we therefore find $\theta_j \in \mathbb{R}$ with $\lambda_j = e^{\theta_j i}$. Define $D \in M_n(\mathbb{C})$ by $Dv_j = i\theta_j v_j$. Since the v_j are orthonormal, $D^\dagger = -D$ (Exercise 2.17). Now $\gamma(t) := e^{tD}$ satisfies $\gamma(1)v_j = gv_j$ for every j, and therefore $g = \gamma(1) = e^D$.

(b) For $g \in \mathrm{O}_n(\mathbb{R})$ we have $gg^\top = \mathbf{1}$ and therefore $1 = \det(gg^\top) = (\det g)^2$. This shows that

$$\mathrm{O}_n(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})_+ \dot\cup \mathrm{O}_n(\mathbb{R})_- \quad \text{with} \quad \mathrm{O}_n(\mathbb{R})_+ = \mathrm{SO}_n(\mathbb{R}),$$

and both sets are closed in $\mathrm{O}_n(\mathbb{R})$ because det is continuous. Therefore $\mathrm{O}_n(\mathbb{R})$ is not connected and hence not arcwise connected. Suppose we knew that $\mathrm{SO}_n(\mathbb{R})$ is arcwise connected and $x, y \in \mathrm{O}_n(\mathbb{R})_-$. Then $\mathbf{1}, x^{-1}y \in \mathrm{SO}_n(\mathbb{R})$ can be connected by an arc $\gamma\colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$, and then $t \mapsto x\gamma(t)$ defines an arc $[0,1] \to \mathrm{O}_n(\mathbb{R})_-$ connecting x to y. So it remains to show that the exponential function of $\mathrm{SO}_n(\mathbb{R})$ is surjective.

From Linear Algebra we know that every orthogonal matrix is conjugate (under an orthogonal matrix) to one in the following normal form

$$\begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \\ & & \ddots \\ & & & \cos\alpha_m & -\sin\alpha_m \\ & & & \sin\alpha_m & \cos\alpha_m \\ & & & & & -1 \\ & & & & & & \ddots \\ & & & & & & & -1 \\ & & & & & & & & 1 \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix}$$

for real numbers $0 < \alpha_j < \pi$. Let $g \in \mathrm{SO}_n(\mathbb{R})$. In the normal form of g, the determinant of each 2×2-block is 1, so that the determinant is the product of all -1-eigenvalues. Hence their number is even, and we can write each consecutive pair as a block

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{pmatrix}.$$

This shows that with respect to some orthonormal basis of $\mathbb{R}^n$, the linear map defined by g has a matrix of the form

$$g = \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \\ & & \ddots \\ & & & \cos\alpha_m & -\sin\alpha_m \\ & & & \sin\alpha_m & \cos\alpha_m \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}.$$

Now we obtain a smooth one-parameter group $\gamma\colon \mathbb{R} \to \mathrm{SO}_n(\mathbb{R})$ with $\gamma(0) = \mathbf{1}$ and $\gamma(1) = g$

by

$$\gamma(t) := \begin{pmatrix} \cos t\alpha_1 & -\sin t\alpha_1 & & & & & & \\ \sin t\alpha_1 & \cos t\alpha_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \cos t\alpha_m & -\sin t\alpha_m & & & \\ & & & \sin t\alpha_m & \cos t\alpha_m & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

□

2.5.2 Non-compact matrix groups

To obtain some information on the topology of non-compact matrix groups as well, we now study the polar decomposition of $\mathrm{GL}_n(\mathbb{R})$ and show that it is inherited by a large class of subgroups. It is an important tool to understand the topology of non-compact Lie groups.

Definition 2.34. We write $\mathrm{Herm}_n(\mathbb{K}) := \{A \in M_n(\mathbb{K}) \colon A^\dagger = A\}$ for the set of *hermitian matrices*. For $\mathbb{K} = \mathbb{C}$ this is not a vector subspace of $M_n(\mathbb{K})$, but it is always a real subspace. A matrix $A \in \mathrm{Herm}_n(\mathbb{K})$ is called *positive definite* if for each $0 \neq z \in \mathbb{K}^n$ we have $\langle Az, z\rangle > 0$, where

$$\langle z, w\rangle := \sum_{j=1}^{n} z_j \overline{w_j}$$

is the natural scalar product on $\mathbb{K}^n$.

Lemma 2.35. *A positive semidefinite matrix A has a unique positive semidefinite square root B, i.e. a matrix B with $B^2 = A$.*

If A is positive definite, then B is also positive definite. In this case there exists a unique hermitian matrix X with $e^X = A$.

In view of the uniqueness of B, it makes sense to write $B := \sqrt{A}$ and $X = \log A$ if A is positive definite.

Proof. We know from Linear Algebra that for each hermitian matrix A there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{K}^n$ consisting of eigenvectors of A, and that all the corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ are real. From that it is obvious that A is positive semidefinite if and only if $\lambda_j \geq 0$ holds for each j.

Existence of a square root: We define B with respect to the basis $(v_1, \ldots, v_n)$ by $Bv_j = \sqrt{\lambda_j} v_j$. Then $B^2 = A$ is obvious and since all λ_j are real and the v_j are orthonormal, B is positive definite because

$$\Big\langle B\Big(\sum_i \mu_i v_i\Big), \sum_j \mu_j v_j\Big\rangle = \sum_{i,j} \mu_i \overline{\mu_j} \langle Bv_i, v_j\rangle = \sum_{j=1}^{n} |\mu_j|^2 \sqrt{\lambda_j} > 0 \quad \text{for} \quad \sum_j \mu_j v_j \neq 0.$$

Uniqueness of a square root: Assume that C is positive definite with $C^2 = A$. Pick an orthonormal basis $w_1, \ldots, w_m$ of C-eigenvectors, so that $Cw_j = \mu_j w_j$ with positive numbers

$\mu_j > 0$. Then $Aw_j = C^2 w_j = \mu_j^2 w_j$ shows that, for $\lambda_j := \mu_j^2$, the matrix C acts on the λ_j-eigenspace of A by multiplication with $\sqrt{\lambda_j} = \mu_j$. This implies $B = C$.

If A is positive definite, then all its eigenvalues are positive, and a similar argument with $\mu_j := \log \lambda_j$ implies the existence of X as well as its uniqueness. $\square$

We have seen already that the unitary group is compact and that its Lie algebra consists of skew-hermitian operators. On the other hand, every matrix $X \in M_n(\mathbb{K})$ has a unique decomposition

$$X = \frac{1}{2}(X + X^\dagger) + \frac{1}{2}(X - X^\dagger)$$

into a hermitian and a skew-hermitian part and now we want to derive a similar multiplicative decomposition of certain matrix groups. Since this does not work without additional hyptheses, we introduce the concept of a real algebraic group.

Definition 2.36. We call a subgroup $G \subseteq \mathrm{GL}_n(\mathbb{R})$ *real algebraic* if there exists a family $(p_j)_{j \in J}$ of real polynomials

$$p_j(x) = p_j(x_{11}, x_{12}, \ldots, x_{nn}) \in \mathbb{R}[x_{11}, \ldots, x_{nn}]$$

in the entries of the matrix $x \in M_n(\mathbb{R})$ such that

$$G = \{x \in \mathrm{GL}_n(\mathbb{R}) \colon (\forall j \in J)\ p_j(x) = 0\}.$$

A subgroup $G \subseteq \mathrm{GL}_n(\mathbb{C}) \subseteq \mathrm{GL}_{2n}(\mathbb{R})$ is called *real algebraic* if it is a real algebraic subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$ (here we use the inclusion $M_n(\mathbb{C}) \hookrightarrow M_{2n}(\mathbb{R})$).

Proposition 2.37. (Polar decomposition for matrix groups) *Let $G \subseteq \mathrm{GL}_n(\mathbb{K})$ be a real algebraic subgroup invariant under $\dagger$, i.e. $G = G^\dagger$. Then $K := G \cap \mathrm{U}_n(\mathbb{K})$ is a compact group and we put $\mathfrak{p} := \mathbf{L}(G) \cap \mathrm{Herm}_n(\mathbb{K})$. Then the map*

$$m \colon K \times \mathfrak{p} \to G, \quad (k, x) \mapsto ke^x$$

is a diffeomorphism.

Proof. (Sketch) The smoothness of the map m is clear.
m **is surjective:** Let $g \in G$. For $0 \neq v \in \mathbb{K}^n$ we then have

$$0 < \langle gv, gv \rangle = \langle g^\dagger g v, v \rangle,$$

showing that $g^\dagger g$ is positive definite. Let $x := \frac{1}{2}\log(g^\dagger g)$ and define $u := ge^{-x}$. Then

$$uu^\dagger = ge^{-x}e^{-x}g^\dagger = ge^{-2x}g^\dagger = g(g^\dagger g)^{-1}g^\dagger = gg^{-1}(g^\dagger)^{-1}g^\dagger = \mathbf{1}$$

implies that $u \in \mathrm{U}_n(\mathbb{K})$, and it is clear that $ue^x = g$. From the assumption that G is real algebraic, on can derive that $x \in \mathfrak{p}$, so that m is surjective.
m **is injective:** If $g = ke^x = he^y$, then $g^\dagger g = e^{2x}$, so that $x = \frac{1}{2}\log(g^\dagger g) = y$ is the unique hermitian logarithm of the positive definite matrix $g^\dagger g$. This implies that $k = ge^{-x} = ge^{-y} = h$.

It is easy to see that the invariance of G under $\dagger$ implies the same for $\mathbf{L}(G)$, so that

$$\mathbf{L}(G) = \big(\mathbf{L}(G) \cap \mathfrak{u}_n(\mathbb{K})\big) \oplus \mathfrak{p} = \mathbf{L}(K) \oplus \mathfrak{p}.$$

Therefore

$$\dim G = \dim \mathbf{L}(G) = \dim \mathbf{L}(K) + \dim \mathfrak{p} = \dim(K \times \mathfrak{p}).$$

Since m is bijective, in view of the Inverse Function Theorem, it suffices to show that all differentials $T_{(u,x)}(m)$ are injective (hence bijective for dimension reasons). This can be done by showing that the exponential map is regular on $\mathrm{Herm}_n(\mathbb{K})$ (see [HN11] for details). $\square$

Corollary 2.38. *The group $\mathrm{GL}_n(\mathbb{C})$ is arcwise connected and the group $\mathrm{GL}_n(\mathbb{R})$ has two arc-components given by*

$$\mathrm{GL}_n(\mathbb{R})_\pm := \{g \in \mathrm{GL}_n(\mathbb{R})\colon\ \pm \det g > 0\}.$$

Proof. If $X = A \times B$ is a product space, then the arc-components of X are the sets of the form $C \times D$, where $C \subseteq A$ and $D \subseteq B$ are arc-components (easy Exercise!). The polar decomposition of $\mathrm{GL}_n(\mathbb{K})$ yields a homeomorphism

$$\mathrm{GL}_n(\mathbb{K}) \cong \mathrm{U}_n(\mathbb{K}) \times \mathrm{Herm}_n(\mathbb{K}).$$

The vector space $\mathrm{Herm}_n(\mathbb{K})$ is arcwise connected. Therefore the arc-components of $\mathrm{GL}_n(\mathbb{K})$ are in one-to-one correspondence with those of $\mathrm{U}_n(\mathbb{K})$ which have been determined in Proposition 2.33. $\square$

Example 2.39. Proposition 2.37 in particular applies to the following groups:

(a) $G = \mathrm{SL}_n(\mathbb{R})$ is $p^{-1}(0)$ for the polynomial $p(x) = \det x - 1$, and we obtain

$$\mathrm{SL}_n(\mathbb{R}) = K \exp \mathfrak{p} \cong K \times \mathfrak{p}$$

with

$$K = \mathrm{SO}_n(\mathbb{R}) \quad \text{and} \quad \mathfrak{p} = \{x \in \mathrm{Sym}_n(\mathbb{R})\colon\ \mathrm{tr}\, x = 0\}.$$

For $\mathrm{SL}_2(\mathbb{R})$, we obtain in particular a homeomorphism

$$\mathrm{SL}_2(\mathbb{R}) \cong \mathrm{SO}_2(\mathbb{R}) \times \mathbb{R}^2 \cong \mathbb{S}^1 \times \mathbb{R}^2.$$

(b) $G = \mathrm{O}_{p,q} := \mathrm{O}_{p,q}(\mathbb{R})$ is defined by the condition $g^\top I_{p,q} g = I_{p,q}$. These are n^2 polynomial equations, one for each entry of the matrix. Moreover, $g \in \mathrm{O}_{p,q}$ implies

$$I_{p,q} = I_{p,q}^{-1} = (g^\top I_{p,q} g)^{-1} = g^{-1} I_{p,q} (g^\top)^{-1}$$

and hence $g I_{p,q} g^\top = I_{p,q}$, i.e. $g^\top \in \mathrm{O}_{p,q}$. Therefore $\mathrm{O}_{p,q}^\top = \mathrm{O}_{p,q}$, and all the assumptions of Proposition 2.37 are satisfied. In this case,

$$K = \mathrm{O}_{p,q} \cap \mathrm{O}_n \cong \mathrm{O}_p \times \mathrm{O}_q,$$

(Exercise 2.10) and we obtain a diffeomorphism

$$\mathrm{O}_{p,q} \cong \mathrm{O}_p \times \mathrm{O}_q \times (\mathfrak{o}_{p,q} \cap \mathrm{Sym}_n(\mathbb{R})).$$

In particular, we see that for $p, q > 0$ the group $\mathrm{O}_{p,q}$ has four arc-components because O_p and O_q have two arc-components (Proposition 2.33).

For the subgroup $\mathrm{SO}_{p,q}$ we have one additional polynomial equation, so that it is also algebraic. Here we have

$$\begin{aligned} K_S := K \cap \mathrm{SO}_{p,q} &\cong \{(a,b) \in \mathrm{O}_p \times \mathrm{O}_q \colon \det(a)\det(b) = 1\} \\ &\cong \big(\mathrm{SO}_p \times \mathrm{SO}_q\big)\dot{\cup}\big(\mathrm{O}_{p,-} \times \mathrm{O}_{q,-}\big), \end{aligned}$$

so that $\mathrm{SO}_{p,q}$ has two arc-components if $p, q > 0$ (cf. the discussion of the Lorentz group in Subsection 2.8.3.

(c) We can also apply Proposition 2.37 to the subgroup $\mathrm{SL}_n(\mathbb{C}) \subseteq \mathrm{GL}_n(\mathbb{C})$ because the equation $\det g - 1 = 0$ in the complex matrix entries can be viewed as a pair of real polynomial equations in the real and imaginary parts of the matrix entries. We have

$$K = \mathrm{SL}_n(\mathbb{C}) \cap \mathrm{U}_n(\mathbb{C}) = \mathrm{SU}_n(\mathbb{C}) \quad \text{and} \quad \mathfrak{p} = \mathfrak{sl}_n(\mathbb{C}) \cap \mathrm{Herm}_n(\mathbb{C}).$$

G	$K = G \cap \mathrm{U}_n(\mathbb{K})$	$\mathfrak{p} = \mathbf{L}(G) \cap \mathrm{Herm}_n(\mathbb{K})$	$\pi_0(G) = G/G_0 = \pi_0(K)$
$\mathrm{GL}_n(\mathbb{R})$	$\mathrm{O}_n(\mathbb{R})$	$\mathrm{Sym}_n(\mathbb{R})$	$\mathbb{Z}/2$
$\mathrm{SL}_n(\mathbb{R})$	$\mathrm{SO}_n(\mathbb{R})$	$X^\top = X, \operatorname{tr} X = 0$	$\mathbf{1}$
$\mathrm{GL}_n(\mathbb{C})$	$\mathrm{U}_n(\mathbb{C})$	$\mathrm{Herm}_n(\mathbb{C})$	$\mathbf{1}$
$\mathrm{SL}_n(\mathbb{C})$	$\mathrm{SU}_n(\mathbb{C})$	$X^\dagger = X, \operatorname{tr} X = 0$	$\mathbf{1}$
$\mathrm{O}_{p,q}(\mathbb{R})$	$\mathrm{O}_p(\mathbb{R}) \times \mathrm{O}_q(\mathbb{R})$		$\mathbb{Z}/2 \times \mathbb{Z}/2$ for $p, q > 0$
$\mathrm{SO}_{p,q}(\mathbb{R})$	$S(\mathrm{O}_p(\mathbb{R}) \times \mathrm{O}_q(\mathbb{R}))$		$\mathbb{Z}/2$ for $p, q > 0$
$\mathrm{SO}_{1,n}(\mathbb{R})$	$\mathrm{O}_n(\mathbb{R})$		$\mathbb{Z}/2$
$\mathrm{Sp}_{2n}(\mathbb{R})$	$\mathrm{U}_n(\mathbb{C})$	$\mathrm{Sym}_n(\mathbb{C})$	$\mathbf{1}$

2.6 Integrating homomorphisms of Lie algebras

In Proposition 2.14 we have seen that every homomorphism of Lie groups $\varphi\colon G \to H$ defines by its derivative in the identity $\mathbf{L}(\varphi) = T_{\mathbf{1}}(\varphi)\colon \mathbf{L}(G) \to \mathbf{L}(H)$ a homomorphism of Lie algebras. In this section we briefly discuss the question to which extent φ is determined by $\mathbf{L}(\varphi)$ and when there exists for a given homomorphism $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$ of Lie algebras a group homomorphism $\varphi\colon G \to H$ with $\mathbf{L}(\varphi) = \psi$.

Proposition 2.40. *For any smooth homomorphism $\varphi\colon G \to H$ of Lie groups, we have*

$$\exp_H \circ \mathbf{L}(\varphi) = \varphi \circ \exp_G, \tag{12}$$

i.e. the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "\varphi"] & H \\ \mathbf{L}(G) \arrow[u, "\exp_G"'] \arrow[r, "\mathbf{L}(\varphi)"] & \mathbf{L}(H). \arrow[u, "\exp_H"'] \end{tikzcd}$$

Proof. For $x \in \mathbf{L}(G)$ we consider the smooth homomorphism

$$\gamma_x \in \mathrm{Hom}(\mathbb{R}, G), \quad \gamma_x(t) = \exp_G(tx).$$

Then $\varphi \circ \gamma_x$ is a smooth one-parameter group of H with infinitesimal generator $(\varphi \circ \gamma_x)'(0) = \mathbf{L}(\varphi)\gamma_x'(0) = \mathbf{L}(\varphi)x$. We conclude that

$$\varphi(\gamma_x(t)) = \gamma_{\mathbf{L}(\varphi)x}(t), \quad t \in \mathbb{R},$$

and for $t = 1$, this proves the lemma. $\square$

Lemma 2.41. *The subgroup $\langle \exp_G(\mathbf{L}(G)) \rangle$ of G generated by $\exp_G(\mathbf{L}(G))$ coincides with the identity component G_0 of G, i.e. the connected component containing $\mathbf{1}$.*

Proof. Since $\exp_G$ is a local diffeomorphism in 0, $\exp_G(\mathbf{L}(G))$ is a neighborhood of $\mathbf{1}$, so that the subgroup $H := \langle \exp_G(\mathbf{L}(G)) \rangle$ generated by the exponential image is a $\mathbf{1}$-neighborhood. According to Lemma 2.27, H is open and closed. Since G_0 is connected and has a non-empty intersection with H, it must be contained in H.

On the other hand, $\exp_G$ is continuous, so that it maps the connected space $\mathbf{L}(G)$ into the identity component G_0 of G, which leads to $H \subseteq G_0$, and hence to equality. $\square$

Proposition 2.42. *For two smooth morphisms $\varphi_1, \varphi_2 \colon G \to H$ of Lie groups we have $\mathbf{L}(\varphi_1) = \mathbf{L}(\varphi_2)$ if and only if φ_1 and φ_2 coincide on the identity component G_0 of G.*

Proof. If $\varphi_1|_{G_0} = \varphi_2|_{G_0}$, then we clearly have $\mathbf{L}(\varphi_1) = T_{\mathbf{1}}(\varphi_1) = T_{\mathbf{1}}(\varphi_2) = \mathbf{L}(\varphi_2)$.

If, conversely, $\mathbf{L}(\varphi_1) = \mathbf{L}(\varphi_2)$, then Proposition 2.40 implies that $\varphi_1(\exp x) = \varphi_2(\exp x)$ for every $x \in \mathbf{L}(G)$, so that the assertion follows from Lemma 2.41. $\square$

Remark 2.43. It is easy to see that Proposition 2.42 is optimal. If the Lie group G is not connected, then its identity component is a proper normal subgroup and we may consider $\pi_0(G) := G/G_0$ as a discrete group. Any discrete group is a 0-dimensional Lie group. Now the trivial homomorphism $\varphi_1 \colon G \to G/G_0$ and the quotient homomorphism $\varphi_2 \colon G \to G/G_0, g \mapsto gG_0$ are different but coincide on G_0.

A more concrete example is obtained from the homomorphism

$$\det \colon \mathrm{O}_n(\mathbb{R}) \to \mathbb{R}^\times$$

which is non-trivial but trivial on the identity component $\mathrm{SO}_n(\mathbb{R})$ (Proposition 2.33).

The preceding proposition shows that the problem to construct a Lie group homomorphism $\varphi \colon G \to H$ from a homomorphism $\psi \colon \mathbf{L}(G) \to \mathbf{L}(H)$ of Lie algebras only makes sense if G is connected. So assume that G is connected. By Lemma 2.41, every element $g \in G$ can be written as a product

$$g = \exp x_1 \cdots \exp x_n,$$

and, whenever φ exists, it must satisfy

$$\varphi(g) = \exp(\psi x_1) \cdots \exp(\psi x_n) \tag{13}$$

(Proposition 2.40). However, we cannot use this relation to define φ because the representation of g as a product of exponentials is highly non-unique. For φ to exist, ψ has to satisfy the condition

$$(\forall n \in \mathbb{N})(\forall x_1, \ldots, x_n \in \mathbf{L}(G))\ \exp x_1 \cdots \exp x_n = \mathbf{1} \quad \Rightarrow \quad \exp(\psi x_1) \cdots \exp(\psi x_n) = \mathbf{1}. \tag{14}$$

If, conversely, (14) is satisfied, then (13) yields a well-defined smooth homomorphism $\varphi \colon G \to H$ (Exercise). Therefore (14) is necessary and sufficient for φ to exist, but this condition is impossible to verify in practise.

The main idea to turn (14) into a verifiable condition is to observe that any relation of the form $\exp x_1 \cdots \exp x_n = \mathbf{1}$ defines a closed piecewise smooth path $\gamma \colon [0,1] \to G$ by

$$\gamma(t) := \exp x_1 \cdots \exp x_{k-1} \exp(nt - (k-1))x_k \quad \text{for} \quad \frac{k-1}{n} \leq t \leq \frac{k}{n}.$$

Now let $\gamma \colon [0,1] \to G$ be any piecewise smooth path and

$$\xi(t) := \gamma(t)^{-1}\dot{\gamma}(t) \in \mathbf{L}(G)$$

be its *logarithmic derivative* (which is, strictly speaking, only defined on each subinterval on which γ is differentiable). Then the initial value problem

$$\eta(0) = \mathbf{1} \quad \text{and} \quad \dot{\eta} = \eta \cdot (\psi \circ \xi)$$

has a unique piecewise smooth solution $\eta_\gamma \colon [0,1] \to H$ (Remark 2.25). If φ exists, then $\eta_\gamma = \varphi \circ \gamma$ follows from

$$(\varphi \circ \gamma)' = T(\varphi)\gamma' = (\varphi \circ \gamma) \cdot \mathbf{L}(\varphi)\xi = (\varphi \circ \gamma) \cdot (\psi \circ \xi).$$

Here we have used that

$$T(\varphi)(g \cdot x) = \varphi(g)\,\mathbf{L}(\varphi)x \quad \text{for} \quad g \in G, x \in TG$$

(Exercise). We thus arrive at the necessary condition

$$\gamma(1) = \mathbf{1} \quad \Rightarrow \quad \eta_\gamma(1) = \mathbf{1}.$$

This looks even worse than (14), because there are even more closed piecewise smooth paths than exponential products representing the identity. However, the value $\eta_\gamma(1)$ does not change if γ is deformed with fixed endpoints. This leads us to the concept of homotopic paths.

Definition 2.44. Let X be a topological space. We call two continuous paths $\alpha_0, \alpha_1 \colon [0,1] \to X$ starting in x_0 and ending in x_1 *homotopic*, written $\alpha_0 \sim \alpha_1$, if there exists a continuous map

$$H \colon I \times I \to X \quad \text{with} \quad H_0 = \alpha_0, \quad H_1 = \alpha_1$$

(for $H_t(s) := H(t,s)$) and

$$(\forall t \in I) \quad H(t,0) = x_0,\ H(t,1) = x_1.$$

It is easy to show that $\sim$ is an equivalence relation, called *homotopy*. The homotopy class of α is denoted by $[\alpha]$.

For $\alpha(1) = \beta(0)$ we define the concatenation product $\alpha * \beta$ as

$$(\alpha * \beta)(t) := \begin{cases} \alpha(2t) & \text{for } 0 \leq t \leq \frac{1}{2} \\ \beta(2t-1) & \text{for } \frac{1}{2} \leq t \leq 1. \end{cases}$$

It turns out that $[\alpha]*[\beta] := [\alpha*\beta]$ is a well-defined product on the set of homotopy classes and that, for any $x_0 \in X$, the set

$$\pi_1(X, x_0) := \{[\alpha] \colon \alpha \in C([0,1], X), \alpha(0) = \alpha(1) = x_0\}$$

is a group with respect to $*$. Here $[\gamma]^{-1}$ is represented by $t \mapsto \gamma(1-t)$ and the identity element is the constant path. The group $\pi_1(X, x_0)$ is called the *fundamental group of X with respect to x_0*. An arcwise connected space X is called *simply connected* if $\pi_1(X, x_0)$ vanishes for an $x_0 \in X$.

After this interlude on homotopy classes, we can formulate the integrability condition for Lie algebra homomorphisms.

Theorem 2.45. *Let G and H be Lie groups and $\psi \colon \mathbf{L}(G) \to \mathbf{L}(H)$ be a homomorphism of Lie algebras. Suppose that G is connected. Then we obtain a well-defined homomorphism*

$$\mathrm{per}_\psi \colon \pi_1(G, \mathbf{1}) \to H, \quad [\gamma] \mapsto \eta_\gamma(1).$$

A smooth homomorphism $\varphi \colon G \to H$ with $\mathbf{L}(\varphi) = \psi$ exists if and only if per_ψ is trivial. In particular, φ always exists if G is simply connected.

Example 2.46. (a) We identify $\mathbf{L}(\mathbb{T})$ with $\mathbb{R}$, so that $\exp_{\mathbb{T}}(x) = e^{ix}$. Then a linear map $\psi \colon \mathbf{L}(\mathbb{T}) \to \mathbf{L}(\mathbb{T})$ is given by multiplication with a real number λ. To see for which λ we have a morphism $\varphi \colon \mathbb{T} \to \mathbb{T}$ of Lie groups with $\mathbf{L}(\varphi) = \psi$, we note that $\varphi(e^{ix}) = e^{i\lambda x}$ is only well-defined if $\lambda 2\pi\mathbb{Z} \subseteq 2\pi\mathbb{Z}$, i.e. if $\lambda = n \in \mathbb{Z}$. Then $\varphi(z) = z^n$ is the corresponding group homomorphism.

(b) For the determinant function

$$\det \colon \mathrm{U}_n(\mathbb{C}) \to \mathbb{T}$$

the relation $\det(e^x) = e^{\operatorname{tr} x} = e^{i(-i \operatorname{tr} x)}$ shows that $\mathbf{L}(\det) = -i \operatorname{tr}$. This a homomorphism of Lie algebras $\mathfrak{u}_n(\mathbb{R}) \to \mathbb{R}$, and since $\mathbb{R}$ is abelian, this simply means that $\operatorname{tr}([x,y]) = 0$ for $x, y \in \mathfrak{u}_n(\mathbb{C})$.

We thus obtain for each $\lambda \in \mathbb{R}$ a homomorphism $\psi := -i\lambda \operatorname{tr} \colon \mathfrak{u}_n(\mathbb{C}) \to \mathbb{R}$ and can ask under which conditions there exists a homomorphism $\varphi \colon \mathrm{U}_n(\mathbb{C}) \to \mathbb{T}$ with $\mathbf{L}(\varphi) = \psi$. Then it would make sense to write $\varphi = \det^\lambda$. This is clearly the case for $\lambda \in \mathbb{Z}$. That this condition is actually necessary follows from the fact that, for $x := iE_{11} \in \mathfrak{u}_n(\mathbb{C})$ with $\exp(2\pi x) = \mathbf{1}$, we have $e^{i\psi(2\pi x)} = e^{\lambda \operatorname{tr}(2\pi x)} = e^{\lambda 2\pi i} = 1$ only if $\lambda \in \mathbb{Z}$. This implies in particular, that the group $\mathrm{U}_n(\mathbb{C})$ is not simply connected.

Remark 2.47. (a) Suppose that the topological space X is contractible, i.e. there exists a continuous map $H \colon I \times X \to X$ and $x_0 \in X$ with $H(0,x) = x$ and $H(1,x) = x_0$ for $x \in X$. Then $\pi_1(X, x_0) = \{[x_0]\}$ is trivial (Exercise).

(b) $\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$.

(c) $\pi_1(\mathbb{R}^n, 0) = \{0\}$ because $\mathbb{R}^n$ is contractible.

More generally, if the open subset $\Omega \subseteq \mathbb{R}^n$ is starlike with respect to x_0, then $H(t,x) := x + t(x - x_0)$ yields a contraction to x_0, and we conclude that $\pi_1(\Omega, x_0) = \{[x_0]\}$.

(d) If $G \subseteq \mathrm{GL}_n(\mathbb{K})$ is a linear Lie group with a polar decomposition, i.e. for $K := G \cap \mathrm{U}_n(\mathbb{K})$ and $\mathfrak{p} := \mathbf{L}(G) \cap \mathrm{Herm}_n(\mathbb{K})$, the polar map $p\colon K \times \mathfrak{p} \to G, (k,x) \mapsto ke^x$ is a homeomorphism, then the inclusion $K \to G$ induces an isomorphism

$$\pi_1(K, \mathbf{1}) \to \pi_1(G, \mathbf{1})$$

because the vector space $\mathfrak{p}$ is contractible.

(e) $\pi_1(\mathbb{S}^1) \cong \pi_1(\mathbb{C}^\times) \cong \mathbb{Z}$ follows from the classification of homotopy classes of loops in the punctured plane by their winding number with respect to the origin.

(f) The group

$$\mathrm{SU}_2(\mathbb{C}) = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C})\colon |a|^2 + |b|^2 = 1 \right\}$$

is homeomorphic to the 3-sphere

$$\{(a,b) \in \mathbb{C}^2\colon \|(a,b)\| = 1\} \cong \mathbb{S}^3$$

which is simply connected (Exercise 2.20). One can show that the sphere $\mathbb{S}^n$ carries a Lie group structure if and only if $n = 0, 1, 3$.

(g) With some more advanced tools from homotopy theory, one can show that the groups $\mathrm{SU}_n(\mathbb{C})$ are always simply connected. However, this is never the case for the groups $\mathrm{U}_n(\mathbb{C})$.

To see this, consider the group homomorphism

$$\gamma\colon \mathbb{T} \to \mathrm{U}_n(\mathbb{C}), \quad z \mapsto \mathrm{diag}(z, 1, \ldots, 1)$$

and note that $\det \circ \gamma = \mathrm{id}_{\mathbb{T}}$. From that one easily derives that the multiplication map

$$\mu\colon\ \mathrm{SU}_n(\mathbb{C}) \times \mathbb{T} \to \mathrm{U}_n(\mathbb{C}), \quad (g,z) \mapsto g\gamma(z)$$

is a homeomorphism, so that

$$\pi_1(\mathrm{U}_n(\mathbb{C})) \cong \pi(\mathrm{SU}_n(\mathbb{C})) \times \pi_1(\mathbb{T}) \cong \pi_1(\mathbb{T}) \cong \mathbb{Z}.$$

G	$K = G \cap \mathrm{U}_n(\mathbb{K})$	$\pi_1(G) = \pi_1(K)$
$\mathrm{GL}_n(\mathbb{R})$	$\mathrm{O}_n(\mathbb{R})$	$\begin{cases} \mathbb{Z}/2 & \text{for } n > 2 \\ \mathbb{Z} & \text{for } n = 2 \\ \mathbf{1} & \text{for } n = 1. \end{cases}$
$\mathrm{SL}_n(\mathbb{R})$	$\mathrm{SO}_n(\mathbb{R})$	$\mathbb{Z}/2$ for $n > 2$
$\mathrm{SL}_2(\mathbb{R})$	$\mathrm{SO}_2(\mathbb{R}) \cong \mathbb{T}$	$\mathbb{Z}$
$\mathrm{GL}_n(\mathbb{C})$	$\mathrm{U}_n(\mathbb{C})$	$\mathbb{Z}$
$\mathrm{SL}_n(\mathbb{C})$	$\mathrm{SU}_n(\mathbb{C})$	$\mathbf{1}$
$\mathrm{SO}_{p,q}(\mathbb{R})$	$S(\mathrm{O}_p(\mathbb{R}) \times \mathrm{O}_q(\mathbb{R}))$	$\mathbb{Z}/2 \times \mathbb{Z}/2$ for $p, q > 2$
$\mathrm{SO}_{1,n}(\mathbb{R})$	$\mathrm{O}_n(\mathbb{R})$	$\begin{cases} \mathbb{Z}/2 & \text{for } n > 2 \\ \mathbb{Z} & \text{for } n = 2 \\ \mathbf{1} & \text{for } n = 1 \end{cases}$
$\mathrm{Sp}_{2n}(\mathbb{R})$	$\mathrm{U}_n(\mathbb{C})$	$\mathbf{1}$

2.7 The adjoint representation

Definition 2.48. If V is a vector space and G a group, then a homomorphism $\varphi\colon G \to \mathrm{GL}(V)$ is called a *representation of G on V*. If $\mathfrak{g}$ is a Lie algebra, then a homomorphism of Lie algebras $\varphi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is called a *representation of $\mathfrak{g}$ on V*.

If V is an n-dimensional vector space, then $\mathrm{GL}(V)$ carries a natural Lie group structure for which it is isomorphic to $\mathrm{GL}_n(\mathbb{R})$ (cf. Exercise 2.3). As a consequence of Proposition 2.14, we therefore obtain:

Corollary 2.49. *If $\varphi\colon G \to \mathrm{GL}(V)$ is a smooth representation of the linear Lie group G, then $\mathbf{L}(\varphi)\colon \mathbf{L}(G) \to \mathfrak{gl}(V) = (\mathrm{End}(V), [\cdot,\cdot])$ is a representation of the Lie algebra $\mathbf{L}(G)$.*

The representation $\mathbf{L}(\varphi)$ obtained in Corollary 2.49 from the group representation φ is called the *derived representation*. This is motivated by the fact that for each $x \in \mathbf{L}(G)$ we have

$$\mathbf{L}(\varphi)x = \frac{d}{dt}\Big|_{t=0} e^{t\,\mathbf{L}(\varphi)x} = \frac{d}{dt}\Big|_{t=0} \varphi(\exp tx).$$

Definition 2.50. Let G be a Lie group and $\mathbf{L}(G)$ its Lie algebra. For $g \in G$ we recall the conjugation automorphism $c_g \in \mathrm{Aut}(G), c_g(x) = gxg^{-1}$, and define

$$\mathrm{Ad}(g) := \mathbf{L}(c_g) \in \mathrm{Aut}(\mathbf{L}(G)).$$

Then

$$\mathrm{Ad}(g_1g_2) = \mathbf{L}(c_{g_1g_2}) = \mathbf{L}(c_{g_1}) \circ \mathbf{L}(c_{g_2}) = \mathrm{Ad}(g_1)\,\mathrm{Ad}(g_2)$$

shows that $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathbf{L}(G))$ is a group homomorphism. It is called the *adjoint representation of G*. To see that it is smooth, we observe that for each $x \in \mathbf{L}(G)$ we have

$$\mathrm{Ad}(g)x = T_{\mathbf{1}}(c_g)x = T_{\mathbf{1}}(\lambda_g \circ \rho_{g^{-1}})x = T_{g^{-1}}(\lambda_g)T_{\mathbf{1}}(\rho_{g^{-1}})x = g \cdot x \cdot g^{-1}$$

in $T(G)$. Since $T(m_G)$ is smooth, the representation Ad of G on $\mathbf{L}(G)$ is smooth (cf. Exercise 2.14), and

$$\mathbf{L}(\mathrm{Ad})\colon \mathbf{L}(G) \to \mathfrak{gl}(\mathbf{L}(G))$$

is a representation of $\mathbf{L}(G)$ on $\mathbf{L}(G)$.

Lemma 2.51. *If G is connected, then*

$$\ker \mathrm{Ad} = Z(G) = \{z \in G \colon (\forall g \in G) gz = zg\}$$

is the center of G.

Proof. In view of Proposition 2.42 and the connectedness of G, the relation $\mathbf{L}(c_z) = \mathrm{Ad}(z) = \mathbf{1}$ is equivalent to $c_z = \mathrm{id}_G$, which means that $z \in Z(G)$. □

The following lemma gives a formula for this representation. Here we use the notation

$$\mathrm{ad}(x)y := [x, y]$$

for elements x, y of a Lie algebra.

Lemma 2.52. $\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}$*, i.e.* $\mathbf{L}(\mathrm{Ad})(x)(y) = [x, y]$.

Proof. Let $x, y \in \mathbf{L}(G)$ and x_l, y_l be the corresponding left invariant vector fields. For $g \in G$ we then have

$$((c_g)_* y_l)(h) = T(c_g) y_l(c_g^{-1}(h)) = g \cdot \big((g^{-1}hg) \cdot y\big) \cdot g^{-1} = hg \cdot y \cdot g^{-1} = (\mathrm{Ad}(g)y)_l(h).$$

On the other hand, the left invariance of y_l leads to

$$(c_g)_* y_l = (\rho_g^{-1} \circ \lambda_g)_* y_l = (\rho_g^{-1})_* (\lambda_g)_* y_l = (\rho_g^{-1})_* y_l.$$

Next we recall that $\Phi_t^{x_l} = \rho_{\exp_G(tx)}$ is the flow of the vector field x_l, so that Theorem 1.59 implies that

$$[x_l, y_l] = \mathcal{L}_{x_l} y_l = \frac{d}{dt}\Big|_{t=0} (\Phi_{-t}^{x_l})_* y_l = \frac{d}{dt}\Big|_{t=0} (c_{\exp_G(tx)})_* y_l = \frac{d}{dt}\Big|_{t=0} \big(\mathrm{Ad}(\exp_G(tx))y\big)_l.$$

Evaluating in $\mathbf{1}$, we get

$$[x, y] = [x_l, y_l](\mathbf{1}) = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}(\exp_G(tx))y = \mathbf{L}(\mathrm{Ad})(x)(y).$$ □

Example 2.53. For a linear Lie group $G \subseteq \mathrm{GL}_n(\mathbb{R})$, the automorphisms $c_g(h) = ghg^{-1}$ are restrictions of linear endomorphisms of the vector space $M_n(\mathbb{R})$, which leads to

$$\mathrm{Ad}(g)x = gxg^{-1} \quad \text{for} \quad g \in G, x \in \mathbf{L}(G).$$

Accordingly, we find for $\mathrm{ad} = \mathbf{L}(\mathrm{Ad})$ the concrete formula

$$\mathrm{ad}\, x(y) = xy - yx \quad \text{for} \quad x, y \in \mathbf{L}(G).$$

Example 2.54. We take a closer look at the adjoint representation of $G = \mathrm{SU}_2(\mathbb{C})$. We recall that

$$\mathfrak{su}_2(\mathbb{C}) = \{x \in \mathfrak{gl}_2(\mathbb{C}) \colon x^\dagger = -x, \operatorname{tr} x = 0\} = \left\{ \begin{pmatrix} ai & b \\ -\bar{b} & -ai \end{pmatrix} : b \in \mathbb{C}, a \in \mathbb{R} \right\}.$$

This is a three-dimensional real subspace of $\mathfrak{gl}_2(\mathbb{C})$. The hermitian matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are called the *Pauli matrices*. The matrices $i\sigma_j$, $j = 1, \ldots, 3$ form a basis of the Lie algebra $\mathfrak{su}_2(\mathbb{C})$ of $\mathrm{SU}_2(\mathbb{C})$. The Pauli matrices satisfy the commutator relations

$$[i\sigma_1, i\sigma_2] = -2i\sigma_3, \quad [i\sigma_2, i\sigma_3] = -2i\sigma_1, \quad [i\sigma_3, i\sigma_1] = -2i\sigma_2,$$

to that

$$\operatorname{ad}(i\sigma_1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & -2 & 0 \end{pmatrix}, \quad \operatorname{ad}(i\sigma_2) = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}, \quad \operatorname{ad}(i\sigma_3) = \begin{pmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

showing that $\operatorname{ad}(i\sigma_j) = -2J_j \in \mathfrak{so}_3(\mathbb{R})$ in the notation of Example 2.31. We conclude that $\operatorname{ad} \colon \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{R})$ is a linear isomorphism. Since the exponential function of $\mathrm{SU}_2(\mathbb{C})$ is surjective (Proposition 2.33), we find with the relation $\operatorname{Ad}(\exp x) = e^{\operatorname{ad} x}$ (Lemma 2.52 and Proposition 2.40) that

$$\operatorname{Ad}(\mathrm{SU}_2(\mathbb{C})) = \operatorname{Ad}(\exp \mathfrak{su}_2(\mathbb{C})) = e^{\operatorname{ad} \mathfrak{su}_2(\mathbb{C})} = \exp(\mathfrak{so}_3(\mathbb{R})) = \mathrm{SO}_3(\mathbb{R})$$

(cf. Proposition 2.33). Next we observe that

$$\ker \operatorname{Ad} = Z(\mathrm{SU}_2(\mathbb{C})) = \{\pm \mathbf{1}\}$$

(see Corollary 2.51 for the first equality and Exercise 2.16 for the center of $\mathrm{SU}_n(\mathbb{C})$), so that

$$\mathrm{SO}_3(\mathbb{R}) \cong \mathrm{SU}_2(\mathbb{C})/\{\pm \mathbf{1}\}.$$

2.8 Semidirect products

In this subsection we introduce the concept of a semidirect product of two Lie groups. This is a construction to create a new Lie group from two given ones that is more general than the direct product construction. Semidirect products of Lie groups arise naturally as groups of isometries of euclidean spaces and groups of automorphisms of affine spaces. Therefore we start with the concept of an affine space. An affine space can be considered as a vector space where no origin has been specified. This is closer to the physical concept of space, where no point plays a preferred role.

2.8.1 Affine spaces

Definition 2.55. Let V be a vector space. An *affine space with translation space* V consists of a set $\mathbb{A}$ and a map

$$+\colon \mathbb{A} \times V \to \mathbb{A}, \quad (a, \mathbf{x}) \mapsto a + \mathbf{x},$$

such that the following conditions are satisfied

(A1) $a + \mathbf{o} = a$ for all $a \in \mathbb{A}$.

(A2) $a + (\mathbf{x} + \mathbf{y}) = (a + \mathbf{x}) + \mathbf{y}$ for $a \in \mathbb{A}, \mathbf{x}, \mathbf{y} \in V$.

(A3) For $a, b \in \mathbb{A}$ there exists a unique $\mathbf{x} \in V$ with $b = a + \mathbf{x}$. The element $b - a := \mathbf{x}$ is called the *translation vector from a to b.*

A map $\varphi\colon A_1 \to A_2$ between affine spaces with translation space V_1, resp., V_2 is called *affine*, if there exists a linear map $\psi\colon V_1 \to V_2$ with

$$\varphi(a + \mathbf{x}) = \varphi(a) + \psi(\mathbf{x}) \quad \text{for} \quad a \in \mathbb{A}, \mathbf{x} \in V.$$

Example 2.56. (a) For every vector space V, we obtain an affine space $\mathbb{A} := V$ with respect to vector addition.

(b) For $V = \mathbb{R}^n$, the corresponding affine space is called *n-dimensional affine space* $\mathbb{A}^n$.

Remark 2.57. Once a point $o \in \mathbb{A}$ is chosen, the map $V \to \mathbb{A}, \mathbf{x} \mapsto o + \mathbf{x}$ is bijective, so that, as a set, the affine space cannot be distinguished from the vector space V. However, conceptually, the notion of an affine space is different from that of a vector space. In view of the preceding remark, we may think of an affine space $\mathbb{A}$ with translation group V as a copy of V, where no origin is distinguished. Conversely, any choice of origin $o \in \mathbb{A}$ leads to an identification with V and hence to a vector space structure of $\mathbb{A}$.

The difference between $\mathbb{A}$ and V is also visible in the fact that the group $\mathrm{Aut}(\mathbb{A})$ of affine automorphisms of $\mathbb{A}$ is larger than the group $\mathrm{GL}(V)$ of linear automorphisms of V. The translations $\tau_{\mathbf{x}}(a) := a + \mathbf{x}$ are also affine automorphisms and they form a subgroup $\tau_{\mathbb{A}} \subseteq \mathrm{Aut}(\mathbb{A})$ isomorphic to V which acts simply transitively on $\mathbb{A}$. On the other hand, for every point $o \in \mathbb{A}$, the stabilizer $\mathrm{Aut}(\mathbb{A})_o$ is isomorphic to $\mathrm{GL}(V)$ because it consists of maps of the form $\widetilde{\psi}(o + \mathbf{x}) = o + \psi(\mathbf{x})$, $\psi \in \mathrm{GL}(V)$. Since every automorphism $\varphi \in \mathrm{Aut}(\mathbb{A})$ can be written in a unique fashion as

$$\varphi = \tau_{\mathbf{x}} \circ \widetilde{\psi} \quad \text{with } \varphi(o) = o + \mathbf{x}, \psi \in \mathrm{GL}(V),$$

we can think of affine automorphism as pairs $(\mathbf{x}, \psi) \in V \times \mathrm{GL}(V)$. Composition of maps then corresponds to

$$(\mathbf{x}, \psi) \circ (\mathbf{x}', \psi') = (\mathbf{x} + \psi(\mathbf{x}'), \psi\psi').$$

To deal with group structures of this form, we introduce the notion of a semidirect product.

2.8.2 Affine automorphism groups as semidirect products

The easiest way to construct a new Lie group from two given Lie groups G and H, is to endow the product manifold $G \times H$ with the multiplication

$$(g_1, h_1)(g_2, h_2) := (g_1 g_2, h_1 h_2).$$

The resulting group is called the *direct product* of the Lie groups G and H. Here G and H can be identified with normal subgroups of $G \times H$ for which the multiplication map

$$(G \times \{\mathbf{1}\}) \times (\{\mathbf{1}\} \times H) \to G \times H, \quad ((g, \mathbf{1}), (\mathbf{1}, h)) \mapsto (g, \mathbf{1})(\mathbf{1}, h) = (g, h)$$

is a diffeomorphism. Relaxing this condition in the sense that only one factor is assumed to be normal, leads to the concept of a semidirect product of Lie groups, introduced below.

Definition 2.58. Let N and G be Lie groups and $\alpha \colon G \to \mathrm{Aut}(N)$ be a group homomorphism defining a smooth action $(g, n) \mapsto \alpha_g(n)$ of G on N.

Then the product manifold $N \times G$ is a group with respect to the product

$$(n, g)(n', g') := (n\alpha_g(n'), gg') \quad \text{with inversion} \quad (n, g)^{-1} = (\alpha_{g^{-1}}(n^{-1}), g^{-1}).$$

Since multiplication and inversion are smooth, this group is a Lie group, called the *semidirect product of N and G with respect to α.* It is denoted by $N \rtimes_\alpha G$.

Example 2.59. A typical example of a semidirect product is the group $\mathrm{Aff}(\mathbb{A}^n)$ of automorphisms of the n-dimensional affine space $\mathbb{A}^n$, resp., the group $\mathrm{Aff}_n(\mathbb{R})$ of affine isomorphisms $\varphi(x) = Ax + b$ of $\mathbb{R}^n$. Writing the elements of this group as pairs (b, A), we have

$$(b, A)(b', A') = (b + Ab', AA'),$$

so that $\mathrm{Aff}_n(\mathbb{R}) \cong \mathbb{R}^n \rtimes_\alpha \mathrm{GL}_n(\mathbb{R})$ with $\alpha(g)x = gx$.

Definition 2.60. The *n-dimensional euclidean space* $\mathbb{E}^n$ is the affine space $\mathbb{A}^n$, endowed with the euclidean metric

$$d(a, b) := \Big(\sum_{j=1}^n x_j^2\Big)^{1/2} \quad \text{for} \quad b = a + \mathbf{x}.$$

The *euclidean group* is the group $\mathrm{ISO}_n(\mathbb{R})$ of affine isometries of $\mathbb{E}^n$; it is also called $E_n(\mathbb{R})$. Example 2.59 implies immediately that

$$\mathrm{ISO}_n(\mathbb{R}) \cong \mathbb{R}^n \rtimes_\alpha \mathrm{O}_n(\mathbb{R})$$

because an affine map is isometric if and only if its linear part is, which means that it corresponds to an orthogonal matrix. Actually one can show that every isometry of a normed space $(V, \|\cdot\|)$ is an affine map (Exercise 2.21). This implies that all isometries of $\mathbb{E}^n$ are affine.

2.8.3 Lorentz and Poincaré group

We define the *n-dimensional Minkowski space* $\mathbb{M}^n$ as the affine space $\mathbb{A}^n$, endowed with the Lorentzian form

$$q(a,b) := x_0^2 - \sum_{j=1}^{n-1} x_j^2 \quad \text{for} \quad b = a + \mathbf{x}, \mathbf{x} = (x_0, \ldots, x_{n-1}).$$

Its group of affine isometries is the *Poincaré group*

$$\mathrm{ISO}_{1,n-1}(\mathbb{R}) \cong \mathbb{R}^n \rtimes_\alpha \mathrm{O}_{1,n-1}(\mathbb{R})$$

of all affine isomorphisms of $\mathbb{A}^n$ preserving the Lorentzian forn q. Accordingly, $L := \mathrm{O}_{1,n-1}(\mathbb{R})$ is called the *Lorentz group*.

We write

$$\beta(x,y) := x_0 y_0 - \sum_{j=1}^{n-1} x_j y_j$$

for the symmetric bilinear form on $\mathbb{R}^n$ with signature $(1, n-1)$ and $q(x) := \beta(x,x)$ for the corresponding quadratic form. The Lorentz group has several subgroups:

$$L_+ := \mathrm{SO}_{1,n-1}(\mathbb{R}) := L \cap \mathrm{SL}_n(\mathbb{R}) \quad \text{and} \quad L^\uparrow := \{g \in L \colon g_{00} \geq 1\}.$$

The condition $g_{00} \geq 1$ comes from

$$1 = \beta(e_0, e_0) = \beta(ge_0, ge_0) = g_{00}^2 - \sum_{j=1}^{n-1} g_{j0}^2,$$

which implies $g_{00}^2 \geq 1$. Therefore either $g_{00} \geq 1$ or $g_{00} \leq -1$. To understand geometrically why $L^\uparrow$ is a subgroup, we observe that q is invariant under L, so that L preserves the double cone

$$C := \{x \in \mathbb{R}^n \colon q(x) \leq 0\} = \{x = (x_0, \mathbf{x}) \in \mathbb{R}^n \colon |x_0| \geq \|\mathbf{x}\|\}.$$

Let

$$C_\pm := \{x \in C \colon \pm x_0 \geq 0\} = \{x = (x_0, \mathbf{x}) \in \mathbb{R}^n \colon \pm x_0 \geq \|\mathbf{x}\|\}.$$

Then $C = C_+ \cup C_-$ with $C_+ \cap C_- = \{0\}$ and the sets $C_\pm$ are both convex cones, as follows easily from the convexity of the euclidean norm function on $\mathbb{R}^{n-1}$ (Exercise). Each element $g \in L$ preserves the set $C \setminus \{0\}$ which has the two arc-components $C_\pm \setminus \{0\}$. The continuity of the map $g \colon C \setminus \{0\} \to C \setminus \{0\}$ now implies that we have two possibilities. Either $gC_+ = C_+$ or $gC_+ = C_-$. In the first case, $g_{00} \geq 1$ and in the latter case $g_{00} \leq -1$. [4]

The *proper orthochronous Poincaré group* is the corresponding affine group

$$P := \mathbb{R}^n \rtimes L_+^\uparrow.$$

[4] In the physics literature one sometimes finds $\mathrm{SO}_{1,3}(\mathbb{R})$ as the notation for the connected group $L_+^\uparrow := L_+ \cap L^\uparrow$ (cf. Example 2.39), which is inconsistent with the standard notation for matrix groups.

This group is the identity component of $\mathrm{ISO}_{1,n-1}(\mathbb{R})$. [5]

The topological structure of the Poincaré- and Lorentz group become transparent with the polar decomposition (cf. Example 2.39). In particular, it shows that the Lorentz group L has four arc-components

$$L_+{}^{\uparrow}, \quad L_+{}^{\downarrow}, \quad L_-{}^{\uparrow} \quad \text{and} \quad L_-{}^{\downarrow},$$

where

$$L_\pm := \{g \in L\colon \det g = \pm 1\}, \quad L^{\downarrow} := \{g \in L\colon g_{00} \leq -1\}$$

and

$$L_\pm{}^{\uparrow} := L_\pm \cap L^{\uparrow}, \quad L_\pm{}^{\downarrow} := L_\pm \cap L^{\downarrow}.$$

The element

$$T = \begin{pmatrix} -1 & 0 \\ 0 & \mathbf{1}_{n-1} \end{pmatrix}$$

is called *time reversal* and

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -\mathbf{1}_{n-1} \end{pmatrix}$$

is the *parity transformation*. Both are contained in L, and if n is even, we have

$$L = \{\mathbf{1}, T, P, TP\} \cdot L_+^{\uparrow}.$$

Evaluating the condition defining the Lie algebra $\mathfrak{so}_{1,n-1}(\mathbb{R})$ in terms of (2×2)-block matrices according to the decomposition $\mathbb{R}^n = \mathbb{R} \oplus \mathbb{R}^{n-1}$, we obtain

$$\mathfrak{so}_{1,n-1}(\mathbb{R}) = \Big\{ \begin{pmatrix} 0 & v^\top \\ v & D \end{pmatrix} : v \in \mathbb{R}^{n-1}, D^\top = -D \Big\}.$$

In particular,

$$\mathfrak{p} = \mathfrak{so}_{1,n-1}(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}) = \Big\{ \begin{pmatrix} 0 & v^\top \\ v & 0 \end{pmatrix} : v \in \mathbb{R}^{n-1} \Big\} \cong \mathbb{R}^{n-1}.$$

To make the polar decomposition more explicit, we calculate $\exp X$ for

$$X = \begin{pmatrix} 0 & v^\top \\ v & 0 \end{pmatrix}.$$

This can be done explicitly because $X^3 = \|v\|^2 X$. This leads to

$$\begin{aligned} \exp X &= \mathbf{1} + \Big(\frac{1}{2!} + \frac{\|v\|^2}{4!} + \frac{\|v\|^4}{6!} + \dots\Big)X^2 + \Big(\mathbf{1} + \frac{\|v\|^2}{3!} + \frac{\|v\|^4}{5!} + \dots\Big)X \\ &= \mathbf{1} + \frac{\cosh\|v\| - 1}{\|v\|^2}X^2 + \frac{\sinh\|v\|}{\|v\|}X \\ &= \begin{pmatrix} \cosh\|v\| & \frac{\sinh\|v\|}{\|v\|}v^\top \\ \frac{\sinh\|v\|}{\|v\|}v & \mathbf{1} + \frac{\cosh\|v\|-1}{\|v\|^2}vv^\top \end{pmatrix} =: L(v). \end{aligned}$$

[5] Some people use the name Poincaré group only for the simply connected covering group of P which is isomorphic to $\mathbb{R}^4 \rtimes \mathrm{SL}_2(\mathbb{C})$ (cf. Exercise 2.19).

The matrix $L(v)$ is called a *Lorentz boost in direction v with rapidity* $\|v\|$. Putting $w := \frac{\sinh\|v\|}{\|v\|}v$, we obtain the slightly simpler form

$$L(v) = \begin{pmatrix} \sqrt{1+\|w\|^2} & w^\top \\ w & \mathbf{1} + \frac{\sqrt{1+\|w\|^2}-1}{\|w\|^2} ww^\top \end{pmatrix}.$$

Let us assume, from now on, that $n = 4$. Then we obtain in particular

$$L(te_1) = \begin{pmatrix} \cosh t & \sinh t & 0 & 0 \\ \sinh t & \cosh t & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad L(te_2) = \begin{pmatrix} \cosh t & 0 & \sinh t & 0 \\ 0 & 1 & 0 & 0 \\ \sinh t & 0 & \cosh t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and

$$L(te_3) = \begin{pmatrix} \cosh t & 0 & 0 & \sinh t \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh t & 0 & 0 & \cosh t \end{pmatrix},$$

resp., for $s = \sinh t$:

$$L(te_1) = \begin{pmatrix} \sqrt{1+s^2} & s & 0 & 0 \\ s & \sqrt{1+s^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{etc.}$$

Let

$$B_1 := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B_2 := \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B_3 := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

denote the generators of the one-parameter groups $L(te_j)$ and note that they form a basis for $\mathfrak{p} \cong \mathbb{R}^3$. To obtain a basis for $\mathfrak{so}_{1,3}(\mathbb{R}) = \mathfrak{so}_3(\mathbb{R}) \oplus \mathfrak{p}$, these elements have to be supplemented by the generators on the rotations in $\{0\} \times \mathbb{R}^3 \subseteq \mathbb{R}^{1,3}$:

$$R_1 := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad R_2 := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad R_3 := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

From

$$\begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}\begin{pmatrix} 0 & v^\top \\ v & 0 \end{pmatrix} - \begin{pmatrix} 0 & v^\top \\ v & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix} = \begin{pmatrix} 0 & -v^\top A \\ Av & 0 \end{pmatrix} = \begin{pmatrix} 0 & (Av)^\top \\ Av & 0 \end{pmatrix}$$

for $v \in \mathbb{R}^3$ and $A = -A^\top \in \mathfrak{so}_3(\mathbb{R})$ and

$$\begin{pmatrix} 0 & v^\top \\ v & 0 \end{pmatrix}\begin{pmatrix} 0 & w^\top \\ w & 0 \end{pmatrix} - \begin{pmatrix} 0 & w^\top \\ w & 0 \end{pmatrix}\begin{pmatrix} 0 & v^\top \\ v & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & vw^\top - wv^\top \end{pmatrix},$$

we obtain the commutator relations

$$[R_i, R_j] = \sum_k \varepsilon_{ijk} R_k, \quad [R_j, B_j] = \sum_k \varepsilon_{ijk} B_k \quad \text{and} \quad [B_i, B_j] = -\sum_k \varepsilon_{ijk} R_k.$$

Exercises for Section 2

Exercise 2.1. Show that the natural group structure on $\mathbb{T} \cong \mathbb{S}^1 \subseteq \mathbb{C}^\times$ turns it into a Lie group.

Exercise 2.2. Let $G_1, \ldots, G_n$ be Lie groups and $G := G_1 \times \ldots \times G_n$, endowed with the direct product group structure

$$(g_1, \ldots, g_n)(g_1', \ldots, g_n') := (g_1 g_1', \ldots, g_n g_n')$$

and the product manifold structure. Show that G is a Lie group with

$$\mathbf{L}(G) \cong \mathbf{L}(G_1) \times \ldots \times \mathbf{L}(G_n).$$

Exercise 2.3. Let V be an n-dimensional real vector space and fix a linear isomorphism $\iota\colon \mathbb{R}^n \to V$. Then we obtain a linear isomorphism

$$\Phi\colon \operatorname{End}(V) \to M_n(\mathbb{R}), \quad \Phi(\varphi)x = \iota^{-1}(\varphi\iota(x))$$

which we consider as a chart of $\operatorname{End}(V)$. Show that we thus obtain on the open subset $\operatorname{GL}(V)$ the structure of a Lie group.

Exercise 2.4. On the tangent bundle TG of the Lie group G, we consider the multiplication

$$T(m_G)\colon T(G \times G) \cong TG \times TG \to TG, \quad (v_g, w_h) \mapsto g \cdot w + v \cdot h$$

(cf. Lemma 2.6). Show that this turns TG into a Lie group with neutral element $0_{\mathbf{1}} \in T_{\mathbf{1}}(G)$ and inversion $T(\eta_G)$.

If this is too abstract, consider the special case $G = \operatorname{GL}_n(\mathbb{R})$ whose tangent bundle we identify with the open subset $T\operatorname{GL}_n(\mathbb{R}) = \operatorname{GL}_n(\mathbb{R}) \times M_n(\mathbb{R})$ of $M_n(\mathbb{R}) \times M_n(\mathbb{R})$.

Exercise 2.5. Let G be an n-dimensional Lie group and (φ, U) be a local chart of G with $\mathbf{1} \in U$ and $\varphi(\mathbf{1}) = 0$. We then obtain a locally defined smooth function

$$x * y := \varphi(\varphi^{-1}(x)\varphi^{-1}(y))$$

defined in an open neighborhood of 0 in $\mathbb{R}^n \times \mathbb{R}^n$. Show that:

(i) The Taylor polynomial of order 2 of $*$ is of the form $x+y+b(x,y)$, where $b\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is bilinear. Hint: Use the relations $x * 0 = x$ and $0 * y = y$ and that every quadratic form

$$q\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \quad q(x,y) = \sum_{i,j=1}^n a_{ij} x_i x_j + \sum_{i,j=1}^n b_{ij} y_i y_j + \sum_{i,j=1}^n c_{ij} x_i y_j$$

vanishing in all pairs $(x, 0)$ and $(0, y)$ is bilinear.

(ii) The first order Taylor polynomial of the left invariant vector field $x_l(z)$ in 0 is $x + b(z, x)$.

(iii) $[x, y] = b(x, y) - b(y, x)$.

(iii) Apply this to the chart $\varphi(g) = g - \mathbf{1}$ of $\mathrm{GL}_n(\mathbb{R})$.

Exercise 2.6. Show that

$$\gamma : (\mathbb{R}, +) \to \mathrm{GL}_2(\mathbb{R}), \quad t \mapsto \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$$

is a continuous group homomorphism with $\gamma(\pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\mathrm{im}(\gamma) = \mathrm{SO}_2(\mathbb{R})$.

Exercise 2.7. Show that:
(a) $\exp(M_n(\mathbb{R}))$ is contained in the identity component $\mathrm{GL}_n(\mathbb{R})_+$ of $\mathrm{GL}_n(\mathbb{R})$. In particular the exponential function of $\mathrm{GL}_n(\mathbb{R})$ is not surjective because this group is not connected.
(b) The exponential function $\exp : M_2(\mathbb{R}) \to \mathrm{GL}_2(\mathbb{R})_+$ is not surjective.

Exercise 2.8. Every matrix $X \in \mathfrak{sl}_2(\mathbb{K})$ satisfies $X^2 = -\det X\mathbf{1}$. Show that

$$e^X = \Big(\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}(\det X)^k\Big)\mathbf{1} + \Big(\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}(\det X)^k\Big)X.$$

Conclude further that:

(i) $e^X = \mathbf{1} + X$ for $\det X = 0$.

(ii) $e^X = \cosh(\sqrt{-\det X})\mathbf{1} + \frac{\sinh(\sqrt{-\det X})}{\sqrt{-\det X}}X$ for $\det X < 0$.

(iii) $e^X = \cos(\sqrt{\det X})\mathbf{1} + \frac{\sin(\sqrt{\det X})}{\sqrt{\det X}}X$ for $\det X > 0$.

(iv) $\exp\left(t\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ and $\exp\left(t\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$.

Exercise 2.9. On $\mathbb{R}^3$ we consider the vector product

$$v \times w = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ -(v_1 w_3 - v_3 w_1) \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

We define a linear map

$$\varphi\colon \mathbb{R}^3 \to M_3(\mathbb{R}), \quad \varphi(x)y = x \times y.$$

Show that

(i) $\mathrm{im}(\varphi) = \mathfrak{so}_3(\mathbb{R})$.

(ii) $\varphi(x \times y) = [\varphi(x), \varphi(y)]$ for $x, y \in \mathbb{R}^3$.

(iii) $(\mathbb{R}^3, \times)$ is a Lie algebra isomorphic to $\mathfrak{so}_3(\mathbb{R})$.

Exercise 2.10. Show that for $n = p + q$ we have

$$\mathrm{O}_{p,q}(\mathbb{K}) \cap \mathrm{O}_n(\mathbb{K}) \cong \mathrm{O}_p(\mathbb{K}) \times \mathrm{O}_q(\mathbb{K}).$$

Exercise 2.11. $\mathrm{SO}_n(\mathbb{K})$ is a closed normal subgroup of $\mathrm{O}_n(\mathbb{K})$ of index 2 and, for every $g \in \mathrm{O}_n(\mathbb{K})$ with $\det(g) = -1$,

$$\mathrm{O}_n(\mathbb{K}) = \mathrm{SO}_n(\mathbb{K}) \cup g\,\mathrm{SO}_n(\mathbb{K})$$

is a disjoint decomposition.

Exercise 2.12. Let $\beta\colon V \times V \to V$ be a symmetric bilinear form on the vector space V and

$$q\colon V \to V, \quad v \mapsto \beta(v, v)$$

the corresponding quadratic form. Then for $\varphi \in \mathrm{End}(V)$ the following are equivalent:

(1) $(\forall v \in V)\ q(\varphi(v)) = q(v)$.

(2) $(\forall v, w \in V)\ \beta(\varphi(v), \varphi(w)) = \beta(v, w)$.

Exercise 2.13. Let $\varphi\colon G \to H$ be a smooth homomorphism of Lie groups. Show that:

(i) $\mathbf{L}(\ker\varphi) \cong \mathbf{L}(\ker\varphi)$.

(ii) φ has discrete kernel if and only if $\mathbf{L}(\varphi)$ is injective.

(iii) φ is a submersion if and only if $\mathbf{L}(\varphi)$ is surjective.

(iv) If G and H are connected and $\mathbf{L}(\varphi)$ is surjective, then φ is surjective.

(v) If G and H are connected of the same dimension and $\ker\varphi$ is discrete, then φ is surjective.

Exercise 2.14. Let M be a manifold and V a finite-dimensional vector space with a basis $(b_1, \ldots, b_n)$. Let $f\colon M \to \mathrm{GL}(V)$ be a map. Show that the following are equivalent:

(1) f is smooth.

(2) For each $v \in V$ the map $f_v\colon M \to V, m \mapsto f(m)v$ is smooth.

(3) For each i, the map $f\colon M \to V, m \mapsto f(m)b_i$ is smooth.

Exercise 2.15. (The exponential function of $\mathrm{SU}_2(\mathbb{C})$) Show that:

(a) $\mathrm{U}_2(\mathbb{C}) = \mathbb{T}\,\mathrm{SU}_2(\mathbb{C}) = Z(\mathrm{U}_2(\mathbb{C}))\,\mathrm{SU}_2(\mathbb{C})$.

(b) If $x \in \mathfrak{su}_2(\mathbb{C})$ with eigenvalues $\pm i\lambda$, $\lambda \geq 0$, we have $\|x\| = \lambda$.

(c) For $x, y \in \mathfrak{su}_2(\mathbb{C})$, there exists an element $g \in \mathrm{SU}_2(\mathbb{C})$ with $y = \mathrm{Ad}(g)x$ if and only if $\|x\| = \|y\|$.

(d) No one-parameter group $\gamma : \mathbb{R} \to \mathrm{SU}_2(\mathbb{C})$ is injective, in particular, the image of $\gamma(\mathbb{R})$ is always circle group.

Exercise 2.16. Show that

$$Z(\mathrm{U}_n(\mathbb{C})) = \mathbb{T}\mathbf{1} \quad \text{and} \quad Z(\mathrm{SU}_n(\mathbb{C})) = \{z\mathbf{1} \colon z^n = 1\} \cong C_n.$$

Hint: Each $g \in \mathbb{Z}(\mathrm{U}_n(\mathbb{C}))$ satisfies $\mathrm{Ad}(g) = \mathbf{1}$. Conclude from $\mathfrak{gl}_n(\mathbb{C}) = \mathfrak{u}_n(\mathbb{C}) + i\mathfrak{u}_n(\mathbb{C})$ that g commutes with all matrices. For $g \in Z(\mathrm{SU}_n(\mathbb{C}))$, use $\mathfrak{gl}_n(\mathbb{C}) = \mathfrak{su}_n(\mathbb{C}) + i\,\mathfrak{su}_n(\mathbb{C}) + \mathbb{C}\mathbf{1}$ is a similar fashion.

Exercise 2.17. (a) Show that a matrix $A \in M_n(\mathbb{C})$ is hermitian if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{C}^n$ and real numbers $\lambda_1, \ldots, \lambda_n$ with $Av_j = \lambda_j v_j$.

(b) Show that a complex matrix $A \in M_n(\mathbb{C})$ is unitary if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{C}^n$ and $\lambda_j \in \mathbb{C}$ with $|\lambda_j| = 1$ and $Av_j = \lambda_j v_j$.

(c) Show that a complex matrix $A \in M_n(\mathbb{C})$ is *normal*, i.e. satisfies $A^\dagger A = AA^\dagger$, if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{C}^n$ and $\lambda_j \in \mathbb{C}$ with $Av_j = \lambda_j v_j$.

Exercise 2.18. Show that the groups $\mathrm{O}_n(\mathbb{C})$, $\mathrm{SO}_n(\mathbb{C})$ and $\mathrm{Sp}_{2n}(\mathbb{R})$ have polar decompositions and describe their intersections with $\mathrm{U}_n(\mathbb{C})$, resp., $\mathrm{O}_{2n}(\mathbb{R})$.

Exercise 2.19. On the four-dimensional real vector space $V := \mathrm{Herm}_2(\mathbb{C})$ we consider the symmetric bilinear form β given by

$$\beta(A, B) := \tfrac{1}{2}(\mathrm{tr}\, A \,\mathrm{tr}\, B - \mathrm{tr}(AB)).$$

Show that:

(1) The corresponding quadratic form is given by $q(A) := \beta(A, A) = \det A$.

(2) Show that the basis σ_j, $j = 0, \ldots, 3$ with $\sigma_0 = \mathbf{1}$ and where σ_j, $j = 1, 2, 3$, are the Pauli matrices, is orthogonal with respect to β and that we thus obtain an isomorphism $(V, \beta) \cong \mathbb{R}^{1,3}$:

$$q(a_0\sigma_0 + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3) = a_0^2 - a_1^2 - a_2^2 - a_3^2.$$

(3) For $g \in \mathrm{GL}_2(\mathbb{C})$ and $A \in \mathrm{Herm}_2(\mathbb{C})$ the matrix $gAg^\dagger$ is hermitian and satisfies

$$q(gAg^\dagger) = |\det(g)|^2 q(A).$$

(4) For $g \in \mathrm{SL}_2(\mathbb{C})$ we define a linear map $\rho(g) \in \mathrm{GL}(\mathrm{Herm}_2(\mathbb{C})) \cong \mathrm{GL}_4(\mathbb{R})$ by $\rho(g)(A) := gAg^\dagger$. Then we obtain a homomorphism

$$\rho \colon \mathrm{SL}_2(\mathbb{C}) \to \mathrm{O}(V, \beta) \cong \mathrm{O}_{3,1}(\mathbb{R}).$$

(5) Show that $\ker \rho = \{\pm \mathbf{1}\}$.

(6) $\mathbf{L}(\rho) \colon \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_{3,1}(\mathbb{R})$ is an isomorphism of Lie algebras. Hint: Use that $\ker \mathbf{L}(\rho) = \mathbf{L}(\ker \rho)$ (Exercise 2.13) and compare dimensions.

(7) $\mathrm{SO}_{1,3}(\mathbb{R})_0 \cong \mathrm{SL}_2(\mathbb{C})/\{\pm\mathbf{1}\}$ (see Example 2.54 for similar arguments).

(8) $\rho(\mathrm{SU}_2(\mathbb{C})) = \mathrm{SO}_3(\mathbb{R})$ consists of those matrices fixing σ_0 (cf. Example 2.54).

Exercise 2.20. Show that for $n > 1$ the sphere $\mathbb{S}^n$ is simply connected. For the proof, proceed along the following steps:
(a) Let $\gamma : [0,1] \to \mathbb{S}^n$ be continuous. Then there exists an $m \in \mathbb{N}$ such that $\|\gamma(t) - \gamma(t')\| < \frac{1}{2}$ for $|t - t'| < \frac{1}{m}$.
(b) Define $\widetilde{\alpha} : [0,1] \to \mathbb{R}^{n+1}$ as the piecewise affine curve with $\widetilde{\alpha}(\frac{k}{m}) = \gamma(\frac{k}{m})$ for $k = 0, \ldots, m$. Then $\alpha(t) := \frac{1}{\|\widetilde{\alpha}(t)\|}\widetilde{\alpha}(t)$ defines a continuous curve $\alpha : [0,1] \to \mathbb{S}^n$.
(c) $\alpha \sim \gamma$.
(d) α is not surjective. The image of α is the central projection of a polygonal arc on the sphere.
(e) If $\beta \in \Omega(\mathbb{S}^n, y_0)$ is not surjective, then $\beta \sim y_0$ (it is homotopic to a constant map).
(f) $\pi_1(\mathbb{S}^n, y_0) = \{[y_0]\}$ for $n \geq 2$ and $y_0 \in \mathbb{S}^n$.

Exercise 2.21. [Isometries of euclidean spaces are affine maps] Let (X, d) be a euclidean space. Show that each isometry $\varphi\colon (X,d) \to (X,d)$ is an affine map by using the following steps:

(1) It suffices to assume that $\varphi(0) = 0$ and to show that this implies that φ is a linear map.

(2) $\varphi(\frac{x+y}{2}) = \frac{1}{2}(\varphi(x) + \varphi(y))$ for $x, y \in X$. Hint: Use that two points $x, y \in X$ has a unique midpoint z with $d(x,z) = d(y,z) = \frac{1}{2}d(x,y)$.

(3) φ is continuous.

(4) $\varphi(\lambda x) = \lambda\varphi(x)$ for $\lambda \in 2^{\mathbb{Z}} \subseteq \mathbb{R}$.

(5) $\varphi(x + y) = \varphi(x) + \varphi(y)$ for $x, y \in X$.

(6) $\varphi(\lambda x) = \lambda\varphi(x)$ for $\lambda \in \mathbb{R}$.

Exercise 2.22. Let G be a group, $N \subseteq G$ a normal subgroup and

$$q\colon G \to G/N, \quad g \mapsto gN$$

be the quotient homomorphism. Show that:

(1) If $G \cong N \rtimes_\delta H$ for a subgroup H, then $H \cong G/N$.

(2) There exists a subgroup $H \subseteq G$ with $G \cong N \rtimes_\delta H$ if and only if there exists a group homomorphism $\sigma\colon G/N \to G$ with $q \circ \sigma = \mathrm{id}_{G/N}$.

Exercise 2.23. Let $N \rtimes_\alpha G$ be a semidirect product of the Lie groups G and N with respect to $\alpha\colon G \to \mathrm{Aut}(N)$. On the manifold $G \times N$ we also obtain a Lie group structure by

$$(g,n)(g',n') := (gg', \alpha_{g'}^{-1}(n)n'),$$

and this Lie group is denoted $G \ltimes_\alpha N$. Show that the map

$$\Phi\colon N \rtimes_\alpha G \to G \ltimes_\alpha N, \quad (n,g) \mapsto (g, \alpha_g^{-1}(n))$$

is an isomorphism of Lie groups.

3 Geometric Structures on Manifolds

3.1 Geometric structures on vector spaces

In this subsection we introduce various types of structures on real vector spaces that will be used below to define corresponding structures on manifolds. It will turn out that fixing an ordered basis B determines the finest possible structure, namely an isomorphism $\iota_B \colon \mathbb{R}^n \to V$, and the only linear automorphism in $\mathrm{GL}(V)$ preserving this structure is the identity. The other geometric structures on V correspond to non-trivial subgroups of $\mathrm{GL}(V)$, such as $\mathrm{GL}(V)_+$ (for orientations), $\mathrm{SL}(V)$ (for volume forms), $\mathrm{O}(V,\beta)$ (for symmetric bilinear forms) and $\mathrm{Sp}(V,\omega)$ (for symplectic structures). All these groups are Lie groups, and we shall take a closer look at their structure and topology below.

Definition 3.1. (Oriented vector spaces)
(a) Let V be an n-dimensional real vector space. If $B = (b_1, \ldots, b_n)$ and $C = (c_1, \ldots, c_n)$ are two ordered bases of V, then we write $M = [\mathrm{id}]^C_B$ for the transition matrix defined by

$$b_j = \sum_{i=1}^{n} m_{ij} c_i.$$

We say that B and C are *equally oriented*, denoted $B \sim_{\mathrm{or}} C$ if $\det M > 0$. Then $\sim_{\mathrm{or}}$ is an equivalence relation on the set of all bases of V. The equivalence classes are called *orientations* and we write $\mathrm{or}(V)$ for the set of orientations on V. We write $[B]$ for the orientation defined by the basis B. Since we either have $\det M > 0$ or $\det M < 0$, there are only two equivalence classes, i.e. V carries two orientations. Accordingly, we write $-[B]$ for the opposite orientation. If O is an orientation on V, then the pair (V, O) is called an *oriented vector space*.

(b) If (V, O) and (V', O') are oriented vector spaces, then an invertible linear map $\varphi \colon V \to V'$ is said to be an *isomorphism* of oriented vector spaces or *orientation preserving* if $\varphi(O) = O'$, where this expression is defined by $\varphi([B]) = [\varphi(B)]$ for an ordered basis B of V.

Remark 3.2. (a) For $\varphi \in \mathrm{GL}(V)$ and an orientation $[B]$ of V, we have $[\varphi(B)] = \mathrm{sgn}(\det(\varphi))[B]$ In particular, φ preserves the orientation if and only if it belongs to the subgroup

$$\mathrm{GL}(V)_+ := \{\varphi \in \mathrm{GL}(V) \colon \det \varphi > 0\}.$$

For the canonical orientation $[e_1, \ldots, e_n]$ of $\mathbb{R}^n$, we thus obtain the matrix group

$$\mathrm{GL}_n(\mathbb{R})_+ := \{g \in \mathrm{GL}_n(\mathbb{R}) \colon \det g > 0\}.$$

This is an open subgroup of $\mathrm{GL}_n(\mathbb{R})$ so that it has the same Lie algebra as $\mathrm{GL}_n(\mathbb{R})$ (Lemma 2.27).

(b) A vector space V has no preferred orientation. The group $\mathrm{GL}(V)$ acts transitively on the set of all orientations by $\varphi[B] = [\varphi(B)]$.

(c) For an n-dimensional real vector space V, we write $\mathrm{bas}(V)$ for the set of all ordered bases $B = (b_1, \ldots, b_n)$ of V. Once an orientation $[B]$ of V is fixed, it defines a function

$$s \colon \mathrm{bas}(V) \to \{\pm 1\}, \quad \varphi(B) \mapsto \mathrm{sgn}(\det(\varphi)).$$

This function satisfies the equation

$$s(\varphi(C)) = \operatorname{sgn}(\det(\varphi))s(C) \quad \text{for} \quad \varphi \in \mathrm{GL}(V), C \in \operatorname{bas}(V).$$

Conversely, any function $\operatorname{bas}(V) \to \{\pm 1\}$ with this transformation behavior defines an orientation of V by $O = \{[B] \colon s(B) = 1\}$.

Definition 3.3. Let V be an n-dimensional real vector space.

(a) A *density on* V is a function $\delta \colon \operatorname{bas}(V) \to \mathbb{R}_+^\times$ with the property

$$\delta(\varphi B) = |\det(\varphi)|\delta(B) \quad \text{for} \quad B \in \operatorname{bas}(V), \varphi \in \mathrm{GL}(V).$$

In particular, the density is preserved by the subgroup

$$\mathrm{VGL}(V) = \{g \in \mathrm{GL}(V) \colon |\det(g)| = 1\}.$$

(b) A *volume form on* V is a non-zero n-linear alternating function $\mu \colon V^n \to \mathbb{R}$. It defines a function

$$\widetilde{\mu} \colon \operatorname{bas}(V) \to \mathbb{R}, \quad B = (b_1, \ldots, b_n) \to \mu(b_1, \ldots, b_n)$$

satisfying

$$\widetilde{\mu}(\varphi B) = \det(\varphi)\widetilde{\mu}(B) \quad \text{for} \quad B \in \operatorname{bas}(V), \varphi \in \mathrm{GL}(V).$$

For vectors $v_j = \sum_i a_{ji} b_i$ and $A = (a_{ij})$, expansion of the n-linear form yields

$$\mu(v_1, \ldots, v_n) = (\det A)\mu(b_1, \ldots, b_n) = (\det A)\widetilde{\mu}(B),$$

so that μ is completely determined by the function $\widetilde{\mu}$. We conclude that a volume form μ on V is preserved by the subgroup

$$\mathrm{SL}(V) := \{g \in \mathrm{GL}(V) \colon \det(g) = 1\}.$$

Definition 3.4. A pair (V, β) of a $\mathbb{K}$-vector space V and a symmetric bilinear form $\beta \colon V \times V \to \mathbb{K}$ is called a *quadratic vector space* and

$$\mathrm{O}(V, \beta) := \{\varphi \in \mathrm{GL}(V) \colon (\forall v, w \in V)\, \beta(\varphi v, \varphi w) = \beta(v, w)\}$$

is called the *isometry group* or the *orthogonal group* of (V, β). The symmetric bilinear form β is called *degenerate* if there exists a vector $v \in V \setminus \{0\}$ with $\beta(v, w) = 0$ for all $w \in V$. Otherwise it is called *non-degenerate*.

Remark 3.5. (a) For every non-degenerate symmetric bilinear form β on an n-dimensional vector space V, there exists an ordered basis $B = (b_1, \ldots, b_n)$ of V such that

$$\beta(b_i, b_j) = \epsilon_i\, \delta_{ij}, \quad \epsilon_i = \begin{cases} 1 & i \in \{1, \ldots, p\} \\ -1 & i \in \{p+1, \ldots, n\}. \end{cases}$$

Such a basis is called an *orthonormal basis* for β. It is obtained from a given ordered basis $C = (c_1, \ldots, c_n)$ by the generalised *Gram–Schmidt process*. The numbers $p, n-p$ are independent of the choice of orthonormal basis, and the pair $(p, n-p)$ is is called the *signature* of β.

(b) A symmetric non-degenerate bilinear form β of signature $(n, 0)$ is called a *scalar product* on V, and a vector space with a scalar product is called a *Euclidean vector space*. A symmetric bilinear form β of signature $(1, n-1)$ on V is called a *Minkowski metric* on V and V a *Lorentzian* vector space.

Definition 3.6. A pair (V,ω) of a real vector space V and a non-degenerate alternating bilinear form $\omega\colon V \times V \to \mathbb{K}$ is called a *symplectic vector space* and

$$\mathrm{Sp}(V,\omega) := \{\varphi \in \mathrm{GL}(V)\colon \omega(\varphi v, \varphi w) = \beta(v,w)\ \forall v,w \in V\}$$

is called the *symplectic group* of (V,ω).

Remark 3.7. (a) Every symplectic vector space (V,ω) is even-dimensional.[6]

(b) Let (V,ω) be a real a symplectic vector space of dimension $2n$. A *Darboux basis* of V is an ordered basis $B = (b_1,\dots.b_{2n})$ such that

$$\omega(b_i,b_j) = \begin{cases} 1 & j = i+n \\ -1 & j = i-n \\ 0 & \text{otherwise} \end{cases} .$$

Every symplectic vector space has a Darboux basis, and there is an algorithm which allows one to transform a given ordered basis $C = (c_1,\dots,c_{2n})$ into a Darboux basis. This can be viewed as the symplectic counterpart of the Gram–Schmidt process (see Exercise 3.1).

We summarise the relevant structures on an n-dimensional vector space V and the associated structure preserving subgroups of $\mathrm{GL}(V)$ in the following table:

Structure	**Data**	**Structure Preserving Subgroup** $G \subset \mathrm{GL}(V)$
orientation	ordered basis on V	$\mathrm{GL}_+(V) = \{\varphi \in \mathrm{GL}(V):\ \det\varphi > 0\}$
density	function $\delta:\ \mathrm{bas}(V) \to \mathbb{R}_+^\times$ $\partial(\varphi B) = \lvert\det\varphi\rvert\delta(B)$	$\mathrm{VGL}(V) = \{\varphi \in \mathrm{GL}(V):\ \lvert\det\varphi\rvert = 1\}$
volume form	non-zero n-linear alternating function $\mu : V^n \to \mathbb{R}$ $\mu \circ (\varphi \times \ldots \times \varphi) = \det(\varphi)\cdot\mu$	$\mathrm{SL}(V) = \{\varphi \in \mathrm{GL}(V):\ \det\varphi = 1\}$ $\cong \mathrm{SL}_n(\mathbb{R})$
quadratic vector space	symmetric bilinear form $\beta : V \times V \to \mathbb{R}$	isometry group $\mathrm{O}(V,\beta) = \{\varphi \in \mathrm{GL}(V) : \beta\circ(\varphi\times\varphi) = \beta\}$
special cases:		
Euclidean vector space	positive definite symmetric bilinear form $\beta : V \times V \to \mathbb{R}$	$\mathrm{O}(V,\beta) \cong \mathrm{O}_n$
Lorentzian vector space	non-degenerate symmetric bilinear form $\beta : V \times V \to \mathbb{R}$ of signature $(1,n-1)$	$\mathrm{O}(V,\beta) \cong \mathrm{O}_{1,n-1}$
symplectic vector space	non-degenerate alternating bilinear form $\omega : V \times V \to \mathbb{R}$	symplectic group $\mathrm{Sp}(V,\omega) = \{\varphi \in \mathrm{GL}(V) : \omega\circ(\varphi\times\varphi) = \omega\} \cong \mathrm{Sp}_n(\mathbb{R})$

[6] This is true more generally for symplectic vector spaces over a field $\mathbb{K}$ as long as $\mathbb{K}$ is of characteristic $\mathrm{char}(\mathbb{K}) \neq 2$.

Exercises for Section 3.1

Exercise 3.1. Show that every symplectic vector space (V, ω) is even-dimensional. Describe an algorithm that transforms an ordered basis $C = (c_1, \ldots, c_{2n})$ of V into a Darboux basis.

Exercise 3.2. Consider $\mathbb{R}^{2n}$ with the standard symplectic form

$$\omega(e_i, e_j) = \begin{cases} 1 & j = i + n \\ -1 & j = i - n \\ 0 & \text{otherwise.} \end{cases}$$

Show that every linear map $\varphi \in \mathrm{Sp}_{2n}(\mathbb{R})$ is orientation preserving and volume preserving. Compute the group $\mathrm{Sp}_{2n}(\mathbb{R})$ explicitly for $n = 1$ and $n = 2$.

Exercise 3.3. Let β be a non-degenerate symmetric bilinear form of signature $(1, n-1)$, $n \geq 1$, on an n-dimensional vector space V . Show that the restriction of β to the orthogonal complement

$$v^{\perp_\beta} = \{w \in V : \ \beta(v, w) = 0\}$$

of any vector v with $\beta(v, v) > 0$ is of signature $(0, n-1)$.

Exercise 3.4. Let (V, β) be a Lorentzian vector space. Show that there exist vectors $v, w \in V$ with

$$|\beta(v + w, v + w)| > |\beta(v, v)| + |\beta(w, w)|.$$

In other words: there is no counterpart of the triangle inequality for Lorentzian vector spaces.

Exercise 3.5. Let (V, β) be a Lorentzian vector space.

(a) Show that the set $M = \{v \in V : \ \beta(v, v) > 0\}$ of timelike vectors has two connected components and that two timelike vectors $v, w \in V$ are in the same connected component if and only if $\beta(v, w) > 0$. Conclude that the relation $v \sim w$ if $\beta(v, w) > 0$ defines an equivalence relation on the set of timelike vectors and that there are exactly two equivalence classes.

(b) Let $v \in V$ be a timelike vector. Show that each vector $w \in V \setminus \{0\}$ with $\beta(w, w) \geq 0$ satisfies either $\beta(v, w) > 0$ or $\beta(v, w) < 0$.

3.2 Geometric structures on manifolds

In this section, we show how the structures on vector spaces can be generalised to corresponding structures on smooth manifolds. The general principle is the same as in Sections 1.1 to 1.4, where we defined the relevant structures locally by means of charts in such a way that they did not depend on the choice of chart and then extended them to the whole manifold.

The only difference is that the structures in the previous subsection are associated with vector spaces. Their generalisations to manifolds should therefore live on the tangent bundle $T(M)$ and be defined in terms of vector fields on M, which take the role of the charts in Sections 1.1 to 1.4.

The basic idea is to use collections of smooth vector fields on open subsets $U \subset M$ which define a basis of $T_p(M)$ for each $p \in U$. By means of these vector fields, we can then identify the tangent spaces $T_p(M)$ with $\mathbb{R}^n$ and transport the structures on vector spaces to the tangent bundle $T(U)$.

Definition 3.8. Let M be a smooth n-dimensional manifold. A *local frame* on M is an open subset $U \subset M$ together with an ordered n-tuple of smooth vector fields $(X_1, \ldots, X_n)$ on U such that $X_1(p), \ldots, X_n(p)$ form an ordered basis of $T_p(M)$ for all $p \in U$. For two local frames $\alpha = (U, X_1, \ldots, X_n)$, $\beta = (V, Y_1, \ldots, Y_n)$ with $U \cap V \neq \emptyset$, we have

$$Y_i(p) = \sum_{j=1}^{n} \theta^{\beta\alpha}_{ji}(p) X_j(p) \qquad \forall p \in U \cap V.$$

with smooth matrix valued functions $\theta^{\alpha,\beta} : U \cap V \to \mathrm{GL}_n(\mathbb{R})$. The functions $\theta^{\alpha,\beta} : U \cap V \to \mathrm{GL}_n(\mathbb{R})$ are called *transition functions* for the local frames α, β.

Remark 3.9. (a) A local frame $\alpha = (U, X_1, \ldots, X_n)$ on M induces a smooth map $\Phi_\alpha : T(U) \to \mathbb{R}^n$ whose restriction to $T_p(M)$ is the linear isomorphism

$$\Phi_\alpha|_{T_p(M)} : T_p(M) \to \mathbb{R}^n, \qquad \sum_{i=1}^{n} v_i X_i(p) \mapsto \sum_{i=1}^{n} v_i e_i.$$

If $\alpha = (U, X_1, \ldots, X_n)$, $\beta = (V, Y_1, \ldots, Y_n)$ are local frames with $U \cap V \neq \emptyset$, then the associated maps $\Phi_\alpha|_{T(U\cap V)}, \Phi_\beta|_{T(U\cap V)} : T(U \cap V) \to \mathbb{R}^n$ are related by the transition functions

$$\Phi_\beta|_{T_p(M)} = \theta^{\beta\alpha}(p) \cdot \Phi_\alpha|_{T_p(M)} \qquad \forall p \in U \cap V.$$

(b) For all local frames $\alpha = (U, X_1, \ldots, X_n)$, $\beta = (V, Y_1, \ldots, Y_n)$ with $U \cap V \neq \emptyset$, we have $\theta^{\alpha\beta} = \iota \circ \theta^{\beta\alpha}$, where $\iota : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$, $g \mapsto g^{-1}$ denotes the inversion map. In particular, we have $\theta^{\alpha\alpha}(p) = \mathrm{id}_{\mathbb{R}^n}$ for all $p \in U$.

(c) If (φ, U) is a chart on M, then the φ-basic vector fields form a local frame $(U, b_1^\varphi, \ldots, b_n^\varphi)$. Remark 1.32 implies that the transition functions between the local frames α, β associated with two charts (U, φ) and (V, ψ) are given by $\theta^{\alpha\beta} = \mathtt{d}_{\psi(p)}(\varphi \circ \psi^{-1})$.

Given a local frame $\alpha = (U, X_1, \ldots, X_n)$ on M and one of the structures from the previous subsection on $\mathbb{R}^n$, we can use the linear isomorphisms $\Phi_\alpha|_{T_p(M)} : T_p(M) \to \mathbb{R}^n$ to define a corresponding structure on each tangent space. To illustrate the general pattern, we consider the example of a quadratic form $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. In this case, we obtain a quadratic form g^α_p on each tangent space $T_p(M)$, $p \in U$, by setting

$$g^\alpha_p(v, w) := g(\Phi_\alpha(v), \Phi_\alpha(w)) \qquad \forall v, w \in T_p(M), p \in U.$$

Clearly, g_p depends smoothly on p due to the smoothness of the vector fields $X_1, \ldots, X_n \in \mathcal{V}(\mathcal{U})$. We can now define quadratic form on $T_p(M)$ for each $p \in M$ by covering M with the open domains of local frames. Given two local frames $\alpha = (U, X_1, \ldots, X_n)$, $\beta = (V, Y_1, \ldots, Y_n)$ and a point $p \in U \cap V$, it is natural to ask how the associated quadratic forms g^α_p, g^β_p on $T_p(M)$ are related on $U \cap V$. From Remark 3.9 we obtain

$$\begin{aligned} g^\beta_p(v, w) &= g(\Phi_\beta(v), \Phi_\beta(w)) = g(\theta^{\beta\alpha}(p)\Phi_\alpha(v), \theta^{\beta\alpha}(p)\Phi_\alpha(w)) \\ g^\alpha_p(v, w) &= g(\Phi_\alpha(v), \Phi_\alpha(w)) \quad \forall v, w \in T_p(M). \end{aligned}$$

To ensure that the quadratic forms on the overlap of two local frames do not depend on the choice of the frame, we have to require that the transition functions $\theta^{\beta\alpha} : U \cap V \to \mathrm{GL}_n(\mathbb{R})$ take values in the isometry group $\mathrm{O}(\mathbb{R}^n, g) \subset \mathrm{GL}_n(\mathbb{R})$. This motivates the following definition.

Definition 3.10. (G-structure) Let M be an n-dimensional smooth manifold and G a subgroup of $\mathrm{GL}_n(\mathbb{R})$. Two local frames $\alpha = (U, X_1, .., X_n)$, $\beta = (V, Y_1, \ldots, Y_n)$ on M are called *G-compatible* if their transition functions $\theta_{\alpha\beta} : U \cap V \to \mathrm{GL}_n(\mathbb{R})$ take values in G. A *G-structure* on M is a maximal family $\mathcal{G} = (\alpha_i)_{i\in I}$ of local frames $\alpha_i = (U_i, X_1^{(i)}, \ldots, X_n^{(i)})$ which are pairwise G-compatible, and whose domains cover M: $M = \bigcup_{i\in I} U_i$. Maximal means that every local frame that is G-compatible with all local frames in $\mathcal{G}$ is already contained in $\mathcal{G}$.

Example 3.11. (a) Every family of pairwise G-compatible local frames on M whose domains cover M defines a unique G-structure on M. The proof is analogous to the one for for C^k-atlases. In practice it is advantageous to use as few local frames as possible.

(b) It follows from Remark 3.9 (c) that every smooth n-dimensional manifold has a unique $\mathrm{GL}_n(\mathbb{R})$-structure, which is defined by the local frames associated to the charts (φ, U) of M.

(c) A $\mathrm{GL}_+(\mathbb{R}^n)$-structure on M is called an *orientation* on M. For each local frame $(U, X_1, \ldots, X_n)$ on M, the tangent vectors $X_1(p), \ldots, X_n(p)$ define an orientation of $T_p(M)$, and the requirement that the transition functions between two local frames $\alpha = (U, X_1, \ldots, X_n)$ and $\beta = (V, Y_1, \ldots, Y_n)$ take values in $\mathrm{GL}_+(\mathbb{R}^n)$ ensures that the orientations for α and β agree on $T_p(M)$ for all $p \in U \cap V$.

(c) An $\{\mathrm{id}_{\mathbb{R}^n}\}$-structure on M is equivalent to the existence of a *global frame*, i. e. a frame with domain M. This implies that the tangent bundle $T(M)$ is diffeomorphic to $M \times \mathbb{R}^n$.

(d) An $\mathrm{SL}_n(\mathbb{R})$-structure on M is equivalent to the existence of a *volume form* on M, i. e. an assignment of an alternating n-form $\mathrm{vol}_p \in \mathrm{Alt}^n(T_p(M), \mathbb{R})$ to each point $p \in M$ such that for each chart (φ, U) of M the function $\mathrm{vol}^\varphi : p \mapsto \mathrm{vol}_p(b_1^\varphi, \ldots, b_n^\varphi)$ is smooth.

The cases $G = \mathrm{O}(p, q)$ and $G = \mathrm{Sp}_{2n}(\mathbb{R})$ are particularly relevant to physics, since they are related, respectively, to the concepts of a semi-Riemannian manifold and an (almost) symplectic manifold.

Definition 3.12. (a) A *semi-Riemannian* manifold is a smooth manifold M together with an assignment $g : p \to g_p$ of a non-degenerate symmetric bilinear form g_p on $T_p(M)$ to each point $p \in M$ such that for all charts (φ, U) of M the coefficient functions $g_{ij}^\varphi : U \to \mathbb{R}$, $p \mapsto g_p(b_i^\varphi, b_j^\varphi)$ are smooth. The map $g : p \to g_p$ is called *(semi-Riemannian) metric* on M.

(b) An *almost symplectic manifold* is a smooth n-dimensional manifold with an assignment $\omega : p \mapsto \omega_p$ of a non-degenerate alternating bilinear form ω_p on $T_p(M)$ to each point of M such that for all charts (φ, U) the coefficient functions ω_{ij}^φ, $p \mapsto \omega_p(b_i^\varphi, b_j^\varphi)$ are smooth. The map $\omega : p \mapsto \omega_p$ is called an *almost symplectic form* on M.

Remark 3.13. (a) The smoothness of the coefficient functions $g_{ij}^\varphi : U \to \mathbb{R}$ implies that the non-degenerate symmetric bilinear form g_p has the same signature for all $p \in M$. The signature of the semi-Riemannian manifold (M, g) is defined as the signature of g_p.

(b) A semi-Riemannian manifold (M, g) of signature $(1, q)$ $(q \geq 1)$ is called a *Lorentzian manifold* and g is called a *Lorentzian metric* on M. A semi-Riemannian manifold of signature $(q, 0)$ is called a *Riemannian manifold* and g is called a *Riemannian metric* on M.

(c) In the physics literature, the metric g of a semi-Riemannian manifold (M, g) is often denoted by ds^2, and on the domain of each chart (φ, U) one writes

$$ds^2 = g_{ij}^\varphi dx^i dx^j \quad \text{for} \quad g_p(b_i^\varphi(p), b_j^\varphi(p)) = g_{ij}^\varphi(p) \qquad \forall p \in U.$$

Example 3.14. (a) We consider $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. For $p \in \mathbb{S}^n$, the tangent space $T_p(\mathbb{S}^n)$ is the orthogonal complement $p^\perp = \{x \in \mathbb{R}^{n+1} : \langle p, x\rangle = 0\}$, where $\langle\cdot,\cdot\rangle$ is the Euclidean metric on $\mathbb{R}^{n+1}$. A Riemannian metric on $\mathbb{S}^n$ is given by the restriction of $\langle\cdot,\cdot\rangle$ to $p^\perp$:

$$g_p(x, y) = \langle x, y\rangle \qquad \forall x, y \in p^\perp.$$

(b) More generally, for any submanifold M of $\mathbb{R}^n$, the restriction of the Euclidean metric on $\mathbb{R}^n$ to $T_p(M)$ defines a Riemannian metric on M (see Exercise 3.7).

(c) We consider the n-dimensional hyperbolic space $\mathbb{H}^n = \{x \in \mathbb{R}^{n+1} : \langle x, x\rangle_M = 1, x_0 > 0\}$, where $\langle\ ,\ \rangle_M$ denotes the Minkowski metric on $\mathbb{R}^{n+1}$: $\langle x, y\rangle_M = x_0y_0 - \sum_{i=1}^n x_iy_i$. Then the tangent space $T_p(\mathbb{H}^n)$ is the orthogonal complement

$$p^{\perp_M} = \{x \in \mathbb{R}^{n+1} : \langle x, p\rangle_M = 0\},$$

and the restriction of the Minkowski metric to $p^{\perp_M}$ defines a metric of signature $(0, -n)$ on $\mathbb{H}^n$ (see Exercise 3.6).

(d) Let $M = \{x \in \mathbb{R}^3 : \langle x, x\rangle_M = 0, x_0 > 0\}$ with $\langle x, y\rangle_M = x_0y_0 - x_1y_1 - x_2y_2$ be a cone in $\mathbb{R}^3$. Then the tangent space $T_p(M)$ can be identified with the plane

$$T_p(M) = \{x \in \mathbb{R}^3 : \langle p, x\rangle_M = 0\} = \mathrm{span}\{p, (0, -p_2, p_1)\}.$$

The Minkowski metric does not induce a metric on M because the restriction $\langle\cdot,\cdot\rangle|_{T_p(M)\times T_p(M)}$ is degenerate: $\langle p, y\rangle_M = 0$ for all $y \in T_p(M)$.

(e) We consider the tangent bundle $M = TN$ of an n-dimensional smooth manifold N. Then in the domain of each chart (φ, U) of N we can identify $TU \cong U \times \mathbb{R}^n$ and the tangent space $T_{(p,v)}(T(U))$ with $T_{(p,v)}(TU) \cong T_p(U) \times \mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n \cong T_p(U) \times T_p(U)$. The pairs of φ-basic vector fields $(b_i^\varphi, b_j^\varphi)$, $i, j \in \{1, ..., n\}$ form a local frame on $TU \subset M$. With the definition

$$\omega((b_i^\varphi, b_j^\varphi), (b_k^\varphi, b_l^\varphi)) = \delta_{il} - \delta_{jk}$$

we obtain a symplectic form on $TU \subset M$. It is easy to show that this symplectic form is independent of the choice of the chart and defines a symplectic form on $M = TN$.

As already suggested by the discussion at the beginning of this subsection, a semi-Riemannian metric of signature (p, q) on M corresponds to an $\mathrm{O}_{p,q}$-structure on M and an almost symplectic form on M to a Sp_{2m}-structure on M. We have the following proposition.

Proposition 3.15. *Let M be a smooth n-dimensional manifold. An $\mathrm{O}_{p,q}$-structure on M with $n = p + q$ corresponds to a semi-Riemannican metric of signature (p, q) on M, and an $\mathrm{Sp}_{2m}(\mathbb{R})$-structure with $n = 2m$ to an almost symplectic form on M.*

Proof. (1) Let M be equipped with a $O_{p,q}$-structure and $(U, X_1, \ldots, X_n)$ a local frame compatible with the $O_{p,q}$-structure. Then we define for each $p \in U$ a symmetric bilinear form of signature (p, q) on $T_p(M)$ by

$$g_p(X_i(p), X_j(p)) = \epsilon_i\, \delta_{ij} \quad \text{where} \quad \epsilon_i = \begin{cases} 1 & i \in \{1, \ldots, p\} \\ -1 & i \in \{p + 1, \ldots, n\} \end{cases}.$$

For each chart (φ, W) of M with $U \cap W \neq \emptyset$, the coefficient functions g_{ij}^φ are given by the expressions for the associated φ-basic vector fields in terms of the vector fields X_i. On

$U \cap W$, we have $b_i^\varphi(p) = \sum_{k=1}^n a_{ij}(p) X_j(p)$ with smooth functions $a_{ij} \in C^\infty(U \cap W, \mathbb{R})$, and the coefficient functions are given by

$$g_{ij}^\varphi(p) = g_p(b_i^\varphi(p), b_j^\varphi(p)) = \sum_{k,l=1}^n a_{ik}(p) a_{jl}(p) g_p(X_k(p), X_l(p)) = \sum_{k=1}^n a_{ik}(p) a_{jk}(p) \epsilon_k$$

This implies in particular that the coefficient functions are smooth. If $(V, Y_1, \ldots, Y_n)$ is another local frame with $U \cap V \neq \emptyset$ that is $O_{p,q}$-compatible to X , then we have for $p \in U \cap V$

$$g_p(Y_i(p), Y_j(p)) = g_p\left(\theta^{\alpha\beta}(p) X(p), \theta^{\alpha\beta}(p) X_j(p)\right) = g_p(X_i(p), X_j(p)) = \epsilon_i\, \delta_{ij}. \qquad (15)$$

The symmetric bilinear form on $T_p(M)$ is thus independent of the choice of the local frame, and we obtain a semi-Riemannian metric on M.

Conversely, given a semi-Riemannian metric on M and a chart (φ, U) on M, we apply for each $p \in U$ the Gram–Schmidt process to the ordered basis $b_1^\varphi(p), \ldots, b_n^\varphi(p)$. As the φ-basic vector fields and the coefficient functions g_{ij}^φ are smooth, this yields smooth vector fields $X_1, \ldots, X_n \in \mathcal{V}(\mathcal{U})$ which satisfy $g_p(X_i(p), X_j(p)) = \epsilon_i\, \delta_{ij}$ for all $p \in U$. On the overlap of the domains of two local frames $(U, X_1, \ldots, X_n)$, $(V, Y_1, \ldots, Y_n)$ with this property, equation (15) then implies that the transition functions $\theta^{\alpha\beta}$ take values in $\mathrm{O}_{p,q}$.

(2) The proof for the almost symplectic case is analogous. Given a $\mathrm{Sp}_{2m}(\mathbb{R})$-structure on M and a frame $(U, X_1, \ldots, X_{2m})$ that is $\mathrm{Sp}_{2m}(\mathbb{R})$-compatible with this structure, we define an almost symplectic form ω_p on $T_p(M)$ by setting

$$\omega_p(X_i(p), X_j(p)) = \begin{cases} 1 & j = i + m \\ -1 & j = i - m \\ 0 & \text{otherwise.} \end{cases} \qquad (16)$$

The required properties of the almost symplectic form $\omega : p \mapsto \omega_p$ then follow as in the semi-Riemannian case.

Conversely, given an almost symplectic form $\omega : p \mapsto \omega_p$ on M and a chart (φ, U) on M, we apply for each $p \in U$ the symplectic counterpart of the Gram–Schmidt process to the ordered basis $b_1^\varphi(p), \ldots, b_n^\varphi(p)$ to obtain a Darboux basis (see Exercise 3.1). The smoothness of the φ-basic vector fields and of the coefficient functions ω_{ij}^φ ensures that the resulting vector fields are smooth. This defines a local frame $(U, X_1, \ldots, X_n)$ which satisfies (16). For any two such frames with overlapping domains, one finds that the transition functions take values in $\mathrm{Sp}_{2m}(\mathbb{R})$. □

Remark 3.16. It is also possible to consider subgroups $G \subset \mathrm{GL}_n(\mathbb{R})$ that are obtained as intersections $G = G_1 \cap G_2$ of two subgroups introduced above. In that case, a manifold M with a G-structure exhibits both structures associated with the subgroups G_1 and G_2. For instance, we have $\mathrm{SO}_n = \mathrm{GL}_+(\mathbb{R}^n) \cap \mathrm{O}_n$. A SO_n-structure on M corresponds to a metric and an orientation on M. Similarly, we have $L_+ = \mathrm{SO}_{1,n-1} \cap \mathrm{GL}_+(\mathbb{R}^n)$ and $L_+^\uparrow = L_+ \cap L^\uparrow$. An L_+-structure on M therefore consists of a Lorentzian metric on M together with an orientation and an $L_+^\uparrow$-structure on M of a Lorentzian metric on M together with an orientation and a time orientation. Some examples are given in the following table.

Group G	Structure on manifold	Structure on tangent spaces
$G = \mathrm{GL}_+(\mathbb{R}^n)$	orientation	oriented bases of $T_p(M)$
$G = \mathrm{O}_{p,q}$	semi-Riemannian manifold	symmetric, non-degenerate bilinear form $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$ of signature (p, q)
$G = \mathrm{O}_{1,n}$	Lorentzian manifold	symmetric, non-degenerate bilinear form $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$ of signature $(1, n)$
$G = \mathrm{O}_n$	Riemannian manifold	scalar product $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$
$G = \mathrm{Sp}_{2m}(\mathbb{R})$	almost symplectic manifold	almost symplectic form $\omega_p : T_p(M) \times T_p(M) \to \mathbb{R}$
$G = \mathrm{SO}_n$	oriented Riemannian manifold	scalar product $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$ and oriented bases of $T_p(M)$
$G = L^{\uparrow}$	Lorentzian manifold with time orientation	symmetric, non-degenerate bilinear form $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$ of signature $(1, n)$ together with choice of connected component of $\{v \in T_p(M) :\ g_p(v, v) > 0\}$.
$G = L^{\uparrow}_{+}$	oriented Lorentzian manifold with time orientation	symmetric, non-degenerate bilinear form $g_p : T_p(M) \times T_p(M) \to \mathbb{R}$ of signature $(1, n)$ together with oriented basis of $T_p(M)$ and with choice of connected component of $\{v \in T_p(M) :\ g_p(v, v) = 0\}$.

After developing the concept of a manifold with a G-structure, it is natural to investigate smooth maps between manifolds with G-structures that preserve G-structures. Let $(M, \mathcal{G}^M)$, $(N, \mathcal{G}^N)$ be manifolds with a G-structures and $f : M \to N$ a smooth map. If f is compatible with the G-structures, it should relate the local frames in $\mathcal{G}^M$ to the ones in $\mathcal{G}^N$. Concretely, for all local frames $(U, X_1, \ldots, X_n)$ of X there is a local frame $(V, Y_1, \ldots, Y_n)$ of N with $V \cap f(U) \neq \emptyset$ and X_i and Y_i are f-related. As local frames are required to form a basis of $T_p(M)$ for all points in their domain, this implies already that $T_p(f) : T_p(M) \to T_{f(p)}N$ is an isomorphism for all $p \in M$. By the Inverse Function Theorem, it follows that $f : M \to N$ is a local diffeomorphism, i. e. that for every point $p \in M$ there exists an open neighbourhood U of p such that $f|_U : U \to f(U)$ is a diffeomorphism.

Definition 3.17. Let $G \subset \mathrm{GL}_n(\mathbb{R})$ be a subgroup and M, N smooth manifolds with G-structures $\mathcal{G}^M$, $\mathcal{G}^N$. A *local isomorphism* of G-structures is a smooth map $f : M \to N$ such that for all points $p \in M$ there exists an open neighbourhood U such that $f|_U : U \to f(U)$ is a diffeomorphism and a local frame $\alpha = (U, X_1, \ldots, X_n)$ in $\mathcal{G}^M$ such that

$$f_*\alpha = (f(U), f_*X_1, \ldots, f_*X_n)$$

is a local frame in $\mathcal{G}^N$. An *isomorphism of G structures* is a local isomorphism of G-structures that is a diffeomorphism. A (local) isomorphism of $O_{p,q}$-structures is called a (local) *isometry*, a (local) isomorphism of Sp_{2m}-structures a *symplectomorphism*.

Remark 3.18. (a) For any smooth manifold M with a G-structure, the isomorphisms $f : M \to M$ of G-structures form a group, denoted $\mathrm{Aut}_{\mathcal{G}}(M)$. If M, N are smooth manifolds with G-structures, then an isomorphism $f : M \to N$ of G-structures defines a group homomorphism $\Phi_f : \mathrm{Aut}_G(M) \to \mathrm{Aut}_G(N)$, $\varphi \mapsto f \circ \varphi \circ f^{-1}$. The group homomorphisms Φ_f are functorial: if $g : N \to P$ is another isomorphism of G structures, then $\Phi_{g\circ f} = \Phi_g \circ \Phi_f$.

(b) If M is a smooth manifold, N a smooth manifold with a G-structure $\mathcal{G}^N$ and $f : M \to N$ a local diffeomorphism, then there exists a unique G-structure on M such that $f : M \to N$ is a (local) isomorphism of G-structures. This G-structure is called the *pull-back* of $\mathcal{G}^N$.

Example 3.19. (a) Let (M, g), (N, h) be semi-Riemannian manifolds. Then it follows from the proof of Proposition 3.15 that a diffeomorphism $f : M \to N$ is a (local) isometry if and only if the linear isomorphism $T_p(f) : T_p(M) \to T_{f(p)}(N)$ is an isometry for all $p \in M$:

$$h_{f(p)}(T_p(f)v, T_p(f)w) = g_p(v, w) \qquad \forall v, w \in T_p(M).$$

The group of isometries of a semi-Riemannian manifold (M, g) is called the *isometry group* of M and denoted $\mathrm{Isom}(M)$. If M is a smooth manifold, (N, h) a semi-Riemannian manifold of signature (p, q) and $f : M \to N$ a local diffeomorphism, then the pull-back of the $\mathrm{O}_{p,q}$-structure on N determines a metric f_*h of signature (p, q) on M. This metric on M is called the pull-back of h by f and given by

$$(f_*h)_p(v, w) = h(T_p(f)v, T_p(f)w) \qquad \forall p \in M, v, w \in T_p(M).$$

(b) Similarly, we find for almost symplectic manifolds (M, ω) and (N, η) that a local diffeomorphism $f : M \to N$ is a symplectomorphism if and only if the linear isomorphism $T_p(f) : T_p(M) \to T_p(N)$ satisfies

$$\eta_{f(p)}(T_p(f)v, T_p(f)w) = \omega_p(v, w) \qquad \forall v, w \in T_p(M).$$

If M is a smooth manifold and (N, η) an almost symplectic manifold, then the pull-back of the Sp_{2m}-structure on N determines an almost symplectic form $f_*\eta$ on M

$$(f_*\eta)_p(v, w) = \eta_{f(p)}(T_p(f)v, T_p(f)w) \qquad \forall v, w \in T_p(M).$$

Example 3.20. (a) We consider the n-Sphere $\mathbb{S}^n$ with the Riemannian metric induced by the Euclidean scalar product on $\mathbb{R}^{n+1}$

$$g_p(x, y) = \langle x, y\rangle \qquad \forall x, y \in T_p(\mathbb{S}^n) \cong p^\perp.$$

Then a smooth map $f : \mathbb{S}^n \to \mathbb{S}^n$ is an isometry if and only if

$$\langle T_p(f)v, T_p(f)w\rangle = \langle v, w\rangle \qquad \forall p \in \mathbb{S}^n, v, w \in p^\perp.$$

We will show later (see Lemma 3.44) that an isometry on a connected semi-Riemannian manifold is determined uniquely by $f(p)$ and $T_p(f) : T_p(M) \to T_p(M)$ for a given point $p \in M$. As O_{n+1} acts transitively on $\mathbb{S}^n$, there is a unique element $A \in \mathrm{O}_{n+1}$ with $f(p) = Ap$ and $T_p(f)x = Ax$ for all $x \in T_p(M)$. This implies $f(q) = Aq$ for all $q \in \mathbb{S}^n$. Conversely, for every $A \in \mathrm{O}_{n+1}$, $f : \mathbb{S}^n \to \mathbb{S}^n$, $q \mapsto Aq$ is an isometry. The isometry group of the sphere $\mathbb{S}^n$ is therefore given by $\mathrm{Isom}(\mathbb{S}^n) \cong \mathrm{O}_{n+1}$.

(b) We consider n-dimensional hyperbolic space $\mathbb{H}^n$ with the metric of signature $(0, -n)$ induced by the Minkowski metric on $\mathbb{R}^{n+1}$. Then a smooth map $f : \mathbb{H}^n \to \mathbb{H}^n$ is an isometry if and only if

$$\langle T_p(f)v, T_p(f)w\rangle_M = \langle v, w\rangle_M \qquad \forall p \in \mathbb{H}^n, v, w \in p^{\perp_M}.$$

As in the case of the sphere, this implies that f is of the form $f(x) = Ax$ with $A \in \mathrm{O}_{1,n}$. To map the hyperboloid $\mathbb{H}^n$ to itself, f must preserve the time orientation, which implies $A \in L^{\uparrow}$. Conversely, for every $A \in L^{\uparrow}$, $f : \mathbb{H}^n \to \mathbb{H}^n$, $f(q) = Aq$ is an isometry. This implies that the isometry group of $\mathbb{H}^n$ is $\mathrm{Isom}(\mathbb{H}^n) \cong L^{\uparrow}$.

Exercises for Section 3.2

Exercise 3.6. Show that the Minkowski metric on $\mathbb{R}^{n+1}$

$$\langle x, y\rangle_M = x_0 y_0 - \sum_{i=1}^{n} x_i y_i$$

induces a metric of signature $(0, -n)$ on n-dimensional hyperbolic space

$$\mathbb{H}^n = \{x \in \mathbb{R}^{n+1} : \langle x, x\rangle_M = 1, x_0 > 0\}.$$

Hint: It is sufficient to show that the restriction of $\langle\ ,\ \rangle_M$ to

$$T_p(\mathbb{H}^n) = \{x \in \mathbb{R}^{n+1} : \langle p, x\rangle_M = 0\}$$

is non-degenerate and of signature $(0, -n)$.

Exercise 3.7. Let M be a submanifold of $\mathbb{R}^n$. Show that the restriction of the Euclidean metric on $\mathbb{R}^n$ to $T_p(M)$ defines a Riemannian metric on M. **Hint:** Recall Example 1.24.

Exercise 3.8. Let (M, g) be a Lorentzian manifold. A *timelike vector field* on M is a smooth vector field $X \in \mathcal{V}(M)$ such that $g_p(X(p), X(p)) > 0$ for all $p \in M$. A Lorentzian manifold is called *time-orientable* if it admits a timelike vector field. A *time orientation* of M is a choice of a timelike vector field on M.

(a) Show that a Lorentzian manifold is time-orientable if and only if it has a $L^{\uparrow}$-structure and that each time orientable manifold has exactly two time orientations.

(b) Construct an example of a Lorentzian manifold (M, g) that is not time-orientable.

Hint: Consider the Möbius strip M and construct a Lorentzian metric on M from the Minkowski metric on $\mathbb{R}^2$.

3.3 Semi-Riemannian geometry

In this subsection we will focus on structures associated with semi-Riemannian manifolds (M, g). The first important concept is the notion of a (torsion-free and metric) connection on M, which lies at the foundation of geometric notions such as parallel transport of vector fields, geodesics and curvature. While torsion-free connections exist in the more general context of smooth manifolds and are in general non-unique, the requirement of compatibility with a semi-Riermannian metric selects a unique torsion-free and metric connection, the Levi–Civita connection.

Definition 3.21. Let M be a n-dimensional smooth manifold. A *connection* on M is a map $\nabla : \mathcal{V}(\mathcal{M}) \times \mathcal{V}(\mathcal{M}) \to \mathcal{V}(\mathcal{M})$, $(X, Y) \mapsto \nabla_X Y$ such that

(C1) is $\mathbb{R}$-linear in both arguments: for all $\lambda_1, \lambda_2 \in \mathbb{R}$ and $X_1, X_2, Y_1, Y_2 \in \mathcal{V}(\mathcal{M})$

$$\nabla_{\lambda_1 X_1 + \lambda_2 X_2} Y = \lambda_1 \nabla_{X_1} Y + \lambda_2 \nabla_{X_2} Y, \qquad \nabla_X(\lambda_1 Y_1 + \lambda_2 Y_2) = \lambda_1 \nabla_X Y_1 + \lambda_2 \nabla_X Y_2.$$

(C2) satisfies $\nabla_{fX} Y = f \cdot \nabla_X Y$ and $\nabla_X(f \cdot Y) = f \cdot \nabla_X Y + \mathcal{L}_X f \cdot Y$ for all $f \in C^\infty(M, \mathbb{R})$ and $X, Y \in \mathcal{V}(\mathcal{M})$.

A connection is called *torsion-free* if for all vector fields $X, Y \in \mathcal{V}(\mathcal{M})$

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

A connection ∇ on a semi-Riemannian manifold (M, g) is called a *metric connection* if

$$\mathcal{L}_Z g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) \qquad \forall X, Y, Z \in \mathcal{V}(\mathcal{M}).$$

An important property of a connection is its *locality*, which is guaranteed under an additional assumption on the smooth manifold, namely the the requirement of *paracompactness*. Smooth manifolds without this property are generally considered pathological, and many authors include this property in the definition of a smooth manifold. In the following, we will assume without further mention that all smooth manifolds under consideration are paracompact.

Definition 3.22. A topological space X is called *paracompact* if every open cover $(U_\alpha)_{\alpha \in A}$ of X has a *locally finite refinement*. This means that there exists an open cover $(V_\beta)_{\beta \in B}$ of X such that for every $\beta \in B$ there is an $\alpha \in A$ with $V_\beta \subset U_\alpha$ and for every $p \in X$ there is an open neighbourhood $W_p \subset X$ such that $\{\beta \in B : V_\beta \cap W_p \neq \emptyset\}$ is finite.

It can be shown, see for instance [Br93] or [HN11] p 397 ff, that any smooth paracompact manifold M has smooth *partitions of unity*:

Lemma 3.23. *Let M be a smooth paracompact manifold. Then for any open cover $M \subset \bigcup_{\alpha \in A} U_\alpha$, there exists a* smooth partition of unity*. This is a set of smooth functions $(f_i)_{i \in I}$, $f_i \in C^\infty(M, \mathbb{R})$ such that*

(P1) For all $i \in I$ there exists an $\alpha \in A$ with $\mathrm{supp}(f_i) \subset U_\alpha$.

(P2) For all $p \in M$, we have $f_i(p) \neq 0$ for only finitely many $i \in I$.

(P3) for all $p \in M$ the functions f_i satisfy $0 \leq f_i(p) \leq 1$ for all $i \in I$ and $\sum_{i\in I} f_i(p) = 1$.

This implies that for all open subsets $U_1, U_2 \subset M$ with $\overline{U_1} \subset U_2$ compact, there exists a bump function*, e. g. a smooth function $f \in C^\infty(M,\mathbb{R})$ with $0 \leq f \leq 1$, $f|_{U_1} = 1$ and $f|_{M\setminus U_2} = 0$. Such a bump function is given by*

$$f = \sum_{i\in I, \mathrm{supp}(f_i)\subset U_2} f_i.$$

The existence of smooth partitions of unity and of bump functions implies that connections on a semi-Riemannian manifold (M, g) have a *locality property*, namely that the value of $\nabla_X Y$ in a point p depends only on $X(p)$ and the behaviour of the vector field Y in a small neighbourhood of p. To prove this, we use the following lemma.

Lemma 3.24. *(a) Let M be a smooth manifold and $F : \mathcal{V}(M) \to C^\infty(M,\mathbb{R})$ or $F : \mathcal{V}(M) \to \mathcal{V}(M)$ a linear map that satisfies $F(f \cdot X) = f \cdot F(X)$ for all $X \in \mathcal{V}(M)$ and $f \in C^\infty(M,\mathbb{R})$. Then $F(X)(p)$ depends only on $X(p)$.*

(b) If (M, g) is a semi-Riemannian manifold and $F : \mathcal{V}(M) \to C^\infty(M,\mathbb{R})$ a linear map with $F(f \cdot X) = f \cdot F(X)$ for all $X \in \mathcal{V}(M)$ and $f \in C^\infty(M,\mathbb{R})$, then there exists a unique vector field $Y \in \mathcal{V}(M)$ with $F(X) = g(Y, X)$ for all $X \in \mathcal{V}(M)$.

Proof. (a) We first show that $F(X)(p)$ depends only on $X|_V$ for any open neighbourhood V of p. Let V be an open neighbourhood of p such that $X|_V = 0$. Then by means of bump functions, we can construct a function $f \in C^\infty(M,\mathbb{R})$ with $f|_U = 0$ for an open neighbourhood $U \subset V$ of p and $f|_{M\setminus V} = 1$. This implies $f \cdot X = X$ and therefore

$$F(X)(p) = F(f \cdot X)(p) = f(p) \cdot F(X)(p) = 0.$$

Due to the linearity of F, this implies that $F(X)(p)$ depends only on $X|_V$.

Let now X be a vector field on M with $X(p) = 0$. Then there is a chart (φ, U) with $p \in U$ and $F(X)(p)$ depends only on $X|_U$. The vector field $X|_U$ is given uniquely in terms of the φ-basic vector fields as $X|_U = \sum_{i=1}^n x_i b_i^\varphi$ with $x_i \in C^\infty(U)$, $x_i(p) = 0$. This yields

$$F(X)(p) = F\left(\sum_{i=1}^n x_i \cdot b_i^\varphi\right)(p) = \sum_{i=1}^n x_i(p) F(b_i^\varphi) = 0.$$

(b) To demonstrate that there is a vector field Y on M with $F(X) = g(X, Y)$ for all $X \in \mathcal{V}(M)$, we consider a chart (φ, U) with $p \in U$. Denoting by $g_\varphi^{ij} \in C^\infty(U,\mathbb{R})$ the components of the matrix inverse of the coefficent matrix of g, we define a smooth vector field $Y \in \mathcal{V}(U)$ by

$$Y(p) = \sum_{i,j=1}^n F(b_j^\varphi) g_\varphi^{ij} b_i^\varphi \qquad \text{where} \quad \sum_{j=1}^n g_\varphi^{ij}(p) g^\varphi_{jk}(p) = \delta_{ik} \qquad \forall i, k \in \{1, \ldots, n\},\ p \in U.$$

Then we have for all vector fields $X \in \mathcal{V}(U)$

$$g(X, Y) = \sum_{i,j=1}^n F(b_j^\varphi) g_\varphi^{ij} g(X, b_i^\varphi) = \sum_{i,j,k=1}^n x_k F(b_j^\varphi) g_\varphi^{ij} g^\varphi_{ki} = \sum_{j=1}^n x_j F(b_j^\varphi) = F(X).$$

As g is non-degenerate, two vector fields $Y \in \mathcal{V}(U)$, $Y' \in \mathcal{V}(V)$ with this property must agree on $U \cap V$. We can therefore cover M with charts to obtain a smooth vector field Y on M. $\square$

Using this lemma, we can now prove that any connection on a semi-Riemannian manifold is local:

Lemma 3.25. *Let (M, g) be a semi-Riemannian manifold. If ∇ is a connection on M, $p \in M$ and $X, Y \in \mathcal{V}(M)$, then $\nabla_X Y(p)$ depends only on $X(p)$ and $Y|_V$ for any neighbourhood V of p. We obtain a bilinear map $\nabla : TM \times \mathcal{V}(M) \to \mathcal{V}(M)$, $(v, Y) \mapsto \nabla_v Y$.*

Proof. The property (C2) of a connection together with Lemma 3.24 implies that $\nabla_X Y(p)$ depends only on $X(p)$. To show the second statement, we consider an open neighbourhood of p such that $Y|_V = 0$. Then by means of bump functions, we can construct a function $f \in C^\infty(M, \mathbb{R})$ with $f|_U = 0$ and $f|_{M\setminus V} = 1$ for an open neighbourhood $U \subset V$. This implies $f \cdot Y = Y$ and therefore

$$\nabla_X Y(p) = \nabla_X (f \cdot Y)(p) = \mathcal{L}_X f(p) \cdot Y(p) + f(p) \cdot \nabla_X Y(p) = 0.$$

Hence $\nabla_X Y(p)$ depends only on $Y|_V$ for any neighbourhood V of p. □

The locality properties of a connection allow one to characterise it as a sum of derivatives of vector fields and a component that is function-linear in both arguments. This leads to the notion of Christoffel symbols.

Definition 3.26. Let M be a smooth manifold and ∇ a connection on M. Then, in any local chart, the *Christoffel symbol* associated with ∇ is the smooth bilinear map $\Gamma : \mathcal{V}(M) \times \mathcal{V}(M) \to \mathcal{V}(M)$ defined by

$$\nabla_X Y = \mathtt{d}Y \cdot X + \Gamma(X, Y).$$

Remark 3.27. (a) Due to the properties of the connection and the derivative, we have for all vector fields $X, Y \in \mathcal{V}(M)$

$$\begin{aligned}
&\Gamma(f \cdot X, Y) = f \cdot \nabla_X Y - f \cdot \mathtt{d}Y \cdot X = f \cdot \Gamma(X, Y), \\
&\Gamma(X, f \cdot Y) = f \cdot \nabla_X Y + \mathcal{L}_X f \cdot Y - f \cdot \mathtt{d}Y \cdot X - \mathcal{L}_X f \cdot Y = f \cdot \Gamma(X, Y).
\end{aligned}$$

It follows from Lemma 3.24 that $\Gamma(X, Y)(p)$ depends only on $X(p)$ and $Y(p)$ and hence defines for each $p \in M$ a bilinear map $\Gamma_p : T_p(M) \times T_p(M) \to T_p(M)$.

(b) This implies in particular that we can characterise the Christffel symbols and hence the connection uniquely in terms of the φ-basic vector fields associated with charts (φ, U) on M. Let M be a smooth manifold and (φ, U) a chart on M. The *Christoffel symbols* associated with (φ, U) are the smooth functions $\Gamma^k_{ij} \in C^\infty(U, \mathbb{R})$ defined by

$$\nabla_{b_i^\varphi} b_j^\varphi(p) = \Gamma(b_i^\varphi, b_j^\varphi) = \sum_{k=1}^n \Gamma^k_{ij}(p) b_k^\varphi(p) \qquad \forall p \in U,$$

where $b_1^\varphi, \ldots, b_n^\varphi$ are the φ-basic vector fields on U. For vector fields

$$X = \sum_{i=1}^n x_i b_i^\varphi, \qquad Y = \sum_{i=1}^n y_i b_i^\varphi$$

we obtain

$$\nabla_X Y = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_i \frac{\partial y_j}{\partial \varphi_i} + \sum_{i,k=1}^{n} \Gamma^j_{ik} x_i y_k \right) b^\varphi_j.$$

In physics textbooks connections are often called *covariant derivatives*, and instead of this formula, one often uses the shorthand notation $y^j_{;i} = y^j_{,i} + \Gamma^j_{ik} y^k$, where $y^j_{;i}$ denotes coefficient functions of $\nabla_{b_i^\varphi} Y$, $y^j_{,i} = \frac{\partial y^j}{\partial \varphi_i}$ and the summation over repeated indices is understood (Einstein summation convention).

(c) A short calculation (see Exercise 3.9) shows that a connection ∇ on M is torsion-free if and only if its Christoffel symbols are symmetric

$$\Gamma^k_{ij} = \Gamma^k_{ji} \qquad \forall i, j, k \in \{1, \ldots, n\}$$

for each chart (φ, U) of M. If (M, g) is a semi-Riemannian manifold, a connection on M is a metric connection if and only if for each chart its Christoffel symbols satisfy

$$\frac{\partial g_{ij}}{\partial \varphi_k} = \sum_{l=1}^{n} (\Gamma^l_{kj} g_{li} + \Gamma^l_{ki} g_{lj}) \qquad \forall i, j, k \in \{1, \ldots, n\}.$$

It turns out that the condition of metricity determines a torsion-free connection on a semi-Riemannian manifold (M, g) uniquely and allows one to express the connection as a function of the semi-Riemannian metric g, the Lie derivatives and the Lie bracket on M.

Theorem 3.28. *Let (M, g) be a semi-Riemannian manifold. Then there exists a unique torsion-free, metric connection on M. It is called the* Levi–Civita connection *and determined by the* Koszul-formula

$$2g(\nabla_X Y, Z) = \mathcal{L}_X g(Y, Z) + \mathcal{L}_Y g(Z, X) - \mathcal{L}_Z g(X, Y) - g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]).$$

Proof. (a) Uniqueness: Let ∇ be a torsion-free, metric connection on M. Then we can verify that it satisfies the Koszul-formula by a direct calculation:

$$\begin{aligned}
&\mathcal{L}_X g(Y, Z) + \mathcal{L}_Y g(Z, X) - \mathcal{L}_Z g(X, Y) - g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]) \\
=&g(\nabla_X Y, Z) + g(Y, \nabla_X Z) + g(\nabla_Y Z, X) + g(Z, \nabla_Y X) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) \\
&- g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]) \\
=&g(X, \nabla_Y Z - \nabla_Z Y - [Y, Z]) + g(Y, \nabla_X Z - \nabla_Z X - [X, Z]) + g(\nabla_X Y + \nabla_Y X + [X, Y], Z) \\
=&2g(\nabla_X Y, Z).
\end{aligned}$$

As g is non-degenerate, $\nabla_X Y(p)$ is determined uniquely by $g_p(\nabla_X Y(p), Z(p))$ for all vector fields $Z \in \mathcal{V}(M)$. The Koszul formula thus characterises $\nabla_X Y(p)$ uniquely for all $p \in M$.

(b) Existence: For $X, Y \in \mathcal{V}(\mathcal{M})$, we define a map $F_{X,Y} : \mathcal{V}(\mathcal{M}) \to \mathcal{C}^\infty(\mathcal{M})$

$$F_{X,Y}(Z) = \mathcal{L}_X g(Y, Z) + \mathcal{L}_Y g(Z, X) - \mathcal{L}_Z g(X, Y) - g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]).$$

It follows from the linearity of g and the Lie derivative that $F_{X,Y}$ is linear. For $f \in C^\infty(M, \mathbb{R})$, we compute

$$\begin{aligned}
F_{X,Y}(f \cdot Z) =& \mathcal{L}_X f \cdot g(Y, Z) + f \cdot \mathcal{L}_X g(Y, Z) + \mathcal{L}_Y f \cdot g(Z, X) + f \cdot \mathcal{L}_Y g(X, Z) - f \cdot \mathcal{L}_Z g(X, Y) \\
&- \mathcal{L}_Y f \cdot g(X, Z) - f \cdot g(X, [Y, Z]) - \mathcal{L}_X f \cdot g(Y, Z) + f \cdot g(Y, [Z, X]) + f \cdot g(Z, [X, Y]) \\
=& f \cdot F_{X,Y}(Z).
\end{aligned}$$

Lemma 3.24 then implies that $F_{X,Y}(Z)(p)$ depends only on $Z(p)$ and that there exists a unique vector field $\nabla_X Y \in \mathcal{V}(M)$ with $g(\nabla_X Y, Z) = F_{X,Y}(Z)$. The properties of the connection then follow by a direct calculation from the definition of $\nabla_X Y$. □

Remark 3.29. The Koszul formula allows one to explicitly compute the Christoffel symbols of the Levi-Civita connection on a semi-Riemannian manifold (M, g) from the coefficient functions of the metric. Let (φ, U) be a chart on M and denote by $g^{ij}_\varphi \in C^\infty(U, \mathbb{R})$ the components of the matrix inverse of the coefficient matrix of g with respect to φ:

$$\sum_{j=1}^n g^{ij}_\varphi(p) g_{jk}(p) = \delta_{ik} \qquad \forall p \in U.$$

Then the Koszul formula implies that the Christoffel symbols of the Levi–Civita connection on M are given by

$$\Gamma^k_{ij} = \frac{1}{2} \sum_{l=1}^n g^{kl}_\varphi \left(\frac{\partial g_{jl}}{\partial \varphi_i} + \frac{\partial g_{il}}{\partial \varphi_j} - \frac{\partial g_{ij}}{\partial \varphi_l} \right).$$

Example 3.30. We consider the Euclidean metric on $\mathbb{R}^2$ and the chart (φ, U) defined by polar coordinates

$$U = \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : y = 0, x \geq 0\}, \qquad \varphi^{-1}(r, \theta) = (r\cos\theta, r\sin\theta).$$

Then we have:

$$b^\varphi_r(r\cos\theta, r\sin\theta) = (\cos\theta, \sin\theta), \qquad b^\varphi_\theta(r\cos\theta, r\sin\theta) = (-r\sin\theta, r\cos\theta)$$

and the coefficient functions of the Euclidean metric with respect to φ are given by

$$g^\varphi_{rr} = \langle b^\varphi_r, b^\varphi_r \rangle = 1, \qquad g^\varphi_{\theta\theta} = \langle b^\varphi_\theta, b^\varphi_\theta \rangle = r^2, \qquad g^\varphi_{r\theta} = \langle b^\varphi_r, b^\varphi_\theta \rangle = 0,$$

where we omitted the argument $(r\cos\theta, r\sin\theta)$ to keep the notation simple. The inverse of the coefficient matrix of g is given by

$$g^{rr}_\varphi = 1, \qquad g^{\theta\theta}_\varphi = \frac{1}{r^2}, \quad g^{r\theta}_\varphi = 0,$$

and the Christoffel symbols of the Levi-Civita connection take the form

$$\Gamma^r_{\theta\theta} = -\frac{1}{2}\frac{\partial g^\varphi_{\theta\theta}}{\partial r} = -r, \qquad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{2r^2}\frac{\partial g^\varphi_{\theta\theta}}{\partial r} = \frac{1}{r}, \qquad \Gamma^k_{ij} = 0 \text{ otherwise.}$$

In the following, we will always implicitly assume that a given connection on a semi-Riemannian manifold is its Levi-Civita connection, and when we speak about Christoffel symbols of a semi-Riemannian manifolds, this refers to the Christoffel symbols of its Levi-Civita connection.

The name "connection" is motivated by the fact that a connection allows one to transport tangent vectors between the tangent spaces in different points on M and hence "connects" different tangent spaces. This leads to the concept of parallel transport. To see this, we need the to introduce the notion of a vector field along a smooth curve $c : I \to M$ and its derivative. This is essentially a vector field on M which is defined only on the image $c(I)$.

Definition 3.31. Let M be a semi-Riemannian manifold and $c : I \to M$ a smooth curve. A *vector field along* c is a smooth map $X : I \to T(M)$ with $\pi_M \circ X = c$. If $X : I \to T(M)$ is a vector field along c, we define its *derivative* $\nabla_{\dot{c}} X : T \to T(M)$ by

$$(\nabla_{\dot{c}} X)(t) = \dot{X}(t) + \Gamma(\dot{c}(t), X(t)),$$

where Γ denotes the Christiffel symbols of (M, g). A vector field $X : I \to T(M)$ along c is called *parallel* if $\nabla_{\dot{c}} X(t) = 0$ for all $t \in I$.

Lemma 3.32. *Let (M, g) be a semi-Riemannian manifold and $c : I \to M$ a smooth curve. Then the derivative along c has the following properties:*
(a) If (φ, U) is a chart of M, $c : I \to U$ a smooth curve and X a vector field along c, then

$$\nabla_{\dot{c}} X(t) = \sum_{k=1}^{n} \left(\dot{x}_k(t) + \sum_{i,j=1}^{n} \Gamma^k_{ij}(c(t)) \dot{c}_i(t) x_j(t) \right) b^{\varphi}_k(c(t)) \qquad \forall t \in I$$

where $c_i = \varphi_i \circ c : I \to \mathbb{R}$ and $X|_{c^{-1}(U)} = \sum_{i=1}^{n} x_i \cdot (b^{\varphi}_i \circ c)$.
(b) It is linear in X: for all vector fields $X, Y : I \to T(M)$ along c and all $\lambda, \mu \in \mathbb{R}$:

$$\nabla_{\dot{c}}(\lambda X + \mu Y) = \lambda \nabla_{\dot{c}} X + \mu \nabla_{\dot{c}} Y.$$

(c) For all vector fields $X : I \to M$ along c and all smooth functions $f : I \to \mathbb{R}$:

$$\nabla_{\dot{c}}(f \cdot X) = \dot{f} \cdot X + f \cdot \nabla_{\dot{c}} X.$$

(d) For all vector fields $X, Y : I \to T(M)$ along c:

$$\frac{d}{dt} g_{c(t)}(X(t), Y(t)) = g_{c(t)}(\nabla_{\dot{c}} X(t), Y(t)) + g_{c(t)}(X(t), \nabla_{\dot{c}} Y(t)).$$

(e) For all vector fields $X \in \mathcal{V}(M)$:

$$\nabla_{\dot{c}}(X \circ c)(t) = (\nabla_{\dot{c}(t)} X)(c(t)).$$

Proof. (a) This follows by a direct computation from the formulas for the Christoffel symbols in terms of the φ-basic vector fields for a chart (φ, U). Properties (b), (c) and (e) follow directly from the definition. To demonstrate (d), we compute $\frac{d}{dt} g_{c(t)}(X(t), Y(t))$. The identity then follows from the chain rule and the fact that ∇ is a metric connection. $\square$

Proposition 3.33. *Let M be a semi-Riemannian manifold and $c : I \to M$ a smooth curve. Then there exists for each $t \in I$ and $v \in T_{c(t)} M$ a unique parallel vector field $X_v : I \to T(M)$ with $X_v(t) = v$. For $t, t' \in I$, we define the* parallel transport *map*

$$P^c_{t',t} : T_{c(t)}(M) \to T_{c(t')} M, \quad v \mapsto X_v(t')$$

The parallel transport map is a linear isometry and has properties analogous to the properties of the flows in Section 1.4:

$$P^c_{t,t} = \mathrm{id}_{T_{c(t)}(M)} \qquad P^c_{t_2,t_1} \circ P^c_{t_1,t} = P^c_{t_2,t} \qquad \forall t, t_1, t_2 \in I,$$

and for every vector field $X : I \to T(M)$ along c:

$$\lim_{t' \to t} \frac{P^c_{t,t'}(X(t')) - X(t)}{t' - t} = \nabla_{\dot{c}} X(t).$$

Proof. (a) Let (U, φ) be a chart of M with $c(t) \in U$. In local coordinates the condition that X is parallel reads

$$\dot{x}_i(t) + \sum_{j,k=1}^{n} \Gamma^i_{jk}(c(t)) x_j(t) \dot{c}_k(t) = 0 \qquad \forall t \in c^{-1}(U),$$

where $c_i = \varphi_i \circ c : c^{-1}(U) \to \mathbb{R}$ are the coefficient functions of the curve c and $x_i \in C^\infty(c^{-1}(U), \mathbb{R})$ the coefficient functions of X. This is a first order linear ODE in the coefficients of X and hence has a unique solution for every set of initial values $x_i(t) = v_i$, $i \in \{1, \ldots, n\}$ on the connected component of $c^{-1}(U)$ which contains t. By covering $c(I)$ with charts we obtain a unique vector field $X_v : I \to T(M)$ along c with $X_v(t) = v$.

(b) That the parallel transport map is a linear isometry can be seen as follows from Lemma 3.32 (d): If $X_v : I \to T(M)$ is the parallel vector field along c with $X_v(t) = v$, then

$$\frac{d}{dt'} g_{c(t')}(X(t'), X(t')) = 2 g_{c(t')}(\nabla_{\dot{c}} X(t'), X(t')) = 0 \qquad \forall t' \in I$$

and therefore $g_{c'(t)}(X(t'), X(t')) = g_{c(t)}(v, v)$ for all $t' \in I$.

(c) The first two properties of the parallel transport follow directly from the uniqueness property of parallel vector fields. To show the last one, we choose a chart (φ, U) of M with $c(t), c(t') \in U$ and denote by $Y = \sum_{i=1}^n y_i \cdot (b_i^\varphi \circ c)$ the unique parallel vector field along c with $Y(t') = X(t')$. Then the mean value theorem implies that there exists an $s \in [t, t']$ with

$$\frac{y_k(t') - y_k(t)}{t' - t} = \dot{y}_k(s) = -\sum_{i,j=1}^{n} \Gamma^k_{ij}(c(s)) y_i(s) \dot{c}_k(s).$$

Using the smoothness of X and Y, we obtain

$$\begin{aligned}
&\lim_{t' \to t} \frac{y_k(t) - x_k(t)}{t' - t} = \lim_{t' \to t} \frac{y_k(t) - y_k(t') + y_k(t') - x_k(t)}{t' - t} \\
&= \lim_{s \to t} \sum_{i,j=1}^{n} \Gamma^k_{ij}(c(s)) y_i(s) \dot{c}_j(s) + \lim_{t' \to t} \frac{y_k(t') - x_k(t)}{t' - t} \\
&= \dot{x}_k(t) + \sum_{i,j=1}^{n} \Gamma^k_{ij}(c(t)) y_i(t) \dot{c}_j(t) = \dot{x}_k(t) + \sum_{i,j=1}^{n} \Gamma^k_{ij}(c(t)) x_i(t) \dot{c}_j(t) = \nabla_{\dot{c}} X(t).
\end{aligned}$$

□

Example 3.34. Consider $M = \mathbb{R}^n$ with either the Euclidean or the Minkowski metric and let $c : I \to M$ be a smooth curve. Then $T_pM \cong \mathbb{R}^n$ and a vector field $X = \sum_{i=1}^n x_i e_i : t \to T(M)$ along c is parallel if and only if

$$\dot{x}_k(t) = 0 \qquad \forall t \in I, k \in \{1, \ldots, n\}.$$

This implies that X is constant and the parallel transport map is given by

$$P^c_{t',t}(v) = v \qquad \forall v \in \mathbb{R}^n.$$

Lemma 3.35. *Let (M,g), (N,h) be semi-Riemannian manifolds, $c : I \to M$ a smooth curve and $f : M \to N$ a local isometry. Then a vector field $X : I \to T(M)$ along c is parallel if and only if the vector field $\bar{X} = T(f) \circ X : I \to T(N)$ along the curve $f \circ c$ is parallel. The parallel transport map satisfies:*

$$T_{c(t')}(f) \circ P^c_{t',t} = P^{f\circ c}_{t',t} \circ T_{c(t)}(f) \qquad \forall t, t' \in I.$$

Proof. It is sufficient to prove this statement for vector fields in the domains of charts. Let (φ, U) be a chart of M with $f|_U$ injective. Then $(V = f(U), \psi = \varphi \circ (f|_U))$ is a chart of N. Since $f|_U : U \to f(U)$ is an isometry, we have for the coefficient functions of the metrics g, h:

$$h^\psi_{ij}(f(p)) = h_{f(p)}(b^\psi_i(f(p)), b^\psi_j(f(p))) = h_{f(p)}(T_p(f)(b^\varphi_i), T_p(f)(b^\varphi_j)) = g_p(b^\varphi_i, b^\varphi_j) = g^\varphi_{ij}(p),$$

where we used that the φ- and ψ-basic vector fields are f-related on U: $b^\psi_i(f(p)) = T_p(f)b^\varphi_i(p)$. Via the Koszul formula, we obtain for the Christoffel symbols Γ^k_{ij} on M and Ω^k_{ij} on N:

$$\Omega^k_{ij}(f(p)) = \Gamma^k_{ij}(p) \qquad \forall p \in U.$$

A vector field $Y = \sum_{i=1}^n y_i \cdot (b^\varphi_i \circ c) : I \to T(M)$ along c is parallel if and only if its coefficient functions satisfy the differential equation

$$\dot{y}_k(t) + \sum_{j,k=1}^n \Gamma^k_{ij}(c(t)) y_i(t) \dot{c}_j(t) = 0.$$

Using the fact that the basic vector fields for φ and ψ are f-related and the relation between the Christoffel symbols, we find that this is the case if and only if the vector field $T(f) \circ Y = \sum_{i=1}^n y_i \cdot (b^\psi_i \circ f \circ c)$ along $f \circ c$ satisfies

$$\dot{y}_k(t) + \sum_{j,k=1}^n \Omega^k_{ij}(f(c(t))) \dot{y}_i(t) \dot{c}_j(t) = 0,$$

with $\psi_i \circ (f \circ c) = (\varphi_i \circ (f|_U)^{-1}) \circ (f \circ c) = \varphi_i \circ c$. This proves the claim. $\square$

The concept of vector fields along curves allows one in particular to consider the derivative $\dot{c} : I \to T(M)$ of each smooth curve $c : I \to M$ as a vector field along c. It is then natural to ask for which curves this derivative is a *parallel* vector field along c. This leads to the concept of a geodesic.

Definition 3.36. Let M be a semi-Riemannian manifold. A smooth curve $c : I \to M$ is called a *geodesic* if the vector field $\dot{c} : I \to T(M)$ along c is parallel.

Remark 3.37. (a) If (φ, U) is a chart of M with $c(I) \subset U$, then c is a geodesic if and only if $\nabla_{\dot{c}}\dot{c} = 0$. This is the case if and only if its component functions $c_i = \varphi_i \circ c : I \to \mathbb{R}$ satisfy the second order differential equation

$$\ddot{c}_i(t) + \sum_{j,k=1}^n \Gamma^i_{jk}(c(t)) \dot{c}_i(t) \dot{c}_j(t) = 0 \qquad \forall t \in I.$$

This implies in particular that for every $p \in M$ and $v \in T_p(M)$ there exists a unique geodesic $c : (-\epsilon, \epsilon) \to M$ with $c(0) = p$, $\dot{c}(0) = v$.

(b) If (M, g) and (N, h) are semi-Riemannian manifolds and $f : M \to N$ is a local isometry, then $c : I \to M$ is a geodesic of M if and only if $f \circ c : I \to N$ is a geodesic of N. This follows directly from Lemma 3.35.

(c) Uo to affine parameter transformations $t \mapsto at + b$ with $a \in \mathbb{R} \setminus \{0\}, b \in \mathbb{R}$, every geodesic $c : I \to M$ on a semi-Riemannian manifold (M, g) is parametrised according to arclength, because Lemma 3.32 implies:

$$\frac{d}{dt} g_{c(t)}(\dot{c}(t), \dot{c}(t)) = 2g_{c(t)}(\nabla_c \dot{c}(t), \dot{c}(t)) = 0 \qquad \forall t \in I.$$

(d) If $M = \mathbb{R}^n$ and g a non-degenerate bilinear form on M, then the geodesics of (M, g) are straight lines that are parametrised according to arclength. If one takes the standard chart $(\mathrm{id}, \mathbb{R}^n)$, all Christoffel symbols vanish, and the geodesic equation reduces to

$$\ddot{c}_i(t) = 0 \quad \forall t \in I.$$

This implies that all geodesics are of the form $c(t) = p + tv$.

Example 3.38. In Einstein's theory of general relativity, a universe is described by a four-dimensional Lorentzian manifold (M, g) that solves Einstein's equations. As $g_{c(t)}(\dot{c}(t), \dot{c}(t))$ is constant for each geodesic, one distinguishes three types of geodesics. A geodesic $c : I \to M$ with $\dot{c}(t) \neq 0$ is called *timelike* if $g_{c(t)}(\dot{c}(t), \dot{c}(t)) > 0$, *spacelike* if $g_{c(t)}(\dot{c}(t), \dot{c}(t)) < 0$ and *lightlike* if $g_{c(t)}(\dot{c}(t), \dot{c}(t)) = 0$ for all $t \in I$. Timelike geodesics describe the motion of point masses in free fall, i. e. point masses that are not subject to external forces other than the gravitational force. Lightlike geodesics describe the motion of light. The fact that $g_{c(t)}(\dot{c}(t), \dot{c}(t))$ is constant along each geodesic ensures that this description is consistent.

Although the differential equations in Remark 3.37 allow one in principle to determine the geodesics of a manifold by covering it with charts and solving the geodesic equation on the domain of each chart, in practice there are often better ways to determine geodesics. One way is to use the fact that isometries map geodesics to geodesics together with the uniqueness property of geodesics.

Example 3.39. We consider $\mathbb{S}^n$ with the metric induced by the Euclidean metric on $\mathbb{R}^n$. Then the unique geodesic $c : \mathbb{R} \to \mathbb{S}^n$ with $c(0) = p \in \mathbb{S}^n$ and $\dot{c}(0) = x \in p^\perp \setminus \{0\}$ is given by

$$c(t) = p \cos(t||x||) + \frac{x}{||x||} \sin(t||x||)$$

and the unique geodesic with $c(0) = p$ and $\dot{c}(0) = 0$ by $c(t) = p$ for all $t \in \mathbb{R}$.

Proof: The case $\dot{c}(0) = 0$ is obvious. Let $c : [-\epsilon, \epsilon] \to \mathbb{S}^n$ be the unique geodesic with $c(0) = p \in \mathbb{S}^n$ and $\dot{c}(0) = x \in p^\perp \setminus \{0\}$. Consider the plane $E_{p,x} = \mathrm{span}\{x, p\} \subset \mathbb{R}^{n+1}$ and the reflection $R_{p,x} : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ on this plane. Then $R_{x,p} \in \mathrm{O}_n = \mathrm{Isom}(\mathbb{S}^n)$ is an isometry and hence maps c to another geodesic $d = R_{x,p} \circ c : [-\epsilon, \epsilon] \to \mathbb{S}^n$. As $R_{p,x}x = x$ and $R_{p,x}p = p$, we have $d(0) = c(0) = p$ and $\dot{d}(0) = \dot{c}(0) = x$. Due to the uniqueness property of geodesics this implies $d(t) = R_{x,p}c(t) = c(t)$ for all $t \in [-\epsilon, \epsilon]$. Hence, for all $t \in [-\epsilon, \epsilon]$ the point $c(t)$ must lie in the intersection $E_{x,t} \cap \mathbb{S}^n$. The geodesic c is then determined uniquely by its initial values and the requirement that it is parametrised according to arclength. The curve

$$c(t) = p\cos(t||x||) + \frac{x}{||x||}\sin(t||x||).$$

satisfies these requirements and hence coincides with the geodesic c.

The fact that for every point $p \in M$ and every tangent vector $v \in T_p(M)$ there is a unique geodesic $c_v : (-\epsilon, \epsilon) \to M$ with $c_v(0) = p$ and $\dot{c}_v(0) = v$ can be used to construct a diffeomorphism from an open neighbourhood of 0 in the tangent space $T_p(M)$ to an open neighbourhood U of p. The associated coordinates have particularly nice properties and can be viewed as coordinates adapted to the geometry of M near p.

Definition 3.40. Let (M, g) be a semi-Riemannian manifold, $p \in M$ and define

$$D_p = \{v \in T_p(M) : \text{ the geodesic } c \text{ on } M \text{ with } c(0) = p \text{ and } \dot{c}(0) = v \text{ is defined on } [0,1]\}.$$

The map $\exp_p : D_p \to M$, $v \mapsto c_v(1)$ is called *exponential map* on M in p. It extends to a smooth map $\exp : D = \bigcup_{p\in M} D_p \to M$ with $\exp|_{T_p(M)} = \exp_p$ for all $p \in M$.

Remark 3.41. (a) The exponential map $\exp_p : D_p(M) \to M$ satisfies $\exp_p(tv) = c_{tv}(1) = c_v(t)$ for all $t \in [0,1]$ and $v \in D_p$. This follows directly from the uniqueness property of the geodesics. In particular, this implies that D_p is star-shaped for all $p \in M$: $v \in D_p \Rightarrow tv \in D_p$ for all $t \in [0,1]$.

(b) The tangent map $T_0(\exp) : T_0T_p(M) \to T_p(M)$ is the canonical isomorphism from Section 1.2.

(c) The Inverse Function Theorem for manifolds implies that there is an open subset $V \subset T_p(M)$ of 0 such that $\exp_p : V \to \exp_p(V)$ is a diffeomorphism. However, in general V_p is smaller than D_p and the map $\exp_p : D_p \to \exp(D_p)$ is *not* a diffeomorphism.

(d) If $\varphi : M \to M$ is an isometry, then the exponential map satisfies

$$\varphi \circ \exp_p = \exp_{\varphi(p)} \circ T_p(\varphi).$$

This is the semi-Riemannian analogue of formula (12) in Proposition 2.40 and can be seen as follows: As φ is an isometry, for every geodesic $\gamma : [0,1] \to M$, $\gamma(t) = \exp_p(tv)$ with $\gamma(0) = p$, $v \in D_p \subset T_p(M)$, the image $\varphi \circ \gamma$ is a geodesic with $\varphi \circ \gamma(0) = \varphi(p)$ and $(\varphi \circ \gamma)'(0) = T_p(\varphi) \circ \gamma'(0) = T_p(\varphi)v$. This implies

$$\varphi \circ \exp_p(v) = \varphi \circ \gamma(1) = \exp_{\varphi(p)}((\varphi \circ \gamma)'(0)) = \exp_{\varphi(p)}(T_p(\varphi)v).$$

Since for every $v \in D_pT_p(M)$ there exists a geodesic $\gamma : [0,1] \to M$ with $\gamma(0) = p$, $\dot{\gamma}(0) = v$ this proves the claim.

We can use the exponential map to obtain a particularly nice set of coordinates around each point $p \in M$. The idea is to exponentiate the coordinate axes associated with an orthonormal basis of g_p.

Definition 3.42. Let (M, g) be a semi-Riemannian manifold, $p \in M$ and $(v_1, \ldots v_n)$ an orthonormal basis of $T_p(M)$. Consider the linear isomorphism $A : \mathbb{R}^n \to T_p(M)$, $x \mapsto \sum_{i=1}^n x_i v_i$. Then there exist open neighbourhoods V of $0 \in \mathbb{R}^n$ and U of p such that $\exp \circ A : V \to U$ is a diffeomorphism and $(\psi := (\exp_p \circ A)^{-1}, U)$ is a chart on M. The associated coordinate functions $\psi_i : U \to \mathbb{R}$ are called *normal coordinates* around p.

Proposition 3.43. *Let (M, g) be a semi-Riemannian manifold, $p \in M$ and (U, ψ) normal coordinates around p. Then the coefficient functions of the metric and the Christoffel symbols for (ψ, U) in p take the form*

$$g^{\psi}_{ij}(p) = \epsilon_i \delta_{ij}, \qquad \Gamma^k_{ij}(p) = 0.$$

Proof. By definition the ψ-basic vector fields are given by $b^{\psi}_i = T_0(\exp_p \circ A)e_i = T_0(\exp_p)v_i$, and we obtain for the coefficient functions of the metric

$$g^{\psi}_{ij}(p) = g_p(b^{\psi}_i, b^{\psi}_j)) = g_p(T_0(\exp_p)v_i, T_0(\exp_p)v_j) = g_p(v_i, v_j) = \epsilon_i \delta_{ij}.$$

By definition, the unique geodesic $c : (-\epsilon, \epsilon) \to M$ with $c(0) = 0$ and $\dot{c}(0) = \sum_{i=1}^n w_i v_i$ is given by $c(t) = \exp_p \circ A(tw)$. We have $c_i(t) = \psi_i(c(t)) = tw_i$ and the geodesic equation reduces to the equation

$$\sum_{i,j=1}^{k} \Gamma^k_{ij}(p) w_i w_j = 0 \qquad \forall w \in \mathbb{R}^n.$$

This implies that all Christoffel symbols in p vanish. □

In particular, we can use the exponential map and normal coordinates to show that every isometry of a connected semi-Riemannian manifold is determined uniquely by its value and derivative in a single point.

Lemma 3.44. *Let (M, g) be a connected semi-Riemannian manifold, $p \in M$ and $\varphi, \psi : M \to M$ isometries with $\varphi(p) = \psi(p)$ and $T_p(\varphi) = T_p(\psi)$. Then the two isometries agree on M: $\varphi(q) = \psi(q)$ for all $q \in M$.*

Proof. We consider the set $A = \{q \in M : \varphi(q) = \psi(q)\}$. By assumption, A is nonempty. Since φ, ψ are continuous, it follows that $A \subset M$ is closed. If we can show that A is also open, then the connectedness of M implies $A = M$. To show that A is open, let $q \in A$ and choose an $\epsilon > 0$ such that $\exp_q |_{B_\epsilon(0)} : B_\epsilon(0) \subset T_q(M) \to \exp_q(B_\epsilon(0))$ and $\exp_{\varphi(q)} |_{B_\epsilon(0)} \subset T_{\varphi(q)} \to \exp_{\varphi(q)}(B_\epsilon(0))$ are diffeomorphisms. As φ and ψ are isometries and every point $w \in \exp_p(B_\epsilon(0))$ can be connected to q by a geodesic, Remark 3.41 implies

$$\begin{aligned}\varphi(w) &= (\exp_{\varphi(q)} |_{B_\epsilon(0)}) \circ T_q(\varphi) \circ (\exp_q |_{B_\epsilon(0)})^{-1}(w)\\ =&\psi(w) = (\exp_{\varphi(q)} |_{B_\epsilon(0)}) \circ T_q(\psi) \circ (\exp_q |_{B_\epsilon(0)})^{-1}(w)\end{aligned}$$

for all $w \in \exp_q(B_\epsilon(0))$ and therefore $\exp_q(B_\epsilon(0)) \subset A$ for all $q \in A$. Since $B_\epsilon(0)$ is open and $\exp_q |_{B_\epsilon(0)} \to \exp_q(B_\epsilon(0))$ is a diffeomorphism, $\exp_q(B_\epsilon(0))$ is open. This implies that A is open and proves the claim. □

In addition to our characterisation of geodesics as smooth curves whose velocity field is a parallel vector field along c, there is an alternative characterisation of geodesics as critical points of a certain energy functional. In analogy to the kinetic energy $E_{kin} = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{x}^2$ in classical mechanics, we assign to each piecewise C^2-curve $c : I \to M$ on a semi-Riemannian manifold an energy which is determined by its velocity field.

Definition 3.45. Let (M, g) be a semi-Riemannian manifold. The *energy* of a piecewise C^1-curve $c : [0,1] \to M$ is defined as

$$E[c] = \frac{1}{2} \int_0^1 g_{c(t)}(\dot{c}(t), \dot{c}(t)) \, dt.$$

The idea is now to vary the energy with respect to the geodesic, i. e. to consider the change of the energy when the curve c is deformed slightly in such a way that its endpoints stay fixed. For this, we require the concept of a variation with fixed endpoints.

Definition 3.46. Let (M, g) be a semi-Riemannian manifold and $c : [0,1] \to M$ a piecewise C^2-curve. A *variation of c with fixed endpoints* is a continuous map

$$h : [0,1] \times [-\epsilon, \epsilon] \to M$$

with $h(t,0) = c(t)$ for all $t \in [0,1]$, $h(0,s) = c(0)$ and $h(1,s) = c(1)$ for all $s \in [-\epsilon, \epsilon]$. A variation of c with fixed endpoints is called *piecewise* C^2 if there exists a subdivision $0 = t_0 < t_1 < \ldots < t_{N-1} < t_N = 1$ such that $h|_{[t_i, t_{i+1}] \times [\epsilon, -\epsilon]}$ is C^2. If h is piecewise C^2, then the vector field

$$V(t) = \frac{\partial h}{\partial s}(t,s)\Big|_{s=0}$$

along c is piecewise C^1 and satisfies $V(0) = V(1) = 0$. It is called *variation field* of h.

Given two points $p, q \in M$, it is natural to assign to each piecewise C^2-curve that connects p and q its energy and to attempt to determine the curves for which the energy is maximal or minimal. Clearly, a curve of maximal or minimal energy should be a critical point of the energy, i. e. the derivative of the energy with respect to the "deformation parameter" s should vanish. Using the concept of a piecewise C^2-variation with fixed endpoints, we can give this intuition a precise meaning.

Definition 3.47. A piecewise C^2-curve $c : [0,1] \to M$ is called *critical point* of the energy, if for all piecewise C^2-variations $h : [0,1] \times [-\epsilon, \epsilon] \to M$ of c with fixed endpoints

$$\frac{d}{ds}\Big|_{s=0} E[c_s] = 0, \qquad \text{where} \quad c_s : [0,1] \to M, \; c_s(t) := h(t,s).$$

Theorem 3.48. *Let (M, g) be a semi-Riemannian manifold and $c : [0,1] \to M$ a piecewise C^2-curve. If c is a critical point of the energy, then c is a geodesic.*

Proof. (1) We consider first the case where c and h are C^2. Let V be the variation field for h

$$V(t) = \frac{\partial h}{\partial s}(t,0).$$

Then the variation of the energy is given by

$$\begin{aligned}
\frac{d}{ds}\bigg|_{s=0} E[c_s] &= \frac{d}{ds}\bigg|_{s=0} \frac{1}{2}\int_0^1 g_{c_s(t)}(\dot{c}_s(t), \dot{c}_s(t))\, dt = \frac{1}{2}\int_0^1 \frac{\partial}{\partial s}\bigg|_{s=0} g_{c_s(t)}(\dot{c}_s(t), \dot{c}_s(t))\, dt \\
&= \frac{1}{2}\int_0^1 g_{c_0(t)}(\nabla_{\dot{c}_0} V(t), \dot{c}_0(t)) + g_{c_0(t)}(\dot{c}_0(t), \nabla_{\dot{c}_0} V(t))\, dt = \int_0^1 g_{c_s(t)}\left(\nabla_{\dot{c}_0} V(t), \dot{c}_0(t)\right)\, dt \\
&= \int_0^1 \frac{d}{dt} g_{c(t)}\left(V(t), \dot{c}(t)\right) - g_{c(t)}\left(V(t), \nabla_{\dot{c}}\dot{c}(t)\right)\, dt \\
&= \left[g_{c(t)}\left(V(t), \dot{c}(t)\right)\right]_{t=0}^{t=1} - \int_0^1 g_{c(t)}\left(V(t), \nabla_{\dot{c}}\dot{c}(t)\right)\, dt = -\int_0^1 g_{c(t)}\left(V(t), \nabla_{\dot{c}}\dot{c}(t)\right)\, dt,
\end{aligned}$$

where we used the metricity of the connection and the fact that the derivatives with respect to s and t commute to obtain the first expression in the second line. If the curve c and the variation h are only piecewise C^2, then there exists a subdivision $0 = t_0 < t_1 < \ldots < t_{N-1} < t_N = 1$ such that $h|_{[t_i,t_{i+1}]\times[\epsilon,-\epsilon]}$ is C^2. Applying the formula to the restriction $c|_{[t_i,t_{i+1}]}$ and summing over the points in the subdivision points yields

$$\frac{d}{ds}\bigg|_{s=0} E[c_s] = -\int_0^1 g_{c(t)}(V(t), \nabla_{\dot{c}}\dot{c}(t))\, dt + \sum_{i=1}^{N-1} g_{c(t_i)}(V(t_i), \dot{c}(t_i^+) - \dot{c}(t_i^-)),$$

where $\dot{c}(t_i^-) = \lim_{\epsilon\downarrow 0} \dot{c}(t_i - \epsilon)$ and $\dot{c}(t_i^+) = \lim_{\epsilon\downarrow 0} \dot{c}(t_i + \epsilon)$.

(2) Let now $W : [0,1] \to T(M)$ be a piecewise C^2-vector field along c and $f : [0,1] \to [0,1]$ a smooth function with $f(0) = f(1) = 0$. Then $h : [0,1] \times [\epsilon, -\epsilon] \to M$,

$$h(t,s) = \exp_{c(t)}(sf(t)W(t))$$

is a piecewise C^2-variation of c with fixed endpoints and

$$\frac{\partial}{\partial s} h(t,s)\bigg|_{s=0} = f(t)W(t).$$

If c is a critical point of the energy, we have

$$-\int_0^1 f(t) g_{c(t)}(W(t), \nabla_{\dot{c}}\dot{c}(t))\, dt + \sum_{i=1}^{N-1} f(t_i)\cdot g_{c(t_i)}(W(t_i), \dot{c}(t_i^+) - \dot{c}(t_i^-)) = 0.$$

By considering general piecewise C^2 vector fields W along c and choosing the smooth function $f : [0,1] \to [0,1]$ in such a way that $\mathrm{spt}(f) \subset [T-\delta, T+\delta]$ for $T \in]t_i, t_{i+1}[$ and $\delta > 0$ suffiently small, we can show that

$$\nabla_{\dot{c}}\dot{c}(T) = 0 \qquad \forall T \in [0,1] \setminus \{t_0, \ldots, t_N\}.$$

By considering general piecewise C^2-vector fields W along c and choosing the smooth function $f : [0,1] \to [0,1]$ in such a way that $\mathrm{spt}(f) \subset [t_i - \delta, t_i + \delta]$ with $\delta > 0$ sufficiently small, we find $\dot{c}(t_i^+) = \dot{c}(t_i^-)$ for all $i \in \{0, \ldots, N\}$. This implies that c is not only piecewise C^2 but C^2 and $\nabla_c \dot{c}(t) = 0$ for all $t \in [0,1]$. Thus c is a geodesic. $\square$

This characterisation of geodesics is particularly intuitive in the Riemannian context, where the metric is positive definite and there is a concept of length for each pieceweise C^1-curve on M. In this case, every curve of minimal length between two points $p, q \in M$ that is parametrised according to arclength is a geodesic.

Corollary 3.49. *If (M, g) is a Riemannian manifold and $c : [0, 1] \to M$ a piecewise C^2-curve of minimal length from $c(0) = p$ to $c(1) = q$*

$$L[c] = \int_0^1 \sqrt{g_{c(t)}(\dot{c}(t), \dot{c}(t))}\, dt = \inf\{L[d] :\ d : [0, 1] \to M \text{ piecewise } C^2, d(0) = p, d(1) = q\}$$

and parametrised according to arclength, then c is a geodesic.

Proof. The Cauchy–Schwarz inequality implies

$$L[c] = \int_0^1 \sqrt{g_{c(t)}(\dot{c}(t), \dot{c}(t))}\, dt \le \sqrt{\int_0^1 g_{c(t)}(\dot{c}(t), \dot{c}(t))\, dt} \cdot \sqrt{\int_0^1 1\, dt} = \sqrt{2E[c]}$$

and $L[c] = \sqrt{2E[c]}$ if and only if $g_{c(t)}(\dot{c}(t), \dot{c}(t))$ is constant. This implies that any piecewise C^2-curve c of minimal length between p and q that is parametrised according to arclength minimises the energy. If c minimises the energy, then for all piecewise C^2-variations with fixed endpoints the map $E : s \mapsto E[c_s]$ is C^1 and has a minimum in $s = 0$. This implies that c is a critical point of the energy. □

Exercises for Section 3.3

Exercise 3.9. Let (M, g) be a semi-Riemannian manifold and ∇ a connection on M. Show that ∇ is torsion free if and only if for all charts (φ, U) the associated Christoffel symbols satisfy

$$\Gamma^k_{ij} = \Gamma^k_{ji} \qquad \forall i, j \in \{1, \ldots, n\}.$$

Show that it is a metric connection if and only if for all charts (φ, U) the Christoffel symbols satisfy

$$\frac{\partial g_{ij}}{\partial \varphi_k} = \sum_{l=1}^{n} (\Gamma^l_{kj} g_{li} + \Gamma^l_{ki} g_{lj}) \qquad \forall i, j, k \in \{1, \ldots, n\}.$$

Exercise 3.10. Let (M, g) be a semi-Riemannian manifold and (φ, U), (ψ, V) charts of M with $U \cap V \neq \emptyset$. Derive a formula that expresses the Christoffel symbols with respect to (V, ψ) in terms of the christiffel symbols with respect to (U, φ).

Exercise 3.11. Fill in the details in the proof of Lemma 3.32 and verify by explicit calculations the properties of the derivative $\nabla_{\dot{c}}$ stated there.

Exercise 3.12. Consider n-dimensional hyperbolic space $\mathbb{H}^n$ with the metric induced by the Minkowski metric on $\mathbb{R}^{n+1}$. Determine all of its geodesics by using suitable isometries.

3.4 Curvature

After investigating the properties of isometries and geodesics, we will now introduce another fundamental concept in semi-Riemannian geometry, namely the notion of curvature. While there are many concepts of curvature, there is a fundamental one from which all other notions can be derived and which determines them uniquely. This is the Riemann curvature tensor.

Definition 3.50. (Riemann curvature tensor) Let (M, g) be a semi-Riemannian manifold. Then the map $R : \mathcal{V}(M) \times \mathcal{V}(M) \times \mathcal{V}(M) \to \mathcal{V}(M)$, $(X, Y, Z) \mapsto R(X, Y)Z$ with

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

is called the *Riemann curvature (tensor)* of M.

Lemma 3.51. *Let (M, g) be a semi-Riemannian manifold. Then the Riemann curvature tensor of M is a tensor: For all vector fields $X, Y, Z \in \mathcal{V}(M)$, the Riemann curvature tensor $R(X,Y)Z(p)$ depends only on $X(p), Y(p), Z(p)$.*

The Riemann curvature tensor has the following symmetries:

- Anti-symmetry in the first two arguments: $R(X,Y)Z(p) = -R(Y,X)Z(p)$.
- first Bianchi identity: $R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0$
- $g(R(X,Y)Z, W) = -g(Z, R(X,Y)W)$
- $g(R(X,Y)Z, W) = g(R(Z,W)X, Y) \qquad \forall X, Y, Z, W \in \mathcal{V}(M)$.

Proof. To show that the Riemann curvature tensor is a tensor, we consider the second covariant derivative

$$\nabla^2_{X,Y} Z = \nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z.$$

Since $\nabla_X Y(p)$ depends only on $X(p)$ and $\nabla_X \nabla_Y Z$ depends only on $X(p)$, it follows that $\nabla^2_{X,Y}$ depends only on the value of X in p. Moreover, we find for any function $f \in C^\infty(M, \mathbb{R})$

$$\nabla^2_{X,f\cdot Y} = \mathcal{L}_X f \cdot \nabla_Y Z + f \cdot \nabla_X \nabla_Y Z - \nabla_{\mathcal{L}_X f \cdot Y + f \cdot \nabla_X Y} Z = f \cdot \nabla^2_{X,Y} Z.$$

By applying Lemma 3.24 to the map $F_{X,Z,W} : \mathcal{V}(M) \to \mathbb{R}$, $Y \mapsto g(\nabla^2_{X,Y} Z, W)$ we then find that $\nabla^2_{X,Y} Z(p)$ depends only on the value of Y in p. The Riemann curvature tensor is given by

$$R(X,Y)Z = \nabla^2_{X,Y} Z - \nabla^2_{Y,X} Z,$$

and it follows directly that $R(X,Y)Z(p)$ depends only on the values of X and Y in p. To determine its dependence on Z, we calculate

$$\begin{aligned}
&R(X,Y)(f \cdot Z) \\
=&\nabla_X(\mathcal{L}_Y f \cdot Z + f \cdot \nabla_Y Z) - \nabla_Y(\mathcal{L}_X f \cdot Z + f \cdot \nabla_X Z) - \mathcal{L}_{[X,Y]} f \cdot Z - f \cdot \nabla_{[X,Y]} Z \\
=&f \cdot (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) + (\mathcal{L}_x \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f - \mathcal{L}_{[X,Y] f}) \cdot Z \\
+&\mathcal{L}_Y f \cdot \nabla_X Z + \mathcal{L}_X f \cdot \nabla_Y Z - \mathcal{L}_X f \cdot \nabla_Y Z - \mathcal{L}_Y f \cdot \nabla_X Z \\
=&f \cdot R(X,Y)Z.
\end{aligned}$$

Applying Lemma 3.24 to the map $F_{X,Y,W} : \mathcal{V}(M) \to \mathbb{R}$, $Z \mapsto g(R(X,Y)Z, W)$ we then find that $R(X,Y)Z(p)$ depends only on the value of Z in p.

The antisymmetry of the Riemann curvature tensor in the first two arguments follows directly from its definition. The first Bianchi identity is obtained from the fact that the Levi–Civita connection is torsion free and from the Jacobi identity for the Lie bracket. To prove the third identity, it is sufficient to show that $g(R(X,Y)Z, Z) = 0$ for all smooth vector fields X, Y, Z. As ∇ is a metric connection, we have

$$g(\nabla_W Z, Z) = \tfrac{1}{2}\mathcal{L}_W g(Z,Z),\ g(\nabla^2_{X,Y} Z, Z) = \tfrac{1}{2}\mathcal{L}_X\mathcal{L}_Y g(Z,Z) - g(\nabla_Y Z, \nabla_X Z) - \tfrac{1}{2}\mathcal{L}_{\nabla_X Y} g(Z,Z).$$

Using the definition of the Riemann curvature tensor in terms of the second covariant derivative, we obtain after some computations the third identity. The forth identity follows from the first three. $\square$

Remark 3.52. (a) If (M,g) is a semi-Riemannian manifold and (φ, U) a chart on M, the Riemann curvature tensor on U is characterised uniquely through its component functions $R^l_{ijk} \in C^\infty(U, \mathbb{R})$

$$R(b^\varphi_i, b^\varphi_j)b^\varphi_k = \sum_{l=1}^n R^l_{ijk} b^\varphi_l,$$

which is given in terms of the Christoffel symbols by the following equation (see exercise 3.16)

$$R^l_{ijk} = \frac{\partial \Gamma^l_{jk}}{\partial \varphi_i} - \frac{\partial \Gamma^l_{ik}}{\partial \varphi_j} + \sum_{m=1}^n \Gamma^m_{jk}\Gamma^l_{im} - \Gamma^m_{ik}\Gamma^l_{jm}.$$

(b) If (M,g) and (N,h) are semi-Riemannian manifolds and $f : M \to N$ is a local isometry, then the Riemann curvature tensor satisfies

$$T_p(f)(R^M(x,y)z) = R^N(T_p(f)x, T_p(f)y)T_p(f)z \qquad \forall p \in M, x, y, z \in T_p(M).$$

This follows directly from the fact that local isometries preserve the metric and hence the Levi–Civita connection.

Besides the Riemann curvature tensor, there are other notions of curvature which play an important role in differential geometry and general relativity. They are all determined by the Riemann curvature tensor and the most important ones are given in the following definition.

Definition 3.53. Let (M,g) be a semi-Riemannian manifold and $p \in M$.

(a) Let $E \subset T_p(M)$ be a plane for which the restriction $g_p|_{E\times E}$ is non-degenerate and and $x, y \in T_p(M)$ two vectors which span E. Then the *sectional curvature* of E is defined as

$$K_p(E) = \frac{g_p(R(x,y)y, x)}{g_p(x,x)g_p(y,y) - g_p(x,y)^2}.$$

It depends only on E and not on the choice of x and y. If $\dim(M) = 2$ the plane E coincides with $T_p(M)$ and the sectional curvature is called *Gauß curvature*.

(b) For $x, y \in T_p(M)$, the *Ricci curvature* is defined as the trace of the linear map $R_{x,y} : T_p(M) \to T_p(M)$, $z \mapsto R(x,y)z$

$$\mathrm{ric}_p(x,y) = \mathrm{Tr}(R_{x,y}).$$

It defines a non-degenerate symmetric bilinear form on $T_p(M)$, and there exists a unique linear map $\mathrm{Ric}_p : T_p(M) \to T_p(M)$ with $g_p(\mathrm{Ric}_p(x), y) = \mathrm{ric}(x, y)$ for all $x, y \in T_p(M)$.

(c) The *scalar curvature* $\mathrm{scal}(p)$ of M in p is defined as the trace of the linear map $\mathrm{Ric}_p : T_p(M) \to T_p(M)$:

$$\mathrm{scal}(p) = \mathrm{Tr}(\mathrm{Ric}_p).$$

Remark 3.54. (a) That the sectional curvature does not depend on the choice of x and y can be shown via a direct calculation: two vectors $x' = ax + by$ and $y' = cx + dy$ with $a, b, c, d \in \mathbb{R}$ form a basis of E if and only if $ad - bc \neq 0$. Using the properties of the curvature tensor, one obtains

$$\begin{aligned}
&g_p(R(x', y')y', x') \\
&= adcbg_p(R(x, y)x, y) + (ad)^2 g_p(R(x, y)y, x) + (bc)^2 g_p(R(y, x)x, y) + adbcg_p(R(y, x)y, x) \\
&= (ad - bc)^2 g_p(R(x, y)y, x)
\end{aligned}$$

and a short computation yields

$$g_p(x', x')g_p(y', y') - g_p(x', y')^2 = (ad - bc)^2 (g_p(x, x)g_p(y, y) - g_p(x, y)^2).$$

(b) The Riemann curvature tensor is determined uniquely by the sectional curvatures $K_p(E)$ for all planes $E \subset T_p(M)$ for which $g_p|_{E\times E}$ is non-degenerate. The sectional curvature is determined by the Ricci curvature only for $\dim(M) \leq 3$ and the Ricci curvature by the scalar curvature only for $\dim(M) = 2$.

(c) For every chart (φ, U) of M, the Ricci curvature on U is described uniquely by its component functions $\mathrm{ric}_{ij} \in C^\infty(U, \mathbb{R})$, $\mathrm{ric}_{ij} = \mathrm{ric}_p(b_i^\varphi, b_j^\varphi)$ which are given in terms of the Riemann curvature tensor by

$$\mathrm{ric}_{ij} = \sum_{k=1}^{n} R^k_{ijk}.$$

and the scalar curvature is given by

$$\mathrm{scal} = \sum_{j,k=1}^{n} R^k_{ijk} g^{ij}_\varphi.$$

Example 3.55. In general relativity, a universe is described by a four-dimensional Lorentzian manifold (M, g). The Lorentzian metric g on M is required to be a solution of the *Einstein equations*

$$\mathrm{ric}_p(x, y) - \tfrac{1}{2} g_p(x, y) \cdot \mathrm{scal}(p) + \Lambda g_p(x, y) = \frac{8\pi G}{c^4} t_p(x, y) \qquad \forall p \in M, x, y \in T_p(M),$$

where $t_p : T_p(M) \times T_p(M) \to \mathbb{R}$ is a symmetric tensor called the *stress-energy tensor*. It is determined by the matter content of the universe (massive matter and radiation). The constant $\Lambda \in \mathbb{R}$ is called the *cosmological constant* and $G \in \mathbb{R}$ the *gravitational constant*.

A solution of Einstein's equations for vanishing stress-energy tensor is called a *vacuum spacetime*.

As the Riemann curvature and hence also the Ricci and scalar curvature depend non-linearly on the metric and its derivatives, Einstein's equations define a complicated system of non-linear differential equations that can be solved only numerically for many configurations.

The situation simplifies considerably if one considers the three-dimensional version of the theory and vacuum spacetimes. In that case, the Ricci tensor determines the Riemann curvature completely and for any vacuum spacetime the Riemann curvature tensor is constant.

Exercises for Section 3.4

Exercise 3.13. Consider the two-sphere $\mathbb{S}^2$ with the metric induced by the Euclidean metric on $\mathbb{R}^3$ and the chart given by

$$U = \mathbb{S}^2 \setminus \{(x, y, z) : \ x \geq 0, y = 0\} \qquad \varphi^{-1}(\psi, \theta) = (\cos\psi \cdot \sin\theta, \sin\psi \cdot \sin\theta, \cos\theta).$$

Determine the coefficient functions of the metric, the Christoffel symbols and the coefficient functions of the Riemann curvature tensor.

Exercise 3.14. Consider n-dimensional hyperbolic space $\mathbb{H}^n$ with the metric induced by the Minkowski metric on $\mathbb{R}^{n+1}$ and the chart given by

$$U = \mathbb{H}^n \qquad \varphi(x_0, \ldots, x_n) = (x_1, \ldots, x_n).$$

Determine the coefficient functions of the metric, the Christoffel symbols and the coefficient functions of the Riemann curvature tensor.

Exercise 3.15. Determine the number of independent components of the Riemann curvature tensor and the of Ricci tensor on an n-dimensional manifold. Use your result to conclude that the Ricci tensor determines the Riemann curvature tensor uniquely only in dimension $d \leq 3$.

Exercise 3.16. Prove the formula for the component functions of the Riemann curvature in terms of the Christoffel symbols from Remark 3.52

Exercise 3.17. The gravitational field of a point of mass m is described by the *Schwartzschild metric* on $\mathbb{R} \times (\mathbb{R}^3 \setminus \{0\})$. In terms of a coordinate t on $\mathbb{R}$ and polar coordinates (r, θ, φ) on $\mathbb{R}^3$, this metric is given by

$$g(t, r, \theta, \varphi) = \left(1 - \tfrac{2m}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2m}{r}} - r^2 g_{\mathbb{S}^2},$$

where g_{S^2} is the metric on $\mathbb{S}^2$ and $r > 2m$

1. Determine the Christoffel symbols and the Riemann curvature tensor of g. Show that the Ricci curvature of g vanishes and that the Schwartzschild metric is a solution of the vacuum Einstein equations with vanishing cosmological constant.

2. Sketch the vector fields b_t and b_r in a plane with $\theta = \varphi = \text{const}$.

4 The Geometric Structures of Classical Mechanics

In this section we study the mathematical, resp., geometric structures underlying classical mechanics. Since the differential equation describing the time evolution of a mechanical system is of second order, we first introduce second order vector fields on a manifold (Section 4.1). A key point in mechanics is the passage between velocities and momenta. This corresponds to the passage from the tangent bundle TQ of our configuration space Q, for which the elements of $T_p(Q)$ are interpreted as velocities, to the cotangent bundle T^*Q, for which the elements of $T^*_p(Q) = T_p(Q)^*$ are interpreted as momenta (Section 4.2). Before we then turn to symplectic geometry, we recall some basic facts on differential forms and their formulation in the context of general manifolds (Section 4.3). Symplectic manifolds are then introduced in Section 4.4. These are almost symplectic manifolds (M, ω) satisfying the 'integrability condition' $\mathtt{d}\omega = 0$. From the perspective of physics, the key motivation for studying Hamiltonian systems, which are certain flows on symplectic manifolds, is that one can give up the distinction between space and momentum, resp., velocity coordinates. This leads to a significant enlargement of the underlying symmetry group from the diffeomorphism group of configuration space to the group of symplectic diffeomorphisms of the cotangent bundle. Accordingly, cotangent bundles are the prototypes of symplectic manifolds. In Section 4.5 we introduce the formalism of symplectic geometry: Hamiltonian vector fields and Poisson brackets. We eventually come full circle by a discussion of Lagrangian mechanics in Section 4.6 and the Legendre transform in Section 4.7. The Legendre transform provides the translation between Lagrangian mechanics based on the Euler–Lagrange equations in TQ to Hamiltonian mechanics in T^*Q. As a byproduct, this provides new insight in the geometry of semi-Riemannian manifolds because it exhibits the velocity curves of geodesics as the solutions of a Hamiltonian system where the Hamiltonian is the function $H(v) := \frac{1}{2}g(v, v)$ on TQ.

4.1 Second order equations on manifolds

The movement of a point particle of mass m in a force field $\vec{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is determined by Newton's Law, i.e., the second order equation

$$m\mathbf{a}(t) = m\ddot{\mathbf{x}}(t) = F(\mathbf{x}(t)).$$

To model such equations on the level of manifolds, i.e., in a form independent of the choice of coordinates leads to the concept of a second order vector field (Subsection 4.1).

Let M be a smooth manifold. If $\gamma\colon I \to M$ is a smooth curve, then its *velocity curve* $\gamma'\colon I \to TM$ is a smooth curve with values in the tangent bundle TM. Taking one more derivative, we arrive at a curve $\gamma''\colon I \to TTM$. To define second order differential equations in the context of manifolds, we therefore have to consider vector fields on the tangent bundle.

Definition 4.1. Let M be a smooth manifold. A *second order vector field on M* is a vector field $F\colon TM \to TTM$ on TM satisfying $T(\pi) \circ F = \mathrm{id}_{TM}$, where $\pi\colon TM \to M$ is the projection map. For the integral curves $\beta\colon I \to TM$ of F, this means that the corresponding curve $\gamma := \pi \circ \beta\colon I \to M$ satisfies

$$\gamma'(t) = T(\pi)\beta'(t) = T(\pi)F(\beta(t)) = \beta(t), \quad t \in I.$$

We thus obtain the relation
$$\gamma''(t) = F(\beta(t)),$$
which justifies the terminology.

Remark 4.2. (a) To visualize the concepts locally, we consider an open subset $U \subseteq \mathbb{R}^n$. Then $TU \cong U \times \mathbb{R}^n$, $\pi(x,v) = x$, $TTU \cong U \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$, and $T(\pi)(x,v,u,w) = (x,u)$. Therefore a second-order vector field $F \colon TU \to TTU$ can be written as

$$F(x,v) = \big(x, v, v, f(x,v)\big),$$

where $f \colon U \times \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map. Therefore it corresponds to the smooth function $U \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n, (x,v) \mapsto (v, f(x,v))$ (cf. Remark 1.30).

(b) In the theory of ODEs of degree 2 on the open subset $U \subseteq \mathbb{R}^n$, one observes that any second order ODE
$$\ddot{\gamma}(t) = f(\gamma(t)) \tag{17}$$
can be reduced to a first order ODE if one replaces the curve $\gamma \colon I \to U$ by the pair $\Gamma := (\gamma, \dot{\gamma}) \colon I \to TU = U \times \mathbb{R}^n$, which can be identified with the velocity curve $\gamma' \colon I \to TU$. Then γ is a solution of (17) if and only if the curve Γ is a solution of

$$\dot{\Gamma}(t) = F(\Gamma(t)), \quad F(x,v) = (v, f(x)). \tag{18}$$

In abstract terms, Γ is a curve in the tangent bundle $TU \cong U \times \mathbb{R}^n$ and the function $F \colon U \times \mathbb{R}^n \to \mathbb{R}^{2n}$ defines a vector field X_F such that the solutions of (18) are the integral curves of X_F.

Definition 4.3. A second order vector field $F \in \mathcal{V}(TM)$ has in local coordinates the form

$$F(x,v) = \big(x, v, v, f(x,v)\big).$$

We call it a *spray* if the maps $f_x(v) := f(x,v)$ are quadratic, i.e., $f_x(sv) = s^2 f_x(v)$ for $s \in \mathbb{R}$, $v \in \mathbb{R}^n$.

Remark 4.4. (Local form of sprays) That the maps f_x are quadratic implies the existence of a unique symmetric bilinear map

$$\Gamma_x \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

with
$$f_x(v) = \Gamma_x(v,v) \quad \text{for} \quad v \in \mathbb{R}^n.$$
This means that
$$\Gamma_x(v,w) = \sum_k \Gamma_{ij}^k e_k,$$
where the Γ_{ij}^k are smooth functions satisfying $\Gamma_{ij}^k = \Gamma_{ji}^k$.

If (φ, U) is the chart we use to obtain the local coordinates and b_j^φ, $j = 1, \ldots, n$, are the corresponding base fields, then

$$\Gamma(b_i^\varphi, b_j^\varphi) = \sum_k \Gamma_{ij}^k b_k^\varphi.$$

Remark 4.5. (Sprays and torsion free connections) (a) If ∇ is a torsion free connection, then it has in local coordinates on $U \subseteq \mathbb{R}^n$ the form

$$(\nabla_X Y)(x) = \mathtt{d}Y(x)X(x) + \Gamma_x(X(x), Y(x)),$$

where the bilinear forms Γ_x are symmetric. This defines a spray

$$F^\nabla(x, v) := (x, v, v, -\Gamma_x(v, v)).$$

Conversely, every spray on U is of this form.

The integral curves $\beta(t) = (\gamma(t), \dot\gamma(t)) \in TU = U \times \mathbb{R}^n$ of F^∇ are determined in local charts by the relation

$$\ddot\gamma(t) = -\Gamma_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t)),$$

which can also be written as $\nabla_{\gamma'}\gamma' = 0$. Here we distinguish between $\gamma'(t) = (\gamma(t), \dot\gamma(t))$ as an element of the tangent bundle and the velocity vector $\dot\gamma(t) \in \mathbb{R}^n$.

Therefore the integral curves $\beta\colon I \to TM$ of F^∇ are precisely the velocity curves $\beta = \gamma'$, where $\gamma\colon I \to M$ is a geodesic for ∇.

(b) This correspondence between sprays and connections can be made global (cf. [La99]). Here the main point is to verify that the transformation rule for the local forms Γ^φ corresponding to a connection ∇ by a chart (φ, U) of M are the same as the transformation rules required for the corresponding local expressions of F^∇ to define a global vector field on M.

So let $\varphi\colon U \to V$ be a diffeomorphism of open subsets of $\mathbb{R}^n$ and suppose that ∇ and ∇' are connections on U, resp., V, related by φ in the sense that

$$\varphi_*\nabla_X^Y = \nabla'_{\varphi_* X}\varphi_* Y \quad \text{for} \quad X, Y \in \mathcal{V}(U) \cong C^\infty(U).$$

Writing

$$\nabla_X Y = \mathtt{d}Y \cdot X + \Gamma(X, Y) \quad \text{and} \quad \nabla'_{X'}Y' = \mathtt{d}Y' \cdot X' + \Gamma'(X', Y'),$$

we obtain for $X' = \varphi_* X$ and $Y' = \varphi_* Y$ the relations

$$(\nabla'_{X'}Y')_{\varphi(x)} = \mathtt{d}\varphi_x((\mathtt{d}Y)_x X_x + \Gamma_x(X_x, Y_x)) = \mathtt{d}\varphi_x(\mathtt{d}Y)_x X_x + \mathtt{d}\varphi_x\Gamma_x(X_x, Y_x)$$

and

$$\begin{aligned}
(\nabla'_{X'}Y')_{\varphi(x)} &= (\mathtt{d}Y')_{\varphi(x)}X'_x + \Gamma'_{\varphi(x)}(X'_{\varphi(x)}, Y'_{\varphi(x)}) \\
&= \mathtt{d}((\mathtt{d}\varphi \cdot Y) \circ \varphi^{-1})_{\varphi(x)}(\mathtt{d}\varphi)_x X_x + \Gamma'_{\varphi(x)}((\mathtt{d}\varphi)_x X_x, (\mathtt{d}\varphi)_x Y_x) \\
&= \mathtt{d}(\mathtt{d}\varphi \cdot Y)_x \mathtt{d}(\varphi^{-1})_{\varphi(x)}(\mathtt{d}\varphi)_x X_x + \Gamma'_{\varphi(x)}((\mathtt{d}\varphi)_x X_x, (\mathtt{d}\varphi)_x Y_x) \\
&= \mathtt{d}(\mathtt{d}\varphi \cdot Y)_x X_x + \Gamma'_{\varphi(x)}((\mathtt{d}\varphi)_x X_x, (\mathtt{d}\varphi)_x Y_x) \\
&= (\mathtt{d}^2\varphi)_x(Y_x, X_x) + (\mathtt{d}\varphi)_x(\mathtt{d}Y)_x X_x + \Gamma'_{\varphi(x)}((\mathtt{d}\varphi)_x X_x, (\mathtt{d}\varphi)_x Y_x).
\end{aligned}$$

This leads to the transformation rule

$$\Gamma'_{\varphi(x)}((\mathtt{d}\varphi)_x v, (\mathtt{d}\varphi)_x w) = (\mathtt{d}\varphi)_x \Gamma_x(v, w) - (\mathtt{d}^2\varphi)_x(v, w)$$

for the Christoffel symbols.

For the corresponding second order vector fields

$$F(x, v, v, -\Gamma_x(v, v)) \quad \text{and} \quad F'(x', v', v', -\Gamma'_{x'}(v', v'))$$

we obtain with the diffeomorphism

$$T\varphi\colon TU \to TV, \quad (T\varphi)(x, v) = (\varphi(x), \mathtt{d}\varphi_x v)$$

the condition for F and F' being $T\varphi$-related:

$$F'(T\varphi(x, v)) = TT\varphi(F(x, v)).$$

As

$$T^2\varphi(x, v, a, b) = (\varphi(x), \mathtt{d}\varphi_x v, \mathtt{d}\varphi_x a, (\mathtt{d}^2\varphi)_x(v, a) + \mathtt{d}\varphi_x b),$$

we have

$$TT\varphi(x, v, v, -\Gamma_x(v, v)) = (\varphi(x), \mathtt{d}\varphi_x v, \mathtt{d}\varphi_x v, -\mathtt{d}\varphi_x \Gamma_x(v, v) + (\mathtt{d}^2\varphi)_x(v, v)),$$

which leads to the condition

$$\Gamma'_{\varphi(x)}(\mathtt{d}\varphi_x v, \mathtt{d}\varphi_x v) = \mathtt{d}\varphi_x \Gamma_x(v, v) - (\mathtt{d}^2\varphi)_x(v, v).$$

Comparing both transformation formulas, we see that the Christoffel symbols of a connection are subject to the same transformation rules as the quadratic components of a spray. Therefore the correspondence under (a) has an invariant meaning on a manifold, which leads to a global one-to-one correspondence between sprays and torsion free connections.

(c) If the connection ∇ on M is given, then the value of the corresponding spray $F^\nabla \in \mathcal{V}(TM)$ in $v \in TM$ can be calculated as follows. Let $\gamma\colon I \to M$ be a geodesic with $\gamma'(0) = v$. Then

$$F^\nabla(v) = F^\nabla(\gamma'(t)) = \gamma''(t) \in T_v(TM).$$

Since the integral curves of F^∇ are the velocity fields of the geodesics, the corresponding local flow on TM is called the *geodesic flow of* ∇.

Example 4.6. If (M, g) is a Riemannian manifold, then M carries a canonical spray, corresponding to the Levi–Civita connection (Theorem 3.28). Its geodesic flow on TM preserves the norm squared function $q(v) = g(v, v)$. Below we shall see other interpretations of this observation.

4.2 The cotangent bundle of a manifold

Let M be a smooth manifold and TM be its tangent bundle (cf. Definition 1.26). We know already that TM is a smooth manifold and that any chart (φ, U) of M leads to a chart $(T\varphi, TU)$ of TM.

In a similar fashion we can treat the *cotangent bundle* T^*M of M. As a set, it is defined as

$$T^*M := \dot{\bigcup}_{p\in M} T_p(M)^*.$$

Here we write $V^* := \mathrm{Hom}(V, \mathbb{R})$ for the dual of the vector space V, i.e., the vector space of linear maps $V \to \mathbb{R}$. If $\varphi\colon V_1 \to V_2$ is a linear map, then the corresponding linear map

$$\varphi^*\colon V_2^* \to V_1^*, \quad \alpha \mapsto \alpha \circ \varphi$$

is called its *adjoint*.

Remark 4.7. (a) We recall some facts from Linear Algebra: If $b_1, \ldots, b_n$ is a basis of the finite-dimensional real vector space V, then the *dual basis* is defined by

$$b_j^*(b_k) := \delta_{jk}.$$

Then every $\alpha \in V^*$ has a unique representation as

$$\alpha = \alpha_1 b_1^* + \cdots + \alpha_n b_n^* \quad \text{with} \quad \alpha_j = \alpha(b_j).$$

Therefore coordinates on V introduced by the choice of the basis automatically lead to coordinates on the dual space V^*.

(b) If $\mathbb{R}^n$ is considered as a space of column vectors $M_{n,1}(\mathbb{R})$, i.e., matrices of size $n \times 1$, then it is most natural to consider the elements of the dual space $(\mathbb{R}^n)^*$ as row vectors $M_{1,n}(\mathbb{R})$, i.e., matrices of the form $1 \times n$. Then the evaluation of an element $\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{R}^n)^*$ on $x \in \mathbb{R}^n$ corresponds to the matrix product

$$\alpha(x) = \alpha \cdot x = \sum_{j=1}^n \alpha_j x_j.$$

Here the entries x_j of x are the coordinates of x with respect to the canonical basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$ and the α_j are the coordinates of α w.r.t. the dual basis $e_1^*, \ldots, e_n^*$.

(c) If $L_A\colon \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Ax$ is the linear map defined by the matrix, then its adjoint is a linear map $L_A^*\colon (\mathbb{R}^n)^* \cong M_{1,n}(\mathbb{R}) \to (\mathbb{R}^n)^*$ corresponds to the map $\alpha \mapsto \alpha \circ A$ given by right multiplication with the matrix A.

On the level of coordinates we then have

$$x_i' = \sum_{j=1}^n a_{ij} x_j \quad \text{and} \quad \alpha_i' = \sum_{j=1}^n \alpha_j a_{ji},$$

so that the matrix of the adjoint map L_A^* with respect to the basis e_j^* is the transposed matrix $A^\top$.

For an open subset $U \subseteq \mathbb{R}^n$ we can identify T^*U with the product set $T^*U = U \times (\mathbb{R}^n)^* = U \times M_{1,n}(\mathbb{R}) \cong U \times \mathbb{R}^n$ which carries a natural product manifold structure. After the discussion in the preceding remark, it is clear how to glue these pieces together to obtain a smooth manifold.

If (φ, U) and (ψ, V) are two charts of M with $U \cap V \neq \emptyset$, then the diffeomorphism $\eta := \psi \circ \varphi^{-1}\colon \varphi(U \cap V) \to \psi(U \cap V)$ defines the diffeomorphism

$$T(\eta)\colon \varphi(U \cap V) \times \mathbb{R}^n \to \psi(U \cap V) \times \mathbb{R}^n, \quad (x, v) \mapsto (\eta(x), \mathtt{d}\eta(x)v),$$

where we identify the linear map $\mathtt{d}\eta(x)$ with the corresponding matrix in $M_n(\mathbb{R})$. We thus obtain transitions functions

$$T^*(\eta)\colon \varphi(U\cap V)\times(\mathbb{R}^n)^* \to \psi(U\cap V)\times(\mathbb{R}^n)^*, \quad (x,\alpha)\mapsto(\eta(x),\alpha\circ(\mathtt{d}\eta)_x^{-1}).$$

Now the same arguments that we used to obtain the manifold structure for the tangent bundle leads to a canonical manifold structure for the cotangent bundle T^*M, for which each chart (φ, U) defines a chart

$$T^*\varphi\colon T^*U \to T^*(\varphi(U)) \cong \varphi(U)\times(\mathbb{R}^n)^* \cong \varphi(U)\times\mathbb{R}^n, \quad \alpha_p \mapsto (\varphi(p), \alpha\circ T_p(\varphi)^{-1})$$

(cf. Definition 1.26).

4.3 Differential forms

Differential forms play a significant role in mathematical physics. In this subsection, we describe a natural approach to differential forms on manifolds by defining them directly as families of alternating multilinear functions on tangent spaces and not as sections of a vector bundle.

Definition 4.8. (a) If M is a smooth manifold, then a (smooth) *p-form* ω on M is a family $(\omega_x)_{x\in M}$ which associates to each $x\in M$ a p-linear alternating map $\omega_x\colon T_x(M)^p\to\mathbb{R}$ such that in local coordinates the map $(x, v_1,\ldots,v_p)\mapsto\omega_x(v_1,\ldots,v_p)$ is smooth. We write $\Omega^p(M)$ for the space of p-forms on M and identify $\Omega^0(M)$ with the space $C^\infty(M)$ of real-valued smooth functions on M.

(b) The wedge product

$$\Omega^p(M)\times\Omega^q(M)\to\Omega^{p+q}(M), \quad (\omega,\eta)\mapsto\omega\wedge\eta$$

is defined by $(\omega\wedge\eta)_x := \omega_x\wedge\eta_x$, where

$$(\omega_x\wedge\eta_x)(v_1,\ldots,v_{p+q}) := \frac{1}{p!q!}\sum_{\sigma\in S_{p+q}}\operatorname{sgn}(\sigma)\omega_x(v_{\sigma(1)},\ldots,v_{\sigma(p)})\eta_x(v_{\sigma(p+1)},\ldots,v_{\sigma(p+q)}).$$

Taking into account that the forms are alternating, this product can also be written with $\binom{p+q}{p}$ summands, which are considerably less than $(p+q)!$:

$$(\omega_x\wedge\eta_x)(v_1,\ldots,v_{p+q}) := \sum_{\sigma\in\mathrm{Sh}(p,q)}\operatorname{sgn}(\sigma)\omega_x(v_{\sigma(1)},\ldots,v_{\sigma(p)})\eta_x(v_{\sigma(p+1)},\ldots,v_{\sigma(p+q)}),$$

where $\mathrm{Sh}(p,q)$ denotes the set of all (p,q)-*shuffles* in S_{p+q}, i.e., all permutations with

$$\sigma(1)<\cdots<\sigma(p) \quad \text{and} \quad \sigma(p+1)<\cdots<\sigma(p+q).$$

An easy calculation shows that

$$\omega\wedge\eta = (-1)^{pq}\eta\wedge\omega \quad \text{for} \quad \omega\in\Omega^p(M), \eta\in\Omega^q(M). \tag{19}$$

If $\omega_1, \ldots, \omega_k$ are forms of degree $p_1, \ldots, p_k$ and $p = p_1 + \ldots + p_k$, then we obtain by induction

$$\begin{aligned}&(\omega_1 \wedge \cdots \wedge \omega_k)(v_1, \ldots, v_p) \\ &:= \sum_{\sigma \in \mathrm{Sh}(p_1, p_2, \ldots, p_k)} \mathrm{sgn}(\sigma) \omega_1(v_{\sigma(1)}, \ldots, v_{\sigma(p_1)}) \cdots \omega_k(v_{\sigma(p-p_k+1)}, \ldots, v_{\sigma(p)}),\end{aligned}$$

where $\mathrm{Sh}(p_1, \ldots, p_k)$ denotes the set of all $(p_1, \ldots, p_k)$-shuffles in S_p, i.e., all permutations with

$$\sigma(1) < \cdots < \sigma(p_1), \quad \sigma(p_1+1) < \cdots < \sigma(p_1+p_2), \quad \ldots, \sigma(p-p_k+1) < \cdots < \sigma(p).$$

For $p_1 = \ldots = p_k = 1$ and $k = p$, we obtain in particular

$$(\omega_1 \wedge \cdots \wedge \omega_k)(v_1, \ldots, v_k) := \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \omega_1(v_{\sigma(1)}) \cdots \omega_k(v_{\sigma(k)}) = \det(\omega_i(v_j)).$$

Remark 4.9. To describe differential forms in local coordinates, we recall that every alternating k-form ω on $\mathbb{R}^n$ has a unique description

$$\omega = \sum_{i_1 < \ldots < i_k} \omega_{i_1 \cdots i_k} e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*, \quad \text{where} \quad \omega_{i_1 \cdots i_k} = \omega(e_{i_1}, \ldots, e_{i_k}).$$

Accordingly, we obtain for a chart (φ, U) of M and the base fields b_j^φ, $j = 1, \ldots, n$, the representation of a k-form $\omega \in \Omega^k(U)$ by

$$\omega(v_1, \ldots, v_k) = \sum_{i_1 < \ldots < i_k} \omega_{i_1 \cdots i_k} \, \mathtt{d}\varphi_{i_1} \wedge \cdots \wedge \mathtt{d}\varphi_{i_k},$$

where

$$\omega_{i_1 \cdots i_k} = \omega(b_{i_1}^\varphi, \ldots, b_{i_k}^\varphi)$$

are smooth functions. Here we use the relation $\mathtt{d}\varphi_j(b_k^\varphi) = \delta_{ij}$.

Example 4.10. Differential forms of degree 1, so called *Pfaffian forms*, are smooth functions $\alpha \colon TM \to \mathbb{R}$ that are fiberwise linear. Typical examples arise as $\alpha = \mathtt{d}f$ for smooth functions $f \colon M \to \mathbb{R}$.

A key property of 1-forms is that they can be integrated over (piecewise) smooth paths $\gamma \colon [a, b] \to M$ via

$$\int_\gamma \alpha := \int_a^b \alpha_{\gamma(t)}(\gamma'(t))\, dt,$$

and it is easy to see that this integral does not change under reparametrization of the path as long as the endpoints are fixed.

If $\alpha = \mathtt{d}f$ for a smooth function on M, then

$$\int_\gamma \mathtt{d}f = \int_a^b \mathtt{d}f_{\gamma(t)}(\gamma'(t))\, dt = \int_a^b (f \circ \gamma)'(t)\, dt = f(\gamma(b)) - f(\gamma(a))$$

depends only on the endpoints of γ. Conversely, it a 1-form α has this property, M is arcwise connected and $x_0 \in M$ is fixed, then we obtain a smooth function

$$f\colon M \to \mathbb{R}, \quad f(x) := \int_\gamma \alpha \quad \text{for} \quad \gamma(0) = x_0, \gamma(1) = x.$$

In classical mechanics this situation arises as follows: The function $U\colon M \to \mathbb{R}$ is called a *potential* and the 1-form $F := \mathtt{d}U$ is interpreted as a force field. For a force field $F \in \Omega^1(M)$, the path integral $\int_\gamma F$ is interpreted as the *work* it requires to move a particle in the force field along the path γ. A potential U with $\mathtt{d}U = F$ exists if and only if this work only depends on the endpoints of the path. Then F is called *conservative*.

Next we briefly discuss the exterior differential for differential forms in the context of manifolds.

Definition 4.11. The *exterior differential* $\mathtt{d}\colon \Omega^p(M) \to \Omega^{p+1}(M)$ is determined uniquely by the property that we have for $X_0, \ldots, X_p \in \mathcal{V}(M)$ in the space $C^\infty(M)$ the identity

$$\begin{aligned}(\mathtt{d}\omega)(X_0, \ldots, X_p) \quad := \quad & \sum_{i=0}^p (-1)^i X_i \omega(X_0, \ldots, \widehat{X}_i, \ldots, X_p) \\ & + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_p). \qquad (20)\end{aligned}$$

To show the existence of $\mathtt{d}$ for general manifolds, the main point is to show that in a point $x \in M$ the right hand side only depends on the values of the vector fields X_i in x.

Remark 4.12. In local coordinates the exterior differential takes a rather simple form. For every chart (φ, U), the basic fields b_j^φ commute: $[b_j^\varphi, b_k^\varphi] = 0$, which leads to

$$(\mathtt{d}\omega)_{j_0,\ldots,j_p} = \sum_{i=0}^p (-1)^i \frac{\partial \omega_{j_0,\ldots,\widehat{i},\ldots,j_p}}{\partial \varphi_i}. \qquad (21)$$

For $p = 0$, this means that

$$(\mathtt{d}\omega)_j = \frac{\partial \omega}{\partial \varphi_j},$$

and for $p = 1$, we obtain for $j < k$:

$$(\mathtt{d}\omega)_{j,k} = \frac{\partial \omega_k}{\partial \varphi_j} - \frac{\partial \omega_j}{\partial \varphi_k}.$$

For $\omega = f\mathtt{d}\varphi_{i_1} \wedge \cdots \wedge \mathtt{d}\varphi_{i_k}$, formula (21) can also be written as

$$\mathtt{d}\omega = \sum_{j=1}^n \frac{\partial f}{\partial \varphi_j} \mathtt{d}\varphi_j \wedge \mathtt{d}\varphi_{i_1} \wedge \cdots \wedge \mathtt{d}\varphi_{i_k} = \mathtt{d}f \wedge \mathtt{d}\varphi_{i_1} \wedge \cdots \wedge \mathtt{d}\varphi_{i_k}. \qquad (22)$$

Example 4.13. If α is a form of degree 0, i.e., a function f, then the definition of the exterior derivative gives $\mathtt{d}f(X) = Xf$.

If α is of degree 1, then

$$\mathtt{d}\alpha(X,Y) = X\,\alpha(Y) - Y\,\alpha(X) - \alpha([X,Y]).$$

If α is of degree 2, then

$$\begin{aligned}
&\mathtt{d}\alpha(X,Y,Z)\\
&= X\alpha(Y,Z) - Y\alpha(X,Z) + Z\alpha(X,Y) - \alpha([X,Y],Z) + \alpha([X,Z],Y) - \alpha([Y,Z],X)\\
&= X\alpha(Y,Z) + Y\alpha(Z,X) + Z\alpha(X,Y) - \alpha([X,Y],Z) - \alpha([Z,X],Y) - \alpha([Y,Z],X)\\
&= \sum_{cyc.} X\alpha(Y,Z) - \alpha([X,Y],Z).
\end{aligned}$$

Lemma 4.14. *For every $\omega \in \Omega^p(M)$ we have $\mathtt{d}(\mathtt{d}\omega) = 0$.*

Proof. In local coordinates, this is an easy consequence of (22) and the Schwarz Lemma on the symmetry of second order partial derivatives. □

Definition 4.15. Extending $\mathtt{d}$ to a linear map on the space $\Omega(M) := \bigoplus_{p\in\mathbb{N}_0} \Omega^p(M)$ of all differential forms on M, the relation $\mathtt{d}^2 = 0$ implies that the space

$$Z^p_{\mathrm{dR}}(M) := \ker(\mathtt{d}|_{\Omega^p(M)})$$

of *closed p-forms* contains the space $B^p_{\mathrm{dR}}(M) := \mathtt{d}(\Omega^{p-1}(M))$ of *exact p-forms*, so that we may define the *de Rham cohomology space* by

$$H^p_{\mathrm{dR}}(M) := Z^p_{\mathrm{dR}}(M)/B^p_{\mathrm{dR}}(M).$$

Lemma 4.16. *If α and β are differential forms of degree p and q, respectively, then*

$$\mathtt{d}(\alpha \wedge \beta) = \mathtt{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathtt{d}\beta.$$

Proof. In local coordinates, this is an easy consequence of (22) and the Product Rule for functions:

$$\mathtt{d}(fg) = \mathtt{d}f \cdot g + f \cdot \mathtt{d}g.$$ □

Definition 4.17. Let M and N be smooth manifolds and $\varphi\colon M \to N$ be a smooth map. Given a differential form $\omega \in \Omega^k(N)$ we can define the *pull-back* $\varphi^*\omega$ by $(\varphi^*\omega)_p := T_p(\varphi)^*\omega_{f(p)}$, i.e.,

$$(\varphi^*\omega)_p(v_1,\ldots,v_k) := \omega_{f(p)}(T_p(\varphi)v_1,\ldots,T_p(\varphi)v_k).$$

The pull-back of ω is a differential form $\varphi^*\omega \in \Omega^k(M)$.

Proposition 4.18. *The pull-back of differential forms is compatible with products and the exterior differential. For a smooth map $\varphi\colon M \to N$, we have:*

(i) $\varphi^*(\alpha \wedge \beta) = \varphi^*\alpha \wedge \varphi^*\beta$ *for* $\alpha, \beta \in \Omega(N)$.

(ii) $\mathtt{d}(\varphi^*\omega) = \varphi^*(\mathtt{d}\omega)$ *for* $\omega \in \Omega^k(N)$.

Proof. (i) is obvious from the definitions.

(ii) If $\omega = f \in C^\infty(N) = \Omega^0(N)$, then

$$\varphi^*(\mathtt{d}f) = \mathtt{d}f \circ T\varphi = \mathtt{d}(f \circ \varphi) = \mathtt{d}(\varphi^* f).$$

Now suppose that we can write ω as an exterior product of 1-forms:

$$\omega = f \cdot \mathtt{d}y_{i_1} \wedge \ldots \wedge \mathtt{d}y_{i_k}. \tag{23}$$

Then we obtain with (i):

$$\begin{aligned}
\mathtt{d}(\varphi^*\omega) &= \mathtt{d}\left((f\circ\varphi)\varphi^*\mathtt{d}y_{i_1} \wedge \ldots \wedge \varphi^*\mathtt{d}y_{i_k}\right) = \mathtt{d}\left((f\circ\varphi)\mathtt{d}(\varphi^* y_{i_1}) \wedge \ldots \wedge \mathtt{d}(\varphi^* y_{i_k})\right) \\
&= \mathtt{d}(f\circ\varphi) \wedge \mathtt{d}(\varphi^* y_{i_1}) \wedge \ldots \wedge \mathtt{d}(\varphi^* y_{i_k}) = \varphi^*\mathtt{d}f \wedge \mathtt{d}(\varphi^* y_{i_1}) \wedge \ldots \wedge \mathtt{d}(\varphi^* y_{i_k}) \\
&= \varphi^*\mathtt{d}f \wedge \varphi^*\mathtt{d}y_{i_1} \wedge \ldots \wedge \varphi^*\mathtt{d}y_{i_k} = \varphi^*(\mathtt{d}f \wedge \mathtt{d}y_{i_1} \wedge \ldots \wedge \mathtt{d}y_{i_k}) \\
&= \varphi^*(\mathtt{d}\omega).
\end{aligned}$$

Locally all differential forms can be written as sums of terms of the type (23). Moreover, if ω vanishes on an open subset U of N, then $\mathtt{d}\omega$ vanishes on U, and $f^*\omega$ vanishes on $f^{-1}(U)$. But then also $\mathtt{d}(f^*\omega)$ vanishes on U. Therefore the claim can indeed be checked locally. $\square$

Definition 4.19. (a) (The insertion operator) If $X \in \mathcal{V}(M)$ is a vector field and $\alpha \in \Omega^k(M)$ is a differential form of degree $k > 0$ on a manifold M, we define the *insertion operator*, or *contraction*, of α with respect to X to be a form $i_X\alpha$ of degree $k-1$ given by

$$(i_X\alpha)(X_1, \ldots, X_{k-1}) := \alpha(X, X_1, \ldots, X_{k-1})$$

for $k > 0$. For $k = 0$ we put $i_X\alpha := 0$.

One also finds the notation $X \lrcorner \alpha$ for $i_X\alpha$.

(b) (Lie derivative of differential forms) One defines the Lie derivative of a differential form $\omega \in \Omega(M)$ in the direction of a vector field $X \in \mathcal{V}(M)$ by using its local flow Φ_t^X:

$$\mathcal{L}_X\omega := \frac{d}{dt}\bigg|_{t=0} (\Phi_t^X)^*\omega.$$

Proposition 4.20. *We have the following relations for the operators on differential forms:*

(i) $\mathcal{L}_X(\alpha\wedge\beta) = \mathcal{L}_X\alpha \wedge \beta + \alpha \wedge \mathcal{L}_X\beta$ *for* $\alpha, \beta \in \Omega(M)$.

(ii) $i_X(\alpha\wedge\beta) = i_X\alpha \wedge \beta + (-1)^k \alpha \wedge i_X\beta$ *for* $\alpha \in \Omega^k(M)$, $\beta \in \Omega(M)$.

(iii) $(\mathcal{L}_X\alpha)(X_1, \ldots, X_k) = X\,\alpha(X_1, \ldots, X_k) - \sum_{i=1}^k \alpha(X_1, \ldots, [X, X_i], \ldots, X_k)$ *for* $\alpha \in \Omega^k(M)$ *and* $X, X_1, \ldots, X_k \in \mathcal{V}(M)$.

(iv) *Any vector field X satisfies the Cartan formula*

$$\mathtt{d} \circ i_X + i_X \circ \mathtt{d} = \mathcal{L}_X \quad \text{on} \quad \Omega(M).$$

(v) $[\mathcal{L}_X, i_Y] = i_{[X,Y]}$ *for* $X, Y \in \mathcal{V}(M)$.

Proof. (i) follows from
$$(\Phi_t^X)^*(\alpha \wedge \beta) = (\Phi_t^X)^*\alpha \wedge (\Phi_t^X)^*\beta$$
and the product rule.

(ii) is an easy calculation.

(iii) follows directly from
$$(\Phi_t^X)^*(\omega(X_1, \dots, X_k)) = ((\Phi_t^X)^*\omega)((\Phi_{-t}^X)_*X_1, \dots, (\Phi_{-t}^X)_*X_k)),$$
the product rule and $\mathcal{L}_X Y = [X, Y]$ for $X, Y \in \mathcal{V}(M)$.

(iv) With (iii), this can directly be verified with the formula defining $\mathtt{d}$.

(v) follows directly from (iii). □

4.4 Symplectic manifolds

In Definition 3.12, an *almost symplectic manifold* was defined as a pair (M, ω), where M is a smooth n-dimensional manifold and $\omega \in \Omega^2(M)$ is such that the alternating forms $\omega_p \colon T_p(M)^2 \to \mathbb{R}$ are non-degenerate for every $p \in M$.

Definition 4.21. An almost symplectic manifold (M, ω) is said to be *symplectic* if ω is closed, i.e., $\mathtt{d}\omega = 0$.

Example 4.22. Let (V, ω) be a *symplectic vector space*, i.e. ω is a non-degenerate skew-symmetric form on V. We claim that the constant 2-form Ω, defined by $\Omega_p := \omega$ for every $p \in V$, is closed. In fact, for constant vector fields X, Y and Z, we obtain
$$\begin{aligned}\mathtt{d}\Omega(X, Y, Z) &= X\Omega(Y, Z) - Y\Omega(X, Z) + Z\Omega(X, Y)\\ &\quad - \Omega([X, Y], Z) + \Omega([X, Z], Y) - \Omega([Y, Z], X) = 0\end{aligned}$$
because the Lie brackets of constant vector fields vanish and the functions $\Omega(X, Y)$ etc. are constant, so that all terms of the form $Z\Omega(X, Y)$ also vanish. This proves that $\mathtt{d}\Omega = 0$, hence that (V, Ω) is a symplectic manifold.

Example 4.23. The most direct construction of symplectic vector spaces is to start with a finite-dimensional vector space W and endow $V := W \oplus W^*$ with the symplectic form given by
$$\omega\big((v, \alpha), (w, \beta)\big) := \beta(v) - \alpha(w).$$
For $W = \mathbb{R}^n$, $q = (q_1, \dots, q_n) \in \mathbb{R}^n$ and $p = (p_1, \dots, p_n) \in (\mathbb{R}^n)^*$ this leads to the canonical symplectic form
$$\omega((q, p), (q', p')) = \sum_{j=1}^n q_j p'_j - p_j q'_j$$
on $\mathbb{R}^{2n} \cong \mathbb{R}^n \oplus (\mathbb{R}^n)^*$.

The corresponding 2-form is given in terms of the coordinates (q_i, p_i) for elements $(q, p) \in \mathbb{R}^{2n}$ by
$$\Omega = \sum_i \mathtt{d}q_i \wedge \mathtt{d}p_i.$$
We thus obtain by restriction for each open subset $U \subseteq \mathbb{R}^n$ a natural symplectic form on $T^*U = U \times (\mathbb{R}^n)^*$

Remark 4.24. There are many reasons for assuming the closedness of the form ω in the definition of a symplectic manifold. One is that it implies, for each $p \in M$, the existence of an open neighborhood U and a chart (φ, U) into $\mathbb{R}^{2n}$ for which $\varphi^*\omega_{\mathbb{R}^{2n}} = \omega|_U$, i.e., $(U, \omega|_U)$ is symplectically isomorphic to an open subset of $\mathbb{R}^{2n}$ (Darboux Theorem) (cf. Example 4.26 below). For this to hold it is clearly necessary for ω to be closed because the form $\omega_{\mathbb{R}^{2n}}$ is closed. Hence one may think of the closedness condition $\mathtt{d}\omega = 0$ as necessary for the existence of canonical coordinates (cf. Remark 4.46).

Remark 4.25. (Metric connections in the symplectic context) (a) Another reason is related with the existence of a metric connection. Comparing with the semi-Riemannian context, where the Levi–Civita connection plays a fundamental role, one could ask when a metric torsion free connection ∇ exists for a presymplectic manifold (M, ω). Here the compatibility with ω means that

$$Z\omega(X,Y) = \mathcal{L}_Z\omega(X,Y) = \omega(\nabla_Z X, Y) + \omega(X, \nabla_Z Y) \qquad \forall X,Y,Z \in \mathcal{V}(M) \tag{24}$$

and that it is torsion free means that

$$\nabla_X Y - \nabla_Y X = [X,Y] \quad \text{for } X, Y \in \mathcal{V}(M).$$

This leads to the condition

$$\begin{aligned}
&X\omega(Y,Z) + Y\omega(Z,X) + Z\omega(X,Y)\\
&= \omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z) + \omega(\nabla_Y Z, X) + \omega(Z, \nabla_Y X) + \omega(\nabla_Z X, Y) + \omega(X, \nabla_Z Y)\\
&= \omega(\nabla_X Y - \nabla_Y X, Z) + \omega(\nabla_Z X - \nabla_X Y, Y) + \omega(\nabla_Y Z - \nabla_Z Y, X)\\
&= \omega([X,Y], Z) + \omega([Z,X], Y) + \omega([Y,Z], X).
\end{aligned}$$

which no longer includes ∇. That it is satisfied means that $\mathtt{d}\omega = 0$ (cf. Example 4.13).

(b) (Existence of metric connections: The Hess trick) For a 2-form ω and a connection ∇ on M, we write

$$(\nabla_X \omega)(Y,Z) := X\omega(Y,Z) - \omega(\nabla_X Y, Z) - \omega(X, \nabla_Y Z),$$

so that ∇ is a metric connection (w.r.t. ω) if and only if $\nabla\omega = 0$.

If (M, ω) is symplectic and $\widetilde{\nabla}$ is a torsion free connection on M, then there exists a unique connection ∇ on M defined by the relation

$$\omega(\nabla_X Y, Z) = \omega(\widetilde{\nabla}_X Y, Z) + \frac{1}{3}(\widetilde{\nabla}_X \omega)(Y,Z) + \frac{1}{3}(\widetilde{\nabla}_Y \omega)(X,Z)$$

(see also Lemma 4.29 below). Then an easy calculation shows that $\nabla\omega = 0$, i.e., ∇ is a metric connection, and since $\widetilde{\nabla}_X Y - \widetilde{\nabla}_X Y$ is a symmetric function in X and Y, ∇ is also torsion free.

(c) If ∇' and ∇ are two metric torsion free connections for (M, ω), then $B_X(Y) := \nabla'_X Y - \nabla_X Y$ is $C^\infty(M)$-bilinear, so that $B_X(Y)(p) = B_{X(p)}(Y(p))$ holds for endomorphisms $B_v \in \mathrm{End}(T_p(M))$, $v \in T_p(M)$. We then have

$$\omega(B_X Y, Z) + \omega(Y, B_X Z) = 0 \quad \text{for} \quad X, Y, Z \in \mathcal{V}(M),$$

i.e., $B_{X(p)} \in \mathfrak{sp}(T_p(M), \omega_p)$ for every $p \in M$. The uniqueness of the Levi–Civita connection in the Riemannian case implies that, in the semi-Riemannian case, any such B vanishes. However, in the symplectic case such fields may exist, which implies that metric connections are not unique.

For $V = \mathbb{R}^2$, non-zero maps of this type exist. They correspond to $B_1, B_2 \in \mathfrak{sl}_2(\mathbb{R})$ with $B_1 e_2 = B_2 e_1$, so that

$$B_1 = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \quad \text{and} \quad B_2 = \begin{pmatrix} b & d \\ -a & -b \end{pmatrix}.$$

Therefore we obtain a 4-parameter family of linear maps $B \colon \mathbb{R}^2 \to \mathfrak{sl}_2(\mathbb{R}) \cong \mathrm{sp}_2(\mathbb{R})$ defining symmetric maps $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$.

In Example 4.23 we have seen that, for any open subset $U \subseteq \mathbb{R}^n$, the cotangent bundle $T^*U \subseteq \mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{R}^{2n}$ carries a canonical symplectic structure. This structure is actually natural, i.e., compatible with diffeomorphisms, so that we can use it to obtain a canonical symplectic structure on any cotangent bundle.

Example 4.26. (T^*Q as a symplectic manifold) Let Q be a smooth manifold and $\pi \colon T^*Q \to Q$ be the canonical projection. We define the *Liouville* 1*-form* Θ on T^*Q by

$$\Theta_\alpha(v) := \alpha(T\pi(v)) \quad \text{for} \quad v \in T_\alpha(T^*Q)$$

and consider the 2-form

$$\Omega := -\mathtt{d}\Theta \in \Omega^2(T^*Q).$$

We claim that (T^*Q, Ω) is symplectic.

To verify this claim, we take a closer look at Θ and Ω in a local chart. From a chart (φ, U), $U \subseteq Q$ open, we obtain the corresponding cotangent chart $T^*\varphi \colon T^*U \to \mathbb{R}^{2n} \cong \mathbb{R}^n \times \mathbb{R}^n$. Writing elements of $\mathbb{R}^{2n}$ as pairs (q, p), we obtain coordinates $q_1, \ldots, q_n; p_1, \ldots, p_n$ on U.

In these coordinates we have

$$\pi(q, p) = q \quad \text{and} \quad T\pi(q, p, v, w) = (q, v).$$

For $\alpha = (q, p) \in T^*\mathbb{R}^n$ we thus obtain

$$\Theta_\alpha(v, w) = \langle p, v \rangle = \sum_{i=1}^n v_i p_i,$$

which can also be written as

$$\Theta_\alpha = \sum_{j=1}^n p_i \mathtt{d}q_i,$$

which leads to

$$\Omega_\alpha = \sum_{j=1}^n \mathtt{d}q_i \wedge \mathtt{d}p_i.$$

This shows in particular that Ω_α is non-degenerate in every $\alpha \in T^*Q$, and hence that (T^*Q, Ω) is a symplectic manifold.

Lemma 4.27. (Naturality of the Liouville form) *If $\varphi\colon M \to N$ is a diffeomorphism, then*

$$T^*\varphi\colon T^*M \to T^*N, \quad \alpha_p \mapsto \alpha_p \circ T_p(\varphi)^{-1}$$

also is a diffeomorphism and the Liouville forms $\Theta_M \in \Omega^1(T^*M)$ and $\Theta_N \in \Omega^1(T^*N)$ satisfy

$$(T^*\varphi)^*\Theta_N = \Theta_M.$$

Proof. We write $\pi_N : T^*N \to N$ and $\pi_M : T^*M \to M$ for the canonical projections, so that we have

$$\pi_N \circ T^*\varphi = \varphi \circ \pi_M.$$

For $\alpha \in T_p(M)^*$ we have

$$\begin{aligned}((T^*\varphi)^*\Theta_N)_\alpha(v) &= (\Theta_N)_{T^*\varphi(\alpha)}(T(T^*\varphi)v) = (T^*\varphi)(\alpha)(T(\pi_N)T(T^*\varphi)v) \\ &= (T^*\varphi)(\alpha)(T(\pi_N \circ T^*\varphi)v) = (T^*\varphi)(\alpha)(T(\varphi \circ \pi_M)v) \\ &= (\alpha \circ (T\varphi)^{-1} \circ T\varphi \circ T\pi_M)v = \alpha(T(\pi_M)v) = \Theta_M(v).\end{aligned}$$

□

Remark 4.28. Let M be a smooth manifold, $\mathrm{Diff}(M)$ be its group of diffeomorphisms and $\mathrm{Symp}(T^*M, \Omega_M)$ be the group of symplectic diffeomorphisms of T^*M. The preceding lemma implies that

$$(T^*\varphi)^*\Omega_M = -(T^*\varphi)^*\mathtt{d}\Theta_M = -\mathtt{d}(T^*\varphi)^*\Theta_M = -\mathtt{d}\Theta_M = \Omega_M,$$

so that $T^*\varphi \in \mathrm{Symp}(T^*M, \Omega_M)$. Moreover, for $\varphi, \psi \in \mathrm{Diff}(M)$, we have

$$\begin{aligned}T^*(\varphi \circ \psi)\alpha_p &= \alpha_p \circ T_p(\varphi \circ \psi)^{-1} = \alpha_p \circ T_p(\psi)^{-1} \circ T_{\psi(p)}(\varphi)^{-1} \\ &= T^*(\varphi)(\alpha_p \circ T_p(\psi)^{-1}) = T^*(\varphi)T^*(\psi)\alpha_p,\end{aligned}$$

so that

$$\mathrm{Diff}(M) \to \mathrm{Symp}(T^*M, \Omega_M), \quad \varphi \mapsto T^*\varphi$$

is a group homomorphism. That it is injective follows from the relation $\pi_M \circ T^*\varphi = \varphi \circ \pi_M$.

4.5 Hamiltonian vector fields and Poisson brackets

We now turn to the formalism of Hamiltonian vector fields and Poisson brackets. Here a key point is the passage from cotangent bundles T^*Q on which the group $\mathrm{Diff}(Q)$ acts naturally to symplectic manifolds (M, Ω). It provides an environment with more symmetries, represented by the group $\mathrm{Symp}(M, \Omega)$ of symplectic diffeomorphisms, also called *canonical transformations* (cf. Remark 4.28). In the context of physics, where one is interested in the solutions of the Hamilton equations

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial H}{\partial q_j},$$

this allows one to work with coordinates in which these equations take a simpler form. Sometimes this even leads to explicit solutions of the equations of motion. Following this idea systematically leads to the notion of a completely integrable system, a concept which is connected to many branches of mathematics and physics. Here we shall simply scratch the surface.

Let (M, Ω) be a symplectic manifold. Then each $v \in T_p(M)$ defines an element

$$v^\flat \in T_p^*(M), \quad v^\flat(w) := \Omega_p(v, w).$$

This leads to an isomorphism $T_p(M) \to T_p^*(M)$. Its inverse is denoted

$$T_p^*(M) \ni \alpha \mapsto \alpha^\sharp \in T_p(M), \quad \omega_p(\alpha^\sharp, w) = \alpha(w), \quad w \in T_p(M).$$

Lemma 4.29. *The maps $\flat$ and $\sharp$ define diffeomorphisms*

$$\flat\colon T(M) \to T^*(M) \quad \textit{and} \quad \sharp\colon T^*(M) \to T(M)$$

restricting to linear maps on each tangent, resp., cotangent space.

Proof. It clearly suffices to verify this locally, so that we may assume that M is an open subset of $\mathbb{R}^n$. Then ω is represented by a smooth function

$$\Omega\colon M \to M_n(\mathbb{R}) \quad \text{by} \quad \omega_x(v, w) = v^\top \Omega_x w.$$

For $v \in T_x(M) \cong \mathbb{R}^n$ we then have $v^\flat = v^\top \Omega_x \in M_{1,n}(\mathbb{R})$ (row vector) and, accordingly, $\alpha^\sharp = (\alpha \Omega_x^{-1})^\top = -\Omega_x^{-1} \alpha^\top$ for $\alpha \in M_{1,n}(\mathbb{R}) \cong (\mathbb{R}^n)^*$. It is clear that the map

$$\flat\colon TM = M \times \mathbb{R}^n \to T^*M \cong M \times (\mathbb{R}^n)^*, \quad (x, v) \mapsto (x, v^\top \Omega_x)$$

is smooth with inverse

$$\sharp\colon T^*M \cong M \times (\mathbb{R}^n)^* \to TM = M \times \mathbb{R}^n, \quad (x, \alpha) \mapsto (x, -\Omega_x^{-1} \alpha^\top). \qquad \square$$

With the diffeomorphism $\flat\colon TM \to T^*M$ from Lemma 4.29 which is linear on each tangent space, we obtain linear bijections

$$\flat\colon \mathcal{V}(M) \to \Omega^1(M), \quad X^\flat(p) := X(p)^\flat \quad \text{and} \quad \sharp\colon \Omega^1(M) \to \mathcal{V}(M), \quad \alpha^\sharp(p) := \alpha(p)^\sharp$$

satisfying

$$i_{\alpha^\sharp} \Omega = \alpha. \tag{25}$$

Definition 4.30. Let (M, Ω) be a symplectic manifold (M, Ω).

(a) For $H \in C^\infty(M)$, the vector field

$$X_H := (\mathtt{d}H)^\sharp$$

is called the *Hamiltonian vector field* associated to the function H. It is uniquely determined by the relation

$$i_{X_H} \Omega = \mathtt{d}H.$$

The corresponding local flow $\Phi_t^{X_H}$ is called the *Hamiltonian flow* and

$$\dot\gamma(t) = X_H(\gamma(t))$$

the corresponding *Hamiltonian equation*.

(b) A vector field $X \in \mathcal{V}(M)$ is called *symplectic* if $\mathcal{L}_X\Omega = 0$, i.e., if Ω is invariant under the corresponding local flow (Exercise 4.4).

(c) For $F, G \in C^\infty(M)$ we define the *Poisson bracket* by

$$\{F, G\} := \Omega(X_F, X_G) = \mathtt{d}F(X_G) = X_G F$$

(cf. (25)).

Remark 4.31. The Poisson bracket of two functions can be used to describe the change of the values of a function F on the integral curves γ of the Hamiltonian vector field X_H:

$$\frac{d}{dt}F(\gamma(t)) = \mathtt{d}F_{\gamma(t)}\gamma'(t) = \mathtt{d}F_{\gamma(t)}X_H(\gamma(t)) = (X_H F)(\gamma(t)) = \{F, H\}(\gamma(t)),$$

i.e.,

$$\dot{F} = \{F, H\}, \tag{26}$$

in the sense that the change rate of the function F on any integral curve is given by this differential equation.

Example 4.32. For an open subset $U \subseteq \mathbb{R}^{2n}$, endowed with the canonical symplectic form

$$\Omega = \sum_{j=1}^n \mathtt{d}q_i \wedge \mathtt{d}p_i$$

we obtain for a vector field $X = (Y, Z)$ the relation

$$i_X\Omega = \sum_{j=1}^n Y_j \mathtt{d}p_j - Z_j \mathtt{d}q_j.$$

In view of

$$\mathtt{d}H = \sum_{j=1}^n \frac{\partial H}{\partial q_i}\,\mathtt{d}q_i + \frac{\partial H}{\partial p_i}\,\mathtt{d}p_i,$$

we thus obtain for $H \in C^\infty(\mathbb{R}^{2n})$:

$$X_H = \Big(\frac{\partial H}{\partial p_1}, \ldots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \ldots, -\frac{\partial H}{\partial q_n}\Big).$$

In coordinates, we thus obtain for the Poisson bracket

$$\{F, G\} = X_G F = \sum_{j=1}^n \frac{\partial F}{\partial q_j}\frac{\partial G}{\partial p_j} - \frac{\partial F}{\partial p_j}\frac{\partial G}{\partial q_j}.$$

For the coordinate functions q_i and p_j we obtain in particular the canonical Poisson relations

$$\{q_j, q_i\} = \{p_j, p_i\} = 0 \quad \text{and} \quad \{q_j, p_i\} = \delta_{ij}. \tag{27}$$

For a curve $\gamma(t) = (q(t), p(t))$, the *Hamiltonian equation* associated to $H \in C^\infty(U)$ now has the form:

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial H}{\partial q_j}.$$

The following proposition clarifies some of the relations between these concepts.

Proposition 4.33. *The following assertions hold:*

(i) $\mathcal{L}_{X_f}\Omega = 0$ *for every* $f \in C^\infty(M)$*, i.e., every Hamiltonian vector field is symplectic.*

(ii) *The Poisson bracket is a Lie bracket on* $C^\infty(M)$ *and satisfies the Leibniz rule*

$$\{f, gh\} = \{f, g\}h + g\{f, h\}, \quad f, g, h \in C^\infty(M).$$

(iii) $[X_f, X_g] = X_{\{g,f\}}$ *for* $f, g \in C^\infty(M)$*, so that*

$$(C^\infty(M), \{\cdot,\cdot\}) \to \mathcal{V}(M), \quad f \mapsto -X_f$$

is a homomorphism of Lie algebras.

Proof. (i) From the Cartan formula $\mathcal{L}_X = \mathtt{d} \circ i_X + i_X \circ \mathtt{d}$ (Proposition 4.20), we derive

$$\mathcal{L}_{X_f}\Omega = \mathtt{d}\big(i_{X_f}\Omega\big) + i_{X_f}\mathtt{d}\Omega = \mathtt{d}(\mathtt{d}f) = 0.$$

(iii) From the other Cartan formula $[\mathcal{L}_X, i_Y] = i_{[X,Y]}$ (Proposition 4.20), we obtain with (i)

$$\begin{aligned} i_{[X_f,X_g]}\Omega &= [\mathcal{L}_{X_f}, i_{X_g}]\Omega = \mathcal{L}_{X_f}\big(i_{X_g}\Omega\big) = \mathcal{L}_{X_f}\mathtt{d}g \\ &= \mathtt{d}\big(i_{X_f}\mathtt{d}g\big) + i_{X_f}\mathtt{d}(\mathtt{d}g) = \mathtt{d}\big(i_{X_f}\mathtt{d}g\big) = \mathtt{d}\{g, f\} = i_{X_{\{g,f\}}}\Omega. \end{aligned}$$

Since Ω is non-degenerate, this implies (iii).

(ii) It is clear that $\{\cdot,\cdot\}$ is bilinear and skew-symmetric, and from $\mathtt{d}(fg) = f\mathtt{d}g + g\mathtt{d}f$ we conclude that it satisfies the Leibniz rule. So it remains to check the Jacobi identity. This is an easy consequence of (iii):

$$\begin{aligned} \big\{f, \{g, h\}\big\} &= X_{\{g,h\}}f = -[X_g, X_h]f \\ &= -X_g(X_h f) + X_h(X_g f) = \big\{h, \{g, f\}\big\} - \big\{g, \{h, f\}\big\}. \end{aligned}$$

□

As a corollary, we obtain Jacobi's great insight from about 1830 that was based on his discovery of the Jacobi identity for the Poisson bracket.

Corollary 4.34. *For* $H \in C^\infty(M)$ *a function* $F \in C^\infty(M)$ *is constant on the integral curves of* X_H *if and only if* $\{F, H\} = 0$*. The set of all these functions is a Lie subalgebra of* $(C^\infty(M), \{\cdot,\cdot\})$*.*

Proof. This follows from

$$\{\{F_1, F_2\}, H\} = \{F_1, \{F_2, H\}\} + \{\{F_1, H\}, F_2\}$$

for $F_1, F_2 \in C^\infty(M)$. □

Symmetries and conserved quantities

Definition 4.35. We consider the flow of the Hamiltonian vector field X_H on the symplectic manifold (M, Ω) and call (M, Ω, H) the corresponding *Hamiltonian system.*

A smooth function $F \in C^\infty(M)$ is called a *constant of motion* or *conserved quantity* if F is constant along the integral curves of X_H.

A vector field X is called *an infinitesimal symmetry* of the Hamiltonian system (M, Ω, H) if $\mathcal{L}_X \Omega = 0$ and $XH = 0$.

Theorem 4.36. (Hamilton Version of E. Noether's Theorem) *For each conserved quantity F of the Hamiltonian system (M, Ω, H) the corresponding vector field X_F is an infinitesimal symmetry. Conversely, a Hamiltonian vector field X_F is an infinitesimal symmetry if and only if F is a conserved quantity.*

Proof. This is a direct consequence of Corollary 4.34. If F is a conserved quantity, then the Hamiltonian vector field X_F satisfies $X_F H = \{H, F\} = 0$, so that X_F is a symmetry.

If, conversely, X_F is a symmetry, then $0 = X_F H = \{H, F\} = -X_H F$ implies that F is a conserved quantity. □

Examples 4.37. We consider some examples of Hamiltonian systems on open subsets $U \subseteq \mathbb{R}^{2n} = T^*(\mathbb{R}^n)$ and a Hamiltonian $H \in C^\infty(U)$, corresponding to the energy of the related mechanical system.

(a) The relation $\{H, H\} = 0$ means that H itself is a conversed quantity, i.e., energy is preserved. The corresponding symmetry is represented by the vector field X_H generating the dynamics of the system. In this sense one can also say that the time-independence of X_H corresponds to the preservation of energy.

(b) If H is invariant under the translations $\tau_v \colon \mathbb{R}^n \to \mathbb{R}^n, q \mapsto q + v$, resp., the induced diffeomorphisms

$$\tau_v^* \colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}, \quad (q, p) \mapsto (q + v, p),$$

then the vector field $X(q, p) = (v, 0)$ is a symmetry. As $X = X_{P_v}$ for $P_v(q, p) := \sum_{j=1}^n v_i p_i$, we see that the corresponding conserved quantity is the linear momentum P_v in direction v.

This means that translation invariance corresponds to the conservation of linear momenta.

A typical situation where all linear momenta are preserved is the force free motion of a particle in $\mathbb{R}^3$, where

$$H(q, p) = \frac{\|p\|^2}{2m} \quad \text{or, more generally,} \quad H(q, p) = f(p).$$

(c) On $\mathbb{R}^3$ we consider the vector fields

$$L_j \colon \mathbb{R}^3 \to \mathbb{R}^3, \quad L_j(x) = e_j \times x,$$

where $\times$ denotes the vector product on $\mathbb{R}^3$. These vector fields generate the rotations $(R_i^t)_{t \in \mathbb{R}}$, $i = 1, 2, 3$, around the coordinate axes (cf. Example 2.31).

For each $R \in \mathrm{O}_3(\mathbb{R})$ the corresponding diffeomorphism of $\mathbb{R}^3$ induces a natural symplectic diffeomorphism

$$T^*(R) \colon T^*(\mathbb{R}^3) \cong \mathbb{R}^6 \to \mathbb{R}^6, \quad T^*(R)\alpha = \alpha \circ T(R)^{-1}, \quad T^*(R)(q, p) := (Rq, Rp)$$

(cf. Lemma 4.27). The infinitesimal generator of the flow $T^*(R_i^t)$ on $T^*(\mathbb{R}^3)$ is therefore given by the vector field

$$\widetilde{L}_i(q,p) = (L_i(q), L_i(p)), \quad q,p \in \mathbb{R}^3.$$

The corresponding Hamiltonian function on $T^*Q \cong \mathbb{R}^3 \times \mathbb{R}^3$ is

$$J_{L_i}(q,p) := \langle e_i, q \times p \rangle = \langle e_i \times q, p \rangle = \det(e_i, q, p) = \sum_{j,k} \varepsilon_{ijk} q_j p_k.$$

These functions are called the *angular momenta*.

That the vector field $\widetilde{L}_i$ is an infinitesimal symmetry is equivalent to the invariance of the Hamiltonian H under the rotation group

$$H(R_i^t q, R_i^t p) = H(q,p) \quad \text{for all} \quad t \in \mathbb{R}.$$

Therefore the angular momentum J_{L_i} with respect to the q_i-axis is a conserved quantity if and only if H is invariant under the corresponding rotations.

If H is invariant under the full group $\mathrm{SO}_3(\mathbb{R})$, resp., under all $T^*(R)$, $R \in \mathrm{SO}_3(\mathbb{R})$, then all angular momenta are conserved. In this sense the vector

$$(J_{L_1}, J_{L_2}, J_{L_3})$$

is conserved and in particular its square length, the *total angular momentum*

$$J := J_{L_1}^2 + J_{L_2}^2 + J_{L_3}^2$$

is conserved.

A typical situation where all angular momenta are preserved is the motion of a particle in $\mathbb{R}^3$ under a central force field

$$H(q,p) = \frac{\|p\|^2}{2m} + V(\|q\|), \quad \text{or, more generally,} \quad H(q,p) = f(\|q\|, \|p\|).$$

Remark 4.38. If the symplectic manifold (M, Ω) is *exact* in the sense that there exists a 1-form $\Theta \in \Omega^1(M)$ with $-\mathtt{d}\Theta = \Omega$, then it is easy to find for each $X \in \mathcal{V}(M)$ with $\mathcal{L}_X\Theta = 0$ a corresponding Hamiltonian function:

$$\mathtt{d}(i_X\Theta) = \mathcal{L}_X\Theta - i_X(\mathtt{d}\Theta) = i_X\Omega$$

implies that $f := \Theta(X) \in C^\infty(M)$ is a smooth function with $X_f = X$.

Definition 4.39. Let Q be a smooth manifold. Any vector field $X \in \mathcal{V}(Q)$ defines a smooth function

$$P_X \colon T^*Q \to \mathbb{R}, \quad P_X(\alpha_q) := \alpha_q(X(q)),$$

called the *momentum of* X.

Remark 4.40. The momentum P_X of a vector field is a smooth function on the symplectic manifold T^*Q. To understand the structure of the corresponding Hamiltonian flow, note that the local flow Φ_t^X of X on Q induces a local flow on T^*Q by the maps $T^*(\Phi_t^X)$ (cf. Lemma 4.27). We claim that this is the local flow defined by P_X.

Let

$$\widetilde{X}_\alpha := \frac{d}{dt}\Big|_{t=0} T^*(\Phi_t^X)\alpha = \frac{d}{dt}\Big|_{t=0} \alpha \circ T(\Phi_{-t}^X)$$

denote the generator of this flow. It is called the *canonical lift of* X to T^*Q. Then Lemma 4.27 implies that $\mathcal{L}_{\widetilde{X}}\Theta = 0$, so that Remark 4.38 shows that $\Theta(\widetilde{X})$ is a Hamiltonian function for $\widetilde{X}$. As the construction of $\widetilde{X}$ implies that $T(\pi)\widetilde{X}_{\alpha_p} = X_p$, it follows that $\Theta(\widetilde{X})(\alpha_p) = \alpha_p(X_p) = P_X(\alpha_p)$, and this proves our claim.

Now we show that

$$\{P_X, P_Y\} = P_{[Y,X]} \quad \text{for} \quad X, Y \in \mathcal{V}(Q). \tag{28}$$

This implies in particular, that

$$[\widetilde{X}, \widetilde{Y}] = [X_{P_X}, X_{P_Y}] = X_{\{P_Y, P_X\}} = X_{P_{[X,Y]}} = \widetilde{[X,Y]}.$$

In other words, $\mathcal{V}(Q) \to \mathcal{V}(T^*Q), X \mapsto \widetilde{X}$ is a homomorphism of Lie algebras.

To verify (28), we first observe that, for $\varphi \in \mathrm{Diff}(Q)$, we have

$$P_{\varphi_* X}(\alpha_q) = \alpha_q(T_{\varphi^{-1}(p)}(\varphi)X_{\varphi^{-1}(p)}) = (T^*(\varphi)^{-1}\alpha_q)(X_{\varphi^{-1}(p)}) = P_X(T^*(\varphi)^{-1}\alpha_q),$$

which implies that

$$P_{\varphi_* X} = P_X \circ T^*(\varphi)^{-1}.$$

Therefore

$$\{P_X, P_Y\} = \widetilde{Y} P_X = \frac{d}{dt}\Big|_{t=0} P_X \circ \Phi_t^{\widetilde{Y}} = \frac{d}{dt}\Big|_{t=0} P_X \circ T^*(\Phi_t^Y) = \frac{d}{dt}\Big|_{t=0} P_{(\Phi_{-t}^Y)_* X} = P_{[Y,X]}.$$

Examples 4.41. (a) (Linear momenta) For $Q = \mathbb{R}^3$ and the constant vector fields $P_j = e_j$ the corresponding momentum function on $T^*Q \cong \mathbb{R}^3 \times \mathbb{R}^3$ is given by

$$J_{P_j}(q,p) := p_j = \langle P_j(q), p\rangle.$$

(b) (Angular momenta) For $Q = \mathbb{R}^3$ and the linear vector fields $L_j(q) = e_j \times q$ generating the rotations around the coordinate axes (cf. Example 2.31), the corresponding momentum function on $T^*Q \cong \mathbb{R}^3 \times \mathbb{R}^3$ is given by

$$J_{L_j}(q,p) := \langle e_j, q \times p\rangle.$$

Darboux charts

In this subsection we briefly discuss Darboux charts of symplectic manifolds (cf. Remark 4.24). Their existence is the main motivation for the introduction of symplectic manifolds in classical mechanics because it shows that they provide a geometric structure which locally looks like open subsets of $\mathbb{R}^{2n}$, endowed with its canonical symplectic structure. As we shall see, Darboux charts correspond to local coordinates (q_i, p_i) satisfying the canonical Poisson relations

$$\{q_j, q_i\} = \{p_j, p_i\} = 0 \quad \text{and} \quad \{q_j, p_i\} = \delta_{ij}. \tag{29}$$

Lemma 4.42. *Let (V, ω) and (W, ω') be symplectic vector spaces and $\psi \colon V \to W$ be a linear isomorphism. Then the following are equivalent:*

(i) *φ is symplectic, i.e., $\varphi^*\omega' = \omega$.*

(ii) $\psi^* \circ \flat_W \circ \psi \circ \sharp_V = \mathrm{id}_{V^*}$.

(iii) $\sharp_W = \psi \circ \sharp_V \circ \psi^*$.

(iv) $\{\psi^*\alpha, \psi^*\beta\} = \omega((\psi^*\alpha)^\sharp, (\psi^*\beta)^\sharp) = \omega'(\alpha^\sharp, \beta^\sharp) = \{\alpha, \beta\}$ *for* $\alpha, \beta \in W^*$.

Proof. (i) $\Leftrightarrow$ (ii): That ψ is symplectic is equivalent to the coincidence of

$$\alpha(v) = \omega(\alpha^\sharp, v) \quad \text{and} \quad \omega'(\psi\alpha^\sharp, \psi v) = (\psi^*\psi(\alpha^\sharp)^\flat)(v)$$

for $\alpha \in V^*$, $v \in V$. This is equivalent to $\psi^* \circ \flat_W \circ \psi \circ \sharp_V = \mathrm{id}_{V^*}$.

(ii) $\Leftrightarrow$ (iii): Since ψ is invertible, (ii) is equivalent to

$$\flat_V \circ \psi^{-1} \circ \sharp_W \circ (\psi^*)^{-1} = \mathrm{id}_{V^*},$$

which in turn is equivalent to (iii).

(iii) $\Rightarrow$ (iv): From (iii) we get

$$\omega((\psi^*\alpha)^\sharp, (\psi^*\beta)^\sharp) = \omega'(\psi \circ \sharp_V \circ \psi^*\alpha, \psi \circ \sharp_V \circ \psi^*\beta) = \omega'(\alpha^\sharp, \beta^\sharp)$$

for $\alpha, \beta \in W^*$.

(iv) $\Rightarrow$ (iii): From (iv) we obtain

$$\alpha\big(\psi((\psi^*\beta)^\sharp)\big) = \omega((\psi^*\alpha)^\sharp, (\psi^*\beta)^\sharp) = \omega'(\alpha^\sharp, \beta^\sharp) = \alpha(\beta^\sharp)$$

for $\alpha, \beta \in W^*$. This means that $\psi \circ \sharp_V \circ \psi^* = \sharp_W$. $\square$

Definition 4.43. A smooth map $\varphi \colon (M, \Omega) \to (M, \Omega')$ between symplectic manifolds is called a *Poisson map* if

$$\varphi^*\{F, G\} = \{\varphi^*F, \varphi^*G\} \quad \text{for} \quad F, G \in C^\infty(M').$$

Lemma 4.44. *If (M, Ω) and (M', Ω') are symplectic manifolds and $\varphi \colon M \to M'$ is a symplectic diffeomorphism, then φ is a Poisson map.*

Proof. Applying Lemma 4.42 to the tangent maps $T_p(\varphi) \colon T_p(M) \to T_p(N)$, we obtain

$$\begin{aligned}\{\varphi^*F, \varphi^*G\}(p) &= \Omega_p((T_p(\varphi)^*\mathtt{d}F_{\varphi(p)})^\sharp, ((T_p(\varphi)^*\mathtt{d}G_{\varphi(p)})^\sharp) \\ &= \Omega'_{\varphi(p)}((\mathtt{d}F)^\sharp_{\varphi(p)}, (\mathtt{d}G)^\sharp_{\varphi(p)}) = \{F, G\}(\varphi(p)).\end{aligned}$$ $\square$

Remark 4.45. One may expect that the preceding lemma holds without the assumption of φ being a diffeomorphism. A closer inspection shows that we actually used that it is a local diffeomorphism, i.e., that all its tangent maps are injective. However, one cannot go beyond that, as the following example shows.

Consider the inclusion

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}^4, \quad (q_1, p_1) \mapsto (q_1, 0, p_1, 0).$$

Then the canonical symplectic forms satisfy

$$\varphi^*\Omega_{\mathbb{R}^4} = \Omega_{\mathbb{R}^2}, \quad \varphi^* \sum_{j=1}^{2} \mathtt{d}q_j \wedge \mathtt{d}p_j = \mathtt{d}q_1 \wedge \mathtt{d}p_1$$

but for the Poisson brackets we find

$$\{\varphi^* q_2, \varphi^* p_2\} = \{0, 0\} = 0 \neq 1 = \varphi^* 1 = \varphi^*\{q_2, p_2\}.$$

Actually, one can fabricate a smaller example by the inclusion

$$\varphi\colon \mathbb{R}^0 = \{0\} \to \mathbb{R}^2, \quad 0 \mapsto (0, 0)$$

which satisfies

$$\{\varphi^* q, \varphi^* p\} = 0 \neq 1 = \varphi^* 1 = \varphi^*\{q, p\}.$$

Remark 4.46. (Darboux charts) A chart (φ, U) of the symplectic manifold (M, Ω) is a chart $\varphi\colon U \to \mathbb{R}^{2n}$ for which $\Omega|_U = \varphi^*\Omega_{\mathbb{R}^{2n}}$. We want to give a criterion for a chart to be a Darboux chart in terms of Poisson brackets of the coordinate functions Q_i and P_i defined by $\varphi = (Q_1, \ldots, Q_n, P_1, \ldots, P_n)$.

If φ is a Darboux chart, then Lemma 4.44 implies that

$$\{Q_j, Q_i\} = \{P_j, P_i\} = 0 \quad \text{and} \quad \{Q_j, P_i\} = \delta_{ij}.$$

Suppose, conversely, that these relations are satisfied. Since the linear functions p_j and q_j on $\mathbb{R}^{2n}$ span the dual space, we obtain

$$\Omega_p(((T_p\varphi)^*\alpha)^\sharp, ((T_p\varphi)^*\beta)^\sharp) = \{\varphi^*\alpha, \varphi^*\beta\} = \{\alpha, \beta\} = \Omega_{\mathbb{R}^{2n}}(\alpha^\sharp, \beta^\sharp)$$

for $\alpha, \beta \in (\mathbb{R}^{2n})^*$. Now Lemma 4.42 implies that $\varphi^*\Omega_{\mathbb{R}^{2n}} = \Omega|_U$, i.e., that Ω is a Darboux chart.

Without proof we state the following key theorem of symplectic geometry (cf. [MR99]). It implies in particular that (M, Ω) has an atlas consisting of Darboux charts, so that $2n$-dimensional symplectic manifolds can be considered as obtained by gluing open subsets of $\mathbb{R}^{2n}$ by symplectic diffeomorphisms of open subsets.

Theorem 4.47. (Darboux Theorem) *Let (M, Ω) be a symplecitc manifold and $p \in M$. Then there exists a Darboux chart (φ, U) with $p \in U$.*

4.6 Lagrangian mechanics

The idea behind the Lagrangian formulation of mechanics is that the equations of motion encoded in Newton's law

$$\mathbf{F} = m\mathbf{a}$$

can be derived from variational principles which are based on the Lagrangian L of the system. If $q = (q_1, \dots, q_n)$ are the coordinates of an element of a configuration space $Q \subseteq \mathbb{R}^n$, then L is a function of the form

$$L(q_i, \dot{q}_i, t) = L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t),$$

where $\dot{q} = \frac{dq}{dt} = (\dot{q}_1, \dots, \dot{q}_n)$ is the system velocity. Hamilton's variational principle

$$\delta \int_a^b L(q_i, \dot{q}_i, t)\, dt = 0$$

then leads to the *Euler–Lagrange equations*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, \dots, n. \tag{30}$$

Example 4.48. (a) For a system of N particles moving in $\mathbb{R}^3$, the configuration space is an open subset $Q \subseteq \mathbb{R}^{3N}$ and L often has the form

$$L(\mathbf{q}_i, \dot{\mathbf{q}}_i, t) = \frac{1}{2}\sum_{i=1}^{N} m_i \|\dot{\mathbf{q}}_i\|^2 - V(\mathbf{q}_1, \dots, \mathbf{q}_n),$$

where $\mathbf{q}_i \in \mathbb{R}^3$ is the location of the ith particle. Then the Euler–Lagrange equations reduce to *Newton's second law*

$$\frac{d}{dt}(m_i \mathbf{q}_i) = -\frac{\partial V}{\partial \mathbf{q}_i}, \quad i = 1, \dots, N,$$

which is $\mathbf{F} = m\mathbf{a}$ for the motion of particles in a potential field V.

(b) If the Lagrangian has the form

$$L(\mathbf{q}_i, \dot{\mathbf{q}}_i, t) = \frac{1}{2}\sum_{i=1}^{N} g_{ij}(q)\dot{q}_i\dot{q}_j,$$

where $(g_{ij}(q))$ is a positive definite matrix, then $g = (g_{ij})$ defines a Riemannian metric on Q and it turns out that the Euler–Lagrange equations coincide with the equations of geodesics with respect to this Riemannian metric (cf. Example 4.57).

4.7 The Legendre Transform

In this subsection Q is a smooth manifold which should be interpreted as the *configuration space* of a mechanical system. Its tangent bundle TQ is called the *velocity phase space*. Since the time development of our system is described by a curve $\gamma \colon I \to Q$ that solves a second order differential equation, it is uniquely determined by the element $\gamma'(t_0) \in TQ$. Therefore the elements of TQ describe the possible *states* of our system.

A *kinetic energy* is a function $T \colon TQ \to \mathbb{R}$ of the form $T(v) := \frac{1}{2}g_x(v, v)$ for $v \in T_x(Q)$, where g is a semi-Riemannian metric on Q. A *potential energy* is a function $U \in C^\infty(TQ)$ of the form $U = \pi^*V = V \circ \pi$, where $V \in C^\infty(Q, \mathbb{R})$ and $\pi \colon TQ \to Q$ is the canonical projection. In other words, $U(v_q) = V(q)$ depends only on the base point of $v_q \in T_q(Q)$. A kinetic and a potential energy are combined in the corresponding *Lagrange function*

$$L := T - U \in C^\infty(TQ).$$

In general, any smooth function on TQ is called a *Lagrange function*. The elements of the cotangent bundle T^*Q are called *momenta* of Q and T^*Q is called the *momentum phase space*.

Example 4.49. For a rigid body rotating freely about its center of mass, the configuration space is the Lie group $G = \mathrm{SO}_3(\mathbb{R})$ of rotations of $\mathbb{R}^3$. Accordingly, the velocity phase space is the tangent bundle $TG = T\,\mathrm{SO}_3(\mathbb{R})$.

The connection between velocities and momenta is established in terms of the fiber derivative of the Legendre function:

Definition 4.50. (Fiber derivative) For a smooth function $L\colon TQ \to \mathbb{R}$ we define its *fiber derivative*

$$\mathbb{F}L\colon TQ \to T^*Q, \quad (\mathbb{F}L)(v)(w) := T_v(L)w = \frac{d}{dt}\Big|_{t=0} L(v + tw).$$

Clearly, $\mathbb{F}L$ is a smooth function. It is also called the *Legendre transform*. We say that L is *hyperregular* if $\mathbb{F}L$ is a diffeomorphism.

The function

$$E\colon TQ \to \mathbb{R}, \quad E(v) := (\mathbb{F}L)_v(v) - L(v)$$

is called the corresponding *energy*.

Example 4.51. Suppose that $L = T - U$, where $T(v) = \frac{1}{2}g(v,v)$ is a kinetic energy and $U = \pi^*V$ is a potential energy. Then $\mathbb{F}U = 0$ because U is constant on the fibers T_qQ and $\mathbb{F}T = 2T$. Therefore

$$E = 2T - L = T + U$$

is the sum of the kinetic and the potential energy, i.e., the *total energy*.

Remark 4.52. (a) For $U \subseteq \mathbb{R}^n$ open and $TU \cong U \times \mathbb{R}^n$, the fiber derivative of a smooth function is given by the partial derivative

$$(\mathbb{F}L)(x,v)w = (x, \mathtt{d}L(x,v)(0,w)) = (x, \mathtt{d}_2L(x,v)(w)),$$

where $\mathtt{d}_2$ denotes the partial differential with respect to the second argument.

In coordinates (q_i, v_i) on TU and (q_i, p_i) on T^*U, it takes the form

$$(\mathbb{F}L)(q_i, v_i) = \Big(q_i, \frac{\partial L}{\partial v_i}\Big), \quad \text{i.e.,} \quad p_i = \frac{\partial L}{\partial v_i}$$

are the momentum variables.

(b) The fiber derivative $\mathbb{F}L$ is fiber preserving, i.e., it maps $T_q(Q)$ into $T_q^*(Q)$. Therefore L is hyperregular if and only if, for each $q \in Q$, the map $T_q(Q) \to T_q^*(Q), v \mapsto \mathbb{F}L(v)$ is a diffeomorphism (Exercise 4.3).

Example 4.53. (Movement of a point particle in $\mathbb{R}^n$) For a point particle of mass m in $\mathbb{R}^n$ moving with velocity v, the kinetic energy is given by

$$T(x,v) := \tfrac{1}{2}m\|v\|^2.$$

The fiber derivative of this function is given by

$$(\mathbb{F}T)(x,v)w = m\langle v, w\rangle = \langle mv, w\rangle,$$

so that we may identify $(\mathbb{F}T)(x,v) \in T^*_x(\mathbb{R}^n)$ with the momentum $p = mv$.

Definition 4.54. We define the *Lagrangian forms* on TQ by

$$\Theta_L = \mathbb{F}L^*\Theta \quad \text{and} \quad \Omega_L = \mathbb{F}L^*\Omega$$

(cf. Example 4.26). Then $\Omega_L = -\mathtt{d}\Theta_L$, so that Ω_L is an exact 2-form.

From the description of Θ and Ω in local coordinates

$$\Theta = \sum_i p_i \mathtt{d}q_i \quad \text{and} \quad \Omega = \sum_i \mathtt{d}q_i \wedge \mathtt{d}p_i,$$

we derive

$$\Theta_L = \sum_i \frac{\partial L}{\partial v_i}\mathtt{d}q_i \quad \text{and} \quad \Omega_L = \sum_{i,j} \frac{\partial^2 L}{\partial v_i \partial q_j}\mathtt{d}q_i \wedge \mathtt{d}q_j + \frac{\partial^2 L}{\partial v_i \partial v_j}\mathtt{d}q_i \wedge \mathtt{d}v_j.$$

Proposition 4.55. ([MR99, Thm. 7.3.3]) *Suppose that L is a hyperregular Lagrangian on TQ. Then the 2-form $\Omega_L := (\mathbb{F}L)^*\Omega$ obtained from the Liouville 2-form on T^*Q is symplectic. Let $X_E \in \mathcal{V}(TQ)$ denote the Hamiltonian vector field corresponding to the energy function*

$$E(v) := (\mathbb{F}L)_v(v) - L(v).$$

Then X_E is a second order vector field on TQ and its integral curves are of the form $\beta = \gamma'$, where $\gamma := \pi \circ \beta$ is a solution of the Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial v_i} = \frac{\partial L}{\partial q_i}, \quad i = 1, \ldots, n.$$

The vector field X_E on TQ is called the *Lagrangian vector field* corresponding to L.

Proof. It clearly suffices to verify all that in local coordinates, so that we may assume that Q is an open subset of $\mathbb{R}^n$. We use coordinates $(q,v) = (q_i, v_i)$ on TQ. Then

$$E(q,v) = (\mathtt{d}_2 L)_{(q,v)}(v) - L(v) = \sum_j \frac{\partial L}{\partial v_j} v_j - L$$

leads to

$$\mathtt{d}E = \sum_{i,j} \frac{\partial^2 L}{\partial v_j \partial v_i} v_j \,\mathtt{d}v_i + \sum_{i,j} \frac{\partial^2 L}{\partial v_j \partial q_i} v_j \,\mathtt{d}q_i - \sum_i \frac{\partial L}{\partial q_i}\mathtt{d}q_i.$$

We also recall from Definition 4.54 that

$$\Omega_L = \sum_{i,j} \frac{\partial^2 L}{\partial v_i \partial q_j} \mathtt{d}q_i \wedge \mathtt{d}q_j + \frac{\partial^2 L}{\partial v_i \partial v_j} \mathtt{d}q_i \wedge \mathtt{d}v_j.$$

Since we want a second order vector field, we start with the ansatz

$$X_E(q,v) = (v,b) = (v_1,\ldots,v_n,b_1,\ldots,b_n).$$

This leads

$$i_{X_E}\Omega_L = -\sum_{i,j} \frac{\partial^2 L}{\partial v_i \partial v_j} b_j \,\mathtt{d}q_i + \sum_{i,j} v_i \cdot \Big(\frac{\partial^2 L}{\partial v_i \partial v_j}\,\mathtt{d}v_j + \Big(\frac{\partial^2 L}{\partial v_i \partial q_j} - \frac{\partial^2 L}{\partial v_j \partial q_i}\Big)\,\mathtt{d}q_j \Big).$$

Comparing $i_{X_E}\Omega_L$ with $\mathtt{d}E$ now leads to the following equation for the b_j:

$$\sum_i \frac{\partial^2 L}{\partial v_i \partial q_j} v_i - \frac{\partial L}{\partial q_j} = -\sum_i \frac{\partial^2 L}{\partial v_i \partial v_j} b_i + \sum_i v_i \Big(\frac{\partial^2 L}{\partial v_i \partial q_j} - \frac{\partial^2 L}{\partial v_j \partial q_i}\Big).$$

This in turn simplifies to

$$\frac{\partial L}{\partial q_j} - \sum_i \frac{\partial^2 L}{\partial v_j \partial q_i} v_i = \sum_i \frac{\partial^2 L}{\partial v_i \partial v_j} b_i.$$

Since X_E is a second order vector field, its integral curves are of the form $\beta(t) = (\gamma(t),\dot\gamma(t))$, and $\ddot\gamma(t) = b$ implies with $\dot\beta(t) = X_E(\beta(t))$

$$\frac{\partial L}{\partial q_j} = \sum_i \frac{\partial^2 L}{\partial v_i \partial v_j} \ddot\gamma_i + \sum_i \frac{\partial^2 L}{\partial v_j \partial q_i} \dot\gamma_i = \frac{d}{dt}\frac{\partial L}{\partial v_j}.$$

We conclude that β is an integral curve of X_E if and only if γ satisfies the Euler–Lagrange equations. □

Corollary 4.56. *The energy E is constant along the solutions of the Euler–Lagrange equations.*

Proof. If $\gamma\colon I \to Q$ is a solution of the Euler–Lagrange equations, then γ' is an integral curve of X_E, so that the assertion follows from

$$\dot E = X_E E = \{E,E\} = 0.$$ □

Example 4.57. (Geodesics and force free motion) Let (Q,g) be a semi-Riemannian manifold and consider the Lagrangian $L(v) := \frac{1}{2} g(v,v)$ on TQ. We claim that L is hyperregular and that the integral curves of the Lagrangian vector field X_E on TQ are the curves γ', where $\gamma\colon I \to Q$ is a geodesic. One can therefore interpret the geodesics as describing the motion of a mechanical system under the absence of external forces (forcefree motion).

From Example 4.51 we know that $E = L$ is the corresponding energy function. It suffices to verify the assertion in local coordinates, where the Lagrangian has the form

$$L(q,v) = \frac{1}{2} \sum_{j,k} g_{jk} v_j v_k.$$

This leads to

$$\frac{\partial L}{\partial q_i} = \frac{1}{2}\sum_{j,k} \frac{\partial g_{jk}}{\partial q_i} v_j v_k \quad \text{and} \quad \frac{\partial L}{\partial v_i} = \sum_j g_{ji} v_j.$$

As $\frac{\partial^2 L}{\partial v_i \partial v_j} = g_{ij}$ is an invertible matrix, L is hyperregular.

On the other hand, we obtain for $v_j = \dot\gamma_j$:

$$\frac{d}{dt}\frac{\partial L}{\partial v_i} = \frac{d}{dt}\sum_j g_{ji}\dot\gamma_j = \sum_j g_{ji}\ddot\gamma_j + \sum_{j,k}\frac{\partial g_{ji}}{\partial q_k}\dot\gamma_k\dot\gamma_j = \sum_j g_{ji}\ddot\gamma_j + \frac{1}{2}\sum_{j,k}\Big(\frac{\partial g_{ji}}{\partial q_k} + \frac{\partial g_{ki}}{\partial q_j}\Big)\dot\gamma_k\dot\gamma_j,$$

so that the Euler–Lagrange equations turn into

$$\sum_j g_{ji}\ddot\gamma_j = \frac{1}{2}\sum_{j,k}\Big(\frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ji}}{\partial q_k} - \frac{\partial g_{ki}}{\partial q_j}\Big)\dot\gamma_k\dot\gamma_j,$$

and thus

$$\ddot\gamma_\ell = \frac{1}{2}\sum_{j,k,\ell} g^{\ell i}\Big(\frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ji}}{\partial q_k} - \frac{\partial g_{ki}}{\partial q_j}\Big)\dot\gamma_k\dot\gamma_j.$$

Comparing with the Koszul formula in local coordinats (Remark 3.29), we see that this is the ODE for the geodesics on Q.

4.8 Exercises for Section 4

Exercise 4.1. Let M be a smooth manifold of dimension $\dim M > 0$. Show that not every smooth curve in M is a solution of a second order differential equation.

Exercise 4.2. Let G be a Lie group. Show that:

(i) There exists a unique connection ∇ such that $\nabla_X Y = 0$ holds for left invariant vector fields. Compute its torsion $T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]$.

(ii) There exists a unique connection ∇ such that $\nabla_X Y = \frac{1}{2}[X,Y]$ holds for left invariant vector fields. Compute its torsion.

(iii) There exists a unique connection ∇ such that $\nabla_X Y = [X,Y]$ holds for left invariant vector fields. Show that $\nabla_X Y = 0$ for right invariant vector fields X, Y on G.

Exercise 4.3. Let X, Y and Z be smooth manifolds and $F\colon X \times Y \to X \times Z$ a smooth map of the form $F(x,y) = (x, G(x,y))$. Show that F is a diffeomorphism if and only if all the maps $G_x\colon Y \to Z, y \mapsto G(x,y)$ are diffeomorphisms. Hint: At some point one should use the Inverse Function Theorem.

Exercise 4.4. Let $X \in \mathcal{V}(M)$ and $\omega \in \Omega^k(M)$. Show that the relation $\mathcal{L}_X\omega = 0$ is equivalent to the invariance of ω under the local flow Φ_t^X generated by X in the sense that

$$(\Phi_t^X)^*\omega = \omega|_{\mathcal{D}_t},$$

where $\mathcal{D}_t \subseteq M$ is the domain of Φ_t^X.

Exercise 4.5. Let $X, Y \in \mathcal{V}(M)$. Show that the associated operators $\mathcal{L}_X$ on $\Omega(M)$ satisfy

$$\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_X\mathcal{L}_Y - \mathcal{L}_Y\mathcal{L}_X.$$

Proceed along the following steps:

(i) $D := [\mathcal{L}_X, \mathcal{L}_Y]$ is a derivation of the algebra $\Omega(M)$, i.e.,

$$D(\alpha \wedge \beta) = D\alpha \wedge \beta + \alpha \wedge D\beta, \quad \alpha, \beta \in \Omega(M).$$

(ii) D commutes with the exterior differential $\mathtt{d}$.

(iii) Verify $D = \mathcal{L}_{[X,Y]}$ on smooth functions f and their differentials $\mathtt{d}f$.

(iv) Now verify the assertion for forms of the type $\omega = f\mathtt{d}\varphi_1 \wedge \cdots \wedge \mathtt{d}\varphi_k$ and argue that this proves the general assertion.

Exercise 4.6. Let $\Theta \in \Omega^1(T^*Q)$ be the canonical 1-form. We consider a 1-form $\alpha \in \Omega^1(Q)$ as a smooth map $\alpha\colon M \to T^*M$. Show that

$$\alpha^*\Theta = \alpha \quad \text{for any} \quad \alpha \in \Omega^1(Q).$$

With a little extra work one can even show that this property determines $\Theta \in \Omega^1(T^*Q)$ uniquely.

Exercise 4.7. A submanifold L of a symplectic manifold (M, Ω) is said to be *Lagrangian* if $\dim M = 2\dim L$ and $\Omega|_L = 0$.

Show that the zero section $\zeta\colon Q \to T^*Q$ is a Lagrangian submanifold.

References

[Br93] Bredon, G. E., "Topology and Geometry," Graduate Texts in Mathematics **139**, Springer-Verlag, Berlin, 1993

[DK00] Duistermaat, J. J., and J. A. C. Kolk, "Lie groups". Springer, 2000

[GW09] Goodman, R., and N. R. Wallach, "Symmetry, Representations, and Invariants," Springer-Verlag, New York, 2009.

[GHV73] Greub, W., Halperin, S., and R. Vanstone, "Connections, Curvature, and Cohomology," Academic Press, New York and London, 1973

[GS84] Guillemin, V., and S. Sternberg, "Symplectic Techniques in Physics," Cambridge University Press, 1984

[He66] Hermann, R., "Lie groups for physicists", Benjamin, Inc., New York, 1966

[He74] Hermann, R., "Physical aspects of Lie group theorey," Les presses de l'université Montréal, 1974

[HN11] Hilgert, J., and K.-H. Neeb, "Structure and Geometry of Lie Groups," Springer, 2011

[KS11] Kosmann-Schwarzbach, Y., "The Noether Theorems," Springer, 2011

[La99] Lang, S., "Fundamentals of Differential Geometry", Grad. Texts Math. **191**, Springer-Verlag, Berlin, 1999

[Ma78] Mackey, G. W., "Unitary group representations in Physics, Probability, and number theory", Addison Wesley, 1978

[MaMa92] Marathe, K. B., and G.Martucci, "The Mathematical Foundations of Gauge Theories," Studies in Math. Physics **5**, North Holland, 1992

[MR99] Marsden, J. E., and T. S. Ratiu, "Introduction to Mechanics and Symmetry," Springer, 1999

[Na00] G. L. Naber, "Topology, Geometry, and Gauge Fields–Interactions," Appl. Math. Sciences **141**, Springer, 2000

[Sch95] Schottenloher, M., "Geometrie und Symmetrie in der Physik," Vieweg Lehrbuch Mathematische Physik, Vieweg Verlag, 1995

[Sch12] Schweigert, Chr., "Mathematical Structures in Physics," Lectures Notes Univ. Hamburg, Winter term 2011/12

[St94] Sternberg, S., "Group Theory and Physics," Cambdrige Univ. Press, 1994

[Va85] Varadarajan, V. S., "Geometry of Quantum Theory," 2nd ed., Springer, 1985

[Wa07] Waldmann, S., "Poisson-Geometrie und Deformationsquantisierung," Springer, 2007

[Wa83] F. Warner, "Foundations of differentiable manifolds and Lie groups," Springer, Grad. Texts in Math. 94, 1983